

С.Н. Павликов
Е.И. Убанкин
Ю.А. Левашов

ОБЩАЯ ТЕОРИЯ СВЯЗИ

С.Н. Павликов, Е.И. Убанкин, Ю.А. Левашов

ОБЩАЯ ТЕОРИЯ СВЯЗИ



Министерство образования и науки Российской Федерации

Владивостокский государственный университет
экономики и сервиса

С.Н. Павликов

Е.И. Убанкин

Ю.А. Левашов

ОБЩАЯ ТЕОРИЯ СВЯЗИ

Учебное пособие

Владивосток
Издательство ВГУЭС
2016

УДК 621.391(075.8)

ББК 32.841

П12

Рецензенты: *В.А. Игнатюк*, д-р физ.-мат. наук, профессор кафедры информационных технологий и систем ВГУЭС
С.В. Глушков, д-р техн. наук, профессор, декан электромеханического факультета МГУ им. адм. Г.И. Невельского

Павликов, С.Н.

П12 **Общая теория связи** [Текст] : учебное пособие / С.Н. Павликов, Е.И. Убанкин, Ю.А. Левашов ; Владивостокский государственный университет экономики и сервиса. – Владивосток : Изд-во ВГУЭС, 2016. – 284 с.

ISBN 978-5-9736-0400-4

Рассматриваются основные сведения о сообщениях, сигналах и помехах, их математические модели; методы формирования и преобразования; алгоритмы цифровой обработки сигналов; каналы электросвязи; теория передачи и кодирования сообщений; помехоустойчивость; многоканальная связь и распределение информации; эффективность систем связи; теоретико-информационная концепция криптозащиты сообщений в телекоммуникационных системах.

Для студентов, обучающихся по направлению подготовки 11.00.00 «Электроника, радиотехника и системы связи».

Табл. 13. Ил. 145. Библиогр. 21 назв.

УДК 621.391(075.8)

ББК 32.841

ISBN 978-5-9736-0400-4

© Владивостокский государственный университет экономики и сервиса, оформление, 2016

© Павликов С.Н., Убанкин Е.И., Левашов Ю.А., текст, 2016

СПИСОК СОКРАЩЕНИЙ

АКФ	автокорреляционная функция
АИМ	амплитудно-импульсная модуляция
АМн	амплитудная манипуляция
АЦП	аналого-цифровое преобразование
АЧХ	амплитудно-частотная характеристика
БГШ	белый гауссовский шум
БИХ	бесконечная импульсная характеристика
БПФ	быстрое преобразование Фурье
БЧХ	Боуза-Чоудхури-Хоквингема
ВАХ	вольтамперная характеристика
ВИМ	время-импульсная модуляция
ВКФ	взаимно-корреляционная функция
ВРК	временное разделение каналов
ГП	генератор поднесущих
ДАМ	дискретная амплитудная модуляция (манипуляция)
ДК	дискретный канал
ДМ	дельта-модуляция
ДНК	дискретно-непрерывный канал
ДПФ	дискретное преобразование Фурье
ДСКБП	дискретный симметричный канал без памяти
ДФМ	дискретная фазовая модуляция (манипуляция)
ДЧМ	дискретная частотная модуляция (манипуляция)
ИКМ	импульсно-кодовая модуляция
ИНЧ	инфранизкие частоты
КИХ	конечная импульсная характеристика
КРК	кодовое разделение каналов
КФ	корреляционная функция
КЧХ	комплексная частотная характеристика
ЛИС	линейная инвариантная к сдвигу
МДВР	многостанционный доступ с временным разделением каналов
МДКР	многостанционный доступ с кодовым разделением каналов
МДЧР	многостанционный доступ с частотным разделением каналов
МСД	многостанционный доступ
МСИ	межсимвольная интерференция
МЧС	многочастотный сигнал

НК	непрерывный канал
НОД	наибольший общий делитель
НОК	наименьшее общее кратное
ОДИМ	односторонняя длительно-импульсная модуляция
ОНЧ	очень низкие частоты
ОСШ	отношение сигнал/шум
ОФМ _н	относительная фазовая манипуляция
ПАКФ	периодическая автокорреляционная функция
ПВКФ	периодическая взаимно-корреляционная функция
ПЛИС	программируемая логическая интегральная схема
ППРЧ	псевдослучайное переключение рабочих частот
ПСП	псевдослучайная последовательность
ПЦХ	плезиохронная цифровая иерархия
ПЭС	первичный электрический сигнал
РС (RS)	Рида-Соломона
СКК	сигнально-кодовые конструкции
СКО	среднеквадратическое отклонение
СПИ	система передачи информации
СПМ	спектральная плотность мощности
СФ	согласованный фильтр
СЦИ	синхронная цифровая иерархия
СЭС	система электрической связи
ТКС	телекоммуникационная система
УГ	управляемый генератор
УМ	угловая модуляция
УРК	устройство разделения каналов
УУК	устройство уплотнения каналов
ФАПЧ	фазовая автоподстройка частоты
ФВЧ	фильтр верхних частот
ФИМ	фазоимпульсная модуляция
ФМ _н	фазовая манипуляция
ФМ-ПСС	фазовая манипуляция псевдослучайными сигналами
ФМ-ШПС	манипуляции широкополосными сигналами
ФНЧ	фильтр нижних частот
ФПВ	функция плотности вероятностей
ФРВ	функция распределения вероятностей
ЦАМ	цифровая амплитудная модуляция (манипуляция)
ЦАП	цифро-аналоговое преобразование
ЦОС	цифровая обработка сигналов
ЦСП	цифровой сигнальный процессор
ЦФ	цифровой фильтр;
ЦФМ	цифровая фазовая модуляция (манипуляция)
ЧИМ	частотно-импульсная модуляция
ЧМ _н	частотная манипуляция
ЧРК	частотное разделение каналов
ШИМ	широкополосная импульсная модуляция
ЭВМ	электронно-вычислительная машина

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

ΔF	– ширина спектра сигнала;
F_{cp}	– средняя частота спектра сигнала;
F_{max}	– максимальная частота в спектре сигнала;
F_{min}	– минимальная частота в спектре сигнала;
$B = \Delta F \cdot \Delta T$	– база сигнала;
D_c	– динамический диапазон;
V_c	– объем сигнала;
$\omega = 2\pi f$	– угловая частота;
$f = 1/T$	– циклическая частота
$\varphi_0 = 2\pi\tau_w/T$	– начальная фаза
$Q = T/\tau_{\text{н}}$	– скважность периодической последовательности импульсов
$\sigma(t)$	– функция Хевисайда (ступенчатая функция);
$M = \{s_1(t), s_2(t), \dots\}$	– множество сигналов;
$\ s\ = \sqrt{\int_{-\infty}^{+\infty} s^2(t) dt}$	– норма вещественного аналогового сигнала;
$\ s\ = \sqrt{\int_{-\infty}^{+\infty} s(t)s^*(t) dt}$	– норма комплексного сигнала;
$E_s = \ s\ ^2 = \int_{-\infty}^{+\infty} s^2(t) dt$	– квадрат нормы или энергия сигнала;
$C_k = \frac{1}{T} \int_0^T s(t) e^{-jk\Omega t} dt$	– коэффициенты разложения функции в комплексный ряд Фурье;

$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(j\Omega) \exp(j\Omega t) d\Omega$	– обратное преобразование Фурье;
$G(j\Omega) = \int_{-\infty}^{\infty} s(t) \exp(-j\Omega t) dt$	– прямое преобразование Фурье (спектральная плотность сигнала);
$\delta(t)$	– дельта-функция;
ω_d	– частота дискретизации;
$F_n(x_1 t_1, \dots, x_n t_n)$	– n -мерная функция распределения вероятностей;
$W(x_1, t_1)$	– одномерная функция плотности вероятностей случайного процесса;
m_1	– математическое ожидание или первый начальный момент;
m_2	– полная средняя мощность случайного процесса на единичном сопротивлении (второй начальный момент);
σ^2	– дисперсия (второй центральный момент);
σ	– среднеквадратическое отклонение;
$\Phi_0(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{y^2}{2}\right] dy$	– интеграл вероятности;
$s_a(t) = s(t) + js^*(t)$	– аналитический сигнал;
$s^*(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau$	– преобразование Гильберта;
$s(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s^*(\tau)}{t - \tau} d\tau$	– обратное преобразование Гильберта;
$A(t) = \sqrt{s^2(t) + s^{*2}(t)}$	– огибающая сигнала;
$\varphi(t) = \text{arctg} \frac{s^*(t)}{s(t)}$	– мгновенная фаза сигнала;
$B_s(\tau)$	– автокорреляционная функция;
τ_k	– интервал корреляции;
$m_{AM}, m_{FM}, m_{ЧМ}$	– индексы амплитудной, фазовой и частотной модуляции;
$J_k(m)$	– функции Бесселя;
$\Delta\omega_m$	– девиация частоты;
$\Delta\varphi_m$	– девиация фазы;

$s[n]$	– дискретный сигнал;
$h[n, k]$	– отклик цепи в момент времени n на δ – последовательность, имеющую единичное значение в момент времени k ;
$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h[n]e^{-j\omega n}$	– комплексная частотная характеристика линейной цепи инвариантной к стволу;
$\{s_i\}, i = \overline{1, L_S}$	– множество входных символов (сигналов) дискретного канала;
$\{y_j\}, j = \overline{1, L_Y}$	– множество выходных символов (сигналов) дискретного канала;
$I(Y, X)$	– количество информации, переданной по каналу;
$H(Y/X)$	– энтропия шума (ложная информация, создаваемая помехами);
$H(Y)$	– энтропия на выходе канала;
$X = \{x_i\}$	– алфавит источника сообщений;
$Y = \{y_j\}$	– алфавит получателя сообщений;
$P_{ош}$	– вероятность ошибки;
C	– пропускная способность канала;
χ	– коэффициент избыточности источника;
$K_{сж}$	– коэффициент сжатия источника;
d	– кодовое (хэммингово) расстояние;
$S = (s_1, s_2, s_3)$	– синдром (трёхсимвольная последовательность);
$RS(n, m)$	– код Рида-Соломона;
$\zeta(t)$	– Аддитивная помеха;
$a(t)$	– сообщение;
$b(t)$	– первичный сигнал;
$b_u(t)$	– цифровой сигнал;
$\Lambda(y)$	– отношение правдоподобия;
h_2^2	– отношение пиковой мощности сигнальной составляющей к средней мощности шума на выходе цепи;
h_1^2	– отношение средних мощностей сигнала P_s и помехи P_ξ на входе фильтра;

$L(\tilde{\lambda} - \lambda)$	– функция потерь, зависящая от разности оценки и истинного значения параметра;
$r(\tilde{\lambda}, z)$	– апостериорный (условный) риск, описывает потери, связанные с получением оценки $\tilde{\lambda}$ при наблюдении реализации z ;
$R(\tilde{\lambda})$	– средний риск (усреднённый по всем реализациям апостериорный риск);
$\eta = R/C$	– коэффициент использования канала по пропускной способности (информационная эффективность), характеризует реальную скорость передачи информации R по отношению к пропускной способности C канала связи;
$\beta = R/(P_c/N_0)$	– коэффициент использования канала по мощности (энергетическая эффективность);
$\gamma = R/\Delta F$	– коэффициент использования канала по полосе частот (частотная эффективность);
g_s	– энергетический выигрыш.

ВВЕДЕНИЕ

Современные системы электрической связи постоянно и стремительно развиваются. Широкое применение находят мобильные и волоконно-оптические системы связи, характеристики которых постоянно совершенствуются, в том числе и за счет применения новых видов сигналов и методов их обработки. Помочь современному специалисту в области телекоммуникаций ориентироваться в этих вопросах может хорошее знание основ теории электрической связи.

Общая теория связи относится к числу фундаментальных дисциплин подготовки инженеров-связистов самого разного профиля, и имеет цель сформировать знания основ теорий передачи и кодирования сообщений, методов передачи и приема дискретных и непрерывных сообщений, цифровой обработки, принципов построения многоканальных систем передачи и методов повышения эффективности систем электросвязи, а также умений использовать методы анализа систем электрической связи для количественной оценки их эффективности.

Предметом изучения дисциплины являются закономерности процессов преобразования и передачи информации в системах электросвязи.

Знания и умения по дисциплине являются составной частью общепрофессиональной подготовки к самостоятельной инженерно-эксплуатационной деятельности.

В учебном пособии раскрываются дидактические единицы (логически самостоятельная часть учебного материала) в соответствии с государственным образовательным стандартом дисциплины «Теория электрической связи»: сообщения, сигналы и помехи, их математические модели; методы формирования и преобразования сигналов; алгоритмы цифровой обработки сигналов; каналы электросвязи; теория передачи и кодирования сообщений; помехоустойчивость; многоканальная связь и распределение информации; эффективность систем связи; теоретико-информационная концепция криптозащиты сообщений в телекоммуникационных системах.

Дисциплина базируется на предшествующем изучении физики, математики, дискретной математики, теории вероятностей, математической статистики и информатики. В свою очередь «Общая теория элек-

трической связи» является базовой для дисциплин «Формирование и передача сигналов», «Устройства приема и обработки радиосигналов», «Цифровая обработка сигналов», «Устройства преобразования и обработки информации в системах подвижной радиосвязи», «Основы теории систем связи с подвижными объектами», «Устройства и системы морской радиосвязи».

Пособие базируется на материалах лекционного курса читаемого авторами на протяжении ряда лет и рассчитано на студентов всех форм обучения.

I. СООБЩЕНИЯ, СИГНАЛЫ И ПОМЕХИ

1. ОБЩИЕ СВЕДЕНИЯ О СИСТЕМАХ ЭЛЕКТРИЧЕСКОЙ СВЯЗИ

1.1. Информация, сообщения, сигналы и помехи

Системы связи предназначены для передачи информации. Информация передается посредством сообщений. Таким образом, сообщение – форма представления информации.

Примерами сообщений могут служить текст телеграммы, фраза в телефонном разговоре, последовательность цифр при передаче данных, изображение в системе фототелеграфии, последовательность изображений (кадров) в системе телевидения и т.п. Сообщение представляет собой совокупность знаков (символов). Например, текст телеграммы состоит из букв, цифр, пробелов и специальных знаков, а готовое для передачи по каналу связи телеграфное сообщение – из канальных символов (например, из «точек», «тире» и пауз при использовании «азбуки Морзе»).

В системе черно-белого телевидения сообщением является последовательность кадров, каждый из которых, в свою очередь, представляет собой последовательность значений яркости, упорядоченных согласно схеме телевизионной развертки. В телефонии сообщение – непрерывная последовательность значений напряжения (тока), отображающая изменение во времени звукового давления на мембрану микрофона.

Из приведенных примеров становится ясно, что сообщения могут быть дискретными (состоящими из символов, принадлежащих конечному множеству – алфавиту) или непрерывными (континуальными, аналоговыми), описываемыми функциями непрерывного времени.

Для передачи сообщения необходим материальный носитель, называемый сигналом. Сигналом может быть свет костра, удар барабана, звук речи или свистка, предмет, находящийся в условленном месте, взмах флажка или шпаги и т.п.

В радиотехнике и электрической связи используются электрические сигналы, которые благодаря простоте их генерирования и преобразования наилучшим образом приспособлены для передачи больших объемов данных на большие расстояния. Заметим, что в современных каналах связи и устройствах хранения данных электрические сигналы зачастую преобразуются в оптические или магнитные, но, как правило, предполагается их обратное преобразование.

Естественной формой представления сигнала считается его описание некоторой функцией времени (зависимой переменной чаще всего является напряжение или ток).

В современных системах связи используются разнообразные сигналы с различными свойствами. Эти сигналы могут быть классифицированы, хотя любая классификация достаточна условна. На рис. 1.1. представлена классификация, в основу которой положен принцип математического описания сигналов, используемый для теоретического изучения и проведения расчетов.



Рис. 1.1. Классификация сигналов

Математическое описание и представление сигналов позволяет создать математическую модель сигнала.

Если математическая модель позволяет точно описать сигнал, то такой сигнал называется детерминированным. В случае невозможности точного описания сигнала в любые моменты времени, сигнал называется случайным.

Высокочастотное модулированное колебание называется радиосигналом.

Сигнал без высокочастотного заполнения является видеосигналом.

Если сигнал может быть описан функцией $s(t) = s(t + T)$, где T – период, он называется периодическим.

При невозможности такого представления, сигнал является непериодическим.

Сигнал, описывающий во времени непрерывно изменяющийся процесс, называется аналоговым. Сигнал конечной длительности является импульсным.

Иногда удобно передавать только значения непрерывного сигнала (отсчеты или выборки), взятые в отдельные моменты времени. Такой квантованный по времени сигнал называется дискретным. Если же передавать не сами выборки в виде коротких импульсов, а их числовые значения, то сначала необходимо эти значения получить. Эта процедура в технике связи называется квантованием по уровню. Таким образом, сигнал, квантованный по времени и уровню, называется цифровым.

Интересно отметить, что детерминированные сигналы не несут в себе никакой информации. Однако с их помощью возможно передавать информацию, если случайным будет расположение сигналов на временной оси. Например, телеграфный сигнал состоит из семи импульсов прямоугольной формы с заданными параметрами рис. 1.3г. Первый (стартовый) и последний (стоповый) импульсы обозначают начало и конец посылки. Информационное содержание посылки зависит от передаваемой в данный момент буквы алфавита и представляет собой соответствующую этой букве комбинацию токовых и безтоковых посылок.

На рис. 1.2. представлена другая возможная классификация сигналов.

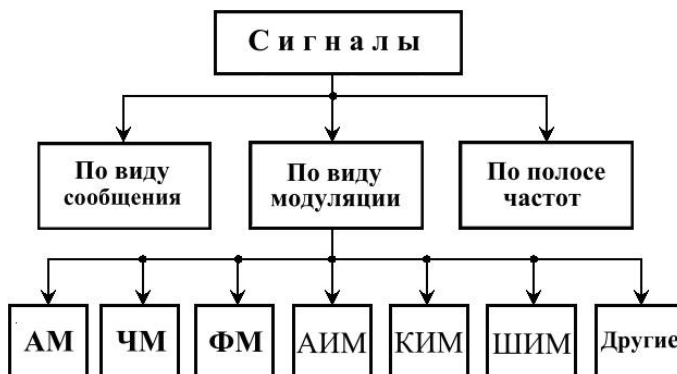


Рис. 1.2. Классификация сигналов

По виду передаваемых сообщений сигналы, например, можно разделить на радиовещательные, телевизионные, телеграфные и т.д.

По полосе частот сигналы обычно подразделяются на узкополосные и широкополосные.

Для широкополосных сигналов $\Delta F/F_{\text{cp}} \gg 1$, где

$\Delta F = F_{\text{max}} - F_{\text{min}}$ – абсолютная ширина спектра сигнала,

$F_{\text{cp}} = (F_{\text{max}} + F_{\text{min}})/2$ – средняя частота спектра сигнала,

F_{max} – максимальная частота в спектре сигнала,

F_{min} – минимальная частота в спектре сигнала.

Для узкополосных сигналов $\Delta F/F_{\text{cp}} < 1$.

Сигналы так же делятся на сложные и простые в зависимости от величины базы сигнала B (произведение длительности сигнала на ширину полосы его спектра).

Для сложных сигналов $B > 1$,

где $\Delta F \cdot \Delta T$ – база сигнала, ΔF – абсолютная ширина спектра сигнала, ΔT – длительность сигнала.

Для простых сигналов $B = 1$.

По виду модуляции сигналы различаются по признаку того параметра, который изменяется по закону передаваемого сообщения. Так как любое гармоническое колебание характеризуется амплитудой, частотой и мгновенной фазой, то и радиосигналы бывают с амплитудной модуляцией (АМ), с частотной (ЧМ) и фазовой модуляцией (ФМ). В настоящее время в системах связи используется большое разнообразие сигналов со сложными видами модуляции, например, с амплитудно-импульсной модуляцией (АИМ), Кодово-импульсной модуляцией (КИМ), широтно-импульсной модуляцией (ШИМ). К настоящему времени разработан не один десяток сложных видов модуляции и, естественно, большое количество соответствующих сигналов с различными характеристиками.

На рис. 1.3 приведены осциллограммы различных, широко применяемых в системах связи, сигналов.

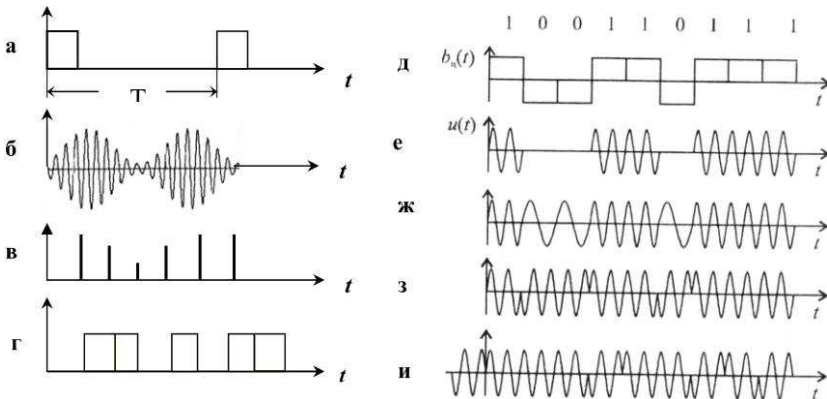


Рис. 1.3. Осциллограммы сигналов, применяемых в системах связи

На этом рисунке изображены следующие сигналы: *a* – периодический импульсный, *б* – непрерывный (аналоговый) радиосигнал с АМ, *в* – дискретный, *г* – случайный, *д* – цифровой кодированный, *е* – с амплитудной манипуляцией, *ж* – с частотной манипуляцией, *з* – с фазовой манипуляцией, *и* – с фазовой манипуляцией.

Необходимо также отметить, что жесткую классификацию к реальным сигналам применить невозможно. Например, сигнал (рис. 1.3а) можно классифицировать как детерминированный периодический импульсный видеосигнал, а сигнал (рис. 1.3з) как случайный радиосигнал с фазовой манипуляцией.

Кроме перечисленных, используются и другие критерии классификации сигналов, например, иногда различают информационные и управляющие сигналы (колебания) и т.д. Некоторые из перечисленных типов сигналов будут в дальнейшем рассмотрены подробнее.

В теории электрической связи принято рассматривать сигнал как «объект транспортировки». С этой точки зрения сигнал можно описать тремя «габаритными характеристиками», подобными длине, ширине и высоте груза, перевозимого, скажем, по железной дороге. Первая из таких характеристик – длительность сигнала T_c , измеряемая в секундах (с). Любой сигнал можно представить суммой (суперпозицией) гармонических колебаний с определенными частотами, поэтому вторая «габаритная характеристика» – ширина спектра, или полоса частот сигнала ΔF_c , равная разности наивысшей и низшей частот его гармонических составляющих и измеряемая в герцах (Гц). Третьей «габаритной» характеристикой служит динамический диапазон, измеряемый в децибелах (дБ) и определяемый формулой

$$D_c = 20\lg(X_{\max} / X_{\min}),$$

где X_{\max} и X_{\min} – соответственно максимальное и минимальное возможные значения сигнала (напряжения или тока). Произведение этих трех величин называется объемом сигнала:

$$V_c = T_c \Delta F_c D_c$$

Полезные сигналы отличаются от мешающих тем, что полезные сигналы служат для передачи сообщений, в то время как мешающие являются причиной их искажения (потери информации).

Часто полезный сигнал называют просто сигналом, а мешающий – помехой. Сигналы и помехи, рассматриваемые в совокупности, будем называть колебаниями.

Помехи могут быть естественными и преднамеренными (искусственными), шумовыми (флюктуационными) и импульсными, активными и пассивными и т.д.

Необходимо отметить, что одно и то же колебание может быть полезным сигналом по отношению, например, к одной системе связи или радиолокации и помехой – по отношению к другой.

Стоит также отметить, что все помехи, как и все сигналы, являются случайными (если помеха детерминированна, то её можно исключить из наблюдаемого колебания и таким образом избавиться от её вредного воздействия на сообщение).

На рис. 1.4 приведены примеры случайного сигнала и случайной (шумовой) помехи.

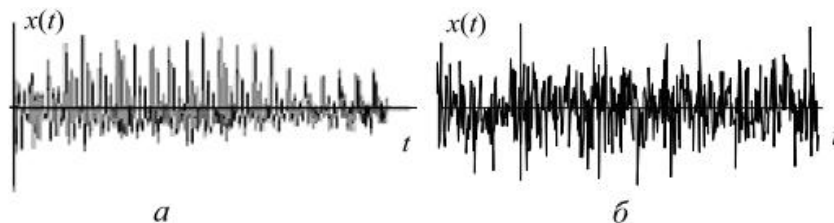


Рис. 1.4. Случайный (речевой) сигнал (а) и случайная помеха (шум) (б)

По способу взаимодействия с сигналом помехи подразделяются на аддитивные (от английского *add* – складывать), мультипликативные (от английского *multiply* – умножать) и смешанные (все взаимодействия, не сводимые к аддитивному или мультипликативному).

1.2. Общие принципы построения систем связи

Современная система связи представляет собой сложную совокупность устройств, выполняющих преобразования сообщений и сигналов с целью наиболее эффективной передачи информации.

К показателям эффективности относятся достоверность и скорость передачи информации, а также некоторые другие величины.

Сообщения – это совокупность сведений об окружающих нас предметах, явлениях. Сообщения могут быть звуковыми (речь, музыка), световыми (изображения неподвижных и подвижных объектов), текстовыми (буквенно-цифровые сообщения).

Обобщенная структурная схема системы связи (рис. 1.5) отражает наиболее типичные преобразования, которым подвергается сообщение в системе связи, она справедлива для любых видов сообщений. Рассмотрим назначение основных блоков системы связи.

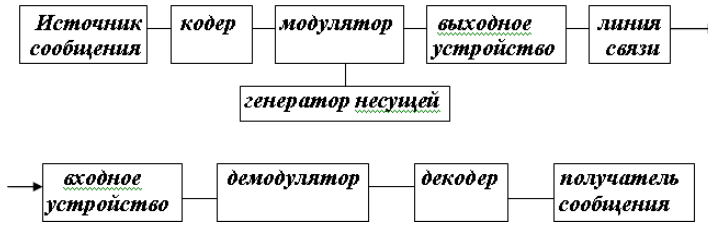


Рис. 1.5. Обобщенная структурная схема системы связи

Источник информации – источник сообщения подлежащего передаче (человек, окружающая среда и т.п.).

Сообщение – речь, музыка, текст, изображение, параметры некоторых объектов и т.п.

Кодер – а) преобразует неэлектрическое сообщение в электрический сигнал (сигнал – это электрическая копия сообщения).

б) преобразует аналоговый (непрерывный) сигнал в дискретный (цифровой);

в) осуществляет эффективное кодирование с целью уменьшения необходимой скорости передачи информации при заданном качестве (устранение избыточности сообщения);

г) осуществляет помехоустойчивое кодирование, позволяющее улучшить качество принимаемого сообщения.

Генератор несущей – генерирует колебания с постоянной амплитудой, частотой, фазой.

Модулятор – изменяет амплитуду, частоту или фазу переносчика в соответствии с модулирующим сигналом, поступающим от кодера.

Выходное устройство – усиливает сигнал, для обеспечения заданного качества связи и ограничивает спектр излучаемого сигнала до полосы частот, отведённой для заданной системы связи.

Кодер, модулятор, генератор несущей и выходное устройство образуют передатчик.

Линия связи – совокупность технических устройств (кабель, двухпроводная линия, оптическая линия связи) или эфир, по которым сигнал поступает от передатчика к приёмнику. Напряжение на входе приёмника можно записать как:

$$U_{\text{прм}}(t) = K(t)U_{\text{прд}} + x(t),$$

где $U_{\text{прм}}(t)$ – напряжение на входе приёмника;

$K(t)$ – мультипликативная помеха (это переменный коэффициент передачи линии связи);

$U_{\text{прд}}$ – напряжение на выходе передатчика;

$x(t)$ – аддитивная помеха (тепловой шум, помеха от соседних передатчиков, помехи от различных технических устройств и т.п.).

Входное устройство – выделяет сигнал своего передатчика, отфильтровывает (не пропускает) сигналы соседних по частоте передатчиков и часть помех, усиливает сигнал.

Демодулятор – преобразует ВЧ модулированный сигнал в НЧ модулирующий (сигнал на выходе демодулятора, примерно, соответствует тому, что было на входе модулятора).

Декодер – а) принимает решение по каждой посылке (1 или 0);

б) декодирует кодовые комбинации, исправляет часть ошибок;

г) преобразует кодовые комбинации в сообщения удобные для получателя.

Получатель сообщения – человек, компьютер или другие технические устройства.

Входное устройство, демодулятор и декодер образуют приемник.

КОДЕР + ДЕКОДЕР = КОДЕК

МОДУЛЯТОР + ДЕМОДУЛЯТОР = МОДЕМ

КОДЕР+МОДУЛЯТОР+ДЕКОДЕР+ДЕМОДУЛЯТОР=КОДЕМ

С кодированием не следует путать шифрование сообщений. Цель шифрования состоит в предотвращении несанкционированного извлечения или преднамеренного изменения информации. При шифровании производится замена открытого сообщения шифрограммой (шифр-текстом), а при расшифровании происходит обратное преобразование. Шифрование выполняется до преобразования сообщения в первичный сигнал или в кодовую последовательность.

Таким образом, для модуляции в зависимости от сложности системы применяется первичный сигнал или последовательность кодовых символов.

В качестве переносчика часто используют гармоническое колебание $A \cdot \cos(\omega t + \varphi)$, которое имеет три параметра: амплитуду A , круговую частоту $\omega = 2\pi f$ и начальную фазу φ . Поэтому возможны три вида модуляции гармонического переносчика аналоговым сигналом: амплитудная модуляция (АМ), частотная модуляция (ЧМ) либо фазовая модуляция (ФМ) (рис. 1.6).

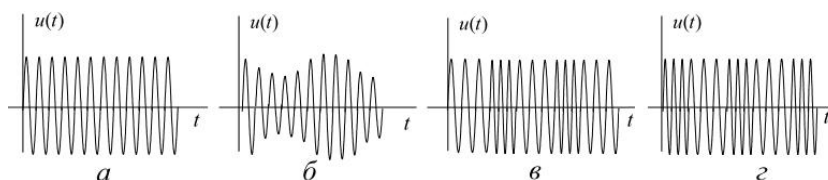


Рис. 1.6. Несущее гармоническое колебание (а) и получаемые на его основе модулированные сигналы: АМ (б), ЧМ (в) и ФМ (г)

Во многих случаях роль переносчика в системах связи играет периодическая последовательность импульсов одинаковой формы (часто импульсы считают в первом приближении прямоугольными).

При заданной форме импульсов последовательность характеризуется амплитудным (пиковым) значением, длительностью импульсов и периодом повторения. Поэтому при аналоговом первичном сигнале различают:

- амплитудно-импульсную модуляцию (АИМ), при которой по закону изменения первичного сигнала изменяется амплитуда импульсов;
- широтно-импульсную модуляцию (ШИМ), при которой изменяется длительность («ширина») импульсов;
- время-импульсную модуляцию (ВИМ), при которой изменяется время задержки импульсов относительно среднего положения;
- частотно-импульсную модуляцию (ЧИМ), когда в такт с первичным сигналом изменяется частота следования импульсов.

Широко применяют также модуляцию гармонического колебания квантованным (цифровым) первичным сигналом. Различают три вида дискретной (цифровой) модуляции (манипуляции): амплитудную (ДАМ, ЦАМ), частотную (ДЧМ, ЦЧМ) и фазовую (ДФМ, ЦФМ) (рис. 1.7). Участок манипулированного сигнала, в течение которого модулируемый параметр постоянен, называется элементарной посылкой, или просто посылкой.

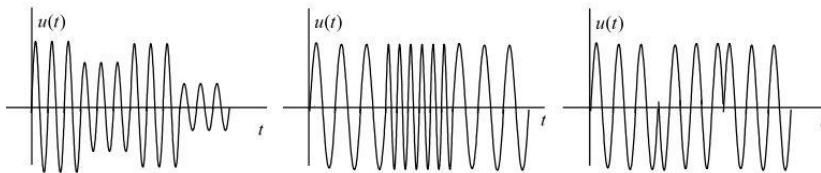


Рис. 1.7. Виды дискретной модуляции (манипуляции) гармонического колебания: ДАМ (а), ДЧМ (б), ДФМ (в)

Колебание при дискретной модуляции характеризуют технической скоростью (скоростью модуляции, скоростью телеграфирования), равной количеству элементарных посылок в секунду. Единицей измерения скорости модуляции является бод (1 бод соответствует одной посылке в секунду).

Демодуляция заключается в восстановлении первичного сигнала по принятому искаженному колебанию, а декодирование – в восстановлении дискретного сообщения по демодулированному сигналу.

Часто перед демодуляцией применяют дополнительное преобразование с целью повышения достоверности (уменьшения вероятности ошибки). Такое преобразование называют обработкой.

Оптимальной называется обработка, обеспечивающая наивысшую достоверность решения.

Если оптимальная обработка оказывается слишком сложной и/или дорогостоящей, применяют квазиоптимальную (субоптимальную) обработку, которая проще и дешевле и при этом обеспечивает достоверность, близкую к предельной. Часто квазиоптимальная обработка представляет собой фильтрацию принятого колебания с целью подавления помех.

1.3. Классификация систем связи

По виду передаваемых сообщений различают:

- 1) телеграфию (передача текста),
- 2) телефонию (передача речи),
- 3) фототелеграфию (передача неподвижных изображений),
- 4) телевидение (передача подвижных изображений),
- 5) телеметрию (передача результатов измерений),
- 6) телеуправление (передача управляющих команд),
- 7) передачу данных (в вычислительных системах и АСУ).

По диапазону частот – в соответствии с декадным делением диапазонов электромагнитных волн от мириаметровых ($3 \div 30$) кГц до децимиллиметровых ($300 \div 3000$) ГГц.

По назначению – вещательные (высококачественная передача речи, музыки, видео от малого числа источников сообщений большому количеству их получателей) и профессиональные (связные), в которых число источников и получателей сообщений одного порядка.

Различают следующие режимы работы СС:

- 1) симплексный (передача сигналов в одном направлении),
- 2) дуплексный (одновременная передача сигналов в прямом и обратном направлениях),
- 3) полудуплексный (поочередная передача сигналов в прямом и обратном направлениях).

Каналом связи называется комплекс радиотехнических устройств, при помощи которых передается и принимается информация, плюс среда между ними. В зависимости от вида сигналов на входе и выходе различают каналы: непрерывные; дискретные; дискретно-непрерывные; непрерывно-дискретные.

Каналы связи можно характеризовать по аналогии с сигналами следующими тремя параметрами:

- временем доступа T_k ,
- шириной полосы пропускания ΔF_k ,
- динамическим диапазоном $D_k = 10 \cdot \lg \frac{P_{k, \text{доп.}}}{P_{\text{ш}}}$ [дБ],

где $P_{k, \text{доп.}}$ – максимально допустимая мощность сигнала в канале,

$P_{\text{ш}}$ – мощность собственных шумов канала.

Обобщенным параметром канала является его емкость

$$V_k = T_k \cdot \Delta F_k \cdot D_k.$$

Очевидным необходимым условием согласования сигнала и канала является выполнение неравенства $V_c < V_k$.

2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СИГНАЛОВ

2.1. Математическое описание сигнала

Математическое описание и представление сигналов позволяет создать математическую модель сигнала.

Математическим описанием детерминированного сигнала служит детерминированная функция времени $S(t)$. Это означает, что любому моменту времени t_i соответствует определенное значение функции $S(t_i)$.

Математическое описание случайных (или нерегулярных) сигналов осуществляется с помощью случайных функций. Случайными сигналами являются, например, напряжения или токи, соответствующие речи, музыке, последовательности телеграфных знаков и т.п. Для случайной функции ее значение при фиксированном аргументе t_i – случайная величина.

В электросвязи находят применение гармонические и импульсные сигналы. Гармонический сигнал (рис. 2.1), записывается в виде:

$$S(t) = A_0 \cos(\omega t + \varphi_0), \quad (2.1)$$

где A_0 – максимальное значение (амплитуда); $\omega = 2\pi f$ – угловая частота; $f = 1/T$ – циклическая частота; $\varphi_0 = 2\pi\tau_n/T$ – начальная фаза.

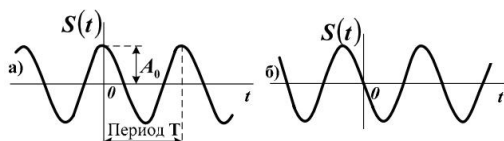


Рис. 2.1. Гармонический сигнал

Для представленных на рис. 2.1. гармонических сигналов значения начальной фазы принимают значения: $\varphi_0 = 0$ (рис. 2.1а); $\varphi_0 = 90^\circ$ (рис. 2.1б).

Импульсными являются сигналы, отличные от нуля в течение ограниченного времени. Эти сигналы существуют лишь в пределах конечного отрезка (t_1, t_2) . При этом различают видеоимпульсы (рис. 2.2а) и радиоимпульсы (рис. 2.2б). Если $s_B(t)$ – видеоимпульс, то соответствующий ему радиоимпульс описывается выражением: $S_P(t) = s_B(t) \cos(\omega t + \varphi_0)$ (частота ω и начальная фаза φ_0 могут быть произвольными). В радиоимпульсе $s_B(t)$ называется огибающей, а функция $\cos(\omega t + \varphi_0)$ – заполнением. Параметрами видеоимпульса принято считать его амплитуду A_0 , длительность τ_n , длительность фронта t_f , длительность спада t_c . Происхождение термина «видеоимпульс» связано с тем, что впервые такие импульсы начали применять для описания сигналов в телевидении.

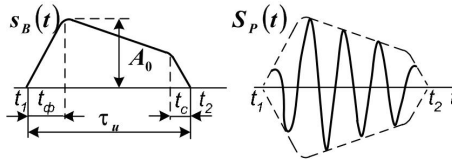


Рис. 2.2. Импульсные сигналы: а) видеоимпульс; б) радиоимпульс

В электросвязи наибольшее применение находят одиночные импульсы или периодические последовательности импульсов, форма которых приближается к прямоугольной. Для периодической последовательности импульсов, вводится понятие скважности, определяемой как отношение периода к длительности импульса: $Q = T / \tau_{и}$, T – период (временной интервал), так что $S(t_i + kT) = S(t_i)$, $k = 0, \pm 1, \pm 2, \dots$

2.2. Математическое представление сигналов

Многие задачи теории связи и радиотехники, например вычисление отклика физической системы на известное входное воздействие, требуют специфической формы представления сигналов.

Способ получения таких моделей сигналов состоит в следующем. Реальный сигнал приближенно представляется суммой некоторых элементарных сигналов, возникающих в последовательные моменты времени. Если теперь устремить к нулю длительность отдельных элементарных сигналов, то, естественно, в пределе будет получено точное представление исходного сигнала. В литературе этот способ описания сигнала получил название динамического представления, подчеркивающее развивающийся во времени процесс.

Широкое применение нашли два способа динамического представления. Согласно первому из них в качестве элементарных сигналов используются ступенчатые функции, возникающие через равные промежутки времени Δ .

В качестве таких функций используются функции включения (рис. 2.3.) или функции Хевисайда $\sigma(t)$, которые описываются следующим образом

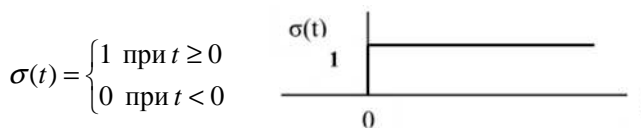


Рис. 2.3. Функция включения

Другая возможность представления сигнала заключается в использовании стандартных прямоугольных функций длительностью Δ . На рис. 2.4 показаны возможные способы представления сигналов.

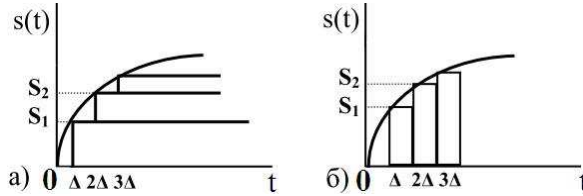


Рис. 2.4. Динамическое представление сигналов

Как видно (рис. 2.4а), текущее значение сигнала при любом t равно сумме ступенчатых функций

$$s(t) \approx s_0 \sigma(t) + (s_1 - s_0) \sigma(t - \Delta) + (s_2 - s_1) \sigma(t - 2\Delta) + \dots \quad (2.2)$$

В случае представления аналогового сигнала суммой примыкающих к друг другу прямоугольных импульсов, элементарный импульс с номером k представляется в виде

$$u_k(t) = s_k [\sigma(t - t_k) - \sigma(t - t_k - \Delta)]. \quad (2.3)$$

Тогда исходный сигнал является суммой элементарных импульсов

$$s(t) = \sum_{k=-\infty}^{+\infty} u_k(t). \quad (2.4)$$

Важное значение при динамическом представлении сигнала играет и другая функция, которая называется дельта-функцией $\delta(t)$ или функцией Дирака. Такой функцией называется импульсный сигнал, площадь которого, например, $A_m \cdot \tau$ равна 1, причем длительность импульса τ стремится к нулю, а амплитуда импульса A_m стремится к бесконечности.

Если в выражении (2.4) Δ устремить к нулю, то получим формулу динамического представления сигнала

$$s(t) = \int_{-\infty}^{+\infty} s(\tau) \delta(t - \tau) d\tau. \quad (2.5)$$

Таким образом, если непрерывную функцию умножить на дельта-функцию и произведение проинтегрировать по времени, то результат будет равен значению функции в той точке, где существует δ -функция. В этом заключается фильтрующее свойство дельта-функции.

2.3. Геометрическое представление сигналов

Идеи функционального анализа дали возможность создать теорию сигналов, в основе которой лежит представление сигнала как вектора в некотором бесконечномерном пространстве.

Если имеется некая совокупность сигналов $s_1(t)$, $s_2(t)$ и т.д., имеющих некоторые общие свойства, то их можно объединить в некоторое множество сигналов $M = \{s_1(t), s_2(t), \dots\}$.

Задача разложения сигнала сложной формы на простейшие составляющие сходна с разложением обычного вектора x трехмерного пространства на его составляющие по координатному базису единичных ортогональных векторов i, j, k . Такое представление можно записать как

$$x = x_1 i + x_2 j + x_3 k \quad (2.6)$$

Составляющими вектора x по базису (i, j, k) будут векторы $x_1 \cdot i$, $x_2 \cdot j$, $x_3 \cdot k$. Коэффициенты x_1, x_2, x_3 представляют собой проекции вектора x на координатные оси i, j, k и называются координатами вектора x . Иначе говоря, вектор x в трехмерном пространстве полностью определяется совокупностью его координат $x = (x_1, x_2, x_3)$.

Чтобы перейти к обобщению понятия вектора трехмерного пространства для случая n -мерного пространства, функцию $x(t)$ по аналогии с (2.6) можно представить в виде суммы

$$x(t) = \sum_{i=0}^{n-1} x_i \cdot \psi_i(t), \quad (2.7)$$

где ψ_i – элементарные базисные функции.

Множество векторов $\{\psi_i\}$ называется линейно независимым (базисом), если условие $\sum_{i=0}^{n-1} x_i \cdot \psi_i(t) = 0$ выполняется лишь тогда, когда все $x_i = 0$.

Линейно независимые векторы $\{\psi_i\}$ можно рассматривать как координатные оси пространства.

Метрическим называется линейное пространство, в котором определено расстояние между элементами (векторами) пространства (метрика), т.е. каждой паре элементов, скажем, x и y может быть поставлено в соответствие некоторое вещественное неотрицательное число $d(x, y)$ и способ, в соответствии с которым находится это число.

Среди линейных метрических пространств важное место занимают нормированные пространства.

Для этого вводится новое понятие, соответствующее длине вектора. В математике длину вектора называют его нормой. Пространство сигналов называется нормированным, если каждому вектору $s(t)$ одно-

значно сопоставлено число $\|s\|$, называемое нормой. Для вещественных аналоговых сигналов в теории сигналов норму сигнала вводят в виде

$$\|s\| = \sqrt{\int_{-\infty}^{+\infty} s^2(t) dt} \quad (2.8)$$

Для комплексных сигналов норма сигнала представляется

$$\|s\| = \sqrt{\int_{-\infty}^{+\infty} s(t) s^*(t) dt} \quad (2.9)$$

Квадрат нормы называется энергией сигнала E_s

$$E_s = \|s\|^2 = \int_{-\infty}^{+\infty} s^2(t) dt \quad (2.10)$$

Такая энергия сигнала выделяется на резисторе с сопротивлением 1 Ом. Выражение (2.10) представляется очень удобным, так как отпадает необходимость расшифровывать размерность сигнала, т.е. сигнал задан в виде тока или напряжения.

2.4. Представление сигналов в виде рядов ортогональных функций

При передаче сообщений одновременно существует большое количество разнообразных сигналов. Допустим, что имеются два сигнала s_i и s_j и определим энергию суммарного сигнала

$$E = \int_{-\infty}^{+\infty} (s_i + s_j)^2 dt = E_i + E_j + 2 \int_{-\infty}^{+\infty} s_i \cdot s_j dy.$$

Видно, что в отличие от самих сигналов, их энергии неаддитивны. Энергия суммарного сигнала содержит так называемую взаимную энергию, которая определяется как скалярное произведение двух вещественных сигналов

$$E_{ij} = 2 \int_{-\infty}^{+\infty} s_i \cdot s_j dy. \quad (2.11)$$

Если взаимная энергия сигналов s_i и s_j равна нулю, то они называются ортогональными.

Для исследования различных свойств сообщений, сигналов и помех удобно использовать разложение этих процессов в ряды.

Любой процесс (с некоторыми математическими ограничениями) можно представить в виде ряда:

$$x(t) = \sum_{k=-\infty}^{\infty} C_k \varphi_k(t). \quad (2.12)$$

где $\varphi_k(t)$ – ортогональные функции, т.е.:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \varphi_k(t) \varphi_n(t) dt = \begin{cases} E_k, & k = n \\ 0, & k \neq n \end{cases}$$

C_k – коэффициенты разложения, E_k – энергия ортогональных функций.

$$C_k = \frac{1}{E_k} \int_{-T}^T x(t) \varphi_k(t) dt$$

Если выбрать в качестве ортогональных функций:

$$\varphi_k(t) = \begin{cases} \cos k\Omega t \\ \sin k\Omega t \\ e^{jk\Omega t} \end{cases}$$

то этот ряд (2.12) называется рядом Фурье (тригонометрический ряд).

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\Omega t + b_k \sin k\Omega t) \quad (2.13)$$

$$a_0 = \frac{2}{T} \int_0^T x(t) dt; \quad a_k = \frac{2}{T} \int_0^T x(t) \cos k\Omega t dt; \quad b_k = \frac{2}{T} \int_0^T x(t) \sin k\Omega t dt$$

$\Omega = 2\pi/T$ – частота первой гармоники, определяемая периодом T (T – период функции $x(t)$).

Другая, эквивалентная формула записи тригонометрического ряда:

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} A_k \cos(k\Omega t - \varphi_k), \quad (2.14)$$

$$A_k = \sqrt{a_k^2 + b_k^2}, \quad \varphi_k = \arctg \frac{b_k}{a_k}, \quad a_k = A_k \cos \varphi_k, \quad b_k = A_k \sin \varphi_k.$$

На рис. 2.5 приведены графики, иллюстрирующие представление периодической последовательности прямоугольных импульсов $s(t)$ конечным числом слагаемых ($k = 5$) ряда Фурье.

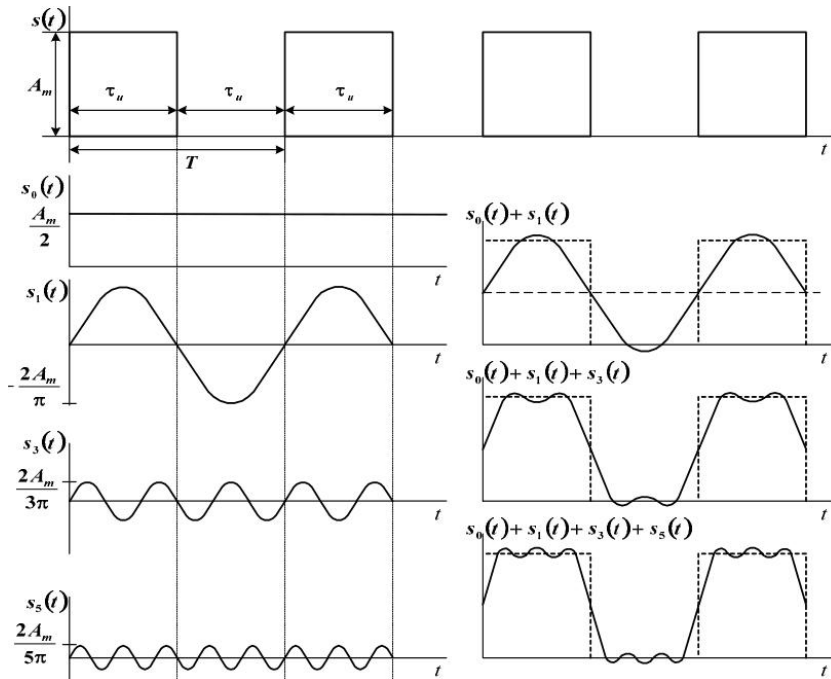


Рис. 2.5. Аппроксимация прямоугольных импульсов суммой гармоник

Для спектрального представления последовательности прямоугольных импульсов начало отсчета целесообразно брать в середине импульса.

Действительно, в этом случае в разложении останутся только косинусоидальные составляющие, так как интегралы от нечетных функций за период равны нулю $b_k = 0$.

Для четного сигнала $x(t) = x(-t)$, коэффициенты $a_k \neq 0$, $b_k = 0$.

Для нечетного сигнала $x(t) = -x(-t)$, коэффициенты $a_k = 0$, $b_k \neq 0$.

Для функции $s(t)$ (рис. 2.5) разложение имеет вид

$$s(t) = \frac{A_m}{2} + \frac{2A_m}{\pi} + \frac{2A_m}{3\pi} \sin 3\Omega_1 t + \frac{2A_m}{5\pi} \sin 5\Omega_1 t + \dots \quad (2.15)$$

Периодическая последовательность прямоугольных импульсов $s(t)$ представляется как результат сложения постоянной составляющей $A_m/2$ и синусоидальных сигналов с частотами $F_1, 3F_1, 5F_1, \dots$, причем период синусоиды с частотой F_1 совпадает с периодом последовательности импульсов $s(t)$. Для удобства F_1 можно представить в виде $F_1 = \Omega_1/2\pi = 1/T$.

Легко заметить, что график суммы двух первых слагаемых разложения (2.15) воспроизводит форму графика функции $s(t)$ очень грубо, только в основных чертах. Учет третьего слагаемого существенно улучшает совпадение суммы с функцией $s(t)$. Таким образом, с увеличением числа учитываемых гармоник точность представления $s(t)$ возрастает.

Совокупность всех гармонических составляющих разложения функции в ряд Фурье называется спектром функции.

Наличие отдельных гармонических составляющих спектра и величины их амплитуд можно наглядно показать с помощью спектральной диаграммы (рис. 2.6), у которой горизонтальная ось служит осью частот, а вертикальная – осью амплитуд.

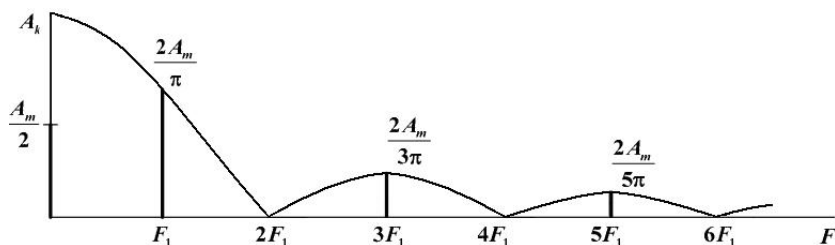


Рис. 2.6. Спектр амплитуд прямоугольных импульсов

В точках оси частот $F_1, 3F_1, 5F_1, \dots$, отображаются амплитуды соответствующих гармонических составляющих разложения функции.

На практике спектральные диаграммы называют более кратко – амплитудный спектр, фазовый спектр. Чаще всего интересуются амплитудным спектром (рис. 2.6). По нему можно оценить процентное содержание гармоник, наличие и уровни отдельных гармонических составляющих спектра.

3. СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ СИГНАЛОВ

3.1. Спектральное представление периодических сигналов

Как известно, разложение периодического сигнала по базису тригонометрических функций – это разложение его в ряд Фурье.

Разложение сигнала в ряд Фурье называется спектром сигнала.

В общем случае периодический сигнал содержит независимую от времени постоянную составляющую и бесконечный набор гармонических колебаний, или гармоник, с частотами, кратными основной частоте последовательности.

Графическое изображение коэффициентов ряда Фурье для конкретного сигнала называется спектральной диаграммой. По горизонтальной оси откладываются частоты гармоник, а по вертикали – амплитуды (амплитудная диаграмма) или начальные фазы (фазовая диаграмма).

При разложении в комплексный ряд Фурье:

$$s(t) = \sum_{k=-\infty}^{\infty} C_k \cdot \exp(jk\Omega t), \quad (3.1)$$

где $C_k = \frac{1}{T} \int_0^T s(t) e^{-jk\Omega t} dt$.

Спектр сигнала содержит компоненты на отрицательной полуоси частот, причём $C_{-k} = C_k^*$ (* обозначено комплексно-сопряжённое число).

Между коэффициентами комплексного и тригонометрического ряда существует связь:

$$A_k = 2|C_k|, \quad \varphi_k = \arg C_k. \quad (3.2)$$

Шириной спектра сигнала ΔF_s называется полоса частот, в пределах которой заключена основная доля энергии сигнала.

В качестве примера рассчитаем спектр периодической последовательности прямоугольных импульсов с амплитудой A :

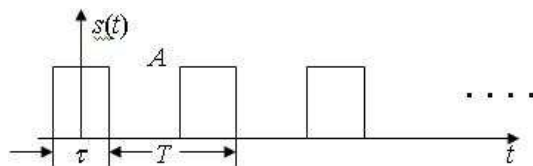


Рис. 3.1. Периодическая последовательность прямоугольных импульсов

Определим коэффициенты разложения в ряд Фурье C_k :

$$a_k = \frac{2}{T} \int_{-\tau/2}^{\tau/2} A \cos k\Omega t dt = \frac{2A}{T} \frac{\sin k\Omega t}{k\Omega} = \frac{4A}{k\Omega T} \sin \frac{k\Omega \tau}{2}$$

$$b_k = \frac{2}{T} \int_{-\tau/2}^{\tau/2} A \cdot \sin k\Omega t dt = 0, \text{ т.к. подынтегральная функция – нечетная.}$$

Пусть $T = 2\tau$, тогда коэффициенты a_k равны:

$$a_0 = A, a_k = 2A/k\pi (\sin k\pi/2), \text{ при } k > 0.$$

Итак, временная диаграмма периодической последовательности импульсов показана на рис. 3.1. Спектр этой последовательности дискретный и показан на рис. 3.2.

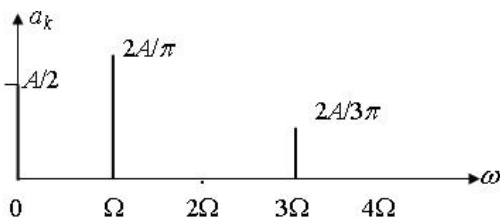


Рис. 3.2. Спектр периодическая последовательность прямоугольных импульсов

Ширина спектра сигнала равна, в данном случае, $\Delta F_3 = 2\pi/\tau$.

3.2. Спектральное представление непериодических сигналов

Для спектрального представления непериодических (импульсных) сигналов $s(t)$, заданных на конечном интервале (t_1, t_2) (рис. 3.3), непосредственно воспользоваться рядом Фурье нельзя. Для гармонического разложения сигнала мысленно дополняют его такими же импульсными сигналами до периодического с некоторым интервалом T (рис. 3.3).

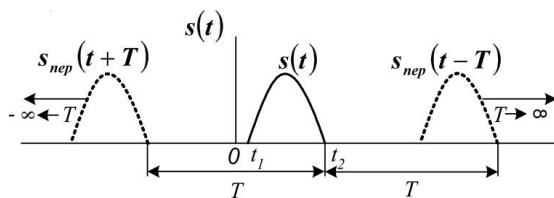


Рис. 3.3. Импульсный сигнал $s(t)$ и его периодическое продолжение $s_{nep}(t+kT)$

Для того чтобы вне искусственно введенного интервала исходный сигнал был равен нулю, необходимо увеличить период повторения импульсов.

В пределе, при увеличении периода $\infty \rightarrow T$ все импульсы уйдут право и влево в бесконечность и периодическая последовательность вновь станет одиночным импульсом.

Для вычисления спектра удобна симметричная комплексная форма ряда Фурье, но в нем вместо суммы будет интеграл с бесконечными пределами (преобразование Фурье):

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(j\Omega) \exp(j\Omega t) d\Omega. \quad (3.3)$$

При таком предельном переходе основная частота сигнала $\Omega = 2\pi/T$ стремится к нулю, бесконечно увеличивается число спектральных составляющих, частоты соседних гармоник $k\Omega$ и $(k + 1)\Omega$ становятся неразличимыми, а спектр будет сплошным.

Функция $G(j\Omega)$ называется спектральной плотностью сигнала $x(t)$.

Функции $G(j\Omega)$ и $s(t)$ представляют собой две математические модели одного и того же физического процесса: одна из них отражает частотный состав сигнала, а другая описывает изменение сигнала с течением времени.

Спектральная плотность сигнала определяется с использованием прямого преобразования Фурье:

$$G(j\Omega) = \int_{-\infty}^{\infty} s(t) \exp(-j\Omega t) dt. \quad (3.4)$$

Таким образом, формулы (3.3) и (3.4) называются соответственно обратным и прямым преобразованиями Фурье. Они показывают взаимосвязь между сигналом $s(t)$ и его комплексной спектральной плотностью $G(j\Omega)$.

Для одиночного прямоугольного импульса с амплитудой A и длительностью τ на рис. 3.4 получим спектр $S(j\Omega)$ на рис. 3.5.

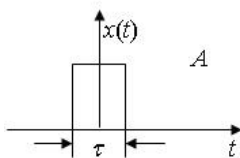


Рис. 3.4. Одиночный прямоугольный импульс

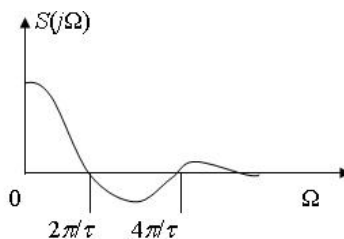


Рис. 3.5. Спектр прямоугольного импульса

$$S(j\Omega) = \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} A e^{-j\Omega t} dt = \frac{A}{-j\Omega} \left[\exp\left(-\frac{j\Omega\tau}{2}\right) - \exp\left(\frac{j\Omega\tau}{2}\right) \right].$$

Это выражение с учетом формулы Эйлера $\sin a = \frac{\exp(ja) - \exp(-ja)}{2j}$ можно переписать в виде

$$S(j\Omega) = A\tau \frac{\sin\left(\frac{\Omega\tau}{2}\right)}{\left(\frac{\Omega\tau}{2}\right)}. \quad (3.5)$$

Спектр непериодического сигнала сплошной, бесконечный, ширина спектра определяется длительностью сигнала и, приближённо, равна $\Delta F_3 \approx 2\pi/\tau$.

3.3. Основные свойства преобразования Фурье:

Свойствами преобразований Фурье определяется взаимное соответствие трансформации сигналов и их спектров.

1. Линейность. Преобразование Фурье относится к числу линейных интегральных операций, т.е. спектр суммы сигналов равен сумме спектров этих сигналов.

$$\sum_n a_n s_n(t) \Leftrightarrow \sum_n a_n s_n(\omega) \quad (3.6)$$

Пример суммирования сигналов и его отображения в суммирования спектров приведен на рис. 3.6:

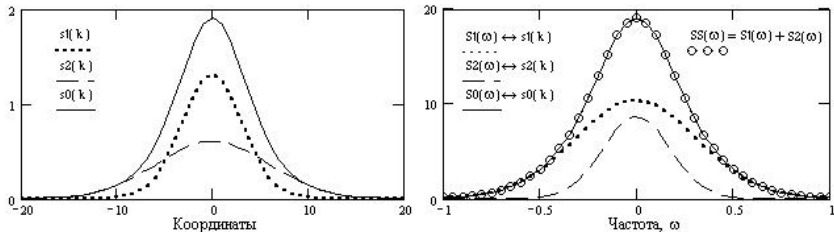


Рис. 3.6. Сигналы и их спектры. $s_0(k) = s_1(k) + s_2(k) \Leftrightarrow S_1(\omega) + S_2(\omega) = S_0(\omega)$

2. Свойства четности преобразования определяются косинусными (четными, действительными) и синусными (нечетными, мнимыми) частями разложения и подобием прямого и обратного преобразований.

На рис. 3.7. приведены примеры, поясняющие свойства четности преобразования. Сигнал $s_1(k)$ является четным, $s_1(k) = s_1(-k)$, и имеет только вещественный четный спектр (мнимая часть спектральной функции представлена нулевыми значениями). Сигнал $s_2(k) = -s_2(-k)$ нечетный и имеет мнимый нечетный спектр, а нулевыми значениями представлена его действительная часть. Сигнал $s_3(k)$ образован суммой сигналов $s_1(k)$ и $s_2(k)$. Соответственно, спектральная функция сигнала представлена и действительной четной частью (принадлежащей $s_1(k)$), и мнимой нечетной частью (принадлежащей $s_2(k)$). При обратном преобразовании Фурье отдельно действительной и мнимой части спектра $S_3(\omega)$, равно как и любых других комплексных спектров, будут восстановлены четная и нечетная части исходного сигнала.

Заметим, что произвольный исходный сигнал может быть задан в одностороннем варианте (в интервале $0 - T$), но четная и нечетная части этого сигнала занимают интервал от $-T$ до T , при этом на левой половине числовой оси (от $-T$ до 0) эти два сигнала компенсируют друг друга, давая нулевые значения.

Сигнал $s(t)$, спектр $S(\omega)$. При этом если:

$s(t)$ – четный, то $S(\omega)$ – вещественный, четный;

$s(t)$ – нечетный, то $S(\omega)$ – мнимый, нечетный

$s(t)$ – произвольный, то $S(\omega)$ – действительная часть – четная,
а мнимая часть – нечетная.

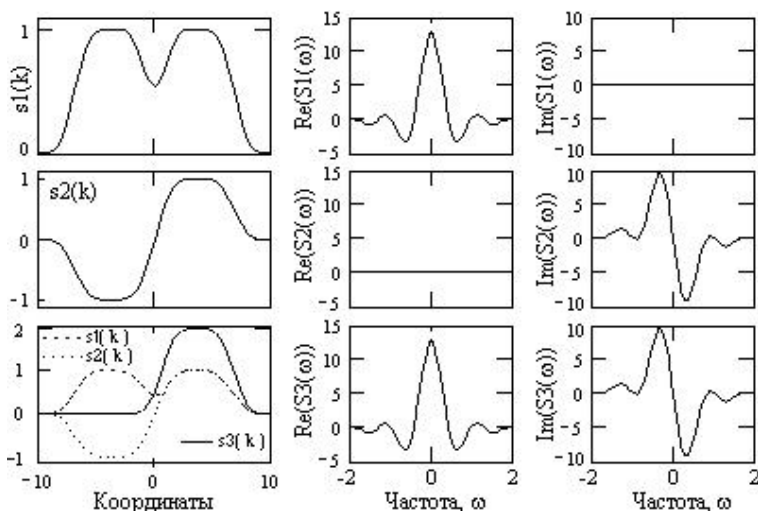


Рис. 3.7. Свойства четности преобразования

3. Изменение аргумента функции (сжатие или расширение сигнала) приводит к обратному изменению аргумента ее Фурье-образа и об-

ратно пропорциональному изменению его модуля. Действительно, если $s(t) \Leftrightarrow S(\omega)$, то при изменении длительности сигнала с сохранением его формы (растяжении сигнала по временной оси), т.е. для сигнала с новым аргументом $s(x) = s(a \cdot t)$ при $x = a \cdot t$, получаем:

$$s(a \cdot t) \Leftrightarrow \int_{-\infty}^{\infty} s(a \cdot t) \exp(-j\omega t) dt = \frac{1}{a} \int_{-\infty}^{\infty} s(x) \exp(-jx \frac{\omega}{a}) dx \quad (3.7')$$

$$s(a \cdot t) \Leftrightarrow \left(\frac{1}{|a|}\right) \cdot S\left(\frac{\omega}{a}\right)$$

Выражение (3.7') действительно при $a > 0$. При $a < 0$ происходит зеркальный поворот сигнала относительно вертикальной оси, а замена переменной $t = x/a$ вызывает перестановку пределов интегрирования и, соответственно, изменение знака спектра:

$$s(a \cdot t) \Leftrightarrow -(1/a)S(\omega/a). \quad (3.7'')$$

Обобщенная формула изменения аргумента:

$$s(a \cdot t) \Leftrightarrow -(1/|a|)S(\omega/a), a \neq 0 \quad (3.7)$$

Если под аргументом функции и ее спектра понимать определенные физические единицы, например, время – частота, то отсюда следует: чем короче по своей длительности сигнал, тем шире по частоте его спектр, и наоборот. Это можно наглядно видеть на рис. 3.6. для сигналов $s_1(k)$ и $s_2(k)$ и их спектров $S_1(\omega)$ и $S_2(\omega)$.

От изменения аргумента функций следует отличать изменение масштаба представления функций. Изменение масштаба аргументов изменяет только оцифровку числовых осей отображения сигналов и их спектров, но не изменяет самих сигналов и спектров. Так, при масштабе оси времен $t = 1$ секунда, масштаб оси частот $f = 1/t = 1$ герц, а при $t = 1$ мксек $f = 1/t = 1$ МГц ($t = a \cdot t, f = 1/a \cdot t, a = 10^{-6}$).

4. Теорема запаздывания. Запаздывание (сдвиг, смещение) сигнала по аргументу функции на интервал t_0 приводит к изменению фазочастотной функции спектра (фазового угла всех гармоник) на величину $-\omega t_0$ без изменения модуля (амплитудной функции) спектра. Применяя замену переменной $t - t_0 = x$, получаем:

$$s(t - t_0) \Leftrightarrow \int_{-\infty}^{\infty} s(t - t_0) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} s(x) \exp(-j\omega x) \exp(-j\omega t_0) dx = \quad (3.8)$$

$$= S(\omega) \exp(-j\omega t_0)$$

Совершенно очевидно, что амплитуды гармоник сигнала при его сдвиге изменяться не должны. С учетом того, что $|\exp(-j\omega t_0)| = 1$, это следует и из (3.8):

$$|S(\omega) \exp(-j\omega t_0)| = |S(\omega)|.$$

Фазовый спектр сдвигается на $-\omega t_0$ с линейной зависимостью от частоты:

$$S(\omega) \exp(-j\omega t_0) = R(\omega) \exp[j(j(\omega))] \exp(-j\omega t_0) = R(\omega) \exp[j(j(\omega) - \omega t_0)]. \quad (3.9)$$



Рис. 3.8. Изменение спектра сигнала при его сдвиге.

Пример двух одинаковых сигналов, сдвинутых относительно друг друга на $t_0 = 1$, и соответствующих данным сигналам спектров приведен на рис. 3.8.

5. Преобразование производной (дифференцирование сигнала):

$$s(t) = d[y(t)]/dt = d \left[\int_{-\infty}^{\infty} Y(\omega) \exp(j\omega t) d\omega \right] / dt = \int_{-\infty}^{\infty} Y(\omega) [d(\exp(j\omega t))/dt] d\omega = \quad (3.10)$$

$$= \int_{-\infty}^{\infty} j\omega Y(\omega) \exp(j\omega t) d\omega \Leftrightarrow j\omega Y(\omega)$$

Таким образом, дифференцирование сигнала отображается в спектральной области простым умножением спектра сигнала на оператор дифференцирования сигнала в частотной области $j\omega$, что эквивалентно дифференцированию каждой гармоники спектра. Умножение на $j\omega$ приводит к обогащению спектра производной сигнала высокочастотными составляющими (по сравнению с исходным сигналом) и уничтожает составляющие с нулевой частотой.

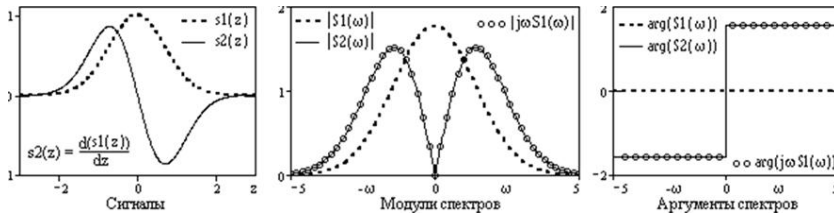


Рис. 3.9. Спектры сигнала и его производной

Пример сигнала, его производной и соответствующих им спектров приведен на рис. 3.9. По изменению аргумента спектра (для четного исходного сигнала он был нулевым) можно видеть, что для всех гармо-

ник спектра появляется сдвиг фаз на $\pi/2$ (90^0) для положительных частот, и на $-\pi/2$ (-90^0) для отрицательных частот.

В общем случае, для кратных производных:

$$d^n[y(t)]/dt^n = (j\omega)^n Y(\omega). \quad (3.11)$$

6. Преобразование интеграла сигнала в частотной области при известном спектре сигнала может быть получено из следующих простых соображений. Если имеет место $s(t) = d[y(t)]/dt \Leftrightarrow j\omega Y(\omega) = S(\omega)$, то должна выполняться и обратная операция:

$$y(t) = \int_{-\infty}^{\infty} s(t)dt \Leftrightarrow Y(\omega) = S(\omega)/j\omega. \text{ Отсюда следует:}$$

$$\int_{-\infty}^{\infty} s(t)dt \Leftrightarrow (1/j\omega) \cdot S(\omega) \quad (3.12)$$

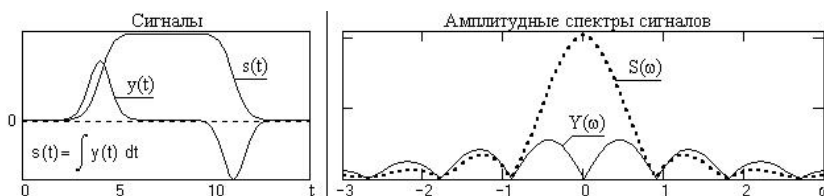


Рис. 3.10. Сигналы и амплитудные спектры сигналов

Оператор интегрирования в частотной области ($1/j\omega$) при $\omega > 1$ ослабляет в амплитудном спектре высокие частоты и при $\omega < 1$ усиливает низкие. Фазовый спектр сигнала смещается на -90^0 для положительных частот и на 90^0 для отрицательных. Пример модуля спектра сигнала и его интегральной функции приведены на рис. 3.10.

Формула (3.12) справедлива для сигналов с нулевой постоянной составляющей. При интегрировании сигналов с определенным значением постоянной составляющей $C = const$ в правой части выражения (3.12) появляется дополнительное слагаемое преобразования Фурье постоянной составляющей C , которое представляет собой, как будет показано ниже, дельта-функцию на нулевой частоте с весовым коэффициентом, равным значению C :

$$y(t) \Leftrightarrow Y(\omega) = (1/j\omega)S(\omega) + C \cdot d(\omega_0).$$

7. Преобразование свертки сигналов $y(t) = s(t) * h(t)$:

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\tau) h(t-\tau) \exp(-j\omega t) d\tau dt$$

$$Y(\omega) = \int_{-\infty}^{\infty} s(\tau) d\tau \int_{-\infty}^{\infty} h(t-\tau) \exp(-j\omega t) dt$$

По теореме запаздывания (3.8):

$$\int_{-\infty}^{\infty} h(t-\tau) \exp(-j\omega t) dt = H(\omega) \exp(-j\omega \tau)$$

Отсюда: $Y(\omega) = \int_{-\infty}^{\infty} H(\omega) s(\tau) \exp(-j\omega \tau) d\tau = H(\omega) \cdot S(\omega)$

$$s(t) * h(t) \Leftrightarrow S(\omega) \cdot H(\omega). \quad (3.13)$$

Пример выполнения свертки в частотной области приведен на рис. 3.11.

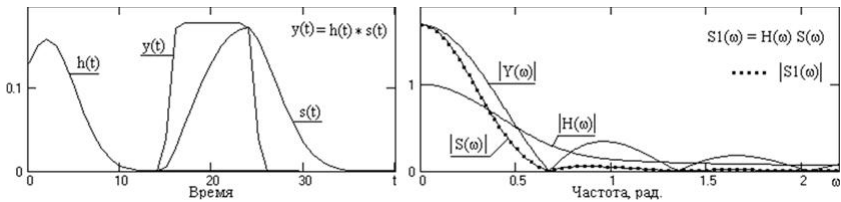


Рис. 3.11. Сигналы и амплитудные спектры сигналов

Отметим, что частотное представление $H(\omega)$ импульсного отклика $h(t)$ линейной системы (или соответствующей линейной операции) имеет смысл частотной передаточной функции системы и позволяет определить сигнал на выходе системы (в частотной форме представления) при задании произвольного сигнала (в частотной форме) на ее входе. По существу, функция $H(\omega)$ представляет собой распределение по частоте коэффициента пропускания частотных составляющих сигнала с входа на выход системы (операции).

Таким образом, свертка функций в координатной форме отображается в частотном представлении произведением Фурье-образов этих функций.

Это положение имеет фундаментальное значение в практике обработки данных.

Любая линейная система обработки данных (информационных сигналов) реализует определенную операцию трансформации сигнала, т.е. выполняет операцию свертки входного сигнала $s(t)$ с оператором системы $h(t)$. С использованием преобразования свертки эта операция может производиться как с динамической, так и с частотной формой представления сигналов. При этом обработка данных, представленных в

цифровой форме, производится, как правило, в частотной области, т.к. может быть на несколько порядков выше по производительности, чем во временной области. Она представляет собой последовательность следующих операций.

1) Перевод сигнала в частотную область: $s(t) \Leftrightarrow S(\omega)$.

2) Умножение спектра сигнала на передаточную функцию системы:

$$Y(\omega) = H(\omega) \cdot S(\omega).$$

Передаточная функция системы определяется аналогичным преобразованием $h(t) \Leftrightarrow H(\omega)$ или задается непосредственно в частотном представлении, что позволяет задавать передаточные функции сколь угодно сложной формы, в том числе с разрывами и скачками, для которых во временной области потребуются операторы $h(t)$ с бесконечной импульсной характеристикой.

3) Перевод спектра обработанного сигнала во временную область:

$$Y(\omega) \Leftrightarrow y(t).$$

8. Преобразование произведения сигналов $y(t) = s(t) \cdot h(t)$:

Произведение функций в координатной форме отображается в частотном представлении сверткой Фурье-образов этих функций, с нормировочным множителем $(1/2\pi)$, учитывающим несимметричность прямого и обратного преобразования Фурье функций $s(t)$ и $h(t)$ при использовании угловых частот.

$$\begin{aligned} Y(\omega) &= \int_{-\infty}^{\infty} s(t)h(t) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} s(t) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega') \exp(j\omega' t) d\omega' \right] dt = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(t)H(\omega') \exp(-j(\omega - \omega')t) d\omega' dt = \quad (3.14) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega') d\omega' \int_{-\infty}^{\infty} s(t) \exp(-j(\omega - \omega')t) dt = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega') S(\omega - \omega') d\omega' = \frac{1}{2\pi} H(\omega) * S(\omega). \end{aligned}$$

9. Спектральная плотность (прямое преобразование Фурье)

а) гармонической функции $s(t) = \cos(\omega_0 t)$

$$G(j\Omega) = \pi(\delta(\Omega - \omega_0) + \delta(\Omega + \omega_0));$$

б) радиоимпульса (свойство смещения спектра) позволяет рассчитать спектральную плотность сигнала $s(t)$, умноженного на гармоническое колебание $s_1(t) = s(t)\cos(\omega_0 t + \varphi_0)$.

$$G(j\Omega) = \frac{1}{2} (G_A(j(\Omega - \omega_0))e^{j\varphi} + G_A(j(\Omega + \omega_0))e^{-j\varphi}),$$

где $G_A(j\Omega)$ – спектральная плотность огибающей $A(t)$.

Следовательно возникает расщепление спектра $G(j\Omega)$ на две части максимумы которых возникают на частотах $(+\omega_0)$ и $(-\omega_0)$.

в) δ -функции $s(t) = \delta(t)$:

$$G(j\Omega) = 1$$

10. Спектры мощности.

Временная функция мощности сигнала в общей форме определяется выражением:

$$w(t) = s(t) s^*(t) = |s(t)|^2.$$

Спектральная плотность мощности, соответственно, равна преобразованию Фурье произведения $s(t) \cdot s^*(t)$, которое отобразится в спектральном представлении сверткой Фурье-образов этих функций:

$$W(f) = \int_{-\infty}^{\infty} S(f) * S^*(f - \nu) d\nu. \quad (3.15)$$

Но для всех текущих значений частоты f интеграл в правой части этого выражения равен произведению $S(f) \cdot S^*(f)$, так как для всех значений сдвига $\nu \neq 0$ в силу ортогональности гармоник $S(f)$ и $S^*(f - \nu)$ значения их произведения равны нулю. Отсюда:

$$W(f) = S(f) * S^*(f) = |S(f)|^2. \quad (3.16)$$

Спектр мощности – вещественная неотрицательная четная функция, которую очень часто называют энергетическим спектром. Спектр мощности, как квадрат модуля спектра сигнала, не содержит фазовой информации о частотных составляющих, а, следовательно, восстановление сигнала по спектру мощности невозможно. Это означает также, что сигналы с различными фазовыми характеристиками могут иметь одинаковые спектры мощности. В частности, сдвиг сигнала не отражается на его спектре мощности.

Для функций мощности взаимодействия сигналов в частотной области соответственно имеем частотные спектры мощности взаимодействия сигналов:

$$\begin{aligned} W_{xy}(f) &= X(f) Y^*(f), \\ W_{yx}(f) &= Y(f) X^*(f), \\ W_{xy}(f) &= W_{yx}^*(f). \end{aligned}$$

Функции мощности взаимодействия сигналов комплексные, даже если обе функции $x(t)$ и $y(t)$ вещественны, при этом $\text{Re}[W_{xy}(f)]$ – четная функция, а $\text{Im}[W_{xy}(f)]$ – нечетная. Отсюда полная энергия взаимодействия сигналов при интегрировании функций мощности взаимодействия определяется только реальной частью спектра:

$$E_{xy} = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_{xy}(\omega) d\omega = \frac{1}{\pi} \int_{-\infty}^{\infty} \text{Re}[W_{xy}] d\omega,$$

и всегда является вещественным числом.

11. Равенство Парсеваля. Полная энергия спектра сигнала:

$$E_s = \int_{-\infty}^{\infty} W(f)df = \int_{-\infty}^{\infty} |S(f)|^2 df. \quad (3.17)$$

Так как координатное и частотное представление по существу только разные математические отображения одного и того же сигнала, то равной должна быть и энергия сигнала в двух представлениях, откуда следует равенство Парсеваля:

$$\int_{-\infty}^{\infty} |S(t)|^2 dt = \int_{-\infty}^{\infty} |S(f)|^2 df,$$

т.е. энергия сигнала равна интегралу модуля его частотного спектра – сумме энергий всех частотных составляющих сигнала. Аналогично для энергии взаимодействия сигналов:

$$\int_{-\infty}^{\infty} x(t)y^*(t)dt = \int_{-\infty}^{\infty} X(f)Y^*(f)df.$$

Из равенства Парсеваля следует инвариантность скалярного произведения сигналов и нормы относительно преобразования Фурье:

$$(x(t), y(t)) = (X(f), Y(f)), \|x(t)\|^2 = \|X(f)\|^2.$$

Не следует забывать, что при представлении спектров в круговых частотах (по ω) в правой части приведенных равенств должен стоять множитель $1/2\pi$.

4. СИГНАЛЫ С ОГРАНИЧЕННЫМ СПЕКТРОМ. ТЕОРЕМА КОТЕЛЬНИКОВА

4.1. Разложение непрерывных сигналов в ряд Котельникова

Как отмечено ранее, любые сигналы конечной длительности теоретически имеют бесконечно широкий спектр частот. В то же время доля энергии, передаваемая на высоких частотах, очень мала и ею при расчете полной энергии сигнала можно пренебречь. Следовательно, сигналы с ограниченным спектром являются удобными математическими моделями реальных сигналов.

В 1933 году В.А. Котельников доказал, что сигнал $s(t)$ с ограниченной полосой частот, не имеющий спектральных компонент с частотами, которые превышают значение $\omega_b = 2\pi F_b$, однозначно определяется значениями, выбранными через равные промежутки времени [1]

$$\Delta t = \pi/\omega_b = 1/2F_b.$$

Известно, что при аналогово-цифровом преобразовании, чем меньше частота оцифровки (или больше период дискретизации) и грубее квантование сигнала, тем меньше данных необходимо для представления аналогового сигнала в цифровом виде. С другой стороны с уменьшением объема данных увеличивается вероятность потери информации содержащейся в сигнале.

Чтобы продемонстрировать искажение информации при неправильном выборе частоты дискретизации сигнала рассмотрим примеры.

Пример.

Гармонический сигнал имеет частоту f (период $T = 1/f$). Проведем дискретизацию сигнала с периодом дискретизации T_d меньшим половины периода входного сигнала T (рис. 4.1).

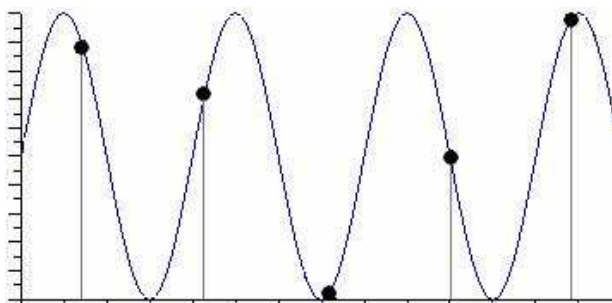


Рис. 4.1. Дискретизация сигнала с периодом $T_d < T/2$

Очевидно, что дискретные отсчеты сигнала однозначно не отображают форму исходного сигнала, в частности по получившимся точкам можно построить гармонический сигнал с периодом $T_{\text{искаж.}}$, отличающимся от периода исходного сигнала T . Период $T_{\text{искаж.}}$ больше периода исходного сигнала T , соответственно частота меньше, частоты исходного сигнала f (рис. 4.2).

Данный эффект называется стробоскопическим эффектом или алиасингом. Он заключается в появлении ложной низкочастотной составляющей при дискретизации сигнала с частотой меньше удвоенной частоты исходного сигнала (или с периодом большим половины периода исходного сигнала), отсутствующей в исходном сигнале.

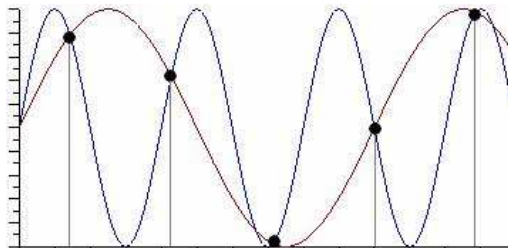


Рис. 4.2. Стробоскопический эффект дискретизации

При дискретизации с периодом равным половине исходного аналогового сигнала ($f_d = 2f$) возникает неопределенность начальной фазы и амплитуды сигнала, т.е. возможно зеркальное искажение (противофаза), при этом частота исходного сигнала не искажается. В крайнем случае, мы можем получить отсчеты сигнала равные нулю (рис. 4.3).

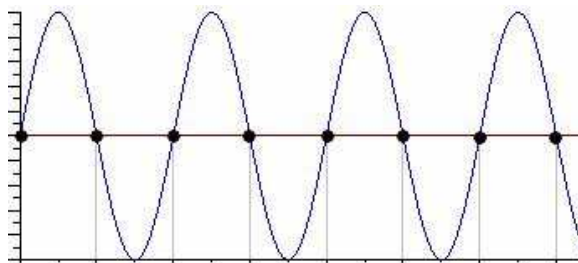


Рис. 4.3. Дискретизация сигнала с периодом $T_d = T/2$

Если период дискретизации меньше половины периода исходного сигнала, то очевидно, что через получившиеся после оцифровки точки можно построить только один гармонический сигнал, соответствующий

исходному, без искажения начальной фазы, амплитуды и частоты (рис. 4.4).

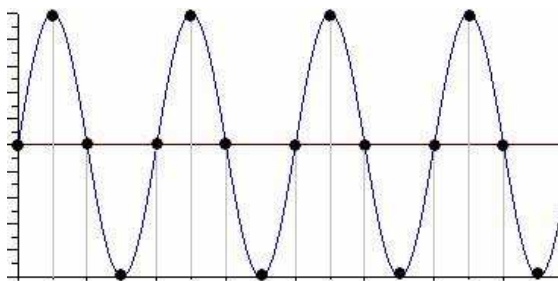


Рис. 4.4. Дискретизация сигнала с периодом $T_d < T/2$

Таким образом, для адекватного восстановления гармонического сигнала по дискретным отсчетам, период дискретизации должен быть не меньше половины периода сигнала. Частота равная половине частоты дискретизации называется частотой Найквиста $f_N = f_d/2$.

Таким образом, аналоговый сигнал с ограниченным спектром может быть восстановлен однозначно и без искажений по своим дискретным отсчетам, взятым с частотой большей удвоенной максимальной частоты его спектра $F_d > 2 \cdot F_{\max}$.

Данное утверждение известно как теорема Котельникова (в западной литературе теорема Найквиста-Шеннона) или теорема отсчетов.

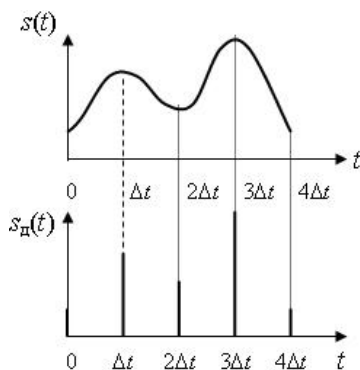


Рис. 4.5. Временные диаграммы непрерывного сигнала $s(t)$ и дискретизированного $s_d(t)$

Важно, что не надо передавать непрерывно исходный сигнал $s(t)$, достаточно передавать отсчеты $s(k\Delta t)$. Это первый шаг перехода от не-

прерывного сигнала к цифровому. С точки зрения математики теорема Котельникова означает представление сигнала в виде ряда:

$$s(t) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \frac{\sin \omega_g(t - k\Delta t)}{\omega_g(t - k\Delta t)}, \quad (4.1)$$

где $s(k\Delta t)$ – отсчёты;

$(\sin \omega_g(t - k\Delta t)) / \omega_g(t - k\Delta t)$ – функции отсчётов.

Ряд Котельникова – это разложение сигнала $s(t)$ в ряд по ортогональным функциям $\varphi_k(t)$.

$$\varphi_k(t) = (\sin \omega_g(t - k\Delta t)) / \omega_g(t - k\Delta t) \quad (4.2)$$

Теоретически дискретизация осуществляется с помощью δ -импульсов.

$$\delta(t) = \begin{cases} 0, & t \neq 0 \\ \infty, & t = 0 \end{cases}; \quad \delta(t-a) = \begin{cases} 0, & t \neq a \\ \infty, & t = a \end{cases}$$

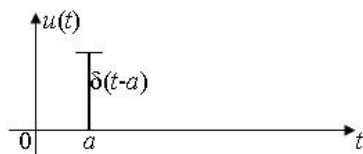


Рис. 4.6. Временная диаграмма одиночного δ -импульса

Спектр одиночного δ -импульса получим, используя преобразование Фурье:

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

Использовано «фильтрующее» свойство дельта-функций:

$$\int_{-\infty}^{\infty} \delta(t-a) f(t) dt = f(a)$$

Следовательно, спектр одиночного дельта-импульса имеет вид:

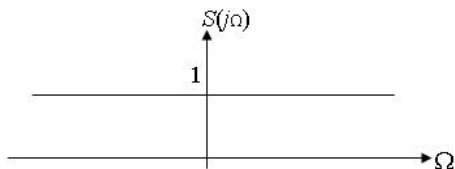


Рис. 4.7. Спектр одиночного δ -импульса

Чтобы получить отсчёты функции $s(t)$ перемножим функцию $s(t)$ на периодическую последовательность дельта-импульсов с периодом $T = \Delta t$.

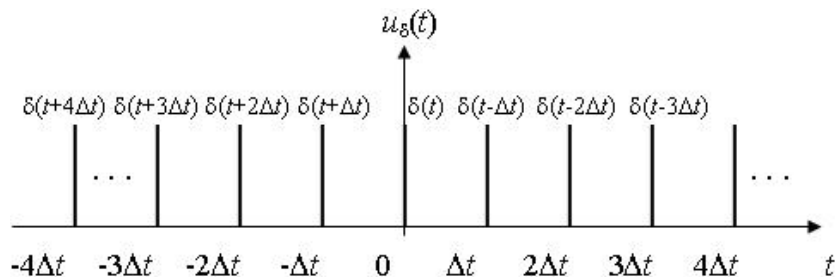


Рис. 4.8. Временная диаграмма периодической последовательности δ -импульсов

Так как сигнал периодический, то его спектр будет дискретным.

$$u_{\delta}(t) = \sum_{k=-\infty}^{\infty} C_k e^{jk\Omega t} = \dots + \frac{1}{\Delta t} e^{-2jk\Omega t} + \frac{1}{\Delta t} e^{-jk\Omega t} + \frac{1}{\Delta t} + \frac{1}{\Delta t} e^{jk\Omega t} + \dots \quad (4.3)$$

$$C_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \delta(t) e^{-jk\Omega t} dt = \frac{1}{\Delta t}; \quad \Omega = \frac{2\pi}{T} = \frac{2\pi}{\Delta t} = \frac{2\pi\omega_s}{\pi} = 2\omega_s = \omega_o$$

$T = \Delta t$; ω_d – частота дискретизации.

Спектр периодической последовательности дельта-импульсов в соответствии с формулой для $U(t)$ имеет следующий вид:

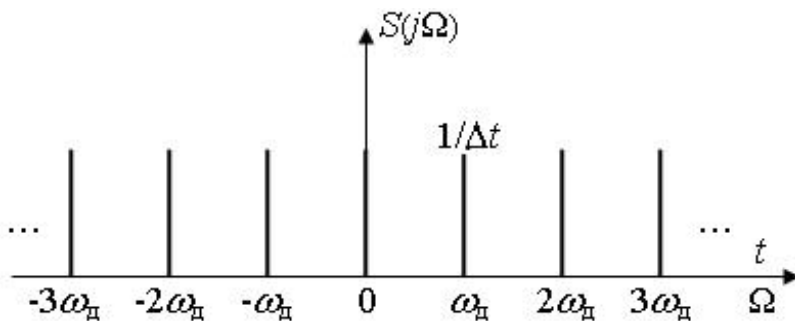


Рис. 4.9. Спектр периодической последовательности δ -импульсов

4.2. Спектр дискретизированного сигнала

Рассмотрим временные диаграммы исходного непрерывного и дискретизированного сигналов:

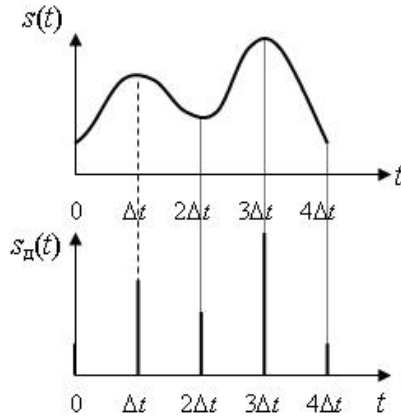


Рис. 4.10. Временные диаграммы непрерывного и дискретизированного сигналов

$s_{\text{д}}(t) = s(t) \cdot U_{\delta}(t)$ – дискретизированный сигнал

$s(t)$ – исходный сигнал.

$U_{\delta}(t)$ – периодическая последовательность δ -импульсов

Разложим периодическую последовательность δ -импульсов в ряд Фурье, как мы это делали выше:

$$U_{\delta}(t) = \dots + \frac{1}{\Delta t} e^{-j\omega_{\text{д}} t} + \frac{1}{\Delta t} + \frac{1}{\Delta t} e^{j\omega_{\text{д}} t} + \dots$$

$$s_{\text{д}}(t) = s(t) U_{\delta}(t) = s(t) \left[\dots + \frac{1}{\Delta t} e^{-j\omega_{\text{д}} t} + \frac{1}{\Delta t} + \frac{1}{\Delta t} e^{j\omega_{\text{д}} t} + \dots \right]$$

Найдём спектр дискретизированного сигнала.

$$\begin{aligned} \dot{S}_{\text{д}}(\Omega) &= \int_{-\infty}^{\infty} s_{\text{д}}(t) e^{-j\Omega t} dt = \dots + \frac{1}{\Delta t} \int_{-\infty}^{\infty} s(t) e^{-j(\Omega + \omega_{\text{д}})t} dt + \frac{1}{\Delta t} \int_{-\infty}^{\infty} s(t) e^{-j\Omega t} dt + \\ &+ \frac{1}{\Delta t} \int_{-\infty}^{\infty} s(t) e^{-j(\Omega - \omega_{\text{д}})t} dt + \dots = \dots + \frac{1}{\Delta t} \dot{S}_x(\Omega + \omega_{\text{д}}) + \frac{1}{\Delta t} \dot{S}_x(\Omega) + \\ &+ \frac{1}{\Delta t} \dot{S}_x(\Omega - \omega_{\text{д}}) + \frac{1}{\Delta t} \dot{S}_x(\Omega - 2\omega_{\text{д}}) + \dots \end{aligned} \quad (4.4)$$

Т.о. мы видим, что спектр дискретизированного сигнала содержит спектр исходного сигнала $S_x(\Omega)$, спектр исходного сигнала смещенный на величину частоты дискретизации вправо $S_x(\Omega - \omega_d)$, тот же спектр смещенный на величину частоты дискретизации влево $S_x(\Omega + \omega_d)$, тот же спектр смещенный на величину $2\omega_d$ и т.д.

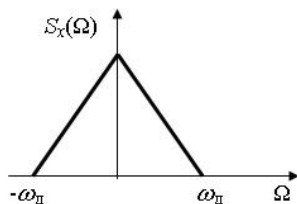


Рис. 4.11. Спектр исходного непрерывного сигнала

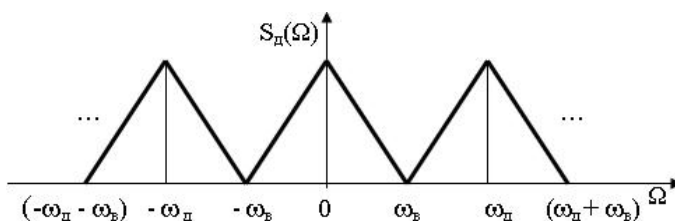


Рис. 4.12. Спектр дискретизированного сигнала

4.3. Спектр сигнала дискретизированного импульсами конечной длительности (амплитудно-импульсно модулированный (АИМ) сигнал)

Очевидно, что реально мы располагаем не последовательностью дельта-импульсов, а последовательностью импульсов конечной длительности.

В результате процесса дискретизации мы получим не последовательность дельта-импульсов, амплитуда которых соответствует значению непрерывного сигнала в тактовые моменты времени, а последовательность реальных, например, прямоугольных импульсов, амплитуда которых соответствует значениям непрерывного сигнала в тактовые моменты времени.

АИМ сигнал можно записать в виде:

$$s_{\text{АИМ}}(t) = s(t)U(t) = s(t) \left[\dots + \frac{a_{-1}}{2} e^{-j\omega_a t} + \frac{a_0}{2} + \frac{a_1}{2} e^{-j\omega_a t} + \dots \right],$$

где $U(t)$ – периодическая последовательность импульсов.

В квадратных скобках – ряд Фурье для последовательности импульсов конечной длительности.

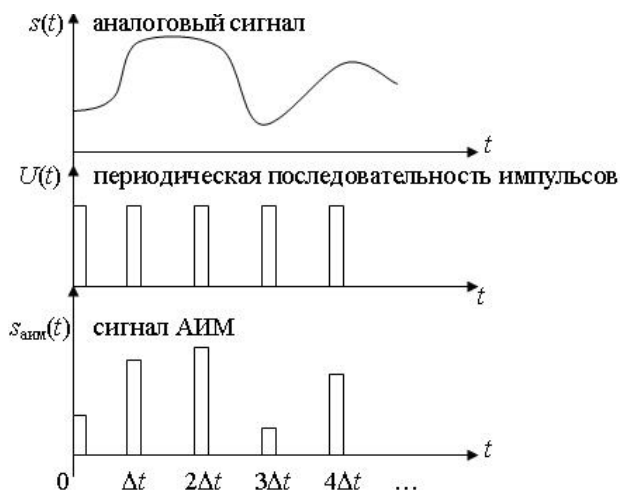


Рис. 4.13. Временные диаграммы формирования АИМ сигнала

Спектр АИМ сигнала, следовательно, похож на спектр дискретизованного сигнала при дискретизации дельта-импульсами, но амплитуда составляющих спектра убывает с ростом номера гармоники:

$$\begin{aligned} \dot{S}_d(\Omega) = & \dots + \frac{a_2}{2} \dot{S}_x(\Omega + 2\omega_d) + \frac{a_1}{2} \dot{S}_x(\Omega + \omega_d) + \frac{a_0}{2} \dot{S}_x(\Omega) + \\ & + \frac{a_1}{2} \dot{S}_x(\Omega - \omega_d) + \frac{a_2}{2} \dot{S}_x(\Omega - 2\omega_d) + \dots \end{aligned} \quad (4.5)$$

Спектр АИМ сигнала в соответствии с формулой (4.5) принимает вид, показанный на рис. 4.14.

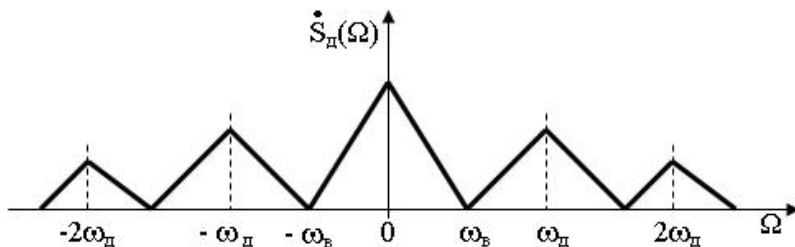


Рис. 4.14. Спектр АИМ сигнала

4.4. Восстановление непрерывного сигнала из отсчётов

В линию связи передаются импульсы-отсчёты, которые поступают на вход приёмника.

Для восстановления исходного непрерывного сигнала из импульсов-отсчётов надо эти импульсы подать на вход идеального фильтра низких частот (ИФНЧ), который имеет следующие характеристики.

Амплитудно-частотная характеристика идеального ФНЧ (АЧХ ИФНЧ) имеет вид:

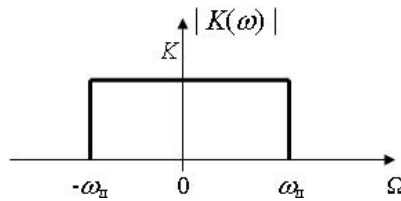


Рис. 4.15. АЧХ идеального фильтра низких частот

Импульсная реакция ИФНЧ, т.е. реакция на дельта-импульс имеет вид:

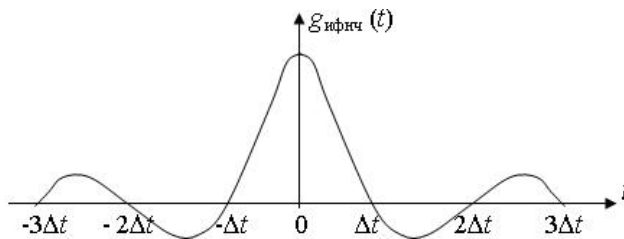


Рис. 4.16. Импульсная реакция ИФНЧ на δ -импульс

$$g_{\text{ифнч}}(t) = K \frac{\sin \omega_b t}{\omega_b t}; \omega_b t = k\pi; t = k \frac{\pi}{\omega_b} = k\Delta t \quad (4.6)$$

Первая формула – это выражение для импульсной реакции ИФНЧ, вторая и третья формулы определяют моменты времени, для которых $g_{\text{ифнч}}(t)$ обращается в ноль.

При прохождении дискретизированного сигнала, вид спектра которого представлен на рисунке 4.12 или 4.14 через ИФНЧ, АЧХ которого имеет вид представленный на рис. 4.15, на выходе получим спектр:

$$S(\omega) = KS_d(\omega) = K S_x(\omega)/\Delta t;$$

для АИМ сигнала, соответственно:

$$S(\omega) = KS_d(\omega) = K a_0 S_x(\omega)/2.$$

Таким образом, с точностью до постоянного множителя мы получили на выходе ИФНЧ спектр исходного сигнала $s(t)$. С временной точки зрения мы получили исходный непрерывный сигнал $s(t)$.

4.5. Погрешности дискретизации и восстановления непрерывных сигналов

Теорема Котельникова точно справедлива только для сигналов с финитным (конечным) спектром. На рис. 4.17 показаны некоторые варианты финитных спектров:

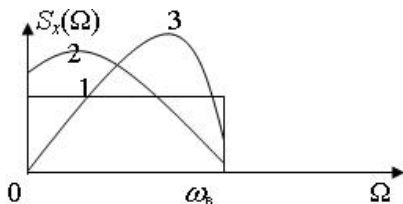


Рис. 4.17. Примеры финитных спектров

Однако спектры реальных информационных сигналов бесконечны. В этом случае теорема Котельникова справедлива с погрешностью.

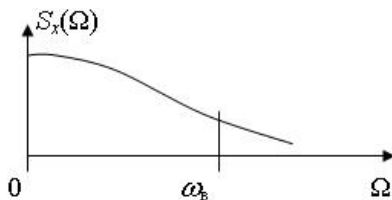


Рис. 4.18. Ограничение спектра реального сигнала

Погрешность дискретизации определяется энергией спектральных составляющих сигнала, лежащих за пределами частоты ω_b .

$$\overline{\Delta E}_d^2 = \int_{\omega_b}^{\infty} |S_x(\Omega)|^2 d\Omega \quad (4.7)$$

Вторая причина возникновения погрешностей – неидеальность восстанавливающего ФНЧ.

Таким образом, погрешность дискретизации и восстановления непрерывного сигнала определяется следующими причинами:

- 1) Спектры реальных сигналов не финитны.
- 2) АЧХ реальных ФНЧ неидеальны.

Вывод: чем выше ω_b и чем ближе характеристики ФНЧ к идеальным, тем ближе восстановленный сигнал к исходному.

5. СЛУЧАЙНЫЕ ПРОЦЕССЫ

5.1. Характеристики случайных процессов

Процессы, рассматриваемые в теории связи, могут быть детерминированными или случайными.

Детерминированные процессы – это процессы, течение которых во времени известно заранее и практически точно.

Например, гармонический сигнал $U(t) = U_m \cos(\omega_0 t + \varphi_0)$, где U_m , ω_0 , φ_0 – заданы.

Это простейшая модель информационного сигнала, она не обеспечивает точного описания его основных преобразований, что приводит к погрешностям в расчетах.

Поэтому вводится новая, более сложная модель – случайные процессы (СП). Описание случайных (или нерегулярных) процессов осуществляется с помощью случайных (вероятностных) функций.

Пример – тепловой шум, процесс случайный, т.к. его параметры неизвестны.

СП описывается своими реализациями, или выборками, совокупность реализаций образует ансамбль (полная, но очень сложная характеристика СП).

Функция распределения вероятностей СП (ФРВ)

Функция распределения вероятностей обозначается $F(x)$, характеризует вероятность того, что случайный процесс в некоторый момент времени t_1 принимает значение меньше x_1 . Полное обозначение одномерной ФРВ

$$F(x_1, t_1) = P(x < x_1, t = t_1).$$

Двумерная ФРВ

$$F_2(x_1 t_1, x_2 t_2) = P(x < x_1, t = t_1, x < x_2; t = t_2).$$

Наиболее полная характеристика **n- мерная ФРВ**:

$$F_n(x_1 t_1 \dots x_n t_n) = P(x < x_1; t < t_1; \dots x < x_n; t = t_n).$$

Функция плотности вероятностей случайного процесса (ФПВ)

В простейшем случае одномерная ФПВ равна:

$$W(x_1 t_1) = \lim_{\Delta x \rightarrow 0} \frac{P(x_1 < x < x_1 + \Delta x; t = t_1)}{\Delta x}.$$

Одномерная ФПВ равна пределу отношения вероятности попадания случайного процесса в интервал от x_1 до $x_1 + \Delta x$, при $t = t_1$, к Δx при Δx стремящемся к нулю.

Наиболее полной характеристикой является n -мерная ФПВ.

ФРВ и ФПВ связаны друг с другом. ФПВ – это первая производная ФРВ по x_1 . Соответственно, ФРВ равна интегралу от $-\infty$ до x_1 от ФПВ:

$$F(x_1 t_1) = \int_{-\infty}^{x_1} W(x_1 t_1) dx_1.$$

Условие нормировки:

$$\int_{-\infty}^{\infty} W(x_1 t_1) dx_1 = 1$$

Числовые характеристики случайного процесса

1. Среднее значение (математическое ожидание или первый начальный момент)

$$m_1 = \bar{x} = \int_{-\infty}^{+\infty} x W(x, t) dx.$$

Физический смысл m_1 – это постоянная составляющая случайного процесса.

2. Второй начальный момент.

$$m_2 = \bar{x}^2 = \int_{-\infty}^{+\infty} x^2 W(x, t) dx.$$

Физический смысл m_2 – это полная средняя мощность случайного процесса на единичном сопротивлении.

3. Дисперсия (второй центральный момент)

$$\sigma^2 = M_2 = \overline{(x - m_1)^2} = \int_{-\infty}^{+\infty} (x - m_1)^2 W(x, t) dx.$$

Физический смысл σ^2 – это средняя мощность переменной составляющей случайного процесса на единичном сопротивлении.

Числовые характеристики связаны между собой:

$$\sigma^2 = m_2 - m_1^2$$

Стационарность

1. Нестационарный случайный процесс – ФПВ и ФРВ зависят от начала отсчета времени.

2. Стационарный в узком смысле – ФПВ и ФРВ не зависят от начала отсчета времени.

3. Стационарный в широком смысле одно- и двумерные ФПВ и ФРВ не зависят от начала отсчета времени.

Для стационарного случайного процесса m_1 , m_2 , σ^2 – не зависят от времени.

Рассмотрим тепловой шум на выходе включенного усилителя. После включения усилитель прогревается и шум на его выходе – нестационарный. После «прогрева» шум будет стационарным процессом.

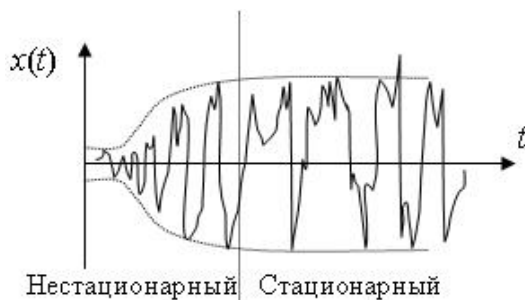


Рис. 5.1. Временная диаграмма теплового шума

Эргодичность

Случайный процесс называется эргодическим, если для него усреднение по времени одной реализации и усреднение по множеству реализаций дает один и тот же результат. Это свойство имеет большое значение на практике, т.к. усреднение по времени одной реализации технически реализовать проще, но оно не всегда дает истинный результат. Поэтому доказательство эргодичности процесса позволяет существенно упростить нахождение его характеристик.

5.2. Нормальный случайный процесс (гауссов процесс)

Процесс называется нормальным или гауссовым, если его одномерная ФПВ имеет вид:

$$W(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m_1)^2}{2\sigma^2}}$$

Графики нормальной ФПВ построены на рис. 5.2:

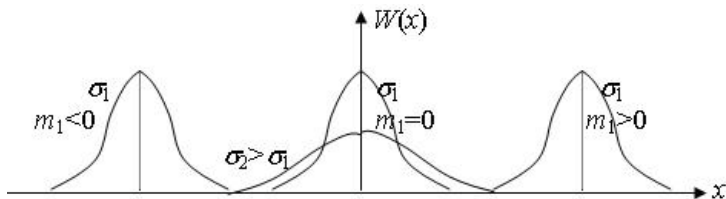


Рис. 5.2. Графики нормальной функции плотности вероятности СП:
 m_1 – среднее значение случайного процесса; σ^2 – дисперсия случайного процесса

Свойства нормального случайного процесса

1. $W(x) \geq 0$
2. Нормальная ФПВ симметрична относительно $x = m_1$
3. $W(x)$ – макс при $x = m_1$
4. Площадь под кривой $W(x)$ равна 1.
5. При изменении m_1 форма кривой не меняется, но кривая смещается вдоль оси x .
6. Чем больше дисперсия σ^2 , тем кривая ниже и шире.
7. С вероятностью близкой к 1 ($P \cong 0,997$) мгновенные значения нормального случайного процесса лежат в пределах: $m_1 - 3\sigma < x < m_1 + 3\sigma$

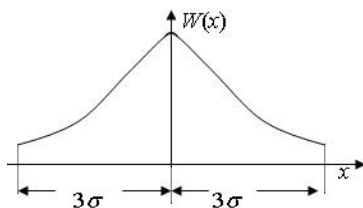


Рис. 5.3. Пределы распределения СП с вероятностью 0,997

Если известна дисперсия и m_1 , то рабочий участок ВАХ должен иметь протяженность $m_1 \pm 3\sigma$.

ФРВ для нормального случайного процесса

$$F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m_1)^2}{2\sigma^2}} dx = \left[y = \frac{x-m_1}{\sigma} \right] = \frac{1\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m_1}{\sigma}} e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m_1}{\sigma}} e^{-y^2} dy =$$

$$= F\left(\frac{x-m_1}{\sigma}\right) - \text{табулированная функция (интеграл вероятности Лапласа)}$$

$$F(0) = 0,5 \quad F(-x) = 1 - F(x)$$

$$F(3,9) = 0,99995 \quad F(-\infty) = 0; \quad F(\infty) = 1.$$

ФРВ для нормального процесса имеет вид:

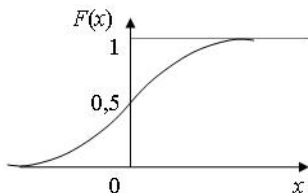


Рис. 5.4. Функция распределения вероятностей нормального процесса

5.3. ФПВ и ФРВ для гармонического колебания со случайной начальной фазой

Рассмотрим случайный процесс в виде гармонического колебания со случайной начальной фазой:

$$X(t) = A \cdot \sin(\omega t + \varphi)$$

φ – случайная величина, равномерно распределенная на интервале $\pm \pi$, т.е. ФПВ мгновенных значений фазы, показанная на рис. 5.5 равна:

$$W(\varphi) = \frac{1}{2\pi}; \quad |\varphi| \leq \pi$$

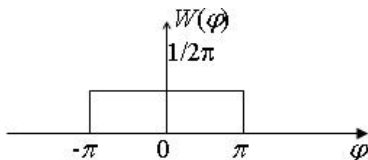


Рис. 5.5. ФПВ мгновенных значений фазы φ гармонического колебания

Вычислим среднее значение φ :

$$m_{1\varphi} = \int_{-\infty}^{+\infty} x W(x) dx = \int_{-\pi}^{\pi} \varphi \frac{1}{2\pi} d\varphi = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi d\varphi = 0.$$

Вычислим дисперсию:

$$\sigma_{\varphi}^2 = \int_{-\infty}^{\infty} (x - m_1)^2 W(x) dx = \int_{-\pi}^{\pi} \varphi^2 \frac{1}{2\pi} d\varphi = \frac{1}{2\pi} \left. \frac{\varphi^3}{3} \right|_{-\pi}^{\pi} = \frac{\pi^3}{3\pi} = \frac{\pi^2}{3}.$$

ФПВ мгновенных значений x гармонического колебания со случайной фазой, изображенная на рис. 5.6, имеет вид:

$$W(x) = \begin{cases} 0, & |x| \geq A \\ \frac{1}{\pi\sqrt{A^2 - X^2}}, & |x| < A \end{cases}$$

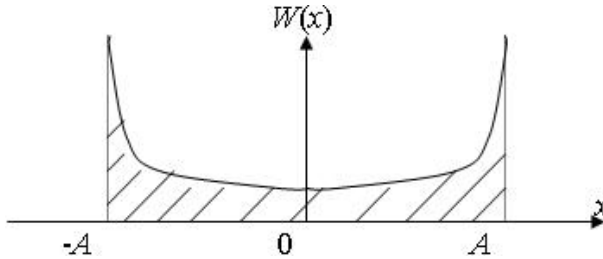


Рис. 5.6. ФПВ мгновенных значений x гармонического колебания со случайной фазой

Чем больше A , тем кривая ниже и шире. Заштрихованная площадь равна единице. Это площадь под кривой $W(x)$ (условие нормировки).

ФРВ мгновенных значений для гармонического колебания со случайной фазой:

$$X(t) = A \cdot \sin(\omega t + \varphi)$$

$$F(x) = \begin{cases} 0, & x \leq -A \\ \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{A}, & |x| \leq A \\ 1, & x > A \end{cases}$$

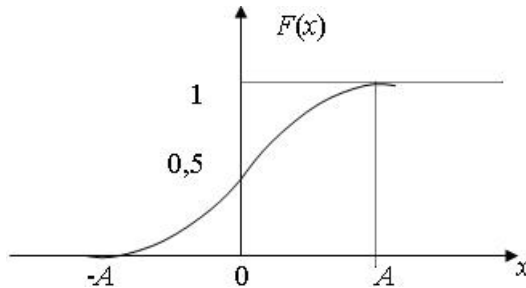


Рис. 5.7. ФРВ мгновенных значений x гармонического колебания со случайной фазой

5.4. ФПВ для суммы нормального случайного процесса и гармонического колебания со случайной начальной фазой

Рассмотрим случайный процесс $z(t)$, равный:

$$z(t) = x(t) + A \cdot \sin(\omega t + \varphi)$$

где $x(t)$ – нормальный случайный процесс;

$A \cdot \sin(\omega t + \varphi)$ – гармоническое колебание со случайной начальной фазой. $W(z)$ в этом случае находится сверткой.

$$W(z) = \int_{-\infty}^{\infty} \frac{1}{\pi\sqrt{A^2 - x^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{(z-x)^2}{2\sigma^2}} dx.$$

Вид ФПВ, т.е. $W(z)$ зависит от параметра:

$$h^2 = \frac{A^2}{2\sigma^2}$$

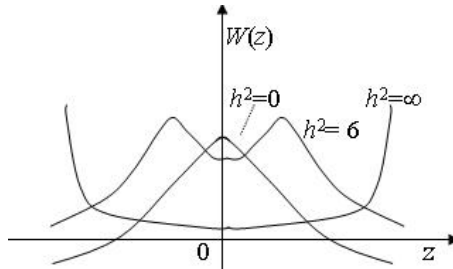


Рис. 5.8. ФПВ для суммы нормального случайного процесса и гармонического колебания со случайной начальной фазой. $h^2 = 0$ – нормальный случайный процесс (чистый шум); $h^2 \rightarrow \infty$ – одно гармоническое колебание

5.5. Огибающая и фаза узкополосного случайного процесса

Случайный процесс $y(t) = U_m(t) \cdot \cos(\omega_0 t + \varphi(t))$ называется узкополосным, если его ширина спектра значительно меньше, чем средняя частота ω_0 (рис. 5.9).

Для нормального случайного процесса фаза $\varphi(t)$ распределена равномерно (см. выше).

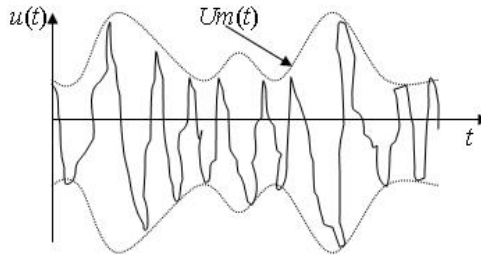


Рис. 5.9. Огибающая и фаза узкополосного случайного процесса:
 $U_m(t)$ – огибающая случайного процесса (случайная амплитуда);
 $\varphi(t)$ – фаза случайного процесса

Огибающая нормального случайного процесса $U_m(t)$ распределена по закону Релея:

$$W(U_m) = \frac{U_m}{\sigma^2} e^{-\frac{U_m^2}{2\sigma^2}}; U_m \geq 0.$$

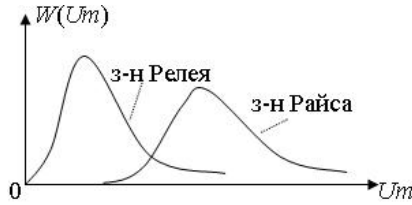


Рис. 5.10. Огибающая распределения нормального случайного процесса и суммы нормального шума и гармонического колебания

Если узкополосный случайный процесс есть сумма нормального шума и гармонического колебания с амплитудой A , то его огибающая распределена по обобщенному закону Релея (закон Райса):

$$W(U_m) = \frac{U_m}{\sigma^2} e^{-\frac{U_m^2 + A^2}{2\sigma^2}} * I_0\left(\frac{U_m A}{\sigma^2}\right) - \text{закон Райса.}$$

$I_0(\cdot)$ – функция Бесселя от мнимого аргумента.

5.6. Флуктуационный шум

Примером случайного процесса является флуктуационный шум, наиболее характерный для большинства каналов электросвязи. Для количественных расчетов воздействия флуктуационного шума на сигнал необходимо знать основные вероятностные характеристики. Поскольку шум образуется как сумма большого числа отдельных независимых ко-

лебаний, он, согласно центральной предельной теореме представляет собой стационарный эргодический случайный процесс с гауссовским (нормальным) распределением вероятности.

ПРВ гауссовского процесса описывается формулой [6, 32]:

$$W(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right],$$

в которую входят два числовых параметра m и σ^2 , имеющие смысл математического ожидания и дисперсии: $m = M(x)$, $\sigma^2 = D(X)$.

График плотности вероятности $W(x)$ представляет собой колоколообразную кривую с единственным максимумом в точке $x = m$ (рис. 5.11). Из графика видно, что с уменьшением σ кривая все более локализуется в окрестности точки $x = m$. Для флуктуационного шума обычно $M(x) = 0$.

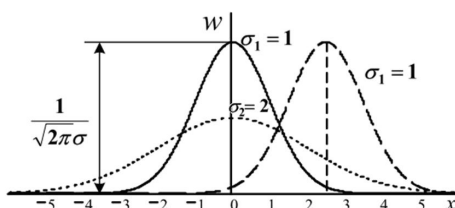


Рис. 5.11. Гауссовское распределение вероятностей

Функция распределения вероятности для гауссовского случайного процесса:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] dx.$$

После замены переменных $y = (x - m)/\sigma$ эта функция приводится к виду:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-m)/\sigma} \exp\left[-\frac{y^2}{2}\right] dy = 0,5 + \Phi_0\left(\frac{x-m}{\sigma}\right),$$

где $\Phi_0(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left[-\frac{y^2}{2}\right] dy$ — интеграл вероятности.

Функция $\Phi_0(z)$ табулирована в математических справочниках. Заметим, что $\Phi_0(-z) = -\Phi_0(z)$, $\Phi_0(0) = 0$, $\Phi_0(\infty) = 0,5$. Для приближенных вычислений можно воспользоваться приближенным выражением:

$$\Phi_0(z) \approx 0,5 - 0,65 \exp[-0,44(z + 0,75)^2]$$

Пример. Вычислим вероятность того, что мгновенное значение флуктуационного шума с нулевым средним и дисперсией $\sigma^2 = 9$ [В²] превысит уровень $x_0 = 6$ [В].

Исходя из определения функции распределения вероятности, вероятность превышения случайным процессом уровня x_0

$$p(X > x_0) = 1 - p(X \leq x_0) = 1 - F(x_0)$$

Подставляя значение $F(x_0)$ для гауссовского случайного процесса, получаем:

$$p(X > x_0) = 1 - 0,5 - \Phi_0[(x_0 - m)/\sigma] = 0,5 - \Phi_0[(x_0 - m)/\sigma]$$

Для заданных числовых значений и $m = 0$, воспользовавшись таблицами или приближенной формулой для $\Phi_0(z)$, получаем:

$$p(X > 6) \approx 2,33 \cdot 10^{-2}.$$

Обычно спектральная плотность мощности $G_x(f)$ флуктуационного шума постоянна в широком диапазоне частот, т.е. можно приближенно считать, что: $G_x(f) = N_0$ при $0 \leq f \leq \infty$. В этом случае шум называют белым. Это название дано по аналогии с белым светом, имеющим все частотные компоненты.

6. КОМПЛЕКСНОЕ ПРЕДСТАВЛЕНИЕ СИГНАЛОВ И ПОМЕХ

6.1. Понятие аналитического сигнала

Представление детерминированных сигналов рядами ортогональных функций полезно при анализе прохождения сигналов через линейные радиотехнические устройства.

При анализе нелинейных преобразований сигналов и, в частности, модуляции и демодуляции, требуется иной подход. Этот подход основывается на понятии аналитического сигнала.

Многие формулы гармонического анализа записываются значительно проще и некоторые задачи решаются легче, если использовать в качестве элементарных функций экспоненциальные функции мнимого аргумента, например комплексная огибающая легко отделяется от множителя с несущей частотой при выражении сигнала в комплексной форме.

$$A \cdot \cos[\omega_0 t + \varphi] = [A \cdot \exp(j\varphi)] \cdot \exp(j\omega_0 t).$$

Если разложить косинус суммы по формуле Эйлера [6, 21], то:

$$\cos(\omega t + \varphi) = 1/2[\exp\{j(\omega t + \varphi)\} + \exp\{-j(\omega t + \varphi)\}]. \quad (6.1)$$

Этой записи можно дать геометрическую трактовку, пользуясь представлением комплексных чисел в виде точек или векторов на плоскости.

Выражение $\exp\{j(\omega t + \varphi)\}$ представляет в данном случае вектор единичной длины, проведенный под углом $\omega t + \varphi$ к действительной оси (рис. 6.1). При изменении времени t этот вектор единичной длины меняет положение, вращаясь в положительном направлении с угловой скоростью ω .

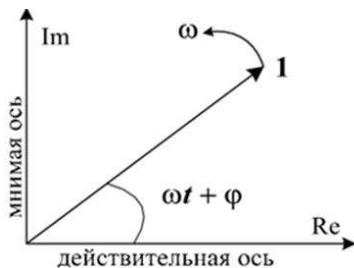


Рис. 6.1. Геометрическая трактовка экспоненциальной функции мнимого аргумента

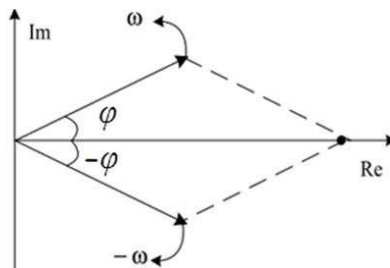


Рис. 6.2. Экспоненциальное представление элементарной функции

Изобразить синусоиду в форме (6.1), это значит представить ее суммой двух векторов, длина каждого из которых равна $1/2$, расположенных в любой момент времени симметрично относительно действительной оси и вращающихся в разных направлениях с угловыми скоростями ω и $-\omega$ (рис. 6.2).

В момент $t = 0$ они занимают положения под углами φ и $-\varphi$ относительно действительной оси. Геометрическая сумма векторов всегда совпадает по направлению с действительной осью и представляет действительную функцию времени $\cos(\omega t + \varphi)$.

При представлении косинусоиды в виде $\cos(\omega t + \varphi) = \text{Re}\{\exp\{j(\omega t + \varphi)\}\}$ можно ограничиться одним вращающимся в положительном направлении вектором и представить косинусоиду его проекцией на действительную ось.

В этом случае нет необходимости вводить отрицательные частоты. Длина вектора представляет амплитуду косинусоиды, а угол, образуемый им в данный момент с действительной осью – полную фазу $(\omega t + \varphi)$.

Проекция этого вектора на мнимую ось равна $\text{Im}\{\exp\{j(\omega t + \varphi)\}\} = \sin(\omega t + \varphi)$, т.е. представляет ту же косинусоиду, сдвинутую по фазе на $\pi/2$ (рис. 6.3).

Значительное количество сигналов применяемых в системах электросвязи можно представлять в виде:

$$s(t) = A(t) \cdot \cos[\omega t + \varphi(t)], \quad (6.2)$$

т.е. как «квазигармоническую» функцию с переменными «амплитудой» и «начальной фазой».

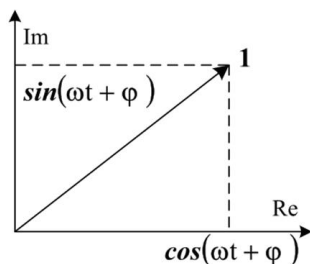


Рис. 6.3. Проекция единичного вектора на действительную и мнимую оси

Такой сигнал можно интерпретировать геометрически как проекцию на действительную ось вращающегося вектора, но при этом изменяющего свою длину и угловую скорость. Для описания свойств сигнала представленного в форме (6.2) вводят понятие комплексного аналитического сигнала.

Аналитический сигнал формируется путем отбрасывания области отрицательных частот спектра вещественного сигнала и удвоения спектра в области положительных частот. Так, если $s(t)$ – вещественный сигнал, записываем $s(t) \leftrightarrow S(f)$ и аналитический сигнал

$$s_a(t) \leftrightarrow 2\sigma(f) \cdot X(f), \quad (6.3)$$

где $\sigma(f)$ – единичная ступенчатая функция в частотной области.

Единичная ступенчатая функция равна нулю в интервале от минус бесконечности до некоторой точки и единице в интервале от этой точки до плюс бесконечности. Ступенчатая функция во временной области имеет вид

$$\sigma(t-t_0) = \begin{cases} 0 & \text{при } t < t_0; \\ 1 & \text{при } t > t_0. \end{cases} \quad (6.4)$$

Преобразование Фурье единичной ступенчатой функции равно

$$\sigma(t) \leftrightarrow \frac{1}{2} \left[\delta(f) + \frac{1}{j\pi f} \right]. \quad (6.5)$$

Таким образом, можно определить единичную ступенчатую функцию в частотной области и ее преобразование Фурье. Пусть

$$\sigma(f) = \begin{cases} 0 & \text{при } f < 0; \\ 1 & \text{при } f > 0, \end{cases}$$

тогда

$$\sigma(f) \leftrightarrow \frac{1}{2} \left[\delta(t) - \frac{1}{j\pi} \right]. \quad (6.6)$$

Заметим, что если $s(t)$ – синусоида постоянной амплитуды, то операция, указанная в выражении (6.3), состоит просто в замене вещественной синусоиды комплексной экспонентой. В более общем случае для нахождения вещественной и мнимой частей $s_a(t)$ можно преобразовать спектр в выражении (6.3) во временную область:

$$\left. \begin{aligned} \left[\delta(t) - \frac{1}{j\pi} \right] &\leftrightarrow 2\sigma(f); \\ s(t) &\leftrightarrow S(f), \\ s_a(t) &= s(t) * \left[\delta(t) - \frac{1}{j\pi} \right]. \end{aligned} \right\} \quad (6.7)$$

Свертка $s(t)$ с импульсной функцией не меняет $s(t)$, поэтому можно записать

$$s_a(t) = s(t) + \frac{j}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau \quad (6.8)$$

Вещественная часть $s_a(t)$ является исходной вещественной функцией, а мнимая часть определяется в формуле (6.8) интегралом от функции, содержащей $s(t)$. Этот интеграл называется *преобразованием Гильберта* функции $s(t)$ и обозначается

$$s^*(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau. \quad (6.9)$$

Существует следующее обратное соотношение:

$$s(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s^*(\tau)}{t - \tau} d\tau. \quad (6.10)$$

Таким образом, $s(t)$ и $s^*(t)$ представляют собой **пару преобразований Гильберта** и аналитический сигнал можно записать как

$$s_a(t) = s(t) + js^*(t). \quad (6.11)$$

Исходя из этого, аналитический сигнал в момент времени t может быть представлен точкой на комплексной плоскости, если по оси абсцисс откладывать значения реального сигнала $s(t)$, а по оси ординат – сопряженного с ним сигнала $s^*(t)$, (рис. 6.4).

Легко показать, что функция $\sin(\omega_0 t)$ является преобразованием Гильберта $\cos(\omega_0 t)$. Поэтому аналитический сигнал, соответствующий $\cos(\omega_0 t)$, имеет вид $s_a(t) = \cos \omega_0 t + j \sin \omega_0 t = \exp(j\omega_0 t)$. Аналитический сигнал общего вида удобно представлять в экспоненциальной форме как

$$s_a(t) = |s_a(t)| \exp [j\Phi(t)], \quad (6.12)$$

где $|s_a(t)| = [s^2(t) + j \cdot s^{*2}(t)]^{1/2}$;

$$\Phi(t) = \text{arctg} [s^*(t)/s(t)].$$

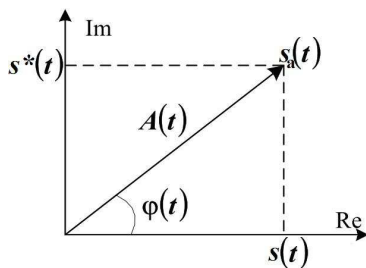


Рис. 6.4. Представление аналитического сигнала точкой

Теперь положим $\Phi(t) = \omega_0 t + \varphi(t)$ и запишем

$$s_a(t) = |s_a(t)| \exp [j\varphi(t)] \exp (j\omega_0 t) = \mu(t) \exp (j\omega_0 t). \quad (6.13)$$

Комплексная огибающая $\mu(t)$ получается удалением комплексного множителя, связанного с несущей, из аналитического сигнала:

$$\mu(t) = s_a(t) \cdot \exp(-j\omega_0 t) = |s_a(t)| \exp[j\varphi(t)]. \quad (6.14)$$

Если $\mu(t)$ – узкополосная относительно f_0 функция, то она будет обладать свойствами, которые мы интуитивно связываем с огибающей. В противном случае это просто удобное математическое представление.

Для получения спектра функции $s^*(t)$ можно применить функцию sign, которая тесно связана с единичной ступенчатой функцией и определяется как

$$\text{sign}(t - t_0) = \begin{cases} -1 & \text{при } t < t_0 \\ 1 & \text{при } t > t_0 \end{cases} = 2\sigma(t - t_0) - 1. \quad (6.15)$$

Соотношение между функцией sign и единичной ступенчатой функцией $\sigma(t)$ показано графически на рис. 6.5.

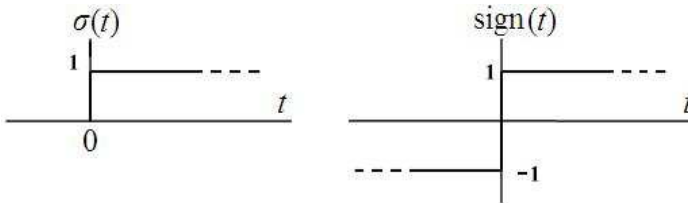


Рис. 6.5. Единичная ступенчатая функция $\sigma(t)$ и функция $\text{sign}(t)$

Функция sign может быть определена и в частотной области $\text{sign}(f)$. С помощью выражений (6.5) и (6.15) получаются следующие соотношения для пар преобразований:

Функция sign может быть определена и в частотной области $\text{sign}(f)$. С помощью выражений (6.5) и (6.15) получаются следующие соотношения для пар преобразований:

$$\left. \begin{aligned} \text{sign}(t) = 2\sigma(t) - 1 &\leftrightarrow \frac{1}{j\pi f}; \\ -\frac{1}{j\pi f} &\leftrightarrow 2\sigma(f) - 1 = \text{sign}(f). \end{aligned} \right\} \quad (6.16)$$

Таким образом, спектр функции $s^*(t)$ имеет вид:

$$\begin{aligned} 1/\pi t &\leftrightarrow -j \cdot \text{sign}(f); \\ s^*(t) = s(t) \otimes \frac{1}{2\pi} &\leftrightarrow -j \cdot \text{sign}(f) \cdot S(f) = \begin{cases} -jS(f) & \text{при } f > 0; \\ jS(f) & \text{при } f < 0. \end{cases} \end{aligned} \quad (6.17)$$

С помощью обобщения теоремы Парсеваля можно показать, что $s(t)$ и $s^*(t)$ ортогональны. Для двух заданных функций $s_1(t)$ и $s_2(t)$ обобщенное соотношение Парсеваля состоит в том, что

$$\int_{-\infty}^{+\infty} s_1(t)s_2^*(t)dt = \int_{-\infty}^{+\infty} S_1(f)S_2^*(f)df \quad (6.18)$$

откуда

$$\int_{-\infty}^{+\infty} s(t)s^*(t)dt = \int_{-\infty}^{+\infty} S(f)[-j \cdot \text{sign}(f)S(f)]^*df, \quad (6.19)$$

но $S(f)/[-j \cdot \text{sign}(f)S(f)]^* = j|S(f)|^2 \cdot \text{sign}(f)$.

В силу того, что эта функция частоты нечетна, интеграл в выражении (6.19) по всему частотному диапазону равен нулю. Поэтому

$$\int_{-\infty}^{\infty} s(t)s^*(t)dt = 0.$$

Спектральные плотности энергии $s(t)$ и $s^*(t)$ одинаковы, следовательно, полная энергия аналитического сигнала в два раза больше энергии вещественного.

6.2. Огибающая, мгновенная фаза и мгновенная частота узкополосного случайного процесса

Комплексный сигнал можно представить в форме [6]:

$$s_a(t) = A(t) \cdot e^{j\varphi(t)} = A(t) \cdot (\cos \varphi(t) + j \sin \varphi(t)), \quad (6.20)$$

где $A(t) = \sqrt{s^2(t) + s^{*2}(t)}$ называют огибающей сигнала, (6.21), а

$$\varphi(t) = \text{Arg } s_a(t) = \text{arctg} \frac{s^*(t)}{s(t)} + \begin{cases} 0, & s^*(t) \geq 0 \\ \pi, & s^*(t) < 0 \end{cases} \text{мгновенной фазой сигнала.}$$

Здесь $s(t) = A(t) \cdot \cos \varphi(t)$; $s^*(t) = A(t) \cdot \sin \varphi(t)$

Функция $\varphi(t)$ называется мгновенной фазой сигнала.

Производная от мгновенной фазы сигнала по времени называется мгновенной частотой сигнала:

$$\omega(t) = \frac{d\varphi(t)}{dt} = \frac{1}{2\pi} \cdot \frac{s^*(t) \cdot s(t) - s'(t) \cdot s^*(t)}{s^2(t) + s^{*2}(t)}. \quad (6.22)$$

Например, для гармонического сигнала [6]:

$$\omega(t) = \frac{\cos \omega_0 t \cdot \cos \omega_0 t \cdot \omega_0 \cdot A^2 + \sin \omega_0 t \cdot \sin \omega_0 t \cdot \omega_0 \cdot A^2}{A^2 \cdot t + A^2 \cdot \sin^2 \omega_0 t} = \omega_0 = \text{const}.$$

В общем случае мгновенная частота изменяется во времени.

Из (6.21) следует, что $A(t) \geq s(t)$, причем равенство достигается в моменты времени, когда $s^{*}(t) = 0$. В этих точках производная $A(t)$ совпадает с производной сигнала $s(t)$:

$$A'(t) = \frac{s(t) \cdot s'(t) + s^{*}(t) \cdot s^{*'}(t)}{\sqrt{s^2(t) + s^{*2}(t)}}; \quad A'(t) = s'(t) \text{ при } s^{*}(t) = 0. \quad (6.23)$$

Следовательно, при $s^{*}(t) = 0$, огибающая $A(t)$ касается сигнала $s(t)$.

Функция $\cos(\varphi(t))$ называется высокочастотным заполнением сигнала.

Процесс формирования сигнала на основе огибающей $A(t)$ и фазы $\varphi(t)$ показан на рис. 6.6.

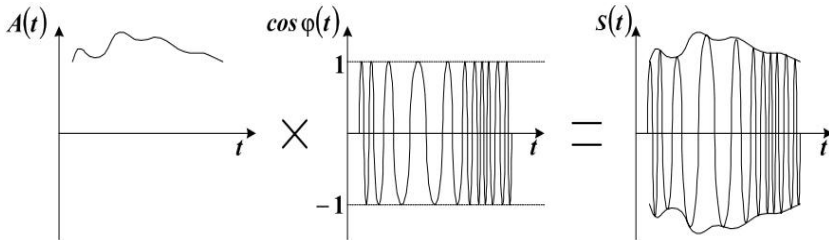


Рис. 6.6. Временное представление огибающей и высокочастотного заполнения

Если мгновенная частота колеблется вокруг среднего значения ω_{cp} , то можно записать:

$$\begin{aligned} \varphi(t) &= \omega_{cp}t + \Theta(t); \\ s(t) &= A(t) \cos[\omega_{cp}t + \Theta(t)], \end{aligned} \quad (6.24)$$

где $\Theta(t)$ – называется мгновенной начальной фазой сигнала.

Выражение (6.24) удобно для описания узкополосных сигналов. В этом случае основная часть спектра амплитуд сосредоточена в относительно узкой, по сравнению с $A(t)$ и $\varphi(t)$, полосе частот. При этом $A(t)$ и $\varphi(t)$ изменяются медленно по сравнению с $\cos(\omega_{cp}t)$. Такие сигналы называются квазигармоническими. У случайных сигналов и помех $A(t)$, $\varphi(t)$, $\omega(t)$, $\omega_{cp}(t)$ и $\Theta(t)$ являются случайными функциями времени.

7. КОРРЕЛЯЦИОННАЯ ФУНКЦИЯ ДЕТЕРМИНИРОВАННЫХ СИГНАЛОВ

7.1. Автокорреляция вещественного сигнала

Греческое слово «корреляция» обозначает связь между различными процессами и явлениями. Автокорреляцией называется согласование сигнала $s(t)$ с собственной запаздывающей версией $s(t-\tau)$. Автокорреляционная функция (АКФ) действительного вещественного сигнала $s(t)$ конечной длительности определяется следующим образом:

$$B_S(\tau) = \int_{-\infty}^{\infty} s(t) \cdot s(t-\tau) dt \quad (7.1.)$$

Автокорреляционная функция $B_S(\tau)$ дает меру схожести сигнала с собственной копией, смещенной на τ единиц времени. Переменная τ играет роль параметра сканирования или поиска.

Если сигнал $s(t)$ является периодическим с периодом T_0 , то автокорреляционную функцию следует выражать следующим образом:

$$B_S(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} s(t) \cdot s(t-\tau) dt \quad (7.2)$$

Таким образом, значение автокорреляционной функции в нуле периодического сигнала равно средней мощности сигнала.

Разделив $B_S(\tau)$ на $B_S(0)$, получим нормированную корреляционную функцию, которая имеет обозначение $R_S(\tau)$.

В системах связи для передачи сигналов широко используются видеоимпульсы прямоугольной формы. Для таких сигналов АКФ имеет вид

$$B_S(\tau) = \begin{cases} A_m^2 \cdot \tau_{И} \left(1 - \frac{|\tau|}{\tau_{И}}\right), & |\tau| \leq \tau_{И} \\ 0, & |\tau| > \tau_{И} \end{cases} \quad (7.3)$$

Энергия такого видеоимпульса равна

$$E_S = A_m^2 \cdot \tau_{И}. \quad (7.4)$$

Важным параметром сигнала $s(t)$ является длительность его АКФ, называемая интервалом корреляции τ_k . Он определяется как отношение площади, ограниченной АКФ сигнала $s(t)$, к энергии сигнала

$$\tau_k = \frac{\int_{-\infty}^{\infty} B_S(\tau) d\tau}{B_S(0)} = \frac{\int_{-\infty}^{\infty} R_S(\tau) d\tau}{B_S(0)}$$

Свойства автокорреляционной функции вещественного сигнала:

1. $B_S(\tau)$ – четная; $B_S(\tau) = B_S(-\tau)$ (симметрия по τ относительно 0);
2. $B_S(0)$ – max; $B_S(0) = \int_{-\infty}^{\infty} s(t)^2 dt$, т.е. в нуле равно энергии сигнала;
3. $B_S(\tau) \leq B_S(0)$, корреляционная функция является убывающей функцией модуля τ , т.е. $\lim_{\tau \rightarrow \pm\infty} B_S(\tau) = 0$;

4. $\tau_k = \frac{\int_0^{\infty} B(\tau) d\tau}{B(0)}$ – интервал корреляции случайного процесса, харак-

теризует ширину графика функции корреляции:

$|\tau| \leq \tau_k$ – то значения коррелированы,

$|\tau| > \tau_k$ – то значения не коррелированы.

5. $R(\tau) = B(\tau) / B(0)$ – коэффициент корреляции, $|R(\tau)| \leq 1$.

Процедура нахождения АКФ представлена на рис. 7.1, где видно, что АКФ прямоугольного видеосигнала является равнобедренным треугольником, и ее длительность равна удвоенной длительности импульса.

Если же сигнал $s(t)$ задается в виде пачки n импульсов, сдвинутых один относительно другого на время T , то максимальное значение $B_S(\tau)$ при $\tau = 0$ равно произведению энергии одного импульса на количество импульсов (рис. 7.2).

Для оценки степени связи между двумя различными сигналами $s_1(t)$ и $s_2(t)$ используется взаимная корреляционная функция (ВКФ) $B_{12}(\tau)$

$$B_{12}(\tau) = \int_{-\infty}^{\infty} s_1(t) \cdot s_2(t - \tau) dt \quad (7.5)$$

Рассмотренная выше автокорреляционная функция $B_S(\tau)$ является частным случаем функции $B_{12}(\tau)$, когда $s_1(t) = s_2(t) = s(t)$.

Необходимо также отметить, что $B_{12}(\tau)$ является асимметричной функцией относительно оси ординат и не всегда достигает максимума при $\tau = 0$.

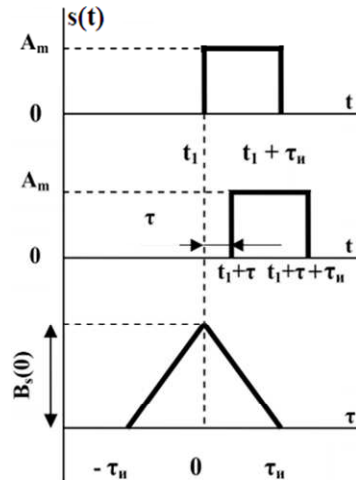


Рис. 7.1. Нахождение АКФ прямоугольного видеоимпульса

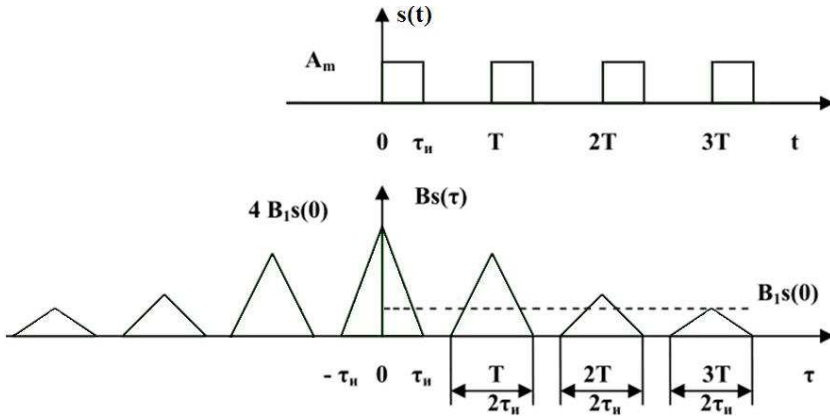


Рис. 7.2. Автокорреляционная функция пачки импульсов

7.2. Автокорреляция дискретного сигнала

По аналогии с формулой (7.1) АКФ дискретного сигнала $\{s_k\}$ и его задержанной копии на время $\tau = n\Delta t_d$ $\{s_{k-n}\}$ может быть представлена в виде

$$\widehat{B}_S(n) = \sum_{k=-\infty}^{\infty} s_k \cdot s_{k-n} \cdot \quad (7.6)$$

Эта функция, зависящая от числа тактов сдвига дискретной последовательности n , как и обычная АКФ, является четной, т.е. при нулевом сдвиге $n = 0$ дискретная АКФ определяет энергию дискретного сигнала

$$\widehat{B}_S(0) = \sum_{k=-\infty}^{\infty} s_k^2 \cdot \quad (7.7)$$

В качестве примера рассчитаем АКФ дискретного аналога импульса прямоугольной формы с единичными амплитудами $\{1, 1, 1\}$. Копии этого сигнала и значения дискретной АКФ имеют вид рис. 7.3:

$$n = 0, \{1, 1, 1\} \quad \widehat{B}_S(0) = 1 + 1 + 1 = 3$$

$$n = 1, \{0, 1, 1, 1\} \quad \widehat{B}_S(1) = 1 + 1 = 2$$

$$n = 2, \{0, 0, 1, 1, 1\} \quad \widehat{B}_S(2) = 1$$

$$n = 3, \{0, 0, 0, 1, 1, 1\} \quad \widehat{B}_S(3) = 0$$

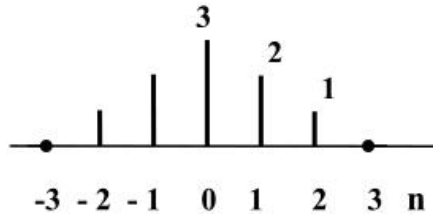


Рис. 7.3. Автокорреляционная функция дискретного сигнала

Как и в случае аналоговых видеоимпульсов, лепестки дискретной АКФ с увеличением сдвига n уменьшаются по линейному закону.

Изменим форму дискретного сигнала, так что он будет описываться в виде последовательности $\{1, 1, -1\}$. Тогда его АКФ будет иметь вид рис. 7.4:

$$\begin{aligned}
 n = 0, \{1, 1, -1\} \quad \widehat{B}_S(0) &= 1 + 1 + 1 = 3 \\
 n = 1, \{0, 1, 1, -1\} \quad \widehat{B}_S(1) &= 1 - 1 = 0 \\
 n = 2, \{0, 0, 1, 1, -1\} \quad \widehat{B}_S(2) &= -1 \\
 n = 3, \{0, 0, 0, 1, 1, -1\} \quad \widehat{B}_S(3) &= 0
 \end{aligned}$$

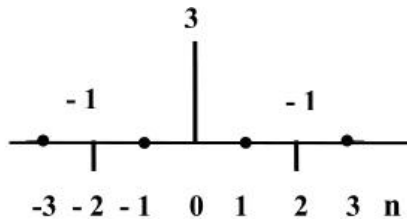


Рис. 7.4. Автокорреляционная функция дискретного сигнала

Сравнивая дискретные АКФ на рис. 7.3 и рис. 7.4, можно отметить, что именно сигнал $\{1, 1, -1\}$ имеет наиболее совершенную с точки зрения уровня боковых лепестков корреляционную функцию. Этот сигнал является простейшим из семейства сигналов Баркера, представляющих из себя M -позиционные сигналы, у которых значения боковых лепестков АКФ при $n \neq 0$ не превышают единицы. Энергия этих сигналов $\widehat{B}_S(0) = E_S$ всегда равна числу позиций M . Доказано, что число M в сигналах Баркера не может превышать числа 13. К настоящему времени известны сигналы, у которых число позиций M равно 3, 4, 5, 7, 11 и 13.

7.3. Связь корреляционной функции с энергетическим спектром

Как отмечалось в 3 разделе, преобразования Фурье обладают частотно-временной дуальностью. Это означает, в частности, что преобразование сигнала, заключающееся в нахождении его автокорреляционной функции, должно иметь дуальное соответствие в частотной области. Для установления этой связи воспользуемся выражением теоремы Парсеваля,

$$\int_{-\infty}^{\infty} s(t) \cdot u(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{U}(\omega) \cdot \dot{S}^*(\omega) d\omega, \quad (7.8)$$

в котором положим $u(t) = s(t + \tau)$ и соответственно $\dot{U}(\omega) = \dot{S}(\omega) \cdot e^{j\omega\tau}$.

Тогда получим

$$B_S(\tau) = \int_{-\infty}^{\infty} s(t) \cdot s(t - \tau) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(\omega) \cdot \dot{S}^*(\omega) \cdot e^{j\omega\tau} d\omega, \quad (7.9)$$

Учитывая, что $\dot{S}(\omega) \cdot \dot{S}^*(\omega) = |\dot{S}(\omega)|^2 = W_S(\omega)$, приходим к соотношению

$$B_S(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_S(\omega) \cdot e^{j\omega\tau} d\omega, \quad (7.10)$$

$$W_S(\tau) = \int_{-\infty}^{\infty} B_S(\omega) \cdot e^{-j\omega\tau} d\tau. \quad (7.11)$$

Итак, прямое преобразование Фурье (7.11) корреляционной функции $B_S(\tau)$ позволяет получить энергетический спектр сигнала $W_S(\omega)$ (спектральную плотность мощности), а обратное преобразование (7.10) определяет корреляционную функцию. Выражения (7.10) и (7.11) составляют суть теоремы, известной как теорема Винера-Хинчина.

Из выражений (7.10) и (7.11) можно сделать следующие выводы.

1. Чем шире спектр $|\dot{S}(\omega)|$ сигнала $s(t)$, тем меньше интервал корреляции τ_k .

2. Чем больше интервал корреляции τ_k заданного сигнала, тем меньше ширина его спектра.

3. Корреляционная функция $B_S(\tau)$ не зависит от ФЧХ спектра сигнала. Но так как форма сигнала $s(t)$ при заданном амплитудном спектре существенно зависит от ФЧХ, то можно заключить, что различным по форме сигналам, имеющим одинаковые амплитудные спектры, соответствуют одинаковые корреляционные функции $B_S(\tau)$.

Необходимо отметить, что полученная связь между автокорреляционной функцией и энергетическим спектром позволяет установить критерий существования сигнала с заданными характеристиками. Известно, что энергетический спектр любого сигнала, по определению, всегда положителен. Это означает, что корреляционная функция не может иметь, например вид прямоугольника, так как в этом случае энергетический спектр должен описываться знакопеременной функцией, что противоречит физическим представлениям.

7.4. Практическое применение корреляционной функции

Приемное устройство любого канала связи, использующего аналоговый сигнал, должно воспроизвести с наилучшей достоверностью первичный сигнал $s(t)$. Поэтому выходной сигнал такого приемника по форме должен совпадать с входным воздействием. При использовании дискретных или цифровых сигналов для передачи сообщений, приемник должен определить наличие или отсутствие импульса на определенной позиции кодовой посылки и не обязательно с сохранением формы входного импульса.

Используя выражение (7.1), можно построить приемное устройство, структурная схема которого (рис. 7.5) состоит из линии задержки 1, с помощью которой образуется копия сигнала, умножителя 2 и интегратора 3. Особенность данного способа приема, называемого корреляционным приемом, состоит в том, что выходным сигналом является корреляционная функция входного сигнала.

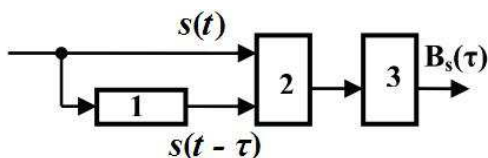


Рис. 7.5. Структурная схема корреляционного приемника

Если входной сигнал представляет последовательность прямоугольных импульсов, то на выходе корреляционного приемника существует последовательность импульсов в виде равнобедренных треугольников.

При задании некоторого уровня, который называют порогом (рассчитывается в зависимости от вида сигнала и уровня помех в канале связи), то в случае превышения его выходным сигналом регистрирующее устройство фиксирует «1», в противном случае фиксируется «0».

Введение порогового уровня позволяет избавиться от импульсных помех, амплитуда которых ниже заданного порога.

Структурная схема корреляционного приемника для приема двоичных сигналов с пассивным нулем представлена на рис. 7.6.

На вход умножителя поступает сигнал из канала связи $z(t)$, равный сумме исходного двоичного сигнала $s(t)$ и шума $n(t)$, а также копия импульсов от генератора Γ_1 , который запускается импульсами битовой синхронизацией. Отклик интегратора поступает на регистрирующее устройство РУ, которое в случае превышения порогового уровня с блока ЗП (здатчик порога) фиксирует «1» и отправляет ее в регистр оперативной памяти.

Если максимальное значение АКФ не превышает уровень порога, то на соответствующую позицию записывается «0».

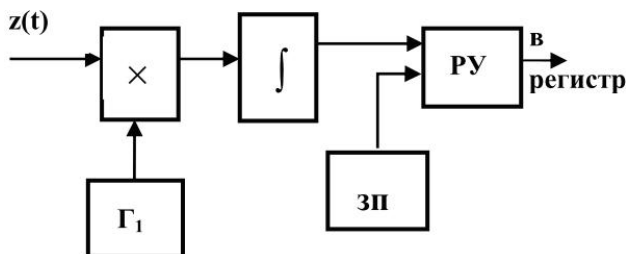


Рис. 7.6. Корреляционный приемник двоичных сигналов

II. МЕТОДЫ ФОРМИРОВАНИЯ И ПРЕОБРАЗОВАНИЯ СИГНАЛОВ

8. МОДУЛЯЦИЯ СИГНАЛОВ

8.1. Общие положения

Формирование модулированных сигналов (модуляция) предполагает взаимодействие двух сигналов: управляющего модулирующего и вспомогательного несущего. Суть управляющего воздействия модулирующего сигнала $s_c(t)$ заключается в том, что некоторые параметры несущего колебания изменяются в соответствии с модулирующим колебанием.

В системах связи в качестве управляющих колебаний используются разнообразные первичные электрические сигналы (ПЭС): телефонные, телеграфные, телевизионные и др.

В качестве несущих широко применяются гармонические сигналы, собственная частота которых ω_0 значительно превосходит верхнюю частоту спектра модулирующего колебания Ω_{\max} . Это означает, что по отношению к несущему колебанию модулирующее колебание медленно изменяет свои значения во времени. Медленность изменения $s_c(t)$ подчеркивает, что на период модулирующего колебания приходится тысячи, сотни тысяч и более периодов несущего колебания. При этом с одной стороны, обеспечивается достаточно полное отображение модулирующего колебания в несущем колебании, а с другой, обуславливается узкополосность спектра модулированного колебания.

Таким образом, для передачи информации, содержащейся в ПЭС, используется вспомогательное несущее колебание, выполняющее роль переносчика сообщения

$$S_n(t) = U \cdot \cos(\omega_n t + \varphi_n), \quad (8.1)$$

Обычно полагают $f_n \gg kF_1$, где F_1 – наивысшая гармоника ПЭС.

Процесс изменения одного или нескольких параметров высокочастотного (несущего) колебания в соответствии с первичным (модулирующим) сигналом называется модуляцией. Дискретную модуляцию обычно называют манипуляцией.

При модуляции информационными параметрами несущего колебания $S_n(t) = U \cdot \cos(\omega_n t + \varphi_n)$ могут быть амплитуда U , частота ω_n или фаза φ_n , которые изменяются в соответствии с модулирующим сигналом $s_c(t)$, поэтому различают амплитудную модуляцию (АМ), частотную модуляцию (ЧМ) и фазовую модуляцию (ФМ).

В модулируемых колебаниях изменяемые параметры имеют вид:

при амплитудной модуляции – $U(t) = U_m + \Delta U(t) = U_m + a \cdot s_c(t)$;

при частотной модуляции – $\omega_n(t) = \omega_0 + \Delta\omega(t) = \omega_0 + a \cdot s_c(t)$;

при фазовой модуляции – $\varphi_n(t) = \varphi_0 + \Delta\varphi(t) = \varphi_0 + a \cdot s_c(t)$;

где $\Delta A(t)$, $\Delta\omega(t)$, $\Delta\varphi(t)$ – приращения, пропорциональные модулирующему колебанию $s_c(t)$; a – коэффициент пропорциональности.

Устройство для получения результирующего (модулированного) сигнала $S_m(s_c, t)$, называется модулятором, (рис. 8.1), на один вход которого подается несущее (модулируемое) колебание $S_n(t)$, на второй вход первичный (модулирующий) сигнал $s_c(t)$.

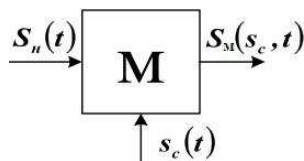


Рис. 8.1. Обобщенная схема модулятора

8.2. Амплитудная модуляция гармонического колебания

Амплитудная модуляция – процесс изменения амплитуды несущего колебания, соответствующего изменению непрерывного информационного сигнала [2].

При амплитудной модуляции мгновенная амплитуда несущего колебания:

$$U(t) = U_m + a \cdot s_c(t), \quad (8.2)$$

где U_m – амплитуда несущей; a – коэффициент пропорциональности, выбираемый так, чтобы амплитуда $U(t)$ всегда была положительной. Частота и фаза несущего гармонического колебания при АМ остаются неизменными.

Для математического описания АМ сигнала в (8.2) вместо коэффициента a , зависящего от конкретной схемы модулятора, вводится индекс модуляции:

$$m_{AM} = \frac{(U_{max} - U_{min})}{(U_{max} + U_{min})} \quad (8.3)$$

т.е. отношение разности между максимальным и минимальным значениями амплитуд АМ сигнала к сумме этих значений. Для симметричного модулирующего сигнала $s_c(t)$ АМ сигнал также симметричный, то есть $A_{\max} = A_{\min} = 2\Delta A$. Тогда индекс модуляции равен отношению максимального приращения амплитуды, к амплитуде несущей.

$$m_{AM} = \Delta U / U_m. \quad (8.4)$$

Физически индекс модуляции характеризует собой глубину амплитудной модуляции и может изменяться в пределах $0 \leq m_{AM} \leq 1$.

Таким образом, для любого АМ сигнала справедливо:

$$S_{AM}(s_c, t) = U_m [1 + m_{AM} \cdot s_c] \cos(\omega_0 t + \varphi_0). \quad (8.5)$$

Амплитудная модуляция гармоническим колебанием. В простейшем случае модулирующий сигнал является гармоническим колебанием с частотой $\Omega \ll \omega_0$. При этом выражение

$$S_{AM}(s_c, t) = U_m [1 + m_{AM} \cdot \cos \Omega t] \cos(\omega_0 t + \varphi_0), \quad (8.6)$$

соответствует однотоновому АМ сигналу, представленному на рис. 8.2в.

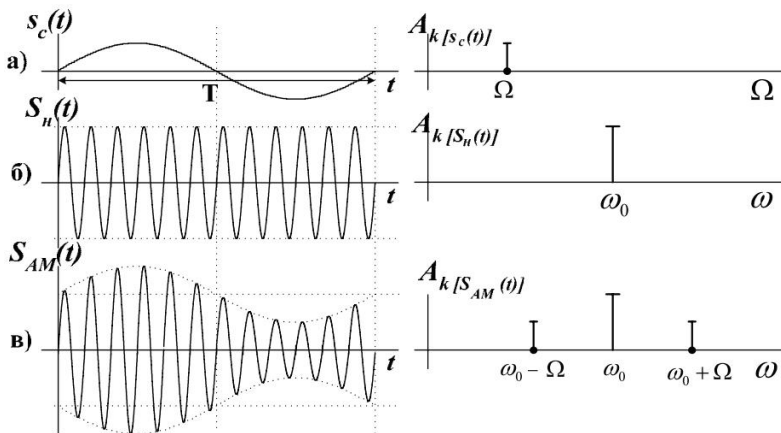


Рис. 8.2. Временные и спектральные диаграммы процесса формирования АМ гармонического колебания

Однотоновый АМ сигнал можно представить в виде суммы трех гармонических составляющих с частотами: ω_0 – несущей; $\omega_0 + \Omega$ – верхней боковой и $\omega_0 - \Omega$ – нижней боковой:

$$S_{AM}(s_c, t) = U_m \cdot \cos(\omega_0 t + \varphi_0) + (U_m \cdot m_{AM} / 2) \cos[(\omega_0 + \Omega)t + \varphi_0] + (U_m \cdot m_{AM} / 2) \cos[(\omega_0 - \Omega)t + \varphi_0] \quad (8.7)$$

Спектральная диаграмма однотонового АМ сигнала, построенная по (8.7), симметрична относительно несущей частоты ω_0 (рис. 8.2б). Амплитуды боковых колебаний с частотами $\omega_0 - \Omega$ и $\omega_0 + \Omega$ одинаковы и даже при $m_{AM} = 1$ не превышают половины амплитуды несущего колебания U_m .

Гармонические модулирующие сигналы и соответственно однотоновый АМ сигнал на практике встречаются редко. В большинстве случаев модулирующие первичные сигналы $s_c(t)$ являются сложными функциями времени (рис. 8.3а). Любой сложный сигнал $s_c(t)$ можно представить в виде конечной или бесконечной суммы гармонических составляющих, воспользовавшись рядом или интегралом Фурье. Каждая гармоническая составляющая сигнала $s_c(t)$ с частотой Ω_i приведет к появлению в АМ сигнале двух боковых составляющих с частотами $\omega_0 \pm \Omega_i$.

Множеству гармонических составляющих в модулирующем сигнале с частотами $\Omega_i, i = 1, 2, \dots, N$ будет соответствовать множество боковых составляющих с частотами $\omega_0 \pm \Omega_i, i = 1, 2, \dots, N$. Для наглядности такое преобразование спектра при АМ показано на рис. 8.3б. Спектр сложномодулированного АМ сигнала, помимо несущего колебания с частотой ω_0 , содержит группы верхних и нижних боковых колебаний, образующих соответственно верхнюю боковую полосу и нижнюю боковую полосу АМ сигнала.

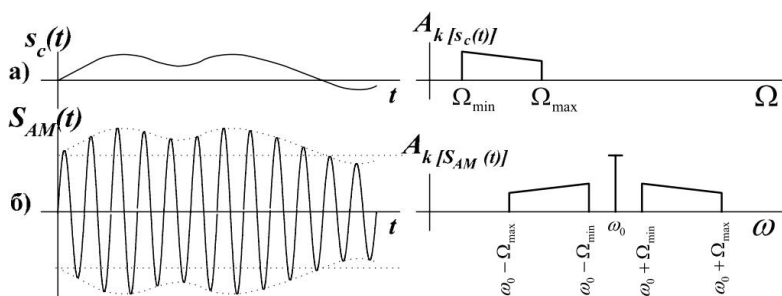


Рис. 8.3. Временные и спектральные диаграммы АМ сигнала

При этом верхняя боковая полоса частот является масштабной копией спектра информационного сигнала, сдвинутого в область высоких частот на величину ω_0 . Нижняя боковая полоса частот также повторяет спектральную диаграмму сигнала $s_c(t)$, но частоты в ней располагаются в зеркальном порядке относительно несущей частоты ω_0 .

Ширина спектра АМ сигнала $\Delta\omega_{AM}$ равна удвоенному значению наиболее высокой частоты Ω_{max} спектра модулирующего низкочастотного сигнала, т. е. $\Delta\omega_{AM} = 2\Omega_{max}$.

Наличие двух боковых полос обуславливает расширение занимаемой полосы частот примерно в два раза, по сравнению со спектром информационного сигнала. Мощность, приходящаяся на колебание несущей

шей частоты, постоянна. Мощность, заключенная в боковых полосах, зависит от индекса модуляции и увеличивается с увеличением глубины модуляции. Однако даже в крайнем случае, когда $m_{AM} = 1$, только 1/3 всей мощности колебания приходится на две боковые полосы.

Полная средняя мощность АМ сигнала на сопротивлении R равна:

$$\frac{U_m^2}{2R} \left(1 + \frac{m_{AM}^2}{2}\right) \quad (8.8.)$$

Если амплитуда несущей в спектре АМ сигнала равна U_m , глубина модуляции равна m_{AM} , то амплитуды боковых частот равны:

$$m_{AM} U_m / 2. \quad (8.9.)$$

8.3. Балансная и однополосная модуляция гармонической несущей

Балансная модуляция. Анализ спектрального состава АМ сигнала показал, что первичный модулирующий сигнал находит свое отображение лишь в составляющих боковых полос спектра АМ сигнала. В процессе отображения первичного сигнала в модулированном колебании составляющая спектра частоты ω_0 выполняет лишь роль своеобразного начала отсчета для частот боковых спектральных составляющих. Поэтому ее можно исключить из спектра передаваемого сигнала и восстановить па приемном конце.

Если модулированное колебание не содержит составляющей несущей частоты ω_0 , то модуляцию называют балансной (БМ). Такой вид модуляции целесообразен с энергетической точки зрения, поскольку на несущую частоту приходится 2/3 всей мощности модулированного колебания. При прочих равных условиях высвободившаяся мощность позволит реализовать большую дальность связи, либо при прежней дальности улучшить ее качество.

Однополосная модуляция. Балансная модуляция позволяет более рационально распределить энергию сигнала, однако ширина спектра $\Delta\Omega_{БМ}$ остается такой же, как и для обычной амплитудной модуляции. В то же время симметрия спектра АМ сигнала означает, что каждая в отдельности верхняя и нижняя боковые полосы полностью отображают модулирующее колебание. При этом вторая боковая полоса не несет никакой дополнительной информации, вдвое расширяя спектр. Вид модуляции, при котором в спектре амплитудно-модулированного колебания сохраняется лишь одна боковая полоса (верхняя или нижняя), называется однополосной модуляцией.

9. МЕТОДЫ УГЛОВОЙ МОДУЛЯЦИИ

9.1. Принципы частотной и фазовой (угловой) модуляции

При фазовой и частотной модуляции сигнал имеет постоянную амплитуду и может быть записан в следующем виде:

$$S_{\text{ФМ(ЧМ)}}(t) = U \cdot \cos(\varphi(t)). \quad (9.1)$$

В отсутствие модуляции аргумент гармонического колебания мгновенная (полная) фаза $\varphi(t) = \omega_0 t$ изменяется с постоянной скоростью ω_0 , т.е. является линейной функцией времени. И фазовая, и частотная модуляция предполагают зависимость изменения фазы $\varphi(t)$ от информационного сигнала $s_c(t)$. Эта общность позволяет объединить оба вида модуляции одним названием – угловая модуляция.

При угловой модуляции линейность изменения $\varphi(t)$ нарушается и в каждый момент времени t скорость изменения $\varphi(t)$ определяется мгновенной частотой $\omega(t)$, причем:

$$\omega(t) = \frac{d\varphi(t)}{dt}; \quad \varphi(t) = \int_0^t \omega(t) dt \cdot$$

Фазовая модуляция – процесс изменения мгновенной фазы несущего колебания пропорционально изменению непрерывного информационного сигнала:

$$\varphi(t) = \omega_0 t + \Delta\varphi(t) = \omega_0 t + a \cdot s_c(t) \quad (9.2)$$

Таким образом

$$S_{\text{ФМ}}(t) = U_m \cdot \cos[\omega_0 t + a \cdot s_c(t)]. \quad (9.3)$$

Максимальное отклонение фазы называется индексом модуляции:

$$a |s_c(t)|_{\max} = m_{\text{ФМ}}. \quad (9.4)$$

Если модуляция осуществляется гармоническим колебанием (тональная модуляция) $s_c(t) = U_{m\Omega} \cdot \cos\Omega t$ с частотой Ω , то

$$S_{\text{ФМ}}(t) = U_m \cdot \cos(\omega_0 t + a \cdot U_{m\Omega} \cdot \cos\Omega t) = U_m \cdot \cos(\omega_0 t + m_{\text{ФМ}} \cdot \cos\Omega t).$$

Заметим, что индекс модуляции $m_{\text{ФМ}} = a \cdot U_{m\Omega}$ пропорционален амплитуде модулирующего колебания.

На рис. 9.1 показано, как изменяются мгновенная частота и фаза при тональной фазовой модуляции.

Информационный однотоновый сигнал $s_c(t) = U_{m\Omega} \cdot \cos\Omega t$ (рис. 9.1а) модулирует несущее колебание $s_{\text{н}}(t)$ (рис. 9.1б), при этом закон изменения мгновенной фазы несущего колебания $\varphi(t) = \omega_0 t + a \cdot s_c(t)$

повторяет закон изменения $s_c(t)$ «косинус» (рис. 9.1б), т.е. на линейное изменение фазы (пунктир на рисунке) накладывается переменное приращение $\Delta\varphi(t) = a \cdot s_c(t)$, а закон изменения мгновенной частоты несущего колебания $\omega(t)$ (рис. 9.1з) определяется производной:

$$\omega(t) = \frac{d\varphi(t)}{dt} = \frac{d}{dt}(\omega_0 t + a \cdot U_{m\Omega} \cos \Omega t) = \omega_0 - a \cdot U_{m\Omega} \sin \Omega t.$$

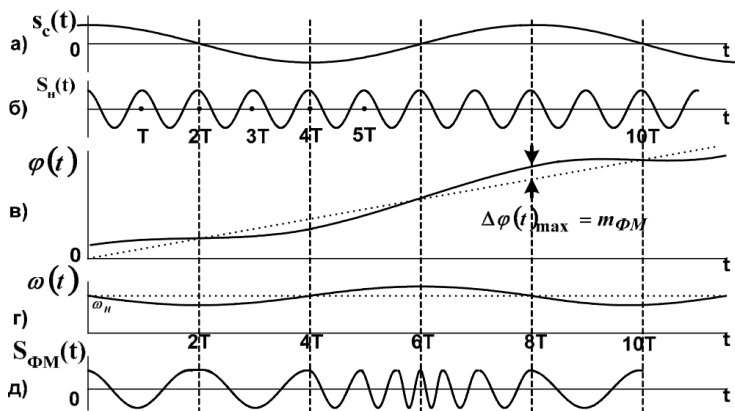


Рис. 9.1. Временные диаграммы процесса формирования ФМ сигналов

Фазомодулированное колебание (рис. 9.1д) построено на основании графика $\omega(t)$; в моменты времени $t = 2T$ и $t = 10T$ сигнал $S_{\text{ФМ}}(t)$ имеет минимальную, а в момент $t = 6T$ максимальную мгновенную частоту.

Частотная модуляция – процесс изменения мгновенной частоты несущего колебания в соответствии с изменением информационного сигнала:

$$\omega(t) = \omega_0 + a \cdot s_c(t).$$

Рассмотрим наиболее простой способ однотоновой частотной модуляции.

На рис. 9.2 изображены временные диаграммы изменения мгновенной частоты и фазы для однотоновой частотной модуляции.

Информационный однотоновый сигнал $s_c(t) = U_{m\Omega} \cdot \cos \Omega t$ (рис. 9.2а) модулирует несущее колебание $s_n(t)$ (рис. 9.2б), при этом закон изменения мгновенной частоты несущего колебания $\omega(t) = \omega_0 + U_m \cdot a \cdot \cos \Omega t$ повторяет закон изменения $s_c(t)$ (рис. 9.2в). Здесь $a \cdot U_{m\Omega} = \Delta\omega_m$ – девиация частоты, пропорциональная амплитуде модулирующего колебания $U_{m\Omega}$. Девиацией частоты называется максимальное отклонение частоты от среднего значения ω_0 :

$$a |s_c(t)|_{\max} = \Delta\omega_m. \quad (9.5)$$

Отношение девиации частоты $\Delta\omega_m$ к частоте модулирующего колебания Ω называется индексом частотной модуляции:

$$m_{\text{ЧМ}} = \Delta\omega_m / \Omega. \quad (9.6)$$

В моменты времени $t = 0, t = 8T$ мгновенная частота максимальна, в момент $t = 4T$ – минимальна. Закон изменения мгновенной фазы несущего колебания $\varphi(t)$ (рис. 9.2з) определяется интегрированием

$$\varphi(t) = \int_0^t \omega(t) dt; \quad \varphi(t) = \omega_0 t + m_{\text{ЧМ}} \sin \Omega t.$$

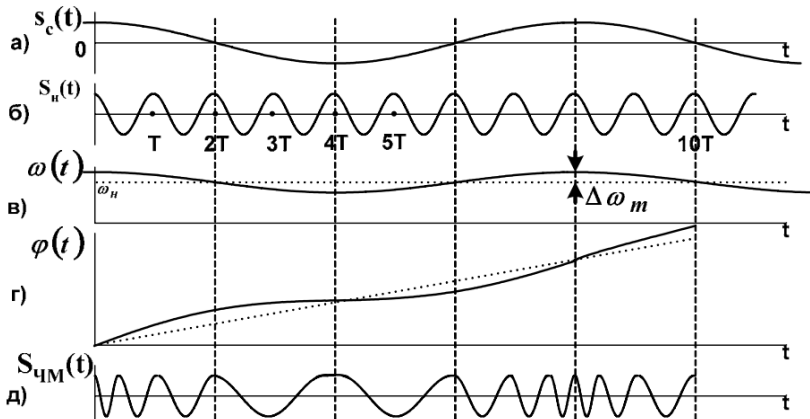


Рис. 9.2. Временные диаграммы процесса формирования ЧМ сигналов

Учитывая связь частоты и фазы, выражение для частотномодулированного сигнала запишется следующим образом:

$$S_{\text{ЧМ}}(t) = U_m \cos\left[\int_0^t \omega(t) dt\right] = U_m \cos\left[\omega_0 t + a \int_0^t s_c(t) dt\right]. \quad (9.7)$$

Для тональной частотной модуляции формула (2.14) принимает вид

$$S_{\text{ЧМ}}(t) = U_m \cdot \cos(\omega_0 t + m_{\text{ЧМ}} \cdot \sin \Omega t). \quad (9.8)$$

Сравнение выражений (9.3) и (9.7) показывает, что при ФМ приращение фазы пропорционально модулирующему колебанию $s_c(t)$, а при ЧМ – интегралу от $s_c(t)$. Если сначала проинтегрировать $s_c(t)$, а затем этим колебанием модулировать несущую по фазе, то получится ЧМ сигнал. Такой способ формирования ЧМ сигнала имеет практическое применение. Подобным же образом, если продифференцировать $s_c(t)$ и это колебание использовать для модуляции частоты, то получим ФМ сигнал.

9.2. Спектр сигналов угловой модуляции

Сигналы с угловой модуляцией, как и при АМ, могут быть представлены в виде суммы гармонических колебаний. Сравнительно просто это можно сделать для тональной модуляции. При тональной модуляции спектры ФМ и ЧМ одинаковы, если $m_{\text{ФМ}} = m_{\text{ЧМ}} = m$, поэтому будем рассматривать только спектр ЧМ сигнала.

Преобразуем (9.8) по формуле косинуса суммы двух аргументов:

$$S_{\text{ЧМ}}(t) = U_m \cdot \cos(\omega_0 t + m \cdot \sin \Omega t) = U_m \cdot \cos(\omega_0 t) \cdot \cos(m \cdot \sin \Omega t) - U_m \cdot \sin(\omega_0 t) \cdot \sin(m \cdot \sin \Omega t). \quad (9.9)$$

Из теории бесселевых функций [2] известны следующие соотношения:

$$\cos(m \sin x) = J_0(m) + 2 \sum_{k=1}^{\infty} J_{2k}(m) \cos 2kx; \quad (9.10)$$

$$\sin(m \sin x) = 2 \sum_{k=1}^{\infty} J_{2k-1}(m) \sin((2k-1)x),$$

где $J_k(m)$ – функция Бесселя k -го порядка от аргумента m . Подставляя (9.10) в (9.9), выполняя обычные алгебраические преобразования и раскрывая произведение тригонометрических функций, получаем:

$$S_{\text{ЧМ}}(t) = U_m J_0(m) \cos(\omega_0 t) + \sum_{k=1}^{\infty} U_m J_k(m) \cos(\omega_0 + k\Omega)t + \sum_{k=1}^{\infty} (-1)^k U_m J_k(m) \cos(\omega_0 - k\Omega)t. \quad (9.11)$$

Таким образом, спектр даже для однотоновой угловой модуляции является довольно сложным. В формуле (9.11) первый член – гармоническая составляющая с частотой несущей. Группа гармонических составляющих с частотами $(\omega_0 + k\Omega)$, $k = 1, 2, \dots$, определяет верхнюю боковую полосу частот, а группа составляющих с частотами $(\omega_0 - k\Omega)$, $k = 1, 2, \dots$, нижнюю боковую полосу частот.

Число верхних и нижних гармоник боковых частот теоретически бесконечно.

Боковые гармонические колебания расположены симметрично относительно ω_0 на расстоянии Ω . Амплитуды всех компонент спектра, в том числе и с частотой ω_0 , пропорциональны значениям функций Бесселя $J_k(m)$.

Формулу (9.11) можно представить в более компактном виде. Действительно учитывая $(-1)^k J_k(m) = J_k(m)$, получаем:

$$S_{\text{ЧМ}}(t) = U_m \sum_{k=-\infty}^{\infty} J_k(m) \cos(\omega_0 + k\Omega)t. \quad (9.12)$$

Для построения спектральных диаграмм необходимо знание функций Бесселя $J_k(m)$ при различных значениях k и m . Эти сведения имеются в математических справочниках [3]. На рис. 9.3 приведены графики функций Бесселя при $k = 0, 1, 2, \dots, 7$. Значения функций Бесселя, соответствующих на графиках, можно найти по рекуррентной формуле:

$$J_{k+1}(m) = \left(\frac{2k}{m}\right)J_k(m) - J_{k-1}(m).$$

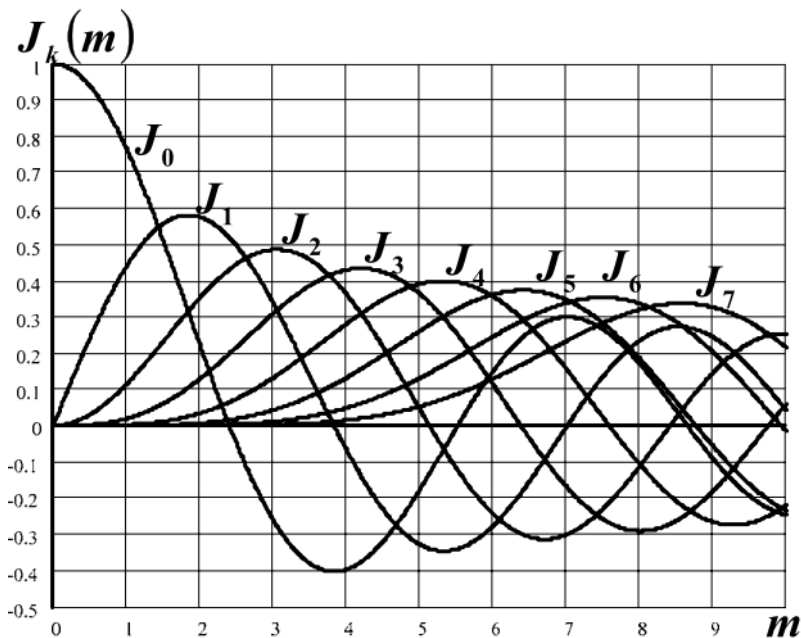


Рис. 9.3. График функций Бесселя

Пример. Задан модулированный сигнал $S(t) = 10 \cdot \cos(2 \cdot 10^6 t + 3 \cdot \cos 10^4 t)$. Построить спектральную диаграмму этого сигнала.

Из математического уравнения сигнала следует, что это однонаправленная угловая модуляция с индексом $m = 3$. Спектральные составляющие сигнала определяем из уравнения (9.11), приняв $k = 0, 1, 2, 3, \dots$, до тех пор, пока амплитуда составляющих не будет заданной, например меньше 2% от U_m . По результатам расчетов построена спектральная диаграмма (рис. 9.4).

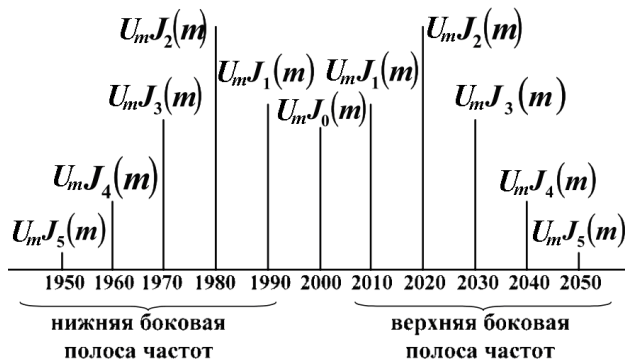


Рис. 9.4. Спектральная диаграмма сигналов с однотоновой угловой модуляцией при $m = 3$

Анализ графиков функций Бесселя показывает, чем больше порядок k функции Бесселя, тем при больших аргументах m наблюдается ее максимум, однако при $k > m$ значения функций Бесселя оказываются малой величиной. Следовательно, малыми будут и соответствующие составляющие спектра; ими можно пренебречь. Поэтому ширину спектра сигналов с угловой модуляцией можно приближенно определить по формуле:

$$\Delta\omega_{\text{УМ}} \approx 2(m + 1)\Omega, \quad (9.13)$$

где Ω – частота модулирующего сигнала. Для передачи модулированного сигнала с высокой точностью иногда считают, что надо учитывать спектральные составляющие с уровнем не менее 1% от уровня несущей. Тогда ширина спектра с угловой модуляцией $\Delta\omega_{\text{УМ}} \approx 2(m + m^{1/2} + 1)\Omega$ [21, 32, 39].

Если $m < 0,6$, то ширина спектра угловой модуляции соизмерима с шириной спектра амплитудной модуляции. Если $m \gg 1$ то при угловой модуляции из (9.13) и (9.7) следует, что ширина полосы частот примерно равна удвоенной девиации частоты.

9.3. Формирование и детектирование сигналов амплитудной и однополосной амплитудной модуляции

Устройства формирования и демодуляции радиосигналов могут быть различными в зависимости от применяемых активных элементов, способа подачи на них несущей и модулирующего сигнала.

Рассмотрим основные принципы построения модуляторов для АМ сигналов. При входном сигнале относительно небольшой мощности одним из методов формирования АМ сигналов $S_{\text{АМ}}(t)$ является операция перемножения двух колебаний: информационного $s_{\text{ПРС}}(t)$ и несущего $S_n(t)$, где в качестве перемножителя может использоваться специальная микросхема.

Операция амплитудного детектирования противоположна амплитудной модуляции. Если детектирование АМ колебания производится без опорного напряжения, то в качестве преобразующего используется нелинейный элемент (НЭ). Второй необходимый элемент детектора – фильтр низких частот (ФНЧ), который осуществляет подавление спектральных составляющих кратных несущей частоте (рис. 9.5).

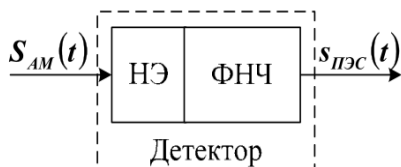


Рис. 9.5. Амплитудный детектор – в режиме больших амплитуд.

Нелинейный элемент амплитудного детектора выполняется обычно на полупроводниковом диоде. В зависимости от амплитуды АМ сигнала и степени нелинейности характеристик НЭ возможны два режима детектирования: квадратичное детектирование при малых амплитудах входного сигнала и линейное детектирование

Наряду с амплитудным, используется синхронное детектирование, которое основано на перемножении АМ сигнала и колебаний опорного генератора, совпадающего по частоте и фазе с несущей АМ сигнала, с последующим выделением низкочастотных составляющих с помощью ФНЧ. Главная трудность при синхронном детектировании заключается в получении синфазного с несущей опорного колебания. Такое колебание формируется с помощью системы фазовой автоподстройки частоты (ФАПЧ) (рис. 9.6).

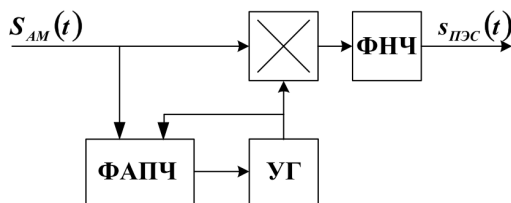


Рис. 9.6. Синхронный детектор

В системе ФАПЧ осуществляется сравнение фазы входного сигнала и управляемого генератора (УГ). При наличии отклонений осуществляется изменение фазы УГ до тех пор, пока не будет обеспечена синфазность входного сигнала и УГ.

При синхронном детектировании сигналов с балансной и однополосной модуляцией возникают принципиальные трудности в получении син-

фазного опорного напряжения. Это связано с тем, что в спектре этих сигналов несущая отсутствует. Находят применение два технических решения.

В первом случае вместе с балансной или однополосной модулированными сигналами передается так называемый пилот-сигнал, представляющий собой остаток несущей. Пилот-сигнал используется в приемнике для системы ФАПЧ опорного генератора.

Второе решение заключается в том, что для детектирования используется высокостабильный генератор несущей, отличающейся по частоте от передаваемой несущей.

В результате несинхронности несущего и опорного колебания в системах передачи с однополосной модуляцией спектр восстановленного сигнала смещается. Эти искажения снижают качество передачи первичных сигналов. Однако, как показывают экспериментальные исследования, небольшой частотный сдвиг заметного влияния на качество не оказывает. При телефонной связи абонент практически не замечает сдвига частот до 10... 20 Гц. При передаче радиовещательных программ допустимым является сдвиг частот до 2 Гц. Примерно такой же сдвиг (± 1 Гц) не сказывается на качестве факсимильной связи. Тем не менее, отсюда следуют весьма жесткие требования к стабильности генераторного оборудования систем связи с ОМ.

9.4. Формирование и детектирование сигналов угловой модуляции

Частотная модуляция (ЧМ) является основным видом модуляции в современных системах передачи информации СВЧ диапазона, в том числе системах спутниковой радиосвязи и телевидения. При ЧМ обеспечивается высокая помехоустойчивость и высокое качество передачи информации, допускается возможность одновременной работы в общем канале связи большого числа корреспондентов и реализуется более полное использование по энергетическим показателям радиопередающего устройства в силу постоянства амплитуды сигнала по сравнению с амплитудной модуляцией.

Способы осуществления частотной и фазовой модуляции можно разделить на две группы: прямые и косвенные (рис. 9.7).

Прямой метод при ЧМ означает непосредственное воздействие на автогенератор или, точнее, – на колебательную систему, определяющую частоту колебаний. Косвенный метод ЧМ состоит в преобразовании фазовой модуляции в частотную [4].

Прямой метод при ФМ означает воздействие на высокочастотный усилитель или умножитель частоты, т.е. на электрические цепи, определяющие фазу высокочастотных колебаний. Косвенный метод ФМ заключается в преобразовании частотной модуляции в фазовую.

Для преобразования фазовой модуляции в частотную на входе фазового модулятора включается интегратор (рис. 9.7з). Для преобразования частотной модуляции в фазовую на входе частотного модулятора включается дифференцирующая цепь (рис. 9.7в).

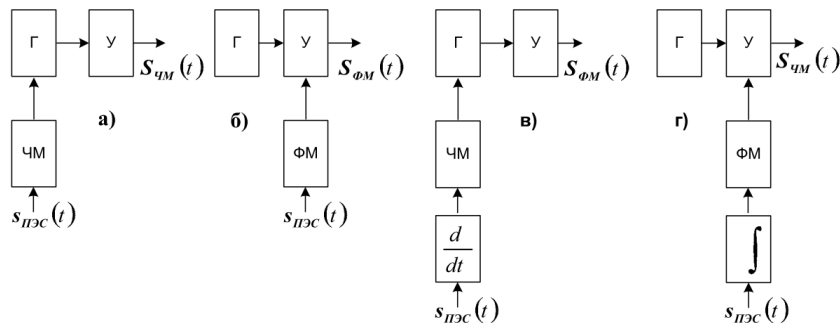


Рис. 9.7. Прямые (а, б) и косвенные (в, г) методы частотной и фазовой модуляции

В зависимости от характера преобразований различают частотно-амплитудные, частотно-фазовые и частотно-импульсные детекторы.

В частотно-амплитудных детекторах изменение частоты сигнала преобразуется в изменение амплитуды которое затем выделяется амплитудным детектором. Для того чтобы на выходе детектора не возникли искажения за счет возможных изменений амплитуды входного напряжения, перед детектированием обычно производят ограничение.

В частотно-фазовых детекторах изменение частоты преобразуется в изменение фазового сдвига между двумя напряжениями с последующим фазовым детектированием.

Фазовые детекторы преобразуют входной фазомодулированный сигнал в выходное напряжение, изменяющееся по закону модулирующего сигнала. Выявить фазовый сдвиг в ФМ сигнале можно путем сравнения с когерентным немодулированным колебанием, которое называют опорным.

Структурная схема фазового детектора аналогична схеме синхронного детектора (рис. 9.6). Все фазовые детекторы различаются по типу используемого перемножителя, наличию или отсутствию ограничителя и методам создания опорного напряжения. В качестве перемножителей можно использовать любые нелинейные или параметрические элементы – диоды, транзисторы, дифференциальные и операционные усилители с управляемой обратной связью, специальные аналоговые перемножители, ключевые схемы и др.

10. МАНИПУЛЯЦИЯ СИГНАЛОВ

10.1. Временные и спектральные характеристики амплитудно-манипулированных сигналов

При дискретном изменении управляющего колебания модулируемые параметры несущей будут изменяться скачком. В этом случае вместо термина «модуляция» применяется термин «манипуляция», а само колебание называется манипулированным. Манипуляция – это модуляция несущего колебания посылками постоянного тока прямоугольной формы.

Дискретное манипулирующее колебание может иметь вид униполярных (рис. 10.2б) или биполярных (рис. 10.2в) прямоугольных импульсов. Для описания двух возможных состояний широко используются термины «посылка» и «пауза». Эти состояния обозначают обычно символами +1 и -1 или 1 и 0.

Амплитудной манипуляцией (АМн) называется процесс изменения амплитуды несущего (высокочастотного, манипулируемого) колебания в соответствии с законом изменения амплитуды дискретного информационного (первичного электрического, манипулирующего) сигнала.

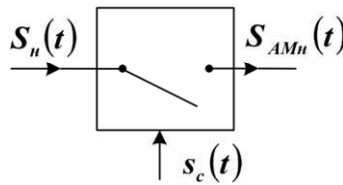


Рис. 10.1. Структурная схема амплитудного модулятора

Структурную схему получения АМн сигнала можно представить как ключ, управляемый первичным сигналом $s_c(t)$, на вход которого поступает несущий сигнал $S_n(t)$ (рис. 10.1). При этом первичный сигнал можно представить в виде отрезка ряда Фурье:

$$S_c(t) = \sum_{k=1}^{\infty} A_k \cdot \cos(\Omega_k t + \varphi_k) - \text{сигнал (рис. 10.2а)}$$

$$S_n(t) = A_m \cdot \cos(\omega_n t + \varphi_0) - \text{несущий сигнал (рис. 10.2б)}$$

Амплитудно-манипулированный сигнал имеет вид последовательности радиоимпульсов с прямоугольной огибающей (рис. 10.2в). Единичные элементы с длительностью интервалов $\tau_{и}$, соответствующих символам кодовой комбинации (1 и 0 или +1 и -1), преобразуются к виду [2]:

$$S_{\text{AM}}(t) = \frac{1}{2} A_m \cdot [1 + x_c(t)] \cdot \cos(\omega_1 t + \varphi_0), \quad (10.1)$$

где $x_c(t)$ – нормированная функция, повторяющая закон изменения $s_c(t)$ (рис. 10.2а) и принимающая значения ± 1 .

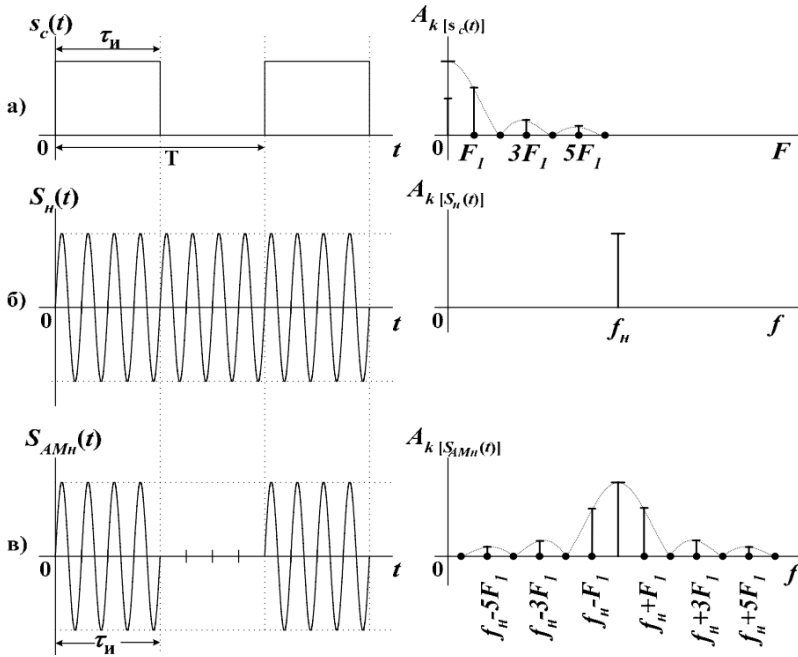


Рис. 10.2. Временные и спектральные характеристики формирования АМн сигнала

Спектральный состав периодической последовательности АМн сигналов определяется следующим выражением [2]:

$$S_{\text{AMn}}(t) = \frac{A_m \cdot \tau_n}{T} [1 + 2 \sum_{k=1}^{\infty} A_k \cdot \cos(k 2\pi F_1 t)] \cdot \cos(2\pi f_n t) = \frac{A_m \cdot \tau_n}{T} \cdot \cos(2\pi f_n t) + \frac{A_m \cdot \tau}{T} \sum_{k=1}^{\infty} \left| \frac{\sin(k\pi F_1 \tau_n)}{k\pi F_1 \tau_n} \right| \times [\cos(f_n + kF_1) 2\pi + \cos(f_n - kF_1) 2\pi]. \quad (10.2)$$

Спектр модулированного сигнала содержит в своем составе: составляющую с амплитудой $A_m \cdot \tau_n / T$ на несущей частоте f_n и две симметричные боковые полосы с частотами составляющих $(f_n + kF_1)$; $(f_n - kF_1)$ – и амплитудами $\frac{A_m \cdot \tau}{T} \cdot \frac{\sin(k\pi F_1 \tau_n)}{k\pi F_1 \tau_n}$.

Для периодических сигналов – спектр дискретный, а при случайном следовании кодовых символов (непериодических сигналов) – спектр становится сплошным.

Ширина спектра АМн колебания:

$$\Delta F_{\text{АМн}} = 2kF_1$$

где k – номер учитываемой гармоники;

$F_1 = 1/T$ – частота первой гармоники информационного сигнала.

В реальных каналах ширину спектра берут с учетом третьей или пятой гармоники, например при необходимости передать цифровой сигнал со скоростью $V = 50$ Бод, ширина спектра $\Delta F_{\text{АМн}} = 2 \cdot 5 \cdot F_1 = 5 \cdot V = 250$ Гц.

В настоящее время двоичная амплитудная манипуляция используется в низкоскоростных системах передачи информации, в многоканальных системах связи с временным разделением, в радиолокационных системах, а также в ряде оптических систем.

10.2. Временные и спектральные характеристики частотно-манипулированных сигналов

При частотной манипуляции (ЧМн) частота высокочастотного колебания изменяется скачком на величину $\pm \Delta f_m$ относительно несущей f_n (рис. 10.3). Таким образом, на выходе ЧМн вырабатываются колебания на частотах f_1 и f_2 .

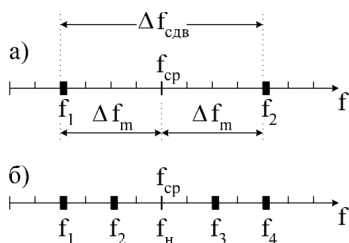


Рис. 10.3. Параметры сигналов ЧМн с разрывом фазы

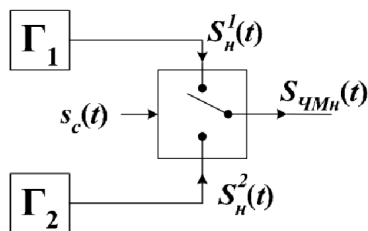


Рис. 10.4. Структурная схема формирования ЧМн колебаний

Разность частот $f_1 - f_2 = \Delta f_{\text{сдв}}$ называют частотным сдвигом. Максимальное отклонение частоты Δf_m от несущей называют девиацией.

Отношение девиации частоты Δf_m к частоте манипулирующего колебания F называется индексом частотной манипуляции. Индекс ЧМн прямо пропорционален девиации и обратно пропорционален частоте информационного сигнала: $m_{\text{ЧМн}} = \Delta f_m / F$

Различают частотную манипуляцию: с разрывом фазы и без разрыва фазы. Общий вид ЧМн сигнала с разрывом фазы можно представить в виде суммы двух АМн сигналов с разными несущими частотами f_1 и f_2 . Технически такой вид манипуляции реализуется с помощью двух генераторов (рис. 10.4), которые управляются ключом под воздействием информационного сигнала: $S_{\text{ЧМн}}(t) = S_{1\text{АМн}}(t) + S_{2\text{АМн}}(t)$.

Это представление позволяет спектр колебания $S_{\text{ЧМн}}(t)$ найти как результат наложения двух спектров колебаний АМн, который будет иметь вид [5]:

$$S_{\text{ЧМн}}(t) = \frac{A_m \cdot \tau_n}{T} \cdot \cos(2\pi F_1 t) + A_m \cdot \left(1 - \frac{\tau_n}{T}\right) \cdot \cos(2\pi F_2 t) + \frac{A_m \cdot \tau_n}{T} \cdot \sum_{k=1}^{\infty} \left| \frac{\sin(k\pi F_1 \tau_n)}{k\pi F_1 \tau_n} \right| \times (10.3)$$

$$\times [\cos(f_1 + kF_1)2\pi + \cos(f_1 - kF_1)2\pi - \cos(f_2 + kF_1)2\pi + \cos(f_2 - kF_1)2\pi].$$

Первое слагаемое определяет составляющую на частоте f_1 , второе – на частоте f_2 . Формирование ЧМн сигнала с разрывом фазы показано на рис. 10.5.

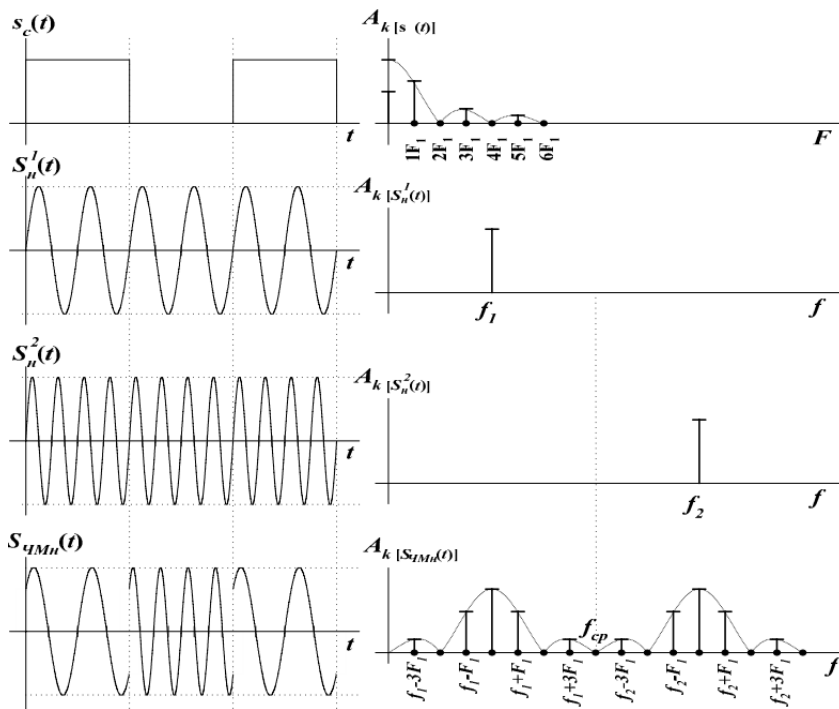


Рис. 10.5. Временные и спектральные характеристики формирования ЧМн сигнала с разрывом фазы

Из рис. 10.5 видно, что ширина спектра ЧМн сигнала отличается от спектра сигнала АМн на величину $2\Delta f_m$: $\Delta F_{\text{ЧМн}} = 2kF_1 + 2\Delta f_m$, где k – номер учитываемой гармоники.

Например, при необходимости передать цифровой сигнал со скоростью $V = 75$ бит/с, $\Delta f_m = 250$ Гц, $k = 3$, ширина спектра

$$\Delta F_{\text{ЧМн}} = 2 \cdot 3 \cdot (75/2) + 2 \cdot 250 = 725 \text{ Гц.}$$

Общий вид ЧМн сигнала без разрыва фазы (рис. 10.6) можно записать в виде [5]: $S_{\text{ЧМн}}(t) = A_m \cos[\omega_c t + \Delta\varphi(t)]$, где $\Delta\varphi(t)$ – приращение фазы, обусловленное приращением частоты $\Delta\omega(t)$.

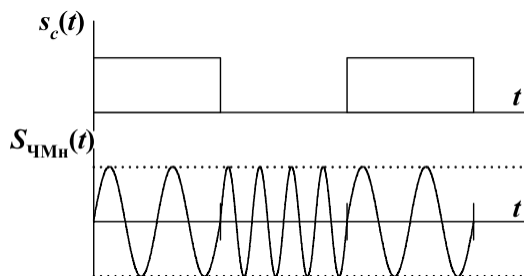


Рис. 10.6. Временные характеристики формирования ЧМн колебаний без разрыва фазы

Этот вид манипуляции предполагает использование одного источника колебаний (рис. 10.7), частота которого изменяется посредством управляемой реактивности (в этом случае фаза изменяется непрерывно – без разрыва).

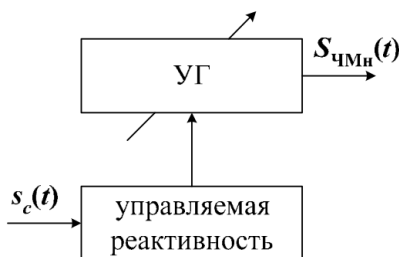


Рис. 10.7. Структурная схема формирования ЧМн колебаний без разрыва фазы

Спектральный состав ЧМн сигнала без разрыва фазы можно получить, раскрывая выражение для $S_{\text{ЧМн}}(t)$:

$$S_{\text{ЧМн}}(t) = A_m [\cos \omega_c t + \cos \Delta\varphi(t) - \sin \omega_c t \cdot \sin \Delta\varphi(t)].$$

Из этой формулы следует, что для нахождения спектра ЧМн сигнала необходимо определить спектр функций $\cos \Delta \varphi(t)$ и $\sin \Delta \varphi(t)$ разложив их в ряд Фурье:

$$S_{\text{ЧМн}}(t) = \frac{2A_m \cdot m}{\pi} \cdot \sum_{k=-\infty}^{\infty} \frac{\sin[0,5\pi(m+k)]}{m^2 - k^2} \cdot \cos(f_i + kF_1)2\pi \cdot \quad (10.4)$$

Из спектральной характеристики (10.4) видно, что для спектра при $m_{\text{ЧМн}} \ll 1$ энергия колебания находится вблизи f_n . Спектр ограничен несущей и двумя боковыми частотами, а ширина спектра равна ширине спектра АМн сигнала [2, 5]:

$$\begin{aligned} S_{\text{ЧМн}}(t) = & A_m \cdot \frac{\sin(0,5\pi m)}{0,5\pi m} \cdot \cos \omega_1 t + \\ & + \frac{2A_m \cdot m}{\pi} \cdot \sum_{k=2,4,\dots}^{\infty} \frac{\sin(0,5\pi m)}{m^2 - k^2} \cdot [\cos(\omega_1 + k\Omega)t + \cos(\omega_1 - k\Omega)t] - \\ & - \frac{2A_m \cdot m}{\pi} \cdot \sum_{k=1,3,\dots}^{\infty} \frac{\sin(0,5\pi m)}{m^2 - k^2} \cdot [\sin(\omega_1 + k\Omega)t + \sin(\omega_1 - k\Omega)t]. \end{aligned} \quad (10.5)$$

По мере увеличения индекса частотной модуляции энергия концентрируется вблизи частот f_1 и f_2 . На рис. 10.8 приведены спектры колебаний при различных $m_{\text{ЧМн}}$.

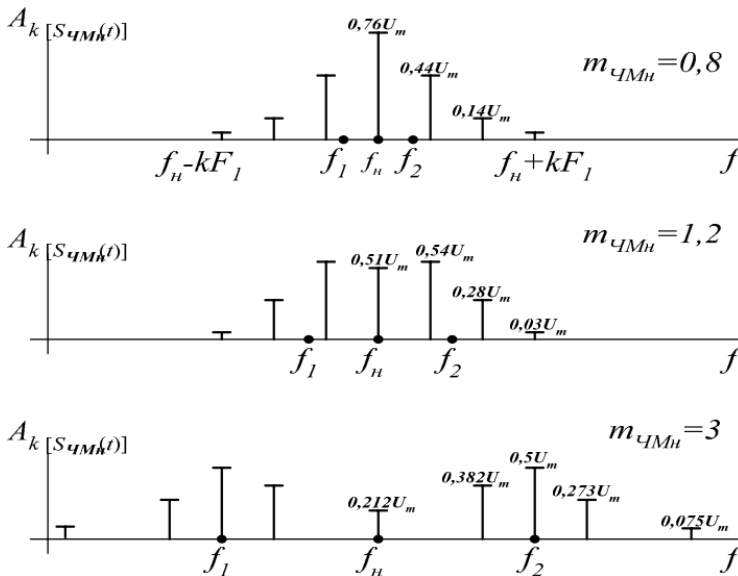


Рис. 10.8. Спектральные характеристики ЧМн сигнала без разрыва фазы для различных индексов модуляции

Ширина спектра определяется по общей формуле [2, 5]:

$$\Delta F_{\text{чМн}} = 2(\Delta f_m + \Delta F) = 2F(m - 2) = 2\Delta f_m(1 + 2/m), \quad (10.6)$$

либо по формулам для различных $m_{\text{чМн}}$:

$$\Delta F_{\text{чМн}} = \begin{cases} (1,3 \cdot m_{\text{чМн}} + 1,4) \cdot V; & 2 \leq m_{\text{чМн}} \leq 8 \\ (1,1 \cdot m_{\text{чМн}} + 1,6) \cdot V; & 8 \leq m_{\text{чМн}} \leq 20 \end{cases} \quad (10.7)$$

где V – скорость телеграфирования в бодах.

10.3. Фазовая (относительно-фазовая) манипуляция сигналов

В настоящее время разработано несколько вариантов двухпозиционной (бинарной) и многопозиционной фазовой манипуляции. В радиосистемах передачи информации наиболее часто применяются двоичная, четырех позиционная и восьми позиционная фазовая манипуляция (ФМн). Данные сигналы обеспечивают высокую скорость передачи, применяются в радиосвязи, в системах фазовой телеграфии, при формировании сложных сигналов.

Временные и спектральные характеристики фазоманипулированных сигналов

Наиболее простой является бинарная ФМн, при которой изменение фазы несущего колебания происходит скачком в определенные моменты первичного сигнала (рис. 10.9а) на 0 или 180°; при этом его амплитуда и частота несущей остаются неизменными.

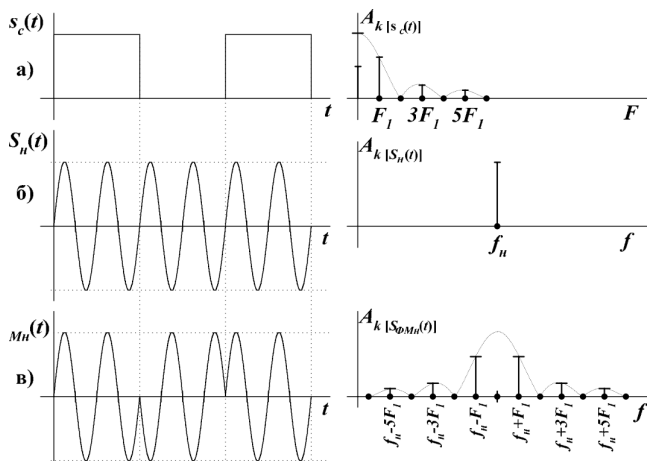


Рис. 10.9. Временные и спектральные характеристики формирования ФМн сигнала

ФМн сигнал имеет вид последовательности радиоимпульсов (отрезков гармонических колебаний) с прямоугольной огибающей (рис. 10.9в) [5, 6]:

$$S_{\text{ФМн}}(t) = A_m \cos[\omega_c t + (1 + x_c(t)) \cdot \Delta\varphi_m], \quad (10.8)$$

где $x_c(t)$ – нормированная функция, принимающая значения -1 и 1, и повторяющая изменения информационного сигнала (рис. 10.9а); $\Delta\varphi_m$ – девиация фазы (максимальное отклонение фазы от начальной).

Величина $\Delta\varphi_m$ может быть любой, однако, для лучшего различения двух сигналов на приеме целесообразно, чтобы они максимально отличались друг от друга по фазе, т.е. на 180° ($\Delta\varphi_m = \pi$).

Таким образом, одни из ФМн колебаний будут синфазны с колебаниями несущей, а другие противоположны по фазе на 180° .

Такой сигнал можно представить в виде суммы двух АМн сигналов, с противофазными несущими 0° и 180° : $S_{\text{ФМн}}(t) = S_{1\text{АМн}}(t) + S_{2\text{АМн}}(t)$.

Структурная схема модулятора в этом случае реализуется с помощью двух самостоятельных источников колебаний (генераторов) с разными начальными фазами, выходы которых управляются информационным сигналом с помощью ключа (рис. 10.10).

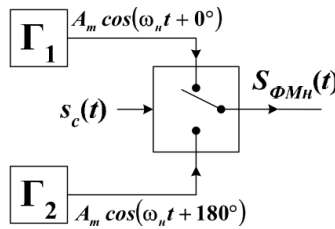


Рис. 10.10. Структурная схема формирования ФМн колебаний

Спектр ФМн колебания находится суммированием спектров колебаний $S_{1\text{АМн}}(t)$ и $S_{2\text{АМн}}(t)$ [2, 5, 6]:

$$S_{\text{ФМн}}(t) = A_m \left(\frac{2\tau_n}{T} - 1 \right) \cdot \cos(2\pi f_n t) + 2A_m \frac{\tau_n}{T} \cdot \sum_{k=1}^{\infty} \left| \frac{\sin(k\pi F_1 \tau_n)}{k\pi F_1 \tau_n} \right| \times \quad (10.9)$$

$$\times [\cos(f_n + kF_1)2\pi t + \cos(f_n - kF_1)2\pi t].$$

Из формулы следует, что спектр колебаний ФМн в общем случае содержит несущее колебание, верхнюю и нижнюю боковые полосы, состоящие из оставяющих частот $(k2\pi f_n \pm k2\pi F_1)t$.

Анализ спектров ФМн сигналов (рис. 10.9) при различных значениях $\Delta\varphi_m$ показывает, что при изменении $\Delta\varphi_m$ от 0 до π происходит перераспределение энергии сигнала между несущим колебанием и боковыми составляющими, а при $\Delta\varphi_m = \pi$ вся энергия сигнала содержится толь-

ко в боковых полосах. Из рис. 10.11 следует, что спектр амплитуд ФМн сигнала содержит те же составляющие, что и спектр АМн сигнала, а для скважности $T/\tau_{\text{н}} = 2$ составляющая на несущей частоте отсутствует. Амплитуды боковых составляющих ФМн сигнала в 2 раза больше, чем АМн сигнала.

Это объясняется наложением 2-х спектров – спектра ФМн сигнала и несущей. На интервале, где колебания синфазны, суммарная амплитуда удваивается, а где фазы противоположны, компенсируется, в результате для нахождения спектра ФМн достаточно определить спектр АМн колебания.

Равенство полос частот АМн и ФМн сигнала предполагает также и равенство максимально возможных скоростей модуляции. Большая амплитуда спектральных составляющих ФМн сигнала по сравнению с АМн обуславливает большую помехоустойчивость.

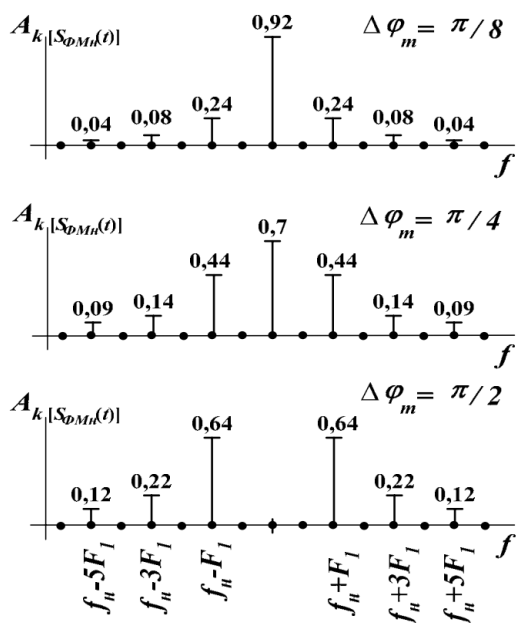


Рис. 10.11. Спектры сигналов фазовой манипуляции при различных значениях девиации фазы

При ФМн начальная фаза является информационным параметром, и в алгоритмах работы фазового демодулятора с целью получения сведений о начальной фазе должны формироваться и храниться образцы вариантов передаваемого сигнала, достаточно точно совпадающие с ним по частоте и начальной фазе. Но на приеме нет признаков по кото-

рым можно точно установить однозначное соответствие между переданными двоичными символами и образцами сигнала на входе демодулятора, в результате возможно явление так называемой «обратной работы».

Неопределенность начальной фазы объясняется с одной стороны тем, что в канале связи к переданной фазе добавляется произвольный и неизвестный фазовый сдвиг. С другой стороны, фаза сигнала всегда приводится к интервалу 2π и сигналы, различающиеся по фазе на 2π , для приемника одинаковы.

Данное свойство неоднозначности решения характерно именно для ФМн. При АМн сигнал, прошедший канал связи, также отличается от переданного, однако если на выходе модулятора сигналу с большей амплитудой соответствовал некоторый двоичный символ, то и на входе демодулятора варианту сигнала с большей амплитудой будет соответствовать тот же самый символ – неоднозначность отсутствует. При ЧМн ситуация аналогична. Если одна из двух частот больше другой на выходе модулятора, то после всех преобразований в канале она останется больше и на входе демодулятора.

Временные характеристики сигналов с относительной фазовой манипуляцией

Неоднозначность характерная для ФМн сигналов, устранена в системах относительно-фазовой манипуляции (ОФМн). У такого метода манипуляции информация заложена не в абсолютном значении начальной фазы, а в разности начальных фаз соседних посылок, которая остается неизменной и на приемной стороне. Для передачи первого двоичного символа в системах с ОФМн необходима одна дополнительная посылка сигнала, передаваемая перед началом передачи информации и играющая роль отсчетной.

Процесс формирования сигнала с ОФМн можно свести к случаю формирования сигнала с ФМн путем перекодирования передаваемой двоичной последовательности. Алгоритм перекодировки прост: если обозначить $s_c^n = \pm 1$ как информационный символ, подлежащий передаче на n -м единичном элементе сигнала, то перекодированный в соответствии с правилами ОФМн символ $s_{\text{отн}}^n$ определяется следующим рекуррентным соотношением: $s_{\text{отн}}^n(t) = s_c^n(t) \cdot s_{\text{отн}}^{n-1}(t)$. Для получения сигнала с ОФМн достаточно умножить полученный (перекодированный) сигнал $s_{\text{отн}}^n(t)$ на несущее колебание. Структурная схема модулятора для ОФМн (рис. 10.12) содержит генератор несущего колебания, перемножитель (ФМ) и перекодирующее устройство (относительный кодер), состоящее из перемножителя и элемента памяти.

Демодулятор сигнала с ОФМн содержит фазовый детектор, состоящий из перемножителя и ФНЧ, на который подается опорное колебание, совпадающее с одним из вариантов принимаемого сигнала. Дальнейшее вычисление разности фаз и определение переданного ПЭС

осуществляется перемножением сигналов на выходе детектора, задержанных друг относительно друга на длительность единичного интервала.

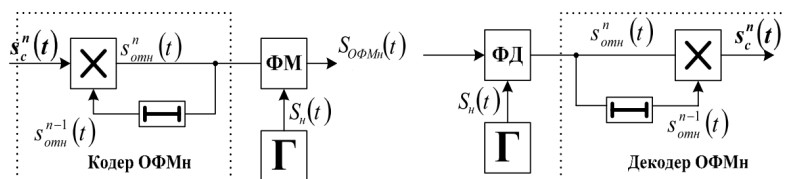


Рис. 10.12. Модулятор и демодулятор ОФМн

На рис. 10.13 представлены временные диаграммы формирования сигналов ОФМн: а) непериодический информационный сигнал; б) информационный сигнал в относительном коде; в) несущее колебание; г) сигнал ОФМн на выходе модулятора.

Алгоритмы демодуляции сигналов с ОФМн в сравнении с ФМн иллюстрируются временными диаграммами на рис. 10.14 и 10.15.

На рис. 10.15 представлены временные диаграммы демодуляции сигналов ОФМн и ФМн при однократной ошибке в принятом радиосигнале, в качестве исходного информационного взят сигнал (рис. 10.14а).

Случай возникновения скачка фазы в опорном колебании представлен на рис. 10.15. При этом в опорное колебание специально введен скачок фазы на 180° между 2-й и 3-й посылками.

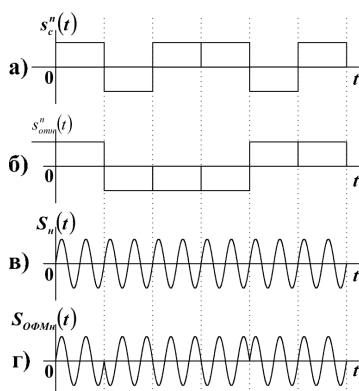


Рис. 10.13. Временные диаграммы формирования сигналов ОФМн: а) сигнал с ОФМн на выходе модулятора; б) сигнал с ОФМн на входе демодулятора, в принятый сигнал специально введена ошибка для 3 посылок; в) опорное колебание; г) принятый информационный сигнал, на выходе относительного декодера; д) принятый информационный сигнал, на выходе демодулятора; е) принятый информационный сигнал, на выходе демодулятора в случае отсутствия ошибки

Это дает возможность проиллюстрировать появление ошибок в системах с ФМн и ОФМн. В системе с ФМн, после изменения полярности опорного колебания, все последующие символы ошибочные (обратная работа), причем ошибка будет оставаться до следующего скачка фазы опорного колебания. В системе с ОФМн скачкообразное изменение полярности опорного колебания приводит к одиночной ошибке, что и определяет преимущества сигналов с ОФМн.

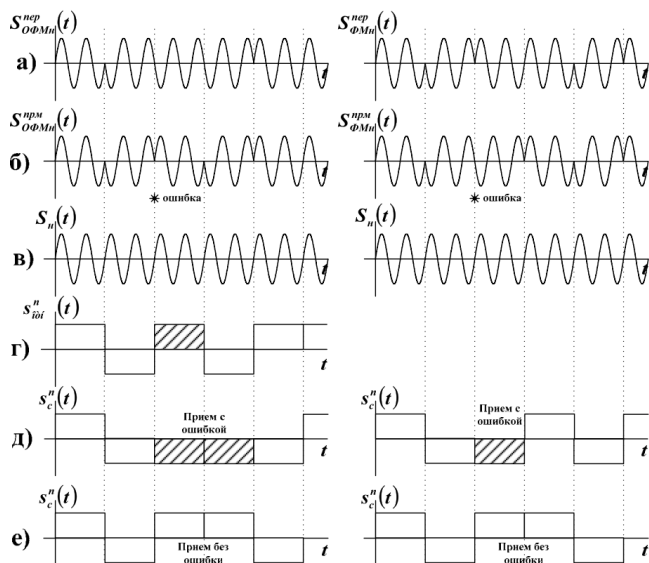


Рис. 10.14. Временные диаграммы демодуляции сигналов ОФМн и ФМн при одной ошибке в принятом радиосигнале

Однако следует отметить недостатки систем с ОФМн, которые следует учитывать при выборе методов модуляции:

- необходимость передачи отсчетной посылки в начале сеанса связи;
- увеличение вероятности ошибки примерно вдвое;
- появление двойных ошибок в цифровом потоке, что усложняет кодек при использовании корректирующих кодов;
- сложность построения модема для ОФМн по сравнению с модемом для ФМн.

Для реализации системы с ФМн необходима передача специального синхросигнала (маркерного сигнала), соответствующего одному из символов, например 0. Другой путь реализации ФМн – применение специальных кодов с избыточностью, позволяющих обнаруживать ошибки типа инвертирования всех символов. Все это ведет к определенным по-

терям – энергетическим, скоростным и аппаратурным. Поэтому при выборе метода модуляции ФМн или ОФМн необходимо учитывать их достоинства и недостатки.

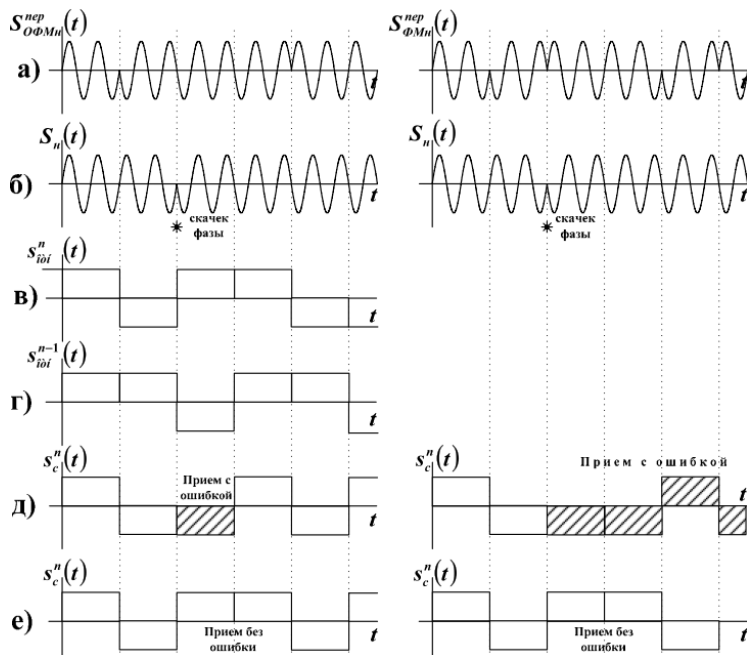


Рис. 10.15. Временные диаграммы демодуляции сигналов ОФМн и ФМн при изменении полярности опорного колебания

III. АЛГОРИТМЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ

11. ОСНОВЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ

11.1. Общие понятия о цифровой обработке

Под обработкой сигналов в широком смысле можно понимать совокупность преобразований, направленную на наиболее эффективную передачу, хранение и извлечение информации. В последние десятилетия все более широко применяется цифровая обработка сигналов (ЦОС), которой свойственны следующие преимущества перед аналоговой обработкой:

- принципиальная возможность реализации практически любых алгоритмов обработки в реальном масштабе времени (возможности аналоговой техники значительно скромнее);
- потенциально сколь угодно высокая точность реализации алгоритмов, определяемая разрядностью цифровых устройств;
- возможность безошибочного воспроизведения сигналов при передаче и хранении на основе помехоустойчивого кодирования, которое применимо только к цифровым сигналам.

Преимущества цифровой обработки основываются на свойствах дискретных сигналов и цепей, которые во многом сходны с аналоговыми, но в то же время имеют и существенные особенности [7].

Под цифровым сигналом понимают дискретный сигнал, квантованный по уровню.

Математической моделью дискретного сигнала служит последовательность $s[n]$ (дискретный аргумент принято заключать в квадратные скобки), где s – аргумент, принимающий значения из дискретного множества, а функция $s[\bullet]$ может принимать значения из непрерывного множества вещественных или комплексных чисел (в пространственно-временной обработке сигналов функция $s[\bullet]$ принимает векторные значения). В соответствии с теоремой отсчетов, аналоговый сигнал $s(t)$ с финитным спектром, сосредоточенным в полосе частот $(-F_B, F_B)$, может без потерь информации быть заменен дискретной последовательностью

$s[n] = s(nT_d)$ своих значений, взятых с шагом $T_d < 1/2F_b$, которая и представляет собой дискретный сигнал.

Таким образом, цифровой сигнал – это последовательность, принимающая значения из дискретного (как правило, конечного) множества. Это связано с тем, что цифровые устройства всегда имеют ограниченную разрядность и отсчеты сигналов, подлежащих цифровой обработке округляются (квантуются). Для изучения ряда вопросов цифровой обработки сигналов удобнее считать, что сигнал принимает значения из непрерывного множества, поэтому используется модель дискретного сигнала. Моделью цифрового сигнала пользуются в тех случаях, когда рассматриваются специфические эффекты, связанные с квантованием сигнала, округлением промежуточных результатов, ограничением разрядной сетки цифрового устройства и т.п.

Для последовательности (дискретного сигнала) $s[n]$, где n принимает целые значения от $-\infty$ до $+\infty$ можно определить дискретное преобразование Фурье

$$S(e^{j\omega}) = \sum_{n=-\infty}^{+\infty} s[n]e^{-j\omega n}, \quad (11.1)$$

Легко заметить, что изменение ω на величину $\pm k \cdot 2\pi$ при любом целом k никак не влияет на результат преобразования. Таким образом, величину ω можно понимать, как угол, а $e^{j\omega}$ – как точку на комплексной плоскости, находящуюся на окружности единичного радиуса (рис. 11.1). Поэтому выражение (11.1) определяет на единичной окружности функцию вещественной переменной ω , которая имеет смысл круговой частоты.

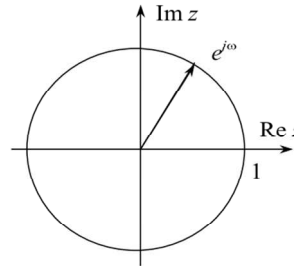


Рис. 11.1. Угловая интерпретация частоты

Вспомним, что при восстановлении аналогового сигнала моделью дискретного сигнала служит идеализированный АИМ-сигнал, состоящий из δ -функций, умноженных на отсчеты сигнала

$$s(t) = \sum_{n=-\infty}^{+\infty} s(nT_d) \delta(t - nT_d).$$

Преобразование Фурье этого аналогового сигнала, обозначив круговую частоту в его спектральном описании буквой Ω , находится как:

$$\begin{aligned} S(\Omega) &= \int_{-\infty}^{+\infty} s(t) \exp(-j\Omega t) dt = \sum_{n=-\infty}^{+\infty} s(nT_d) \int_{-\infty}^{+\infty} \delta(t - nT_d) \exp(-j\Omega t) dt = \\ &= \sum_{n=-\infty}^{+\infty} s(nT_d) \exp(-j\Omega nT_d). \end{aligned} \quad (11.2)$$

Сравнивая выражения (11.1) и (11.2), легко увидеть, что при условии

$$\begin{aligned} s[n] &= s(nT_d), \\ \omega &= \Omega T_d, \quad -\pi \leq \omega \leq \pi \end{aligned} \quad (11.3)$$

их левые части совпадают. Это означает, что выражение (11.1) определяет спектральную плотность дискретного сигнала, совпадающую по форме со спектральной плотностью идеального АИМ-сигнала (который при воздействии на идеальный ФНЧ с П-образной характеристикой позволяет точно восстановить исходный аналоговый сигнал).

Из условия (11.3) следует необходимое ограничение $-\pi/T_d \leq \Omega \leq \pi/T_d$, или, что то же самое, $-\Omega_d/2 \leq \Omega \leq \Omega_d/2$, где $\Omega_d = 2\pi F_d = 2\pi/T_d$ – круговая частота дискретизации. Иными словами, мы снова получили условие выбора частоты дискретизации, как минимум, вдвое выше верхней частоты спектра аналогового сигнала.

Таким образом, при условии финитности и правильного выбора шага дискретизации, любые действия над дискретным сигналом эквивалентны соответствующим действиям над аналоговым сигналом и обработка сигнала может производиться в цифровой форме.

Рассмотрим выражение (11.1) как разложение 2π -периодической функции аргумента ω в комплексный ряд Фурье по базисным функциям $\exp(jn\omega)$, $-\infty \leq n \leq \infty$. Тогда, очевидно, отсчеты $s[n]$ не что иное, как коэффициенты этого ряда и могут быть найдены по общей формуле для вычисления коэффициентов комплексного ряда Фурье:

$$s[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(e^{j\omega}) e^{j\omega n} d\omega, \quad -\infty \leq n \leq \infty, \quad (11.4)$$

Это выражение представляет собой обратное преобразование Фурье для последовательности (дискретного сигнала).

11.2. Квантование сигнала

Операция квантования непрерывной величины состоит в том, что множество ее возможных значений заменяется определённым конечным числом значений. Существующие устройства квантования обычно осуществляют равномерное квантование сигналов, при котором границы интервалов квантования размещаются равномерно в заданном диапазоне значений сигнала, а значения уровней квантования располагаются посередине между этими границами. При равномерном квантовании количество порогов квантования оценивается величиной

$$r = (u_{\max} - u_{\min})/\Delta u,$$

где u_{\max} и u_{\min} – максимальная и минимальная амплитуды дискретизируемого сигнала. Пороги квантования разбивают интервал $(u_{\max} - u_{\min})$ на $(r + 1)$ подинтервалов – уровней квантования.

Отсчет непрерывного процесса в АЦП преобразуется в двоичный код из m разрядов, каждый из которых представлен нулем или единицей. Число разрядов определяется числом уровней квантования:

$$m = [\log_2(r + 1)].$$

При обработке, когда требуется осуществлять цифровую фильтрацию сигналов и компенсацию помех, число уровней квантования нужно увеличивать, чтобы уменьшить по возможности искажения сигналов и помех. На практике часто выбирают $\Delta u = u_{\min} \approx \sigma_{\text{ш}}$, где $\sigma_{\text{ш}}^2$ – дисперсия собственного шума приемника. При этом число порогов квантования равно $r = d - 1$, где $d = u_{\max}/\sigma_{\text{ш}}$ – динамический диапазон аналоговой части приемника. Отсюда получаем требуемое число разрядов кода и соответственно число разрядов АЦП:

$$m = [\log_2 d].$$

Следующий шаг в преобразовании сигнала состоит в переводе квантованного сигнала в цифровой сигнал. Эта операция называется кодированием сигнала.

11.3. Кодирование сигнала

Позиционность. Сначала рассмотрим одно замечательное свойство системы счисления – позиционность. Возьмем какое-нибудь число, например 777. В нем один и тот же знак «7» участвует 3 раза, но обозначает то семь единиц, то – в центре – семь десятков, а слева – семь сотен. Таким образом, при записи числа цифра может иметь одно начертание, а значения в зависимости от места (позиции), разряда, на котором она стоит, – разные. Такой принцип построения чисел называется позиционным. Для записи любых сколь угодно больших чисел достаточно десяти цифр.

Каждая позиция, или разряд, числа имеет определенный «вес» (единицы, десятки, сотни и т. д.), поэтому число 777 можно расписать как

$$777 = 7 \times 10^2 + 7 \times 10 + 7,$$

т.е. семь сотен плюс семь десятков плюс семь единиц.

Если вместо чисел записать буквы, то можно получить общую форму представления числа:

$$M = a_n \times 10^n + a_{n-1} \times 10^{n-1} + \dots + a_1 \times 10 + a_0$$

или сокращенную (опуская степени числа 10) – через коэффициенты:

$$M = (a_n a_{n-1} \dots a_1 a_0)$$

Число 10 является основанием системы счисления. Коэффициенты a_0 (число единиц), a_1 (число единиц второго разряда, т.е. десятков), a_2

(число единиц третьего разряда, т.е. сотен) и т.д. могут принимать значения, не превышающие основания системы: от 0 до 9.

Основанием системы счисления может быть любое целое число, т.е. число можно представить комбинацией степеней основания, например, 7:

$$M = a_n \times 7^n + a_{n-1} \times 7^{n-1} + \dots + a_1 \times 7 + a_0$$

Ясно, что значения коэффициентов a_0, a_1, \dots, a_n должны теперь быть не больше нового основания, т.е. 7: они могут принимать значения от 0 до 6.

Представим число 777 в семеричной системе, разлагая его по степеням основания 7:

$$(777)_{10} = 2 \times 7^3 + 1 \times 7^2 + 6 \times 7.$$

Если опустить степени числа 7, как мы делаем при записи чисел в десятичной системе, то получим семеричную запись этого числа: $(2160)_7$. Здесь цифра 7 в индексе указывает основание системы.

В пятеричной позиционной системе всего пять цифр: 0, 1, 2, 3, 4. В ней число 777 будет представлено количеством «пятерок», «двадцати-пяток» и т.д.:

$$(777)_{10} = 1 \times 5^4 + 1 \times 5^3 + 1 \times 5^2 + 0 \times 5 + 2 = (11102)_5.$$

Посмотрим, как будет представлено число 777 в двенадцатеричной системе. Поскольку в ней должно быть двенадцать цифр, а мы знаем только десять, то придется ввести еще две цифры, обозначив 10, скажем, буквой A, а 11 – буквой B. В результате получим

$$(777)_{10} = 5 \times 12^2 + 4 \times 12 + 9 = (549)_{12}.$$

Как видно, можно придумать много различных позиционных систем счисления, отличающихся только основаниями. И все они, вообще говоря, равнозначны: ни одна из них не имеет явных преимуществ перед другой.

Число 2 – это самое меньшее из чисел, которое можно взять за основание системы счисления. Поэтому в двоичной системе счисления всего две цифры: 0 и 1. Число в двоичной системе запишется так:

$$M = a_n \times 2^n + a_{n-1} \times 2^{n-1} + \dots + a_1 \times 2 + a_0.$$

Если в десятичной системе "вес" каждой позиции (или разряда) числа равен числу 10 в некоторой степени, то в двоичной системе вместо числа 10 используют число 2.

Запишем число $(777)_{10}$ в двоичной системе счисления, представляя его в виде разложения по степеням двойки и отбрасывая потом при записи сами степени:

$$(777)_{10} = 1 \times 2^9 + 1 \times 2^8 + 0 \times 2^7 + 0 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 0 \times 2 + 1 = (1100001001)_2.$$

Итак, в двоичной системе счисления вместо числа 777 приходится писать число 1100001001.

При записи числа в двоичной системе каждая позиция занята двоичной цифрой. Вместо двух слов «двоичная цифра» употребляют одно слово: «бит», составленного из начальных и конечной букв словосочетания «binary digit», что в переводе с английского означает «двоичная цифра».

С помощью одного бита можно записать только числа 0 и 1, двух бит – числа от 0 до 3, трех бит – числа от 0 до 7, четырех бит – числа от 0 до 15 и т.д.

Десятичная запись

0 1 2 3 4 5 6 7 8 9 10 11 12 ... 15 16

Двоичная запись

0 1 10 11 100 101 110 111 1000 1001 1010 1011 1100 ... 1111 100000

Чтобы записать числа от 0 до 1000 в двоичной системе счисления потребуется десять бит, т.е. даже сравнительно небольшое число занимает много позиций.

Перевод дискретных значений сигнала в цифровой двоичный код. Известный математик Леонард Эйлер (XVIII век) показал, что с помощью набора гирь 1, 2, 4, 8, и 16 кг можно взвесить любой груз с точностью до 31 кг. Взвешиваемый груз (обозначим его массу через M , кг) математически можно представить как

$$M = a_4 \times 16 + a_3 \times 8 + a_2 \times 4 + a_1 \times 2 + a_0 = a_4 \times 2^4 + a_3 \times 2^3 + a_2 \times 2^2 + a_1 \times 2^1 + a_0 \times 2^0,$$

где каждый коэффициент $a = 1$, если соответствующую гирю кладем на чашу весов, $a = 0$, если этой гирей не пользуемся при взвешивании. Таким образом, процедура взвешивания сводится к представлению десятичного числа в двоичной системе счисления.

Поясним это на примере. Пусть нам нужно взвесить груз массой 21 кг. Поставим сначала на чашу весов самую большую гирю – массой 16 кг. Поскольку она не перетягивает груз, оставим гирю на чаше ($a_4 = 1$) и добавим следующую – 8 кг. Ясно, что в этом случае чаша весов с гирями перетянет чашу с грузом. Снимем эту гирю ($a_3 = 0$) и установим гирю массой 4 кг. Проведя взвешивание до конца, мы увидим, что на весах остались гири массой 16, 4 и 1 кг. Значения коэффициентов $a_4 \dots a_0$ дают пятиразрядный двоичный код 10101 числа 21.

Механический груз мы взвешивали на механических весах. Считая конкретное значение тока, появляющееся на выходе электронного ключа, как бы «электрическим грузом», можно осуществить аналогичное

взвешивание, но только электронным способом. Такие «электронные весы» назвали кодером (от английского coder – кодировщик).

Допустим, значение тока равно 21 мА. Роль «электрических гирь» в кодере выполняют эталонные токи величиной 16, 8, 4, 2 и 1 мА, которые вырабатываются специальным устройством. Каждая проба (подходит та или иная «гиря» либо нет) производится в строго установленные промежутки времени. Вся процедура «взвешивания» должна закончиться до прихода с электрического ключа следующего значения тока (например, для звуков речи это время составляет всего 125 мкс).

Итак, сначала отсчетное значение тока сравнивается с эталоном, равным 16 мА, и, поскольку оно больше эталона, на выходе кодера появляется импульс тока, что соответствует двоичной цифре 1. В следующий интервал времени к первому эталонному току добавляется второй величиной 8 мА. Теперь суммарный вес «электрической гири» равен 24 мА. Это больше отсчетного значения, поэтому второй эталонный генератор отключается. На данном интервале времени импульс тока на выходе кодера не появляется, что соответствует двоичной цифре 0. Далее добавляется третий эталонный ток величиной 4 мА, таким образом за время «электронного взвешивания» одного отсчётного значения кодер вырабатывает серию импульсов, полностью повторяющую двоичный код отсчетного значения сигнала.

При кодировании появляются искажения. Так, если кодированию подвергается отсчетное значение 21,7 мА, кодер все равно выдает код 10101, как и в случае целого значения 21 мА. Это и понятно, поскольку «взвешивание» проводилось с точностью до 1 мА – веса самой меньшей «электрической гири». Такое округление чисел в технике называется квантованием, а разница между отсчетным значением тока и величиной, набираемой двоичным кодом, – ошибкой квантования.

Перевод целого числа из десятичной системы счисления в *и*ную. Правило перевода основано на последовательном делении числа с остатком на q – новое основание системы счисления:

- Число делится на q и находится первый остаток и первое частное.
- Если первое частное больше или равно q , то оно делится на q , находят второй остаток и второе частное.
- Если второе частное больше или равно q , то оно делится на q .
- Деление и сравнение очередного частного с q продолжается до тех пор, пока последующее частное не станет меньше q . Если где-то произошло деление “на цело”, то остаток считается равным нулю.
- Выписав последнее частное, затем, приписав к нему остатки в обратном порядке по отношению к их появлению, с учетом нулевых остатков, получаем число, записанное в системе счисления с новым основанием q .

Пример. Перевести число 243 в пятеричную систему ($q = 5$).

243: 5 = 48 (ост. 3), $48 > 5$

48: 5 = 9 (ост. 4), $9 > 5$

9: 5 = 1 (ост. 4), $1 < 5$. Процесс деления окончен.

Ответ: 1443_5 – искомая запись числа.

Современный уровень развития схемотехники позволил объединить в корпусе одной микросхемы электронный ключ и кодер. Эта микросхема выполняет преобразование непрерывной (аналоговой) электрической величины в двоичный цифровой код и называется аналого-цифровым преобразователем (АЦП). Выпускаются АЦП с 8-, 10- и 12- и более разрядными двоичными кодами.

Подсчитаем, скорость цифрового потока, полученного из телефонного аналогового сигнала при его дискретизации через 125 мкс и 8-разрядном кодировании. За секунду ток микрофона изменяется 8000 раз. В 8-разрядном кодере каждое измеренное значение тока представляется двоичным словом из 8 бит. Значит, каждую секунду в линию отправляется $8000 \times 8 = 64\,000$ бит, т. е. скорость цифрового потока равна 64 кбит/сек.

Кодовая комбинация из 8 бит, образующая двоичное слово, называется байтом. Символы в каждой кодовой комбинации отделены друг от друга временным интервалом t_T , т.е. следует с частотой $f_T = 1/t_T$. Эта частота называется тактовой. Преобразование отсчетов непрерывного сигнала в двоичный код называется импульсно-кодовой модуляцией (ИКМ). В настоящее время этот способ получения цифровых сигналов из аналоговых нашел наибольшее распространение. Системы передачи, использующие данное преобразование сигналов, называются ИКМ системами. В иностранной литературе используется аббревиатура PCM (от английских слов «pulse code modulation», что в переводе как раз и означает «импульсно-кодовая модуляция»).

11.4. Декодирование сигнала

При приеме сигналов ИКМ для восстановления аналогового сигнала необходимо преобразовать цифровой сигнал (последовательность двоичных импульсов) в квантованный АИМ сигнал (такое преобразование называется декодированием) и затем осуществить операцию демодуляции, т.е. выделения из АИМ-сигнала аналогового сигнала $s(t)$. В процессе использования ИКМ выполняются следующие преобразования аналогового сигнала: в пункте передачи – амплитудно-импульсная модуляция, квантование и кодирование; в пункте приема – декодирование и демодуляция квантованного АИМ сигнала. Полученный на приеме аналоговый сигнал отличается от переданного, так как образуется из квантованных импульсов, амплитуды которых равны не мгновенным значениям сигнала $s(t)$, а ближайшим разрешенным значениям.

Таким образом, операция квантования вносит в процесс передачи сигнала неустранимую ошибку, которая тем меньше, чем больше уровень квантования.

Как узнать, какое десятичное число скрывается под его записью в двоичной системе? Правило перевода: под каждым разрядом двоичного числа записать его «вес». Те «веса», которые соответствуют единичным разрядам, нужно сложить. Полученная сумма и явится десятичным числом.

Вот перед нами число 1001011, записанное в двоичной нумерации. Поступаем согласно сказанному выше:

1	0	0	1	0	1	1
$2^6 = 64$	$2^5 = 32$	$2^4 = 16$	$2^3 = 8$	$2^2 = 4$	$2^1 = 2$	$2^0 = 1$

Как видим, заинтересовавшее нас число складывается из единицы, двойки, восьмерки и шестидесяти четырех ($1 + 2 + 8 + 64$), т.е. оно равно 75.

В состав декодера входит преобразователь последовательного кода в параллельный (рис. 11.2), на выходах которого появляется набор единиц и нулей, соответствующий принятой кодовой комбинации. Каждая единица (токовый импульс) поступает на вход сумматора с весом, где увеличивается в 2^k раз. На выходе сумматора возникает импульс, амплитуда которого определяется кодовой комбинацией на входе декодера. Например, при прохождении кодовой комбинации 0100110 на первый, четвертый, пятый и седьмой входы сумматора напряжение не подается (бестоковые импульсы), а на второй, третий и шестой входы подается напряжение, которое увеличивается соответственно в 2^1 , 2^2 и 2^5 раз. На выходе сумматора появляется напряжение, пропорциональное $2^1 + 2^2 + 2^5 = 38$, т.е. квантованный АИМ сигнал.

На следующем шаге необходимо из отсчетных значений, например, тока получить непрерывный ток. Сделать это нам поможет обычный конденсатор небольшой емкости, который при кратковременном воздействии на него тока (т.е. отсчетного значения) мгновенно зарядится и будет удерживать заряд до следующего кратковременного воздействия.

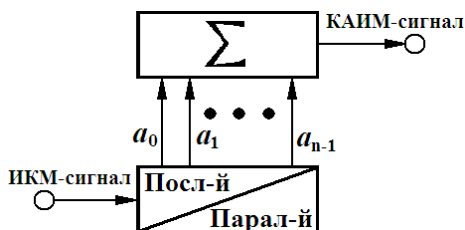


Рис. 11.2. Декодер ИКМ-сигнала.

Отметим еще раз, что восстановленная таким путем кривая непрерывного тока будет несколько отличаться от той, которая была получена, например, на клеммах микрофона: она будет иметь плоские ступеньки между отсчетными значениями. Можно сказать, что процесс взятия отсчетных значений и последующего восстановления непрерывной кривой тока микрофона сопровождается специфическими искажениями, которые могут повлиять на качество воспроизведения звука. Однако на практике для восстановления тока используют не конденсатор, а более сложные схемы, делающие форму восстановленного тока похожей на форму исходного тока и тем самым сводящие на нет действия указанных искажений.

12. ОБРАБОТКА ДИСКРЕТНЫХ СИГНАЛОВ

12.1. Алгоритмы дискретного и быстрого преобразований Фурье

Дискретные сигналы (как и аналоговые) могут быть представлены во временной и частотной областях. В настоящее время обработка дискретных сигналов чаще всего проводится в частотной области, так как при этом значительно сокращается время и упрощается аппаратура.

Подвергнем дискретной обработке аналоговый импульсный сигнал $u(t)$ длительностью T_n , имеющий спектральную плотность $S(\omega)$ (рис. 12.1а, б). Воспользуемся теоретическим представлением дискретизации сигнала периодической последовательностью дельта-функций

$$y(t) = \sum_{k=0}^{N-1} \delta(t - k\Delta t), \quad (12.1)$$

где $N = T_n/\Delta t$ – требуемое число отсчетов, отвечающих теореме Котельникова.

Следовательно выражение для дискретного сигнала (рис. 12.1е) с учетом пределов суммирования (от 0 до $N - 1$) имеет вид

$$u_T(t) = u(t) \sum_{k=0}^{N-1} \delta(t - k\Delta t) = \sum_{k=0}^{N-1} u_k \cdot \delta(t - k\Delta t) \quad (12.2)$$

где $u_k = u(k\Delta t)$.

На основании формулы (12.2) можно сделать вывод, что спектр данного дискретного сигнала имеет периодическую структуру с периодом по оси частот $\omega_1 = 2\pi/\Delta t$ (рис. 12.1з). Мысленно продолжим дискретный сигнал периодически с интервалом T_n (рис. 12.1д) $|C_n|$

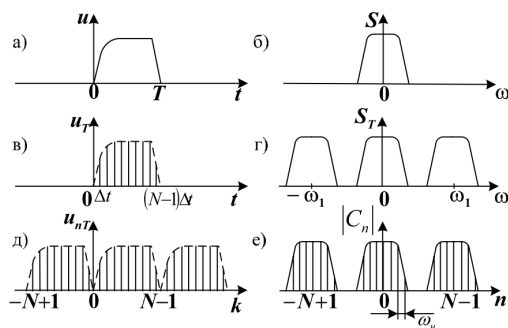


Рис. 12.1. Графики к выводу ДПФ:

а, б – аналоговый сигнал и его спектр; в, г – дискретный сигнал и его спектр; д – периодическая последовательность дискретного сигнала; е – ДПФ сигнала

$$u_{nT}(t + T_n) = u_T(t), \quad n = 0 \pm 1, \pm 2, \dots$$

По аналогии с представлением периодических непрерывных сигналов

$$u(t) = \sum_{n=-\infty}^{\infty} C_n \exp(jn\omega_1 t)$$

где $C_n = \frac{1}{T} \int_{-T/2}^{T/2} u(t) \cdot \exp(jn\omega_1 t) dt$ – комплексная амплитуда n -й гармоники. Дискретную функцию $u_{nT}(t)$ можно разложить в комплексный ряд Фурье:

$$u_{nT}(t) = \sum_{n=-\infty}^{\infty} C_n \exp(jn\omega_1 t), \quad (12.3)$$

где $\omega_n = 2\pi/T_n = 2\pi/(N \cdot \Delta t)$ – частота дискретизации сигнала.

Коэффициенты этого ряда

$$C_n = \frac{1}{T_n} \int_0^{T_n} u_{\Delta}(t) \cdot \exp(jn\omega_1 t) dt = \frac{1}{T_n} \int_0^{T_n} u_{\Delta}(t) \cdot \exp(j2\pi n t / \Delta t) dt. \quad (12.4)$$

Для определения коэффициентов сделаем следующее. Подставим формулу (12.2) в (12.4) и заменим параметр $T_n = N \cdot \Delta t$. Введем безразмерную переменную $y = t/\Delta t$ и запишем

$$\begin{aligned} C_n &= \frac{1}{T_n} \int_0^{T_n} \sum_{k=0}^{N-1} u_k \cdot \delta(t - k\Delta t) \cdot \exp(-j2\pi n t / T_n) dt = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} u_k \int_0^{N-1} \sum_{k=0}^{N-1} \delta(y - k) \cdot \exp(-j2\pi n y / N) dy. \end{aligned}$$

Используя фильтрующее свойство дельта – функции, находим

$$C_n = \frac{1}{N} \sum_{k=0}^{N-1} u_k \exp(-j2\pi n k / N). \quad (12.5)$$

Это соотношение называется дискретным преобразованием Фурье (ДПФ). Дискретное преобразование Фурье по существу представляет собой алгоритм вычисления гармонических составляющих спектра C_n по заданным дискретным отсчетам u_k аналогового сигнала $u(t)$, что значительно сокращает время обработки. Характерный вид модулей коэффициентов C_n показан на рис. 12.1е.

Следует отметить ряд свойств ДПФ, которые вытекают из определения (12.5).

1. Дискретное преобразование Фурье обладает свойством линейности: линейной комбинации дискретных сигналов соответствует линейная комбинация их ДПФ.

2. Коэффициент C_0 представляет собой среднее значение (постоянную составляющую) всех дискретных отсчетов сигнала

$$C_0 = \frac{1}{N} \sum_{k=0}^{N-1} u_k.$$

3. Число различных коэффициентов C_n равно числу отсчетов N за длительность сигнала T_n ; при $n = N$ коэффициент $C_n = C_0$.

Пример. Определить коэффициенты ДПФ дискретизированного прямоугольного импульса единичной амплитуды, заданного четырьмя отсчетами ($N = 4$).

Решение. Используя основную формулу (12.5), вычислим пять первых коэффициентов ДПФ: $C_0 = 4/4 = 1$;

$$C_1 = \frac{1}{4} \sum_{k=0}^{N-1} (1 + \exp(-j\pi/2) + \exp(-j\pi) + \exp(-j3\pi/2)) = 0;$$

$$C_2 = \frac{1}{4} \sum_{k=0}^{N-1} (1 + \exp(-j\pi) + \exp(-j2\pi) + \exp(-j3\pi)) = 0;$$

$$C_3 = \frac{1}{4} \sum_{k=0}^{N-1} (1 + \exp(-j3\pi/2) + \exp(-j3\pi) + \exp(-j9\pi/2)) = 0;$$

$$C_4 = \frac{1}{4} \sum_{k=0}^{N-1} (1 + \exp(-j2\pi) + \exp(-j4\pi) + \exp(-j6\pi)) = 1.$$

При изучении теории ДПФ возникает очевидный вопрос: можно ли по известным коэффициентам ДПФ вычислить отсчетные значения u_k непрерывного сигнала? По аналогии с периодическими сигналами представим заданную периодическую последовательность отсчетов комплексным рядом Фурье, заменив $t = k\Delta t$, $\omega_n = 2\pi/(N \times \Delta t)$ и, учитывая, что суммируется конечное число членов ряда, запишем

$$u_k = \sum_{n=0}^{N-1} C_n \cdot \exp(j2\pi nk / N). \quad (12.6)$$

Данное соотношение определяет алгоритм обратного дискретного преобразования Фурье (ОДПФ). Формулы (12.5) и (12.6) являются аналогами прямого и обратного преобразований Фурье для непрерывных сигналов.

Выражение (12.5) показывает, что для определения одного коэффициента ДПФ сигнальной последовательности из N отсчетов, необходимо выполнить около N операций умножения на комплексное число и столько же сложений, а для нахождения всех коэффициентов объем вычислений составит N^2 . В частности, при $N = 2^{10} = 1024$ надо осуществить более миллиона (1024^2) умножений и сложений. Если длины обрабатываемых массивов превышают тысячу единиц, то дискретная спектральная обработка сигналов в реальном масштабе времени требует высокопроизводительных вычислительных комплексов.

Многokrратно сократить число операций позволяет быстрое преобразование Фурье (БПФ), обеспечивающее вычисление коэффициентов ДПФ за меньшее число операций. В основу БПФ положен принцип раз-

бienia заданной последовательности отсчетов дискретного сигнала на несколько промежуточных последовательностей. Для этого число отсчетов N разделяется на множители (например, $N = 8 = 2 \times 2 \times 2$, $N = 60 = 3 \times 4 \times 5$). Затем определяются спектры этих промежуточных последовательностей и через них находится спектр всего сигнала.

В зависимости от состава, числа и порядка следования указанных множеств можно создать различные алгоритмы БПФ. В цифровой технике удобно обрабатывать сигнальные последовательности со значениями N , являющимся степенью числа два (4, 8, 16 и так далее). Это позволяет многократно делить входную последовательность отсчетов на подпоследовательности.

Пусть требуется вычислить ДПФ дискретного сигнала $\{u(k\Delta t)\} = \{u_k\}$, имеющего четное число отсчетов (рис. 12.2а), причем $N = 2^r$; r – целое число.

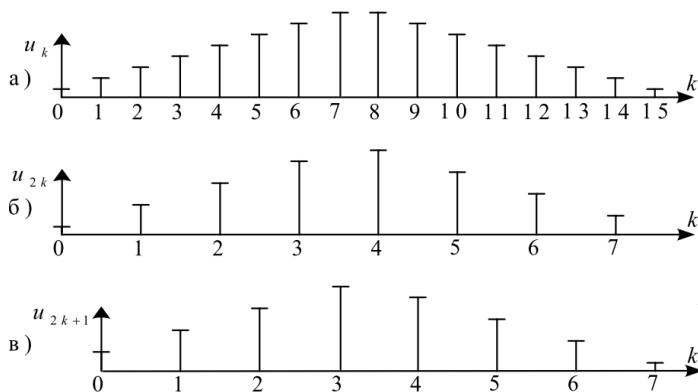


Рис. 12.2. Разбиение последовательности u_k на две подпоследовательности: а – входная; б – с четными номерами; в – с нечетными номерами

Представим входную последовательность в виде двух подпоследовательностей с четными и нечетными номерами и половинным числом членов в каждой (рис. 12.2б, в): $u_{\text{чт}} = u_{2k}$; $u_{\text{нч}} = u_{2k+1}$; $k = 0, 1, 2, \dots, N/2 - 1$.

Коэффициенты ДПФ для последовательностей с четными и нечетными номерами запишем отдельно:

$$\begin{aligned} \frac{1}{N} \sum_{k=0}^{N/2-1} u_{2k} \cdot \exp(-j2\pi mk / N) &= \frac{1}{N} \sum_{k=0}^{N/2-1} u_{2k} \cdot \exp\left(\frac{-j2\pi mk}{N/2}\right) = C_{\text{нчт}}; \\ \frac{1}{N} \sum_{k=0}^{N/2-1} u_{2k+1} \cdot \exp(-j2\pi m(2k+1) / N) &= \\ = \frac{1}{N} \cdot \exp(-j2\pi m / N) \sum_{k=0}^{N/2-1} u_{2k+1} \cdot \exp\left(\frac{-j2\pi mk}{N/2}\right) &= \exp(-j2\pi m / N) C_{\text{нчч}} \end{aligned} \quad (12.7)$$

Коэффициенты C_n результирующего ДПФ входной последовательности можно выразить через параметры $C_{nчт}$ и $C_{nнч}$ двух вновь введенных подпоследовательностей. Анализ (12.7) показывает, что в диапазоне номеров отсчетов от 0 до $N/2 - 1$, ДПФ входной последовательности определяется соотношением:

$$C_n = C_{nчт} + \exp(-j2\pi n/N) \cdot C_{nнч}, \quad n = 0, 12, \dots, N/2 - 1. \quad (12.8)$$

Так как ДПФ четной и нечетной последовательностей являются периодическими, с периодом $N/2$, то $C_{nчт} = C_{(n+N/2)чт}$; $C_{nнч} = C_{(n+N/2)нч}$.

Запишем экспоненциальный множитель в формуле (12.8) при $n \geq N/2$, т.е. для ДПФ $C_{(N/2+n)нч}$, в виде:

$$\exp(-j \frac{2\pi(N/2+n)}{N}) = \exp(-j\pi) \cdot \exp(-j \frac{2\pi n}{N}) = -\exp(-j \frac{2\pi n}{N})$$

С учетом двух последних выражений находим коэффициенты ДПФ входной последовательности для отсчетов с номерами от $N/2$ до $N - 1$:

$$C_{N/2+n} = C_{nчт} - \exp(-j2\pi n/N) \cdot C_{nнч}, \quad n = N/2, N/2 + 1, \dots, N - 1. \quad (12.9)$$

Соотношения (12.8) и (12.9) полностью определяют алгоритмы вычисления коэффициентов с помощью БПФ. Отметим, что экспоненциальные фазовые множители $\exp(-j2\pi n/N)$ в этих алгоритмах учитывают влияние сдвига нечетной подпоследовательности относительно четной.

Чтобы еще уменьшить число вычислений, четную и нечетную подпоследовательности также разбивают каждую на две промежуточные части. Разбиение продолжают вплоть до получения простейших двухэлементных последовательностей. Определив ДПФ данных простейших пар отсчетов, можно вычислить ДПФ четырехэлементных, восьмиэлементных и так далее подпоследовательностей. При объединении ДПФ четной и нечетной подпоследовательностей используют алгоритмы (12.8) и (12.9), подставляя в них соответствующие значения номеров N и n .

Нетрудно заметить, что вычисления по формулам (12.7) не требуют операций умножения, в (12.7) имеются только сложение и вычитание комплексных чисел. Учитываться же должны лишь операции умножения в алгоритмах (12.8) и (12.9) для различных n при разбиениях массива отсчетов на мелкие подпоследовательности. Число этих операций при первом разбиении составляло $N/2$. Такое же число $N/2$ операций требуется выполнить при каждом следующем разбиении. Таким образом, вдвое увеличивается число подпоследовательностей и вдвое сокращается наибольшее число n в формулах вычисления спектральной плотности дискретного сигнала.

Вычисление коэффициентов ДПФ последовательности из N отсчетов по алгоритмам БПФ требует примерно $N \cdot \log_2 N$ операций умножения. Алгоритмы БПФ сокращают число операций по сравнению с алгоритмами ДПФ в $N : \log_2 N$ раз. Например, при количестве отсчетов $N = 2^{10}$, имеем $\log_2 N = 10$ и сокращение числа операций составляет N :

$\log_2 N \approx 100$. При очень больших массивах отсчетов входного сигнала выигрыш в скорости обработки может достигать нескольких тысяч.

Таким образом, в алгоритмах БПФ выполняются операции сложения и вычитания с умножением одного из компонентов на экспоненциальный множитель $\exp(-j2\pi n/N)$. Эту базовую для БПФ операцию очень удобно представлять сигнальным графом, называемым в цифровой технике «бабочкой». БПФ по рассмотренному методу называют методом прореживания отсчетов во времени.

12.2. Стационарные линейные дискретные цепи

Преобразования дискретных сигналов в процессе их обработки могут выполняться специализированными цифровыми устройствами или универсальными вычислителями (процессорами) под управлением программ; в любом случае удобно считать, что преобразование выполняется некоторой дискретной цепью. Таким образом, дискретной цепи соответствует отображение множества входных (дискретных) сигналов на множество выходных сигналов.

Задать отображение – значит задать эти множества и каждому входному сигналу поставить в соответствие единственный выходной. Как и для аналоговых цепей, для упрощения этой задачи на отображение (цепь) накладываются определенные ограничения.

Прежде всего, положим, что множества входных и выходных сигналов совпадают (рассматривается задача фильтрации), тогда понятие отображения сужается до оператора. Будем также считать, что оператор цепи $L\{\bullet\}$ линеен, т.е. удовлетворяет принципу суперпозиции

$$L\{\alpha_1 x_1 + \alpha_2 x_2\} = \alpha_1 L\{x_1\} + \alpha_2 L\{x_2\},$$

где α_1, α_2 – скалярные коэффициенты (вещественные или комплексные)
 $x_1 = x_1[n], x_2 = x_2[n]$ – дискретные сигналы.

Произвольный дискретный сигнал (последовательность) $x[n]$ можно представить в виде обобщенного ряда Фурье относительно базиса, состоящего из сдвинутых δ -последовательностей

$$x(n) = \sum_{k=-\infty}^{+\infty} x(k)\delta(n-k)$$

где отсчеты этого сигнала $x[k]$ рассматриваются как постоянные коэффициенты при базисных функциях $\delta[n-k]$, $-\infty \leq k \leq \infty$. Тогда результат воздействия линейного оператора (линейной цепи) на этот сигнал равен

$$y[n] = L\left\{\sum_{k=-\infty}^{\infty} x[k]\delta[n-k]\right\} = \sum_{k=-\infty}^{\infty} x[k]L\{\delta[n-k]\} = \sum_{k=-\infty}^{\infty} x[k]h[n, k],$$

где $h[n, k]$ представляет собой отклик цепи в момент времени n на δ -последовательность, имеющую единичное значение в момент времени k .

Если кроме линейности потребовать, чтобы весовая последовательность $h[n, k]$ зависела только от разности аргументов, $h[n, k] = h[n - k]$, то цепь станет инвариантной к сдвигу (стационарной), а формула нахождения выходного сигнала примет форму дискретной свертки

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n - k] = \sum_{k=-\infty}^{\infty} h[k]x[n - k]. \quad (12.10)$$

Последовательность $h[n]$ называется импульсной характеристикой линейной инвариантной к сдвигу (ЛИС) цепи и является ее исчерпывающей характеристикой, так как позволяет найти сигнал на выходе данной ЛИС-цепи для произвольного входного сигнала.

Необходимо отметить одно важное свойство дискретных цепей, отличающее их от аналоговых. Дискретная свертка представляет не только метод анализа ЛИС-цепи, подобно интегралу Дюамеля для аналоговых цепей, но также алгоритм работы вычислительного устройства.

Рассмотрим ЛИС-цепь при воздействии на ее вход комплексной экспоненциальной последовательности $x[n] = \exp(j\omega n)$ при, $-\infty \leq n \leq \infty$, тогда выходной сигнал в соответствии с (12.10)

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]e^{j\omega n} e^{-j\omega k} = \exp(j\omega n) \sum_{k=-\infty}^{\infty} h[k] \exp(-j\omega k) = \exp(j\omega n) H(\exp(j\omega)),$$

где $H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h[n]e^{-j\omega n}$ – комплексная частотная характеристика ЛИС-цепи.

Рассматривая выражение (11.4) как представление произвольного дискретного сигнала $x[n]$ суперпозицией несчетного множества комплексных экспоненциальных последовательностей $\exp(j\omega n)$ ($\omega \in [-\pi, \pi]$), умноженных на весовые коэффициенты $(1/2\pi)X(e^{j\omega})$, легко видеть, что выходная последовательность получается домножением каждой из них на значение КЧХ:

$$y[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega}) X(e^{j\omega}) e^{j\omega n} d\omega, \quad -\infty \leq n \leq \infty. \quad (12.11)$$

Сравнивая выражения (12.11) и (11.4), видим, что спектральная плотность выходного сигнала равна $Y(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega})$. Полученное выражение составляет основу спектрального метода анализа ЛИС-цепей.

12.3. Цепи с конечной импульсной характеристикой (КИХ-цепи)

Предположим, что импульсная характеристика некоторой цепи $h[n]$ имеет конечную длину N , т.е. $h[n] \neq 0, 0 \leq n \leq N-1$. Тогда свертка (12.10) принимает вид конечной суммы

$$y[n] = \sum_{k=0}^{N-1} h[k]x[n-k] = \sum_{k=0}^{N-1} b_k x[n-k],$$

и может быть записана в виде разностного уравнения

$$y[n] = b_0 x[n] + b_1 x[n-1] + b_2 x[n-2] + \dots + b_{N-1} x[n-N+1]. \quad (12.12)$$

Вычисление каждого значения выходного сигнала требует учета текущего и $N-1$ предшествующих отсчетов входного сигнала и может быть выполнено цепью, структурная схема которой показана на рис. 12.3. Такие цепи называются трансверсальными, или цепями с конечной импульсной характеристикой (КИХ-цепями).

Комплексная частотная характеристика КИХ-цепи имеет вид полинома порядка $N-1$ относительно $e^{j\omega}$:

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h[n]e^{-jn\omega} = b_0 + b_1 e^{-j\omega} + b_2 e^{-j2\omega} + \dots + b_{N-1} e^{-j(N-1)\omega}.$$

Таким образом, КИХ-цепь умножает спектральную плотность входной последовательности на полином.

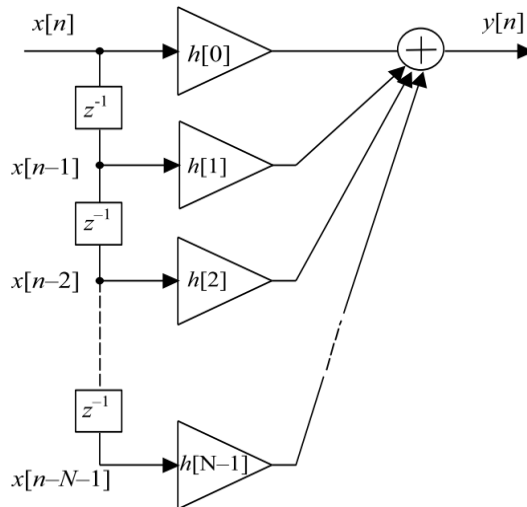


Рис. 12.3. Структура цепи с конечной импульсной характеристикой

12.4. Рекурсивные цепи

Другой важный для практики класс дискретных ЛИС-цепей составляют цепи, которые не умножают, а делят спектральную плотность входной последовательности на полином некоторого порядка $M - 1$ относительно $e^{j\omega}$. Обозначим этот полином $A(e^{j\omega}) = \alpha_0 + \alpha_1 e^{j\omega} + \alpha_2 e^{j2\omega} + \dots + \alpha_{M-1} e^{j(M-1)\omega}$, тогда спектральные плотности входной и выходной последовательностей связаны выражением $Y(e^{j\omega}) = X(e^{j\omega}) / A(e^{j\omega})$, следовательно, $X(e^{j\omega}) = Y(e^{j\omega}) A(e^{j\omega})$, откуда по аналогии с (12.3) можно записать

$$x[n] = \alpha_0 y[n] + \alpha_1 y[n - 1] + \alpha_2 y[n - 2] + \dots + \alpha_{M-1} y[n - M + 1].$$

Решая это уравнение относительно выходного сигнала, получаем

$$y[n] = \frac{1}{\alpha_0} x[n] - \frac{\alpha_1}{\alpha_0} y[n - 1] - \frac{\alpha_2}{\alpha_0} y[n - 2] - \dots - \frac{\alpha_{M-1}}{\alpha_0} y[n - M + 1],$$

откуда, вводя обозначения $b = 1/\alpha_0$, $\alpha_i = -\alpha_i/\alpha_0$, находим окончательно разностное уравнение рекурсивной цепи

$$y[n] = bx[n] + \alpha_1 y[n - 1] + \alpha_2 y[n - 2] + \dots + \alpha_{M-1} y[n - M + 1],$$

структура которой показана на рис. 12.4.

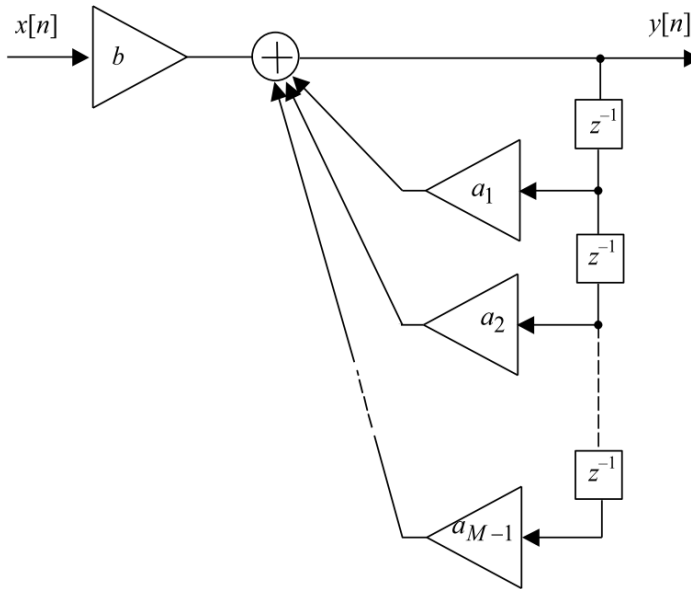


Рис. 12.4. Структура рекурсивной цепи

12.5. Устойчивость ЛИС-цепей

Обычно к ЛИС-цепям предъявляется требование устойчивости. Напомним, что линейная цепь называется устойчивой, если отклик на воздействие, ограниченное по модулю, также ограничен.

Для устойчивости ЛИС-цепи необходимо и достаточно, чтобы ее импульсная характеристика была абсолютно суммируемой, т.е. выполнялось условие [7]

$$\sum_{n=-\infty}^{\infty} |h[n]| < \infty. \quad (12.13)$$

Очевидно, для импульсных характеристик конечной длины это условие выполняется всегда, поэтому КИХ-цепи всегда устойчивы.

Рекурсивные цепи могут быть неустойчивыми из-за наличия обратных связей. Анализ устойчивости ЛИС-цепей основан на использовании z -преобразования, которое формально может быть получено из преобразования Фурье заменой величины $e^{j\omega}$ на комплексное переменное z :

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n}. \quad (12.14)$$

z -преобразование может сходиться для одних значений комплексного переменного z и расходиться для других. Множество точек комплексной z -плоскости, в которых z -преобразование сходится, называется областью сходимости. Для абсолютно суммируемой импульсной характеристики область сходимости ее z -преобразования содержит единичную окружность. Если цепь является физически реализуемой (каузальной), то она устойчива в том и только в том случае, если все полюсы ее передаточной функции

$$H(z) = \sum_{n=0}^{\infty} h[n]z^{-n}$$

по модулю меньше единицы, т.е. находятся внутри единичной окружности.

Самый широкий класс ЛИС-цепей конечного порядка образуют цепи, структура которых может быть сведена к каскадному соединению трансверсальной и рекурсивной частей, что соответствует разностному уравнению вида

$$\begin{aligned} y[n] &= b_0x[n] + b_1x[n-1] + b_2x[n-2] + \dots + b_{N-1}x[n-N+1] + \\ &+ \alpha_1y[n-1] + \alpha_2y[n-2] + \dots + \alpha_{M-1}y[n-M+1] = \\ &= \sum_{k=0}^{N-1} b_k x[n-k] + \sum_{r=1}^{M-1} \alpha_r y[n-r], \end{aligned} \quad (12.15)$$

откуда следует выражение для КЧХ дробно-рационального вида

$$H(e^{j\omega}) = \frac{\sum_{k=0}^{N-1} b_k e^{-j\omega k}}{1 - \sum_{r=1}^{M-1} \alpha_r e^{-j\omega r}}. \quad (12.16)$$

В общем случае ЛИС-цепь конечного порядка с КЧХ вида (12.16) имеет бесконечно длинную импульсную характеристику (БИХ), но если полином-числитель делится на знаменатель без остатка, то результатом деления оказывается полином и импульсная характеристика имеет конечную длину (таковы, например, КИХ-фильтры на основе частотной выборки, см. далее).

13. ЦИФРОВЫЕ ФИЛЬТРЫ

Дискретное преобразование Фурье представляет собой не только инструмент анализа, но и алгоритм ЦОС. На его основе фильтрация сигналов в частотной области может быть реализована следующим образом: для входного сигнала вычисляется ДПФ, полученные спектральные отсчеты умножаются на КЧХ фильтра, а результат умножения подвергается обратному ДПФ. Этот метод фильтрации более экономичный, чем вычисление свертки входного сигнала с импульсной характеристикой фильтра, благодаря существованию очень эффективных (быстрых) алгоритмов, которые получили название быстрого преобразования Фурье (БПФ).

Основное назначение дискретных ЛИС-цепей заключается в фильтрации дискретных сигналов, т.е. в избирательном воздействии на амплитуды и начальные фазы гармонических составляющих различных частот. Это фактически означает, что любая ЛИС-цепь представляет собой фильтр. Однако интерес представляет построение фильтров с заданными частотно-избирательными и фазовыми свойствами. Построить (синтезировать) фильтр – значит найти его разностное уравнение (т.е. алгоритм вычисления выходного сигнала по известному входному) и/или структурную схему.

Таким образом, под синтезом цифрового фильтра (ЦФ) обычно понимается построение дискретной ЛИС-цепи с КЧХ заданной формы. При решении задачи синтеза обычно не делают различия между дискретными и цифровыми цепями, хотя, строго говоря, дискретная ЛИС-цепь становится цифровой в результате квантования коэффициентов ее разностного уравнения.

Ранее было показано, что ЛИС-цепь конечного порядка имеет в общем случае КЧХ дробно-рационального вида (12.7), поэтому, очевидно, задача синтеза ЦФ сводится к аппроксимации желаемой КЧХ функцией дробно-рационального вида, так как, зная эту функцию, легко составить структурную схему цепи или записать разностное уравнение вида (12.6). Указанная аппроксимация сравнительно легко выполняется для КИХ-цепей, когда дробнорациональная функция вырождается в полином, и представляет собой непростую задачу для общего случая. Поэтому методы синтеза ЦФ с конечными и бесконечными импульсными характеристиками различаются.

13.1. Методы синтеза КИХ-фильтров

Фильтры с конечной импульсной характеристикой имеют перед БИХ-фильтрами ряд преимуществ. Во-первых, КИХ-фильтры всегда устойчивы. Во-вторых, только КИХ-фильтр может иметь строго линейную фазочастотную характеристику [7] (фильтр с линейной ФЧХ не

искажает форму сигнала, если его спектр лежит в полосе частот, где амплитудно-частотная характеристика постоянна; при этом сигнал лишь задерживается на время, пропорциональное крутизне ФЧХ). Наконец, для КИХ-фильтров наиболее просто решается задача аппроксимации КЧХ желаемого вида реализуемой функцией (тригонометрическим полиномом). Однако КИХ-фильтры имеют существенный недостаток по сравнению с БИХ-фильтрами: для обеспечения сравнимых частотно-избирательных свойств, в частности крутизны АЧХ в переходной полосе частот, требуется КИХ-фильтр в десятки раз более высокого порядка, чем БИХ-фильтр. На практике в зависимости от конкретных обстоятельств применяются фильтры обоих типов. Ниже вкратце рассматриваются методы синтеза КИХ-фильтров.

Метод взвешивания (метод функций окна)

КЧХ трансверсального дискретного фильтра представляет собой тригонометрический полином, т.е. функцию вида

$$H(e^{j\omega}) = \sum_{n=-M}^M b_n e^{jn\omega}. \quad (13.1)$$

Здесь не предполагается каузальность фильтра; если каузальность необходима, ее легко можно обеспечить умножением (13.1) на фазовый множитель $e^{-jM\omega}$. Если желаемая КЧХ имеет вид $H_x(e^{j\omega})$, то синтез КИХ-фильтра состоит в нахождении тригонометрического полинома, близкого в каком-то смысле к $H_x(e^{j\omega})$. Обычно в качестве критерия близости выбирается среднеквадратическая ошибка аппроксимации

$$\mathcal{E} = \int_{-\pi}^{\pi} |H(e^{j\omega}) - H_x(e^{j\omega})|^2 d\omega.$$

тогда наилучшая аппроксимация обеспечивается, если коэффициентами полинома (13.1) являются коэффициенты разложения желаемой КЧХ в ряд Фурье

$$b_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_x(e^{j\omega}) e^{-jn\omega} d\omega, \quad -M \leq n \leq M. \quad (13.2)$$

Эти коэффициенты представляют собой отсчеты импульсной характеристики КИХ-фильтра, в общем случае некаузального. После соответствующей задержки получается импульсная характеристика каузального фильтра $h[n] = b_{n-M}$, $0 \leq n \leq N-1$, где $N = 2M + 1$. Поскольку всякая ЛИС-цепь однозначно определяется своей импульсной характеристикой, на этом синтез КИХ-фильтра можно было бы считать законченным. Однако если желаемая КЧХ разрывна (например, как часто бывает на практике, требуется АЧХ прямоугольной формы), получаемая

КЧХ, как сумма усеченного ряда Фурье (13.1), содержит гиббсовские осцилляции. Поэтому применяют дополнительное умножение импульсной характеристики на весовую последовательность («окно») подходящей формы.

Причина явления Гиббса заключается в слишком медленном убывании коэффициентов Фурье-разложения разрывной функции, поэтому все применяемые окна убывают от середины к краям [7]. Для достижения приемлемых избирательных свойств длина импульсной характеристики, определяющая объем вычислений, на практике составляет обычно несколько сотен.

Кроме метода взвешивания, иногда применяют другой способ борьбы с гиббсовскими осцилляциями. На этапе формулирования требований к фильтру вводят переходную полосу, в которой задают закон непрерывного изменения АЧХ (например, линейный закон) [7]. Тогда ряд Фурье сходится равномерно и явление Гиббса отсутствует. Это не означает, что исчезает неравномерность АЧХ, просто осцилляции теперь убывают по амплитуде с увеличением порядка фильтра.

Следует также упомянуть машинные методы синтеза КИХ-фильтров на основе численной оптимизации. При этом подбором коэффициентов КИХ-фильтра минимизируется взвешенная среднеквадратическая ошибка

$$\varepsilon = \int_{-\pi}^{\pi} q(\omega) |H(e^{j\omega}) - H_{\text{ж}}(e^{j\omega})|^2 d\omega,$$

где $q(\omega)$ – весовая функция, позволяющая управлять относительной значимостью ошибок на разных участках частотной оси, или максимальная взвешенная погрешность

$$\varepsilon' = \max \left\{ q(\omega) |H(e^{j\omega}) - H_{\text{ж}}(e^{j\omega})| \right\}.$$

Эти методы позволяют получить меньшие погрешности аппроксимации по сравнению с описанным выше методом оконного взвешивания, но их анализ значительно сложнее [7].

Метод частотной выборки

Метод синтеза фильтров с конечной импульсной характеристикой, получивший название метода частотной выборки, основан на задании значений желаемой КЧХ в точках, расположенных равномерно на единичной окружности и соответствующих точкам частотной оси (отсюда название метода) и аппроксимации КЧХ интерполяционным полиномом Лагранжа [7]. Этот метод приводит к построению структуры, содержащей трансверсальную и рекурсивную части, которой, тем не менее, соответствует конечная импульсная характеристика. Благодаря наличию

рекурсии такие фильтры при реализации требуют меньшего числа операций по сравнению с рассмотренными выше КИХ-фильтрами и оказываются предпочтительными.

Метод быстрой свертки

Фильтрация сигналов может быть выполнена в частотной области путем вычисления спектральной плотности входного сигнала, умножения ее на КЧХ фильтра и выполнения обратного преобразования Фурье (на практике для входного сигнала, который представляет собой реализацию случайного процесса, можно вычислить только дискретное преобразование Фурье). Этот на первый взгляд сложный способ нахождения выходного сигнала оказывается на практике более эффективным в вычислительном отношении, чем прямое вычисление свертки, благодаря существованию алгоритмов быстрого преобразования Фурье. Метод КИХ-фильтрации на основе БПФ получил название метода быстрой свертки.

При его реализации необходимо учитывать следующие два обстоятельства. Первое состоит в том, что дискретное преобразование Фурье обладает двойственностью – оно соответствует как последовательностям конечной длины, так и периодическим последовательностям. По этой причине перемножение коэффициентов ДПФ двух последовательностей (входного сигнала и импульсной характеристики) соответствует не обычной (апериодической), а так называемой циклической (круговой) свертке.

Убедимся, что это действительно так.

Пусть $\bar{x}[n]$ и $\bar{h}[n]$ – периодические последовательности с периодом N . Их циклическая свертка определяется выражением

$$\bar{y}[n] = \sum_{m=0}^{N-1} \bar{x}[m] \bar{h}[n-m].$$

ДПФ результирующей последовательности

$$\begin{aligned} \bar{Y}[k] &= \sum_{n=0}^{N-1} \left(\sum_{m=0}^{N-1} \bar{x}[m] \bar{h}[n-m] \right) e^{-j\frac{2\pi}{N}nk} = \\ &= \sum_{n=0}^{N-1} \bar{x}[m] \left(\sum_{m=0}^{N-1} \bar{h}[n-m] e^{-j\frac{2\pi}{N}(n-m)k} \right) e^{-j\frac{2\pi}{N}mk} = \bar{H}[k] \bar{X}[k], \end{aligned} \quad (13.3)$$

$$\bar{H}[k] \bar{X}[k] = H[k] X[k],$$

где $\bar{H}[k]$ и $\bar{X}[k]$ находятся как ДПФ N -периодических последовательностей $\bar{x}[n]$ и $\bar{h}[n]$, а $H[k]$ и $X[k]$ – как ДПФ их конечных фрагментов длины.

При выводе (13.3) учтен тот факт, что сумма в круглых скобках второй строке равна $\bar{H}[k]$ независимо от m в силу периодичности суммируемых членов: при различных m суммируются одни и те же N слагаемых в разном порядке. Таким образом, видно, что циклическая свертка (или свертка периодических последовательностей) соответствует поточечному произведению ДПФ-спектров последовательностей. При фильтрации же должна выполняться обычная аperiodическая свертка, определяемая выражением (13.2).

Преодолеть эту трудность можно следующим образом. Преобразование Фурье последовательности длины N дает полином (относительно $e^{-j\omega}$) степени $N - 1$. Полагая, что $x[n]$ и $h[n]$ – последовательности длины N , видим, что произведение их фурье-образов есть полином степени $2(N - 1)$. Но при поточечном перемножении ДПФ-спектров, имеющих по N отсчетов, получается всего N результирующих отсчетов, что соответствует полиному всего лишь $(N - 1)$ -й степени. Единственный способ получения правильного результата умножения двух полиномов состоит в том, чтобы точек вычисления ДПФ «хватило» для представления результата. Иначе говоря, если используется N -точечное ДПФ, то степень результирующего полинома должна быть не выше $(N - 1)$. Это, в свою очередь, означает, что сумма длин последовательностей $x[n]$ и $h[n]$ (обозначим их M и L) должна удовлетворять очевидному соотношению $M - 1 + L - 1 \leq N - 1$ (или $M + L - 1 \leq N$). Для того чтобы можно было для последовательности $x[n]$ длины $M < N$ получить N отсчетов ДПФ, следует перед вычислением ДПФ последовательность $x[n]$ дополнить $N - M$ нулевыми отсчетами (и то же проделать с другой последовательностью). Тогда результат их циклической свертки, полученный применением БПФ, совпадает с результатом аperiodической свертки.

Второе обстоятельство, которое должно учитываться при реализации КИХ-фильтрации методом быстрой свертки, относится к фильтрации последовательностей большой (в частности, бесконечной) длины. Если длина входной последовательности велика (сотни тысяч отсчетов и более), что типично для обработки сигналов, применяемых в радиотехнике и связи, то необходимое основание БПФ (число вычисляемых отсчетов) оказывается слишком большим. Это влечет за собой высокие требования к объему оперативной памяти вычислителя БПФ, а также приводит к большой задержке результирующего сигнала (результат может быть получен не ранее чем поступит последний отсчет входной последовательности плюс время, необходимое для вычисления прямого БПФ, умножения и обратного БПФ). Для того чтобы снизить требования к памяти и уменьшить задержку, применяют секционирование свертки [7].

Пусть $h[n]$ – импульсная характеристика фильтра, имеющая длину M , а $x[n]$ – сигнальная последовательность бесконечной длины. Представим $x[n]$ в виде

$$x[n] = \sum_{k=-\infty}^{\infty} x_k[n], \text{ где } x_k[n] = \begin{cases} x[n], & kL \leq n \leq (k+1)L \\ 0, & \text{в противном случае.} \end{cases}$$

Нетрудно видеть, что таким образом входная последовательность разбивается (секционируется) на совокупность неперекрывающихся примыкающих друг к другу сегментов длиной L отсчетов каждый.

В силу билинейности свертки (линейности по каждому из операндов)

$$x[n] \otimes h[n] = \left(\sum_{k=-\infty}^{\infty} x_k[n] \right) \otimes h[n] = \sum_{k=-\infty}^{\infty} (x_k[n] \otimes h[n]) = \sum_{k=-\infty}^{\infty} y_k[n].$$

Как видно из полученного выражения, свертка последовательности бесконечной длины с конечной импульсной характеристикой может быть точно заменена бесконечной суммой сверток сегментов фиксированной длины с этой же импульсной характеристикой. Каждая частичная свертка требует для своего вычисления $(L + M - 1)$ -точечного БПФ. Поскольку L выбирается произвольно, секционирование позволяет реализовать КИХ-фильтрацию длинных (потенциально – бесконечно длинных) последовательностей; при этом результат $y_k[n]$ фильтрации сегмента $x_k[n]$ появляется с задержкой, определяемой временем прямого и обратного БПФ и умножения, но эта задержка теперь отсчитывается от момента окончания сегмента, а не всей последовательности.

Результат каждой частичной свертки имеет длину $(L + M - 1)$, а следуют секции с относительным сдвигом L . Таким образом, результаты частичных сверток $y_k[n]$ суммируются с перекрытием носителей $kL \leq n \leq (k+1)L + M - 1$. Этот метод секционирования свертки известен как метод перекрытия с суммированием (overlap-add method) [7].

Альтернативный метод перекрытия с накоплением (overlap-save method) заключается в том, что перекрываются сегменты входной последовательности. Последовательность $x[n]$ разбивается на секции $x_k[n]$ длиной $(L + M - 1)$, причем последние $(M - 1)$ отсчетов каждой секции перекрываются с таким же количеством первых отсчетов следующей секции (M – по-прежнему длина импульсной характеристики, L выбирается произвольно). Каждая секция подвергается БПФ с основанием $(L + M - 1)$. Импульсная характеристика длиной M дополняется нулями до основания БПФ. Результат фильтрации (циклической свертки) содержит всего $(L + M - 1)$ отсчетов, из которых последние $(M - 1)$ отсчетов – «ошибочные», не совпадающие с результатом аперiodической свертки, а остальные L отсчетов являются «правильными». Суммируются результаты циклических частичных сверток после отбрасывания «оши-

бочных» отсчетов, т. е. секции выходной последовательности длиной L следуют друг за другом без перекрытия.

13.2. Синтез БИХ-фильтров на основе аналого-цифровой трансформации

Сравнение БИХ-фильтров с КИХ-фильтрами показывает, что для получения примерно одинаковых частотно-избирательных свойств (имеется в виду крутизна спада АЧХ) КИХ-фильтр должен иметь в $10 \dots 20$ раз более высокий порядок. Это вполне объяснимо, так как известно, что быстро изменяющиеся сигналы имеют широкий спектр, а благодаря двойственности (дуализму) времени и частоты отсюда следует, что круто изменяющейся функции частоты (АЧХ) должна соответствовать функция времени (импульсная характеристика) большой длительности. Но для КИХ-фильтра порядок – это количество отсчетов импульсной характеристики минус 1, в то время как БИХ-фильтр даже первого порядка имеет импульсную характеристику бесконечной длительности. Таким образом, в тех случаях, когда вид фазочастотной характеристики не играет определяющей роли для практического применения разрабатываемого фильтра, следует использовать БИХ-фильтр, так как при этом получается существенный выигрыш в быстродействии (или в аппаратных затратах на реализацию) фильтра. То обстоятельство, что не всякий БИХ-фильтр оказывается устойчивым, не представляет такой заметной опасности, как может показаться на первый взгляд. Во-первых, устойчивость может быть проверена (и обеспечена) на этапе проектирования цифрового фильтра; во-вторых, характеристики цифровых фильтров не подвержены дрейфу с течением времени и при изменении внешних условий, следовательно, устойчивый фильтр останется устойчивым в течение всего времени работы (конечно, нужно учитывать возможность выхода аппаратуры из строя в результате катастрофического отказа; при этом может возникнуть неустойчивость). Следует также отметить, что в цифровых БИХ-фильтрах могут возникать незатухающие паразитные колебания (так называемые предельные циклы) вследствие своеобразного нарушения устойчивости при округлении дробных чисел; подробнее см. [7].

Наиболее широко применяются методы синтеза цифровых БИХ-фильтров, основанные на так называемой аналого-цифровой трансформации, т.е. на преобразовании аналогового фильтра с требуемыми характеристиками в цифровой (дискретный) фильтр.

Это объясняется, во-первых, трудностью решения задачи прямой аппроксимации желаемых характеристик дробно-рациональными передаточными функциями и, во-вторых, наличием развитой теории синтеза

аналоговых фильтров и простотой преобразования аналоговых фильтров-прототипов в дискретные фильтры.

В качестве фильтров-прототипов наиболее часто применяются аналоговые фильтры Баттерворта, Чебышева, Золотарева-Кауэра (эллиптические) и Бесселя. Фильтры Баттерворта имеют при заданном порядке максимально гладкую АЧХ. Фильтры Чебышева имеют АЧХ, пульсирующую в полосе пропускания (фильтры I рода) или в полосе заграждения (фильтры II рода). АЧХ эллиптического фильтра пульсирует как в полосе пропускания, так и в полосе заграждения и поэтому имеет максимальную крутизну спада.

Все перечисленные фильтры характеризуются заметной нелинейностью фазочастотной характеристики. Фильтр Бесселя имеет ФЧХ, близкую к линейной в полосе пропускания.

Аналоговые фильтры принято описывать передаточными функциями, которые связаны с импульсными характеристиками преобразованием Лапласа [8]. Преобразование Лапласа связывает аналоговый сигнал $x(t)$ с его образом (изображением) в виде функции $X(p)$ комплексного переменного p . Мнимая ось комплексной p -плоскости представляет собой ось частот в описании аналогового сигнала. Аналогичную роль в описании дискретных сигналов играет единичная окружность z -плоскости.

Аналого-цифровая трансформация состоит в установлении связи комплексных переменных p и z . Выразив p в виде функции $p = f(z)$ и подставив в выражение передаточной функции $H_a(p)$ аналогового фильтра-прототипа, мы получили бы функцию комплексного переменного z , имеющую смысл передаточной функции $H(z)$ дискретного фильтра. Трудность состоит в том, что, во-первых, мнимая ось p -плоскости имеет бесконечную, а единичная окружность z -плоскости – конечную длину 2π . Во-вторых, реализуемы только ЛИС-цепи конечного порядка, поэтому подстановка $p = f(z)$ в дробно-рациональную функцию $H_a(p)$ должна давать также дробно-рациональную функцию.

Поскольку на единичной окружности $z = e^{j\omega}$, а при дискретизации должно обеспечиваться равенство $\omega = \Omega T_d$, связь комплексных переменных p и z , обусловленная дискретизацией аналоговых сигналов, описывается выражениями

$$z = \exp(pT_d) \quad (13.4)$$

и

$$p = \ln z / T_d, \quad (13.5)$$

которые не являются дробно-рациональными. Различные способы преодоления этой трудности лежат в основе двух рассмотренных ниже методов аналого-цифровой трансформации.

Метод инвариантности импульсной характеристики. Передаточная функция произвольного аналогового фильтра (с сосредоточенными параметрами) имеет вид дробно-рациональной функции комплексного переменного p . Такая функция может быть представлена в виде суммы дробей

$$H_a(p) = \sum_{k=1}^N \frac{A_k}{(p - p_k)}$$

где p_k , $1 \leq k \leq N$ – полюсы передаточной функции, а коэффициенты A_k определяются из условия равенства числителя передаточной функции $H_a(p)$ и числителя правой части после приведения ее к общему знаменателю. (Здесь мы рассматриваем лишь практически важный случай, когда степень числителя $H_a(p)$ меньше степени знаменателя, а все корни знаменателя некрратные.)

Ввиду линейности преобразования Лапласа импульсная характеристика такого фильтра имеет вид суммы экспоненциальных функций непрерывного времени $h_a(t) = \sum_{k=1}^N A_k \exp(p_k t) \cdot \sigma(t)$, $\sigma(t)$ – функция Хевисайда.

Очевидно, для того чтобы импульсная характеристика затухала со временем (т.е. фильтр был устойчивым), необходимо и достаточно, чтобы все полюсы были расположены в p -плоскости слева от мнимой оси.

Метод аналого-цифровой трансформации, известный под названием метода инвариантности импульсной характеристики, основан на прямом применении теоремы отсчетов (теоремы Котельникова). Рассматривая импульсную характеристику аналогового фильтра-прототипа как функцию времени (сигнал), можно заменить ее последовательностью отсчетов, выбранных с достаточно малым шагом дискретизации T_d .

Результатом дискретизации импульсной характеристики аналогового фильтра будет последовательность

$$h[n] = h_a(nT_d) = \sum_{k=1}^N A_k \exp(p_k nT_d) \cdot u[n] = \sum_{k=1}^N A_k r_k^n \cdot u[n],$$

где $r_k = \exp(p_k T_d)$ – полюсы передаточной функции цифрового фильтра, $u[n]$ который $u[n] = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0. \end{cases}$

Из полученного выражения видно, что при дискретизации импульсной характеристики каузального аналогового фильтра с дробно-рациональной передаточной функцией получается сумма каузальных экспоненциальных последовательностей, следовательно, реализуемому аналоговому фильтру соответствует реализуемый цифровой фильтр.

Кроме того, полюсы цифрового фильтра связаны с полюсами фильтра-прототипа соотношением $r_k = \exp(p_k T_d)$, $1 \leq k \leq N$, поэтому устойчивому аналоговому фильтру ($\text{Re}\{p_k\} < 0$) соответствует устойчивый цифровой фильтр того же порядка N (так как $|r_k| = |\exp(p_k T_d)| < 1$). Зная полюсы цифрового фильтра, можно сразу записать его передаточную функцию

$$H(z) = \sum_{k=1}^N \frac{A_k}{(1 - r_k z^{-1})},$$

и на этом аналого-цифровая трансформация заканчивается, так как, зная передаточную функцию, легко составить структурную схему и разностное уравнение цифрового фильтра.

Поскольку импульсная характеристика цифрового фильтра есть продукт дискретизации импульсной характеристики аналогового фильтра, КЧХ цифрового фильтра связана с КЧХ аналогового фильтра соотношением

$$H(e^{j\omega}) = H(e^{j\Omega T_d}) = \frac{1}{T_d} \sum_{k=-\infty}^{\infty} H_a[j(\Omega + k\Omega_d)], \quad (13.6)$$

где $\Omega_d = 2\pi/T_d$ – частота дискретизации. Очевидно, если КЧХ прототипа не финитна, а это всегда так для фильтров конечного порядка, то неизбежно наложение (суммирование) «хвостов» сдвинутых копий $H_a(\bullet)$, и, как следствие, искажение формы получаемой КЧХ дискретного фильтра по отношению к КЧХ фильтра-прототипа. Этот эффект ограничивает практическое применение метода инвариантности импульсной характеристики в основном задачами синтеза цифровых фильтров нижних частот.

Метод билинейного преобразования основан на аппроксимации выражения (13.5), позволяющей сохранить дробную рациональность передаточной функции. Подставив разложение функции логарифма в ряд, ограниченное первым слагаемым, получим

$$p = \frac{2}{T_d} \left(\frac{z-1}{z+1} + \frac{(z-1)^3}{3(z+1)^3} + \frac{(z-1)^5}{5(z+1)^5} + \dots \right) \approx \frac{2}{T_d} \frac{z-1}{z+1},$$

или

$$p = \frac{2}{T_d} \cdot \frac{1 - z^{-1}}{1 + z^{-1}}. \quad (13.7)$$

Это выражение дробно-рационально относительно z^{-1} поэтому после его подстановки в дробно-рациональную передаточную функцию $H_a(p)$ аналогового прототипа получается снова дробно-рациональная, а следовательно, реализуемая передаточная функция $H(z)$ цифрового фильтра.

Выясним, как располагаются в z -плоскости полюсы передаточной функции $H(z)$, если полюсы передаточной функции прототипа $H_a(p)$ находятся в левой части комплексной плоскости (иными словами, является ли устойчивым цифровой фильтр, если устойчив фильтр-прототип).

Выразим на основе (13.7) комплексное переменное z через p :

$$z^{-1}pT_d + pT_d = 2 - 2z^{-1};$$

отсюда

$$z = \frac{2 + pT_d}{2 - pT_d}. \quad (13.8)$$

Чтобы выяснить, в какое множество z -плоскости отображается мнимая ось p -плоскости (ось частоты), подставим в это выражение $j\Omega$ вместо p , тогда получим выражение для образа мнимой оси p -плоскости при отображении, описываемом выражением (13.8):

$$z = \frac{2/T_d + j\Omega}{2/T_d - j\Omega}.$$

Числитель и знаменатель этой дроби суть комплексно-сопряженные числа, поэтому модуль дроби равен единице при всех Ω . Это означает, что мнимая ось p -плоскости отображается преобразованием (13.8) на единичную окружность z -плоскости. Но переменная Ω – это частота, соответствующая описанию аналогового фильтра; роль частотной оси для цифровых цепей играет единичная окружность на z -плоскости (множество точек $e^{j\omega}$ при значениях ω , принимающих значения из интервала от $-\pi$ до π). Заменяя z на $e^{j\omega}$, получим

$$e^{j\omega} = \frac{2/T_d + j\Omega}{2/T_d - j\Omega} = \frac{1 + j\Omega T_d/2}{1 - j\Omega T_d/2}.$$

Тогда

$$\omega = \arg \left\{ \frac{1 + j\Omega T_d/2}{1 - j\Omega T_d/2} \right\} = 2 \operatorname{argtg} \frac{\Omega T_d}{2}$$

следовательно, связь «аналоговой» и «цифровой» частот при билинейном преобразовании описывается выражениями

$$\begin{aligned} \omega/2 &= \operatorname{argtg}(\Omega T_d/2), \\ \Omega T_d/2 &= \operatorname{tg} \omega/2. \end{aligned} \quad (13.9)$$

Поскольку $\pi \leq \omega \leq -\pi$, а $-\infty < \Omega < \infty$, нетрудно видеть, что вся аналоговая частотная ось (бесконечной длины) отображается на единичную окружность (длины 2π), причем это отображение однократно в отличие от (13.6), и вследствие этого различные участки оси Ω испытывают раз-

личное «сжатие» при отображении на ось ω (единичную окружность). Это необходимо учитывать при проектировании цифровых фильтров на этапе определения требований к частотам среза фильтров-прототипов.

Заслуживает внимания вопрос, насколько вредна нелинейная трансформация частотной оси с точки зрения задачи синтеза цифровых фильтров. Очевидно, что при проектировании фильтров с желаемыми АЧХ кусочно-постоянного вида указанная нелинейность трансформации частотной оси не влияет на качество цифрового фильтра, так как приводит лишь к необходимости на этапе построения аналогового фильтра-прототипа учитывать последующее изменение характерных частот фильтра (граничных частот) при билинейном преобразовании. Если же требуется построить фильтр, не являющийся типовым (ФНЧ, ФВЧ, полосовым или режекторным), то в общем случае нелинейность отображения частотной оси приводит к искажению формы АЧХ. Например, интегрирующий аналоговый фильтр имеет амплитудно-частотную характеристику гиперболического вида $\sim 1/\Omega$ и при билинейном преобразовании не приводит к интегрирующему цифровому фильтру.

Для того чтобы устойчивый аналоговый фильтр трансформировался в устойчивый же цифровой фильтр, требуется, чтобы при билинейном преобразовании левая полуплоскость p -плоскости отображалась внутрь единичной окружности. Разлагая p на мнимую и вещественную части, получим для билинейного преобразования (13.8)

$$z = \frac{2 + \operatorname{Re}\{pT_d\} + j \operatorname{Im}\{pT_d\}}{2 - \operatorname{Re}\{pT_d\} - j \operatorname{Im}\{pT_d\}}$$

Поскольку мнимые части числителя и знаменателя одинаковы, модуль дроби будет меньше 1, если $\operatorname{Re}(pT_d) < 0$. Тогда любой полюс функции $H_a(p)$, лежащий слева от мнимой оси, отображается в полюс функции $H(z)$, расположенный внутри единичной окружности. Это означает, что устойчивый аналоговый фильтр трансформируется преобразованием (13.7) в устойчивый дискретный фильтр.

Итак, установлено, что билинейное преобразование трансформирует устойчивый реализуемый аналоговый фильтр в устойчивый реализуемый цифровой фильтр. При этом вследствие однократности отображения частотной оси на единичной окружности отсутствует наложение «хвостов» КЧХ, что является достоинством билинейного преобразования. К недостаткам следует отнести то, что не сохраняются ни импульсная, ни фазочастотная характеристики фильтра (точнее говоря, импульсная и фазочастотная характеристики дискретного фильтра могут сильно отличаться по форме от соответствующих характеристик прототипа).

Порядок расчета цифрового фильтра методом билинейного преобразования состоит в следующем:

1) определение характерных частот ЦФ и пересчет их в частоты аналогового фильтра в соответствии с (13.9);

2) синтез аналогового фильтра, удовлетворяющего заданным условиям;

3) подстановка формулы (13.7) билинейного преобразования в выражение $H_a(p)$ передаточной функции фильтра-прототипа.

Реализация цифровых фильтров (и других алгоритмов цифровой обработки сигналов) возможна на различной элементной базе.

Выбор конкретного воплощения алгоритма ЦОС производится разработчиком с учетом различных показателей, к которым относятся стоимость, массогабаритные характеристики, энергопотребление, быстродействие и т.п. В каждом конкретном случае один или несколько показателей играют наиболее важную роль в выборе способа реализации. Например, в системах подвижной радиосвязи главными показателями являются быстродействие (обработка должна выполняться в реальном времени) и массогабаритные характеристики, при этом желательно обеспечить малое энергопотребление и умеренную цену мобильной станции. Устройство обработки сигналов в таких системах работает по жестким алгоритмам, которые не изменяются в течение всего срока эксплуатации изделия; точность представления данных (разрядность) должна быть достаточна для обеспечения комфортности восприятия речи (разборчивости и возможности узнавания собеседника) и является поэтому сравнительно невысокой. В системах обработки геофизической информации, напротив, обработка в реальном времени не требуется, массогабаритные характеристики не играют доминирующей роли, однако на передний план выступают точность представления данных и гибкость реализуемых алгоритмов.

Поэтому в таких системах обработка сигналов реализуется обычно на универсальных цифровых вычислительных машинах. В системах реального времени (например, в подвижной радиосвязи), когда на обработку отсчета сигнала отводится временной интервал, равный шагу дискретизации, как правило, используются специализированные цифровые устройства, называемые цифровыми сигнальными процессорами (ЦСП). Следует отметить, что в последнее время в устройствах цифровой обработки сигналов широкое распространение получили программируемые логические интегральные схемы (ПЛИС). ПЛИС представляет собой интегральную схему сверхвысокой степени интеграции, содержащую на кристалле порядка 1 миллиона логических вентилях, которые могут быть программным путем соединены в логическую структуру, реализующую заданный алгоритм цифровой обработки сигналов аппаратным способом. Таким образом, ПЛИС сочетают в себе преимущества аппаратной реализации алгоритмов (главное из которых – быстродействие) с достоинствами программируемых устройств.

IV. КАНАЛЫ СВЯЗИ

14. КАНАЛЫ СВЯЗИ

14.1. Основные определения

В общем случае под каналом связи понимают всю совокупность технических средств, обеспечивающих передачу электрических сигналов от источника сообщений к потребителю.

Каналы связи классифицируют по различным признакам:

- по назначению;
- по характеру линий связи;
- по диапазону частот;
- по характеру сигналов на входе и выходе каналов и т.п.

В теории передачи сигналов каналы классифицируют по характеру сигналов на входе и выходе и различают:

- непрерывные;
- дискретные;
- дискретно-непрерывные каналы.

В непрерывных каналах сигналы на входе и выходе непрерывны по уровням;

в дискретных каналах – они соответственно дискретны;

а в дискретно-непрерывных – сигналы на входе дискретны, а на выходе непрерывны, и наоборот.

Возможна также классификация каналов по

- назначению: телеграфные, телефонные, телевизионные, телеметрические и др.;
- виду физической среды распространения: проводные, кабельные, волноводные и др.;
- и диапазону используемых ими частот (табл. 14.1, 14.2).

К радиодиапазону относят частоты в пределах $30 \dots 30 \cdot 10^{12}$ Гц, что соответствует длинам волн от 10^8 м до 0,1 мм.

Кроме радиодиапазона, в настоящее время широкое распространение нашел и оптический диапазон волн.

В силу дискретного характера электромагнитного излучения в оптическом диапазоне волн такие каналы принято называть квантовыми.

Таблица 14.1

Частотные диапазоны проводных каналов связи

Длина волны	Вид линии (диапазон)	Диапазон частот
10^{-1} м	ультрафиолетовый видимый инфракрасный	10^{15} Гц 10^{14} Гц
1 мм	Волновод	100 ГГц
1 см		10 ГГц
1 дм		1 ГГц
1 м		100 МГц
10 м		10 МГц
100 м	Коаксиальный кабель	1 МГц
1 км		100 кГц
10 км	Витая пара	10 кГц
100 км		1 кГц

Таблица 14.2

Частотные диапазоны радиоволн

Ультракороткие волны			
сантиметровые	дециметровые	метровые	
3 - 30 ГГц	300 МГц - 3 ГГц	30 - 300 МГц	
10 - 100 мм	100 мм - 1 м	1 - 10 м	
Короткие волны	Средние волны	Длинные волны	
3 - 30 МГц	300 кГц - 3 МГц	30 - 300 кГц	
10 - 100 м	100 м - 1 км	1 - 10 км	
Сверхдлинные волны			
мириаметровые	гектокилометровые	мегаметровые	декамегаметровые
3 - 30 кГц	300 Гц - 3 кГц	30 - 300 Гц	3 - 30 Гц
10 - 100 км	100 км - 1 Мм	1 - 10 Мм	10 - 100 Мм

По способу распространения радиоволн различают каналы: с открытым и с закрытым распространением.

В каналах с закрытым распространением электромагнитная энергия распространяется по направляющим линиям (кабельные, проводные, волноводные СВЧ тракты и др.): для них характерны малый уровень

помех и постоянство параметров сигнала, что позволяет передавать информацию с высокой скоростью и достоверностью.

Особенности использования радиоволн различных диапазонов в каналах с открытым распространением состоят в следующем. В диапазонах инфранизких (ИНЧ), очень низких (ОНЧ) и низких (НЧ) частот на небольших расстояниях поле в месте приема создается за счет дифракционного огибания волнами выпуклой поверхности Земли. На больших расстояниях радиоволны распространяются в своеобразном сферическом волноводе, внутренняя стенка которого образуется поверхностью Земли, а внешняя – ионосферой. Особенностью этих диапазонов является также способность волн проникать в толщу Земли и воды на глубину в десятки метров.

Принципиальным недостатком таких каналов являются: ограниченная полоса частот (единицы герц) и очень большие линейные размеры антенных устройств, соизмеримых с длиной волны, составляющей километры.

Сверхдлинные волны применяются для навигации и передачи информации на подводные объекты.

В распространении волн диапазона высоких частот (ВЧ) принимает участие ионосфера: если волны длиннее 1 км отражаются от нижнего ее слоя практически зеркально, то декаметровые волны достаточно глубоко проникают в ионосферу, что приводит к эффекту многолучевости, когда в точку приема приходят одновременно несколько сигналов с различным временем запаздывания.

Декаметровые волны широко применяются для глобальной связи и радиовещания. С их помощью можно передавать информацию сравнительно большого объема в пределах всего земного шара при ограниченной мощности передатчика и небольших по размеру антеннах. Полоса частот передаваемых сигналов в декаметровом канале не превышает десяти килогерц.

До появления спутниковых систем связи этот диапазон был единственным пригодным для организации связи между двумя любыми пунктами на Земле без промежуточной ретрансляции.

Гектометровые волны днем распространяются как земные, а ночью – как ионосферные. Дальность распространения земной волны над сушей не превышает 500 км, а над морем – 1000 км. Диапазон средних частот широко используется в радиовещании, связи и радионавигации.

Волны диапазона частот от 30 МГц и выше слабо дифрагируют и поэтому распространяются в пределах прямой видимости. Некоторого увеличения дальности можно достичь, применив поднятые антенны, а для организации связи на расстояния, превышающие прямую видимость, ретрансляцию сигналов. Системы с ретрансляцией сигналов называются радиорелейными линиями.

Одним из основных достоинств высокочастотных диапазонов является большой частотный ресурс, что позволяет создавать радиосистемы передачи информации с высокой скоростью передачи и радиосети с большим числом одновременно работающих радиостанций.

Диапазон миллиметровых волн. Его особенностью является сильное поглощение радиоволн в дожде и тумане, что ограничивает их применение в наземных системах большой дальности. Однако в космических и спутниковых системах они весьма перспективны.

Диапазон рабочих частот 40 МГц... 40 ГГц (спутниковая связь). В настоящее время наибольшее использование находит диапазон 1... 12 ГГц.

Обычно ИСЗ находятся на высоте от 500 до 40 000 км от поверхности Земли и поэтому обеспечивают радиосвязь между земными станциями, удаленными на расстояния до 10... 17 тыс. км.

Длина волн 0,5... 10,6 мкм, диапазон оптической связи (видимый 0,5... 0,76 мкм и инфракрасный 0,76... 10,6 мкм участки спектра электромагнитных колебаний).

Широкая полоса частот оптических каналов связи позволяет создавать каналы и сети связи с огромной пропускной способностью.

Представляет интерес классификация каналов телекоммуникаций по характеру сигналов на входе и выходе. Обычно различают:

- непрерывные каналы, на входе и выходе которых сигналы аналоговые (непрерывные по уровням);
- дискретные каналы, на выход которых поступают дискретные по уровням сигналы, и с выхода также снимаются дискретные сигналы;
- дискретно-непрерывные или полунепрерывные, т.е. дискретные каналы со стороны входа и непрерывные со стороны выхода или наоборот.

Совокупность технических средств, включенных между модулятором и демодулятором (рис. 14.1), т.е. выходные каскады передатчика, передающая антенна, среда распространения, приемная антенна и линейная часть приемника, образуют непрерывный канал, так как входные и выходные радиосигналы непрерывны по своей природе.

Рассматривая часть системы между выходом кодера и входом демодера, получим дискретный канал.

Наконец, часть системы между выходом кодера и входом демодулятора образует дискретно-непрерывный канал.

Таким образом, дискретный канал содержит дискретно-непрерывный, который в свою очередь включает в себя непрерывный канал.

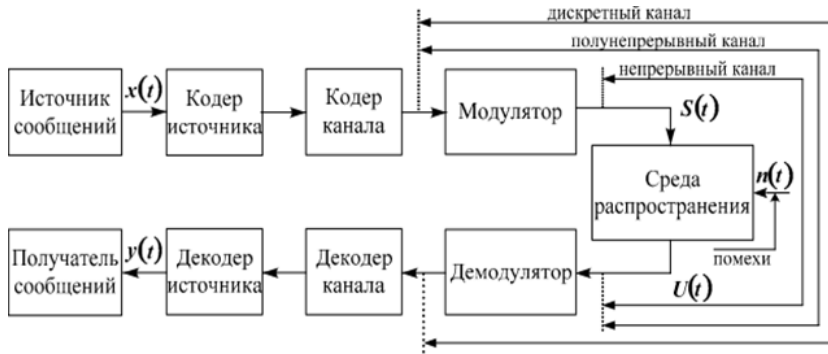


Рис. 14.1. Структурная схема системы электрической связи

14.2. Модели непрерывных каналов

Непрерывными называются каналы, входные и выходные сигналы которых принимают произвольные значения из некоторого интервала.

Непрерывные каналы можно классифицировать по виду преобразования входного сигнала $S_p(t, \lambda_0)$ в полезный принятый $S_p(t, \lambda)$. Если ограничиться предположением, что в канале действует аддитивный нормальный белый шум $n(t)$, то непрерывные каналы подразделяются по виду преобразования $S_p(t, \lambda_0)$ в $S_p(t, \lambda)$, т.е. по виду искажений сигнала. В большинстве радиотехнических систем излученные сигналы являются узкополосными:

$$S_p(t, \lambda_0) = A(t) \cos[\omega_0 t + \varphi(t)],$$

где $A(t)$ и $\varphi(t)$ – функции, отображающие законы амплитудной и угловой модуляции; ω_0 – несущая частота сигнала.

Искажения излученного сигнала $S_p(t, \lambda_0)$ принято рассматривать отдельно для однолучевых и многолучевых каналов. В однолучевых каналах электромагнитные колебания распространяются по одному пути. Однолучевыми каналами являются линии связи на расстояниях прямой видимости: линии ближней радиосвязи на коротких и ультракоротких волнах, линии связи Земля–воздух, воздух–Земля, воздух–воздух и т.п.

Принятый полезный сигнал по отношению к излученному характеризуется дополнительными параметрами: случайным ослаблением $\alpha(t)$, средним временем запаздывания τ_3 , доплеровским смещением частоты Ω и случайной начальной фазой θ и может быть записан в виде

$$S_p(t, \lambda) = \alpha(t) A(t - \tau_3) \cos[(\omega_0 + \Omega)t + \varphi(t - \tau_3) - \theta].$$

Таким образом, совокупность параметров принятого сигнала $\lambda = \{A(t), \varphi(t), \omega_0, \alpha(t), \tau_3, \Omega, \theta\}$. Зная значения дополнительных парамет-

ров $\alpha(t)$, τ_3 , Ω , θ на приемной стороне, можно выделить несколько моделей непрерывных каналов [2, 9].

Гауссовским каналом называется канал, в котором действует аддитивный нормальный белый шум, а искажения полезного сигнала не существенны и могут быть скомпенсированы. Компенсация искажений возможна, если на приемной стороне дополнительные параметры полностью известны или могут быть измерены достаточно точно. Поэтому можно считать, что $S_p(t, \lambda_0) = S_p(t, \lambda)$. Выходной сигнал гауссовского канала

$$Y_p(t) = S_p(t, \lambda_0) + n(t).$$

Представление выходного сигнала в виде суммы полезного сигнала и нормального белого шума $n(t)$ позволяет указать правило принятия решения о переданном сигнале.

Гауссовский канал с неизвестной фазой сигнала определяется параметрами τ_3 , Ω , $\alpha(t) = \alpha$, которые постоянны и известны. Фаза θ считается равномерно распределенной величиной в интервале $0 \div 2\pi$. Такая модель хорошо описывает процессы в линиях радиосвязи на расстояниях прямой видимости.

Канал с амплитудными замираниями является дальнейшим усложнением канала с неизвестной фазой в предположении, что $\alpha(t)$ – случайная функция времени. Выходной полезный сигнал канала с замираниями

$$S_p(t, \lambda) = \alpha(t) A(t) \cos[(\omega_0 + \Omega)t + \varphi(t) - \theta].$$

Случайная функция $\alpha(t)$ перемножается с сигналом и поэтому называется мультипликативной помехой, которую можно рассматривать как функцию, модулирующую по амплитуде излученный сигнал. Модуляция приводит к расширению спектра принятого сигнала относительно спектра излученного сигнала. Поэтому такой канал называют каналом с рассеянием энергии по частоте.

По времени корреляции мультипликативные помехи разделяются на медленные и быстрые [10]. О медленных замираниях говорят в случае, если время корреляции $\alpha(t)$ значительно превышает интервал наблюдения сигнала. Причинами медленных замираний являются изменения свойств среды распространения радиоволн в зависимости от метеорологических условий, времени суток и года, от солнечной активности и т.п. Быстрая мультипликативная помеха имеет время корреляции меньше, чем интервал наблюдения сигнала. Основной причиной быстрых замираний является наличие многих путей, по которым распространяются электромагнитные волны. Многолучевое распространение возникает при передаче информации на дальние расстояния при отражении радиоволн от протяженных поверхностей суши и моря, при отражении от ионосферы и тропосферы. Из-за разных путей распространения время запаздывания отдельных принимаемых сигналов различно.

Поэтому многолучевые каналы называют также каналами с рассеянием энергии во времени.

Результирующий сигнал на выходе многолучевого канала

$$S(t, \lambda) = \sum_{i=1}^l S_{pi}(t, \lambda_{0i}),$$

где $S_{pi}(t, \lambda_{0i}) = \alpha_i(t) A(t) \cos[(\omega_0 + \Omega)t + \varphi(t) - \theta]$.

При большом числе путей можно считать, что $S_p(t, \lambda)$ является реализацией нормального СП. Обычно среднее значение процесса равно нулю, тогда модели многолучевых каналов различаются по виду КФ [11].

Таким образом, непрерывный канал считается заданным, если указаны мощность сигналов, полоса частот, дано описание моделей помех и искажений сигналов.

14.3. Модели дискретных каналов

Дискретными называются каналы, входные и выходные сигналы которых принимают конечное число мгновенных значений. Понятие дискретного канала естественно возникает при передаче дискретных сообщений и определяется как совокупность технических средств, включенных между кодером и декодером канала (рис. 14.1).

Переход от дискретных сигналов к непрерывным осуществляется на передающей стороне при манипуляции параметрами непрерывной несущей. На приемной стороне дискретные сигналы появляются на выходе первой решающей схемы (демодулятора).

Свойства дискретного канала определяются непрерывным каналом и структурой модема. Дискретный канал задается множеством входных $\{s_i\}$, $i = \overline{1, L_s}$ и выходных $\{y_j\}$, $j = \overline{1, L_y}$ символов (сигналов), длительностью

символов τ и условными вероятностями $P(y_j/s_i)$ преобразования входных символов в выходные. Обычно длительности всех входных и выходных символов одинаковы. Объемы алфавитов входных L_s и выходных L_y сигналов в общем случае могут быть разными, причем $L_y \geq L_s$. Однако в большинстве случаев $L_y = L_s$. Для дискретных каналов широко используется представление принятой последовательности символов $Y = (y_1, y_2, \dots, y_n)$ в виде суммы переданной последовательности $S = (s_1, s_2, \dots, s_n)$ и комбинации помехи (вектора ошибки) $E = (e_1, e_2, \dots, e_n)$

$$Y = S \oplus E,$$

где \oplus – понимается как поразрядное сложение S и E по модулю L_s . В случае двоичных последовательностей ($L_s = 2$) нулевой символ вектора ошибки $e_i = 0$ означает, что i -й символ принят правильно ($y_i = s_i$), а $e_i = 1$ указывает на ошибку в приеме ($y_i \neq s_i$).

Классификацию дискретных каналов удобно вести по вектору ошибки E . Разные модели каналов различаются распределением вероятностей вектора E . Наиболее распространены следующие модели [2].

Канал без памяти – это канал, в котором символы e_i являются независимыми СВ. Прием каждого сигнального символа в таком канале не зависит от результата приема предыдущих символов. При наличии такой зависимости имеет место канал с памятью. Дискретный канал называется стационарным, если вероятность ошибочного приема символов не изменяется с течением времени.

В силу простоты технической реализации наибольшее применение находят каналы, сигналы в которых представляются двоичным кодом. Такие каналы называются двоичными (бинарными) и задаются с помощью графа (рис. 14.2). Вероятности $P(0/0)$ и $P(1/1)$ характеризуют правильный прием символов 0 и 1 соответственно, а $P(1/0)$ и $P(0/1)$ – вероятности ошибок при приеме символов 0 и 1.

Симметричным двоичным называется канал, в котором вероятности ошибок при приеме 0 и 1 одинаковы, $P(1/0) = P(0/1)$, а следовательно, равны и вероятности правильного приема символов $P(0/0) = P(1/1) = 1 - p$. Для симметричного стационарного канала без памяти вероятность искажения i -го символа $P(e_i = 1) = p$, а вероятность правильного приема $P(e_i = 0) = 1 - p$.

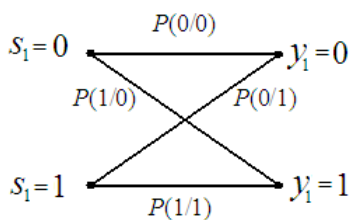


Рис. 14.2. Граф двоичного канала

Двоичный канал без памяти со стиранием отличается от рассмотренного тем, что выходной алфавит помимо 0 и 1 содержит третий символ «?» – символ стирания. Он появляется в тех случаях, когда демодулятор не может надежно опознать переданный символ. Такой канал часто используется в системах передачи информации с обратной связью, когда при приеме символа «?» производится повторение передачи. Это позволяет значительно снизить вероятность ошибочного приема за счет уменьшения скорости передачи.

Марковский канал является простейшей моделью дискретного канала с памятью. Он характеризуется вектором ошибки, символы которого образуют простую цепь Маркова [12]. Вероятность искажения

символа в этом канале зависит от результата приема только предыдущего символа.

Марковская модель задается матрицей переходных вероятностей:

$$p = \begin{bmatrix} 1 - p_1 & p_1 \\ p_2 & 1 - p_2 \end{bmatrix},$$

где p_1 – условная вероятность принять $(i + 1)$ -й символ ошибочно, если i -й принят правильно; $1 - p_1$ – условная вероятность принять $(i + 1)$ -й символ правильно, если i -й принят правильно; p_2 – условная вероятность принять $(i + 1)$ -й символ ошибочно, если i -й принят ошибочно; $1 - p_2$ – условная вероятность принять $(i + 1)$ -й символ правильно, если i -й принят ошибочно.

Безусловная (средняя) вероятность ошибки в рассматриваемом канале должна удовлетворять уравнению:

$$p(x_{i+1} / x_i) = p_2 \cdot p_{\text{ош}}(x_i) + p_1 \cdot p_{\text{прав}}(x_i)$$

или

$$p(x_{i+1} / x_i) = p_1 / (1 + p_1 + p_2).$$

Данная модель имеет достоинство – простоту использования, но не всегда адекватно воспроизводит свойства реальных каналов. Большую точность позволяет получить модель Гильберта для дискретного канала с памятью. В такой модели канал может находиться в двух состояниях S_1 и S_2 . В состоянии S_1 ошибок не происходит; в состоянии S_2 ошибки возникают с вероятностью p_2 .

Также считаются известными вероятности перехода $p(S_1 / S_2)$ из состояния S_1 в S_2 и вероятности перехода $p(S_2 / S_1)$ из состояния S_2 в состояние S_1 . В этом случае простую марковскую цепь образует не последовательность ошибок, а последовательность переходов:

$$p = \begin{bmatrix} 1 - p(S_2 / S_1) & p(S_2 / S_1) \\ p(S_1 / S_2) & 1 - p(S_1 / S_2) \end{bmatrix},$$

При этом достаточно легко выразить безусловные вероятности нахождения канала в состояниях S_1 и S_2 :

$$p(S_1) = \frac{p(S_1 / S_2)}{p(S_1 / S_2) + p(S_2 / S_1)}, \quad p(S_2) = \frac{p(S_2 / S_1)}{p(S_2 / S_1) + p(S_1 / S_2)},$$

Безусловная вероятность ошибки в этом случае может быть определена по формуле:

$$p = p_2 \cdot p(S_2) = p_2 \cdot \frac{p(S_2 / S_1)}{p(S_2 / S_1) + p(S_1 / S_2)}.$$

Наиболее часто при использовании модели Гильберта для двоичного канала полагают $p_2 = 1/2$, т.е. состояние S_2 рассматривается как полный обрыв связи. Это согласуется с представлением о канале, в котором действуют коммутационные помехи.

Из других моделей симметричных двоичных каналов следует отметить канал с пакетами ошибок, который характеризуется тем, что искажающие символы (единицы) вектора ошибки группируются в пакеты. Такое группирование происходит, если в непрерывном канале, входящем в дискретный, действуют сильные замирания сигналов на время длительности нескольких символов или присутствуют импульсные помехи большой длительности. Подобные каналы задаются вероятностями искажений серий из q символов подряд.

V. ТЕОРИЯ ПЕРЕДАЧИ И КОДИРОВАНИЯ СООБЩЕНИЙ

15. ТЕОРИЯ ПЕРЕДАЧИ ИНФОРМАЦИИ

Количество информации $I(x_i)$ в отдельно взятом единичном сообщении x_i определяется величиной, обратной вероятности появления сообщения $p(x_i)$ и вычисляется в логарифмических единицах:

$$I(x_i) = \log_b \left(\frac{1}{p(x_i)} \right) = -\log_b p(x_i). \quad (15.1)$$

При $p(x_i)=1$, $I(x_i)=0$, что соответствует принятому определению информации.

Если источник выдает зависимые сообщения $x_i=x_1, \dots, x_m$, то они характеризуются условными вероятностями $p(x_i/x_1, \dots, x_m)$. И в этом случае количество информации вычисляется по формуле (15.1) с подстановкой в нее условных вероятностей сообщений.

В системах, работающих с двоичными кодами, используется основание логарифма $b = 2$ и информация измеряется в двоичных единицах, обычно вместо двоичных единиц используется эквивалентное название – бит, возникшее как сокращенная запись английских слов binary digit (двоичная цифра).

1 бит это количество информации, которое передается единичным символом сообщения, вероятность передачи которого $p(x_i) = 0,5$:

$$I(x_i) = -\log_2(0,5) = -\log_2 \left(\frac{1}{2} \right) = -\log_2 2^{-1} = 1 \text{ (бит)}.$$

15.1. Количество информации переданной по дискретному каналу

Количество информации, переданной по каналу, это разность между энтропией на выходе и энтропией шума:

$$I(Y, X) = H(Y) - H\left(\frac{Y}{X}\right) = \log_2 m + (1 - p_{\text{ош}}) \cdot \log_2 (1 - p_{\text{ош}}) + p_{\text{ош}} \cdot \log_2 \frac{p_{\text{ош}}}{m-1}, \quad (15.2)$$

$I(Y, X)$ – количество информации, переданной по каналу;

$H(Y/X)$ – энтропия шума (ложная информация, создаваемая помехами);

$H(Y)$ – энтропия на выходе канала;

$X = \{x_i\}$ – алфавит источника сообщений;

$Y = \{y_j\}$ – алфавит получателя сообщений;

$p_{\text{ош}}$ – вероятность ошибки при передаче любого символа;

m – количество символов.

Количество информации, переданной по каналу связи, обладает следующими основными свойствами:

$I(Y, X) \geq 0$, причём $I(Y, X) = 0$ тогда и только тогда, когда входные и выходные сообщения в канале взаимно независимы;

$I(Y, X) \leq H(X)$, причём $I(Y, X) = H(X)$, тогда и только тогда, когда входная последовательность определяется однозначно по выходной последовательности, например, когда в канале нет помех;

$I(Y, X) = I(X, Y) = H(Y) - H(Y/X)$ следует из того, что количество информации не изменится, если входную и выходную последовательность поменять местами.

15.2. Пропускная способность дискретного канала

Пропускной способностью канала, рассчитанной на один входной символ, называется максимальное количество информации, которое может быть передано по каналу, причём максимум ищется по всем возможным источникам X , имеющим различные (произвольные) вероятностные характеристики:

$$C = \max_X I(Y, X) \text{ [бит/символ]}.$$

Часто более удобно пользоваться пропускной способностью канала, рассчитанной не на один входной символ, а на единицу времени:

$$C = C/T \text{ [бит/с]}.$$

Величину C называют пропускной способностью канала в единицу времени или просто пропускной способностью.

Пропускная способность канала обладает следующими основными свойствами:

$C \geq 0$, $C = 0$ тогда и только тогда, когда вход и выход канала статистически независимы;

$C \leq \log_2 m / T$ для канала без помех.

Из определения пропускной способности следует, что характеристика описывает свойства канала, по которому передается информация от определенного источника. Очевидно, никакой источник не способен передать по каналу количество информации большее пропускной способности и данная характеристика описывает потенциальные возможности канала по передаче информации.

15.3. Пропускная способность симметричного дискретного канала без памяти

Пропускная способность дискретного канала с учетом (15.2), по которому передается m дискретных сигналов, вычисляется по формуле

$$C = V_{\text{И}} \left[\log_2 m + (1 - p_{\text{ош}}) \cdot \log_2 (1 - p_{\text{ош}}) + p_{\text{ош}} \cdot \log_2 \frac{p_{\text{ош}}}{m-1} \right], \quad (15.3)$$

где $V_{\text{И}} = 1/T$ – скорость модуляции, бод; T – длительность сигнала; $p_{\text{ош}}$ – вероятность ошибки в канале. Заметим, что пропускная способность дискретного канала без помех (при $p_{\text{ош}} = 0$): $C_{\text{ДК}} = V_{\text{И}} [\log_2 m]$.

В частности пропускная способность двоичного канала ($m = 2$):

$$C_{\text{ДК}} = V_{\text{И}} [1 + (1 - p_{\text{ош}}) \log_2 (1 - p_{\text{ош}}) + p_{\text{ош}} \log_2 (p_{\text{ош}})]. \quad (15.4)$$

Зависимость отношения $C/V_{\text{И}}$ от вероятности ошибки $p_{\text{ош}}$, рассчитанная по формуле (15.4), показана на рис. 15.1.

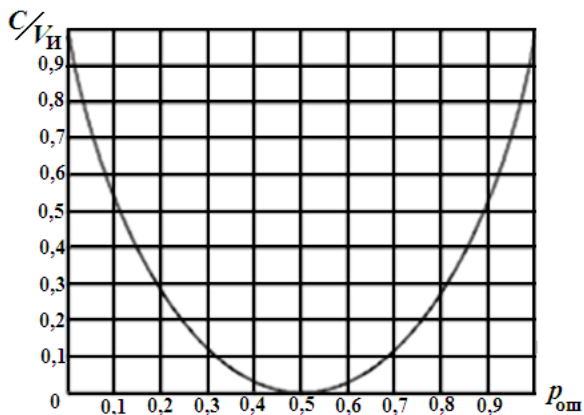


Рис. 15.1. Пропускная способность дискретного канала

Как следует из графика, при $p_{\text{ош}} = 0,5$ пропускная способность двоичного канала равна нулю ($C = 0$). Этот случай называют обрывом канала. Действительно вероятность ошибки $p_{\text{ош}} = 0,5$ можно получить и

без передачи информации по каналу связи. А при $p_{\text{ош}} = 1$ пропускная способность такая же, как и при $p_{\text{ош}} = 0$ (канал без помех). Это объясняется тем, что при $p_{\text{ош}} = 1$ достаточно заменить нули на единицы и единицы на нули, чтобы абсолютно правильно восстановить переданный сигнал.

Пример. Определим пропускную способность двоичного телеграфного канала, если скорость передачи в нем 1000 бит/с и вероятность ошибки 10^{-3} и сделаем вывод о том насколько отличается пропускная способность этого канала от идеального. Согласно формуле (15.4), при заданных параметрах:

$$C_{\text{ДК}} = 1000 [1 + 0,001 \cdot \log_2 0,001 + (1-0,001) \cdot \log_2(1-0,001)] = 989 \text{ [бит/с]}.$$

Для идеального канала при $p_{\text{ош}} = 0$ получаем $C_{\text{ДК}} = V_{\text{и}} = 1000$ бит/с. Сравнение этих величин показывает, что ошибки в канале привели к уменьшению пропускной способности на 11 бит/с (т.е. потери составили 1,1%).

15.4. Методы сжатия дискретных сообщений

Условия существования оптимального неравномерного кода.

При передаче сообщения необходимо согласовывать источник с каналом путем определения правила, по которому каждому элементу сообщения ставится в соответствие некоторый код, преобразуемый далее в сигнал.

В настоящее время существует два основных направления развития теории кодирования:

1) В одном из них рассматриваются задачи повышения достоверности передачи в каналах с помехами, решаемые применением помехоустойчивых кодов, которые позволяют обнаруживать или исправлять ошибки. Такое кодирование называется помехоустойчивым. При этом избыточность кодовой последовательности выше, чем избыточность источника сообщений.

2) Другое направление теории кодирования связано с вопросами устранения избыточности при передаче сообщений в каналах без помех. Цель кодирования при этом состоит в таком преобразовании сообщения, при котором избыточность кодовой последовательности должна стать меньше, чем избыточность сообщений источника. В результате появляется возможность увеличения скорости передачи информации или снижаются требования к пропускной способности канала.

Процесс кодирования с целью уменьшения избыточности источника сообщений называется согласованием источника с каналом, или сжатием сообщения источника (экономного кодирования, энтропийного кодирования).

Количественно избыточность оценивается коэффициентом избыточности:

$$\chi = \frac{H_{\max}(X) - H(X)}{H_{\max}(X)} = 1 - \frac{H(X)}{H_{\max}(X)}, \quad (15.5)$$

где $H(X)$ – энтропия источника;

$H_{\max}(X) = \log_2 m$ – максимальная энтропия источника с алфавитом из m сообщений.

Избыточность (15.5) равна нулю только в том случае, когда элементы сообщения появляются на выходе источника с равными вероятностями $p(x_i) = 1/m$ ($i = 1, 2, 3, \dots, m$) и независимо друг от друга $p(x_i, x_j) = p(x_i) \cdot p(x_j)$. Если же $H(X) < \log_2 m$, то оказывается возможным построение кодов, имеющих меньшую избыточность, чем источник сообщений.

Покажем это на простейшем примере.

Пусть источник имеет алфавит из четырех символов А, Б, В, Г с вероятностями $p(A) = 0,5$; $p(B) = 0,25$; $p(B) = p(\Gamma) = 0,125$.

Энтропия такого источника:

$$\begin{aligned} H(X) &= -\sum_{i=1}^4 p(x_i) \cdot \log_2 p(x_i) = \\ &= -0,5 \cdot \log_2 0,5 - 0,25 \cdot \log_2 0,25 - 0,125 \cdot \log_2 0,125 - 0,125 \cdot \log_2 0,125 = 1,75. \end{aligned}$$

При использовании равномерного кодирования, например, А → 00, Б → 01, В → 10, Г → 11. Тогда среднее число двоичных символов в сообщении, приходящихся на один символ источника, равно двум. Поскольку это на 12,5% больше энтропии источника, то используемый код не является оптимальным.

Рассмотрим теперь неравномерный код: А → 0, Б → 10, В → 110, Г → 111. В этом случае среднее число двоичных символов, приходящихся на один символ источника в сообщении,

$$n_{cp} = \sum_{i=1}^4 p(x_i) \cdot n_i = 0,5 \cdot 1 + 0,25 \cdot 2 + 0,125 \cdot 3 + 0,125 \cdot 3 = 1,75.$$

Таким образом, среднее число двоичных символов, приходящихся на один символ источника, равно энтропии источника, т.е. для указанного источника неравномерный код оказывается более экономичным, чем равномерный.

Важно отметить, что при кодировании неравномерным кодом должна обеспечиваться возможность однозначного декодирования символов сообщения.

Например, для рассмотренного источника, нецелесообразно применять код: А → 0, Б → 1, В → 10, Г → 11, поскольку прием последова-

тельности 10 может означать передачу символа В, или двух символов Б и А. Неоднозначно также декодирование символов 11. Для однозначного декодирования неравномерные коды должны удовлетворять условию префиксности: никакое более короткое слово не должно являться началом более длинного слова. Неравномерные коды, удовлетворяющие этому условию, называют префиксными.

Неравномерные коды позволяют в среднем уменьшить число двоичных символов на единичное информационное сообщение. Однако им присущ существенный недостаток: при возникновении ошибки она распространяется на все последующие элементы сообщения. Возникает ошибка синхронизации, приводящая к резкому ухудшению достоверности приема.

Этот недостаток отсутствует в равномерных кодах. При кодировании равномерными кодами используется одно и то же число двоичных символов – блок; поэтому такие коды называют блоковыми.

Показатели эффективности сжатия. Наряду с коэффициентом избыточности (15.5), часто используется коэффициент сжатия источника:

$$K_{\text{сж.м}} = \frac{H_{\text{max}}(X)}{H(X)} = \frac{1}{1 - \chi}.$$

Коэффициент сжатия источника показывает, во сколько раз можно уменьшить количество двоичных символов для представления единичного символа источника с энтропией $H(X)$ по сравнению со случаем, когда все сообщения источника передаются равномерно.

Например, для источника, рассмотренного выше, коэффициент сжатия

$$K_{\text{сж.м}} = \frac{\log_2 m}{H(X)} = \frac{\log_2 4}{1,75} = \frac{2}{1,75} = 1,14,$$

т.е. скорость передачи информации по каналу связи при использовании экономичного кодирования может быть в 1,14 раза больше, чем при равномерном кодировании.

Кодирование источника дискретных сообщений методом Шеннона-Фано. Кодирование методом Шеннона-Фано рассмотрим на примере. Пусть алфавит источника содержит шесть элементов {А, Б, В, Г, Д, Е}, появляющихся с вероятностями $p(A) = 0,15$; $p(B) = 0,25$; $p(B) = 0,1$; $p(\Gamma) = 0,13$; $p(D) = 0,25$; $p(E) = 0,12$.

Энтропия такого источника:

$$H(X) = -0,15 \cdot \log_2 0,15 - 0,25 \cdot \log_2 0,25 - 0,1 \cdot \log_2 0,1 - 0,13 \cdot \log_2 0,13 - \\ - 0,25 \cdot \log_2 0,25 - 0,12 \cdot \log_2 0,12 = 2,492.$$

Алгоритм построения сжимающего кода Шеннона-Фано заключается в следующем.

1. Все m символов дискретного источника располагаются в порядке убывания вероятностей их появления (табл. 15.1).
2. Образованный столбец символов делится на две группы таким образом, чтобы суммарные вероятности каждой группы мало отличались друг от друга.
3. Верхняя группа кодируется символом «1», а нижняя – «0».
4. Каждая группа делится на две подгруппы с близкими суммарными вероятностями; верхняя подгруппа кодируется символом «1», а нижняя – «0».
5. Процесс деления и кодирования продолжается до тех пор, пока в каждой подгруппе не окажется по одному символу сообщения источника.
6. Записывается код для каждого символа источника; считывание кода осуществляется слева направо.

Таблица 15.1

Построение кода Шеннона-Фано

Элемент сообщения	Вероятность элемента	Деление сообщения на группы и подгруппы	Код	
Б	0,25	} 1 } 1	11	
Д	0,25		} 0	10
А	0,15	} 0 } 1 } 1	011	
Г	0,13		} 0	010
Е	0,12		} 1	001
В	0,1		} 0	000

При использовании простейшего равномерного кода для кодирования шести элементов алфавита источника потребуется по три двоичных символа на каждую букву сообщения. Если же используется код Шеннона-Фано, то среднее число символов на одну букву

$$n_{cp} = \sum_{i=1}^6 p(x_i) \cdot n_i = 0,25 \cdot 2 + 0,25 \cdot 2 + 0,15 \cdot 3 + 0,13 \cdot 3 + 0,12 \cdot 3 + 0,1 \cdot 3 = 2,5,$$

что меньше, чем при простейшем равномерном коде и незначительно отличается от энтропии источника.

Кодирование источника дискретных сообщений методом Хаффмана. Рассмотрим еще один подход к кодированию, предло-

женный Хаффменом, на примере источника сообщений, заданного в табл. 15.2.

Алгоритм построения сжимающего кода Хаффмена включает в себя следующие действия.

1. Все m символов дискретного источника располагаются в таблице в порядке убывания вероятностей.
2. Два символа, имеющих наименьшие вероятности, объединяются в один блок, а их вероятности суммируются.
3. Ветви скобки, идущей к большей вероятности, присваивается символ «1», а идущей к меньшей – символ «0».
4. Операции 2 и 3 повторяются до тех пор, пока не сформируется один блок с вероятностью единица.
5. Записывается код для каждого символа источника; при этом считывание кода осуществляется справа налево.

Таблица 15.2

Построение кода Хаффмена

Элемент сообщения	Вероятность элемента	Деление сообщения на группы и подгруппы	Код
Б	0,25		10
Д	0,25		01
А	0,15		111
Г	0,13		110
Е	0,12		001
В	0,1		000

Среднее число символов на одну букву для полученного кода

$$n_{\text{ср}} = \sum_{i=1}^6 p(x_i) \cdot n_i = 0,25 \cdot 2 + 0,25 \cdot 2 + 0,15 \cdot 3 + 0,13 \cdot 3 + 0,12 \cdot 3 + 0,1 \cdot 3 = 2,5 \cdot$$

Таким образом, для данного примера кодирование методами Хаффмена и Шеннона-Фано одинаково эффективно. Однако опыт кодирования показывает, что код Хаффмена часто оказывается экономичнее кода Шеннона-Фано.

Длина кодовой комбинации таких кодов зависит от вероятности выбора соответствующей буквы алфавита: наиболее вероятным буквам сопоставляются короткие кодовые комбинации, а менее вероятным – более длинные.

15.5. Количество информации, переданной по непрерывному каналу

Вместо последовательностей символов для дискретного канала, в непрерывном канале осуществляется передача последовательности непрерывных величин с дискретным или непрерывным временем (в первом случае эти последовательности можно представить в виде импульсов различной величины, появляющихся в определенные моменты времени, а во втором случае как непрерывные функции времени).

Количество передаваемой информации:

$$I(X, Y) = h(Y) - h\left(\frac{Y}{X}\right) = \int_{-\infty}^{+\infty} w(y) \cdot \log \frac{1}{w(y)} dy - \int_{-\infty}^{+\infty} w(n) \cdot \log \frac{1}{w(n)} dn, \quad (15.6)$$

где $w(y)$ – плотность распределения вероятности выходных случайных величин;

$w(n)$ – плотность распределения вероятности помехи (аддитивной);

$h(Y)$ – дифференциальная энтропия сигнала y ;

$h(Y/X)$ – условная дифференциальная энтропия сигнала y при известном сигнале x ;

Отметим следующие свойства количества информации, передаваемой в непрерывном канале:

$I(Y, X) \geq 0$, причём $I(Y, X) = 0$ тогда, и только тогда, когда вход и выход канала статистически независимы, т.е. $w(y/x) = w(y)$;

$I(Y, X) = I(X, Y)$ – свойство симметрии;

$I(Y, X) = \infty$, если помехи в канале отсутствуют, т.е. $y = x$, $n = 0$

Дифференциальная энтропия $h(Y)$ уже не представляет собой среднее количество информации, выдаваемое источником сигнала (для непрерывного сигнала оно бесконечно). Аналогично $h(Y/X)$ не представляет собой количество информации, потерянной в канале, поскольку эта величина тоже бесконечна. Поэтому дифференциальную энтропию следует понимать лишь формально, как некоторую вспомогательную величину полезную при расчетах.

Если помеха аддитивная $y = x + n$, то нетрудно показать, что

$$h\left(\frac{Y}{X}\right) = \int_{-\infty}^{+\infty} w(n) \cdot \log \frac{1}{w(n)} dn = h(N), \quad (15.7)$$

где $w(n)$ – плотность распределения вероятности помехи, а $h(N)$ – дифференциальная энтропия помехи.

Выражение для определения количества информации, переданной по непрерывному гауссовскому каналу (нормальный закон распределения вероятностей сигнала и помехи):

$$I(Y, X) = h(X) - h\left(\frac{X}{Y}\right) = \frac{1}{2} \log_2 \left(\frac{\sigma_c^2 + \sigma^2}{\sigma^2} \right). \quad (15.8)$$

где

$$h(N) = \log_2 \left(\sqrt{2\pi\sigma^2} \right) + \frac{1}{2} \log_2 e = \log_2 \left(\sqrt{2\pi e \sigma^2} \right) \quad (15.9)$$

$$h(Y) = - \int_{-\infty}^{+\infty} w(y) \cdot \log_2 w(y) dy = \log_2 \left(\sqrt{2\pi e \sigma_y^2} \right)$$

Полученное выражение показывает, что пропускная способность гауссовского канала с дискретным временем определяется отношением дисперсии сигнала σ_c^2 к дисперсии помехи σ^2 . Нередко величину $\sigma_c^2 / \sigma^2 = h^2$ называют отношением сигнал/шум. Чем больше это отношение, тем выше пропускная способность.

15.6. Пропускная способность непрерывного канала

Пусть сигнал $y(t)$ на выходе канала представляет собой сумму полезного сигнала $x(t)$ и шума $n(t)$, т.е. $y(t) = x(t) + n(t)$, причем $x(t)$ и $n(t)$ статистически независимы, и канал имеет ограниченную полосу пропускания шириной $\Delta F_{\text{НК}}$.

Пропускная способность гауссовского канала с дискретным временем, рассчитанная на единицу времени, с учетом (15.8) может быть записана в виде:

$$C = V_{\text{и}} \cdot I(Y, X) = \frac{V_{\text{и}}}{2} \log_2 \left(\frac{\sigma_c^2 + \sigma^2}{\sigma^2} \right) = \frac{V_{\text{и}}}{2} \log_2 (1 + h^2). \quad (15.10)$$

Полученное выражение показывает, что пропускная способность гауссовского канала с дискретным временем определяется числом импульсов, передаваемых в секунду, и отношением сигнал/шум (h).

С учетом взаимосвязи скорости передачи информации и полосы частот непрерывного канала от (15.10) можно перейти к формуле Шеннона, которая устанавливает связь пропускной способности гауссовского канала с полосой пропускания непрерывного канала и отношением мощности сигнала к мощности помехи:

$$C = \Delta F_{\text{НК}} \log_2 (1 + h^2). \quad (15.11)$$

Заметим, что при малом отношении $h^2 \ll 1$, $C \approx \Delta F_{\text{НК}} \cdot 1,442 \cdot h^2$, при большом отношении $h^2 \gg 1$, можно пренебречь единицей и считать, что $C = \Delta F_{\text{НК}} \cdot \log_2(h^2)$, т.е. зависимость пропускной способности непрерывного канала от отношения сигнал/шум логарифмическая.

График отношения $C / \Delta F_{\text{НК}} = \log_2(1+h^2)$ представлен на рис. 15.2.

Пропускная способность канала, как предельное значение скорости безошибочной передачи информации, является одной из основных характеристик любого канала.

Пример. Определим пропускную способность канала тональной частоты, имеющего границы эффективно передаваемых частот 0,3...3,4 кГц, среднюю мощность сигнала на выходе 56 мкВт при средней мощности помехи 69000 пВт.

$$C_{\text{НК}} = 3,1 \cdot 10^3 \cdot \log_2 \left(\frac{56 \cdot 10^{-6}}{69 \cdot 10^{-12}} \right) = 3,0 \cdot 10^4 \left[\frac{\text{бит}}{\text{с}} \right].$$

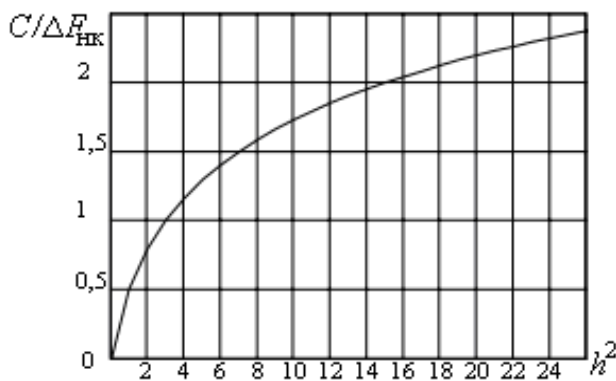


Рис. 15.2. График отношения $C / \Delta F_{\text{НК}}$

Таблица 15.3

Характеристики типовых каналов многоканальной связи

Границы передаваемых частот, Гц	Пропускная способность, бит/с
300 ... 3400	$20,64 \cdot 10^3$
$12,3 \cdot 10^3$... $23,4 \cdot 10^3$	$73,91 \cdot 10^3$
$60,6 \cdot 10^3$... $107,7 \cdot 10^3$	$313,6 \cdot 10^3$
$312,3 \cdot 10^3$... $551,4 \cdot 10^3$	$1,59 \cdot 10^6$
$812,3 \cdot 10^3$... $2043,7 \cdot 10^3$	$8,2 \cdot 10^6$

Зная пропускную способность канала и информационные характеристики сообщений (табл. 15.3), можно определить, какие сообщения (первичные сигналы) можно передавать по заданному каналу.

Например, первичный сигнал телевизионного вещания имеет $C = 208 \cdot 10^6 \text{ бит/с}$ и поэтому не может быть передан ни по одному из типовых непрерывных или цифровых каналов без потери качества. Следовательно, для передачи сигнала телевизионного вещания требуется создание специальных каналов с более высокой пропускной способностью или снижение скорости цифрового потока.

16. ТЕОРИЯ КОДИРОВАНИЯ СООБЩЕНИЙ

16.1. Основные понятия

Решение задачи выбора (отыскания) кода, оптимального по тому или иному критерию, составляет суть теории кодирования.

Заметим, что современные методы кодирования не позволяют близко подойти к потенциальной пропускной способности канала связи при одновременно высокой верности передачи. Однако грамотный выбор кода позволяет, во многих случаях, значительно снизить вероятность ошибочного приема при скорости передачи порядка $10 \div 50\%$ пропускной способности канала.

В настоящее время повышение достоверности передачи в каналах с помехами, осуществляется с помощью кодов, позволяющих обнаруживать или исправлять ошибки. Такое кодирование называется помехоустойчивым. При этом избыточность кодовой последовательности выше, чем избыточность источника сообщений. Благодаря этому и оказывается возможным обнаружение и исправление ошибок приема.

Кодирование называется процесс преобразования сообщений в комбинации из дискретных сигналов. Основными задачами, решаемыми кодированием в процессе передачи сообщений, являются:

- согласование источника сообщений с каналом по объемам алфавитов;
- повышение скорости передачи информации по каналу за счет устранения избыточности в последовательности сообщений;
- повышение помехоустойчивости передачи информации.

Первые две задачи решаются в кодере источника сообщений. Третья задача решается в кодере канала.

Классификация помехоустойчивых кодов. Построение помехоустойчивых кодов в основном связано с добавлением к исходной комбинации (m – символов) контрольных (k – символов). Закодированная комбинация будет составлять n – символов. Эти коды часто называют (n, m) – коды.

16.2. Коды с обнаружением ошибок

Код с проверкой на четность. Такой код образуется путем добавления к передаваемой комбинации, состоящей из m информационных символов, одного контрольного символа k (0 или 1) так, чтобы общее число единиц в передаваемой комбинации было четным.

m	k	n
11011	0	110110
11100	1	111001

Если получено слово с нечётным количеством единичек, то при передаче произошла ошибка.

Например, для кода с $m = 5$ и вероятностью ошибки $p = 10^{-2}$ коэффициент обнаружения составит $K_{обн} = 0,9$. То есть обнаруживаем 90% ошибок, при этом избыточность будет составлять $L = 0,17$ (17 %).

Код с постоянным весом. Этот код содержит постоянное число единиц и нулей. Для примера приведены коды с двумя единицами из пяти и тремя единицами из семи.

$$N = C_5^2 = 10 \qquad N = C_7^3 = 35$$

11000	0000111
10010	1001001
00101	1010100

Число кодовых комбинаций составит: $N = C_n^k = \frac{n!}{k!(n-k)}$

Этот код позволяет обнаруживать любые одиночные ошибки и часть многократных ошибок. Не обнаруживаются этим кодом только ошибки смещения, когда одновременно одна единица переходит в ноль и один ноль переходит в единицу, два нуля и две единицы меняются на обратные символы и т.д.

Для кода C_7^3 при $p = 10^{-2}$ коэффициент обнаружения составит $K_{обн} = 0,985$, избыточность составляет $L = 27\%$.

Корреляционный код (Код с удвоением). Элементы данного кода заменяются двумя символами, единица «1» преобразуется в 10, а ноль «0» в 01.

Вместо комбинации 1010011 передается 10011001011010. Ошибка обнаруживается в том случае, если в парных элементах будут одинаковые символы 00 или 11 (вместо 01 и 10).

Например, при $m = 5$, $n = 10$ и вероятности ошибки $p = 10^{-2}$, $K_{обн} = 0,995$. Но при этом избыточность будет составлять 50%.

Инверсный код. К исходной комбинации добавляется такая же комбинация по длине, т.е. число символов удваивается. Если в исходной комбинации четное число единиц, то добавляемая комбинация повторяет исходную комбинацию, если нечетное, то добавляемая комбинация является инверсной по отношению к исходной.

m	k	n
11011	11011	1101111011
11100	00011	1110000011

Прием инверсного кода осуществляется в два этапа. На первом этапе суммируются единицы в первой основной группе символов. Если число единиц четное, то контрольные символы принимаются без изменения, если нечетное, то контрольные символы инвертируются. На втором этапе контрольные символы суммируются с информационными символами по модулю два. Нулевая сумма говорит об отсутствии ошибок. При ненулевой сумме, принятая комбинация бракуется. Покажем суммирование для принятых комбинаций без ошибок (1, 3) и с ошибками (2, 4):

1	11011 <u>11011</u> 00000	2	11111 <u>00100</u> 11011	3	11100 <u>11100</u> 00000	4	11000 <u>11100</u> 00100
---	--------------------------------	---	--------------------------------	---	--------------------------------	---	--------------------------------

Данный код обнаруживает практически любые ошибки, кроме редких ошибок смещения, которые одновременно происходят как среди информационных символов, так и среди соответствующих контрольных.

При $m = 5$, $n = 10$ и $p = 10^{-2}$, $K_{обн} = 1 \cdot 10^{-5}$.

16.3. Корректирующие коды

Корректирующими называются коды позволяющие обнаруживать и исправлять ошибки. Они характеризуются кодовым (или хэмминговым) расстоянием, обозначаемым буквой d .

Кодовое расстояние – это то минимальное число элементов, в которых одна кодовая комбинация отличается от другой. Для определения кодового расстояния достаточно сравнить две кодовые комбинации по модулю 2.

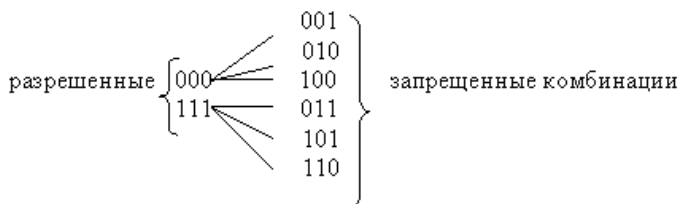
Так, сложив две комбинации определим, что расстояние между ними $d = 7$.

$$\begin{array}{r} 10110101101 \\ \underline{11001010101} \\ 01111111000 \end{array}$$

Код с $n = 3$ и $d = 1$ для передачи используются все восемь кодовых комбинаций 000, 001,..., 111. Такой код является не помехоустойчивым, он не в состоянии обнаружить ошибку.

Если выберем комбинации с кодовым расстоянием $d = 2$, например, 000, 110, 101, 011, то такой код позволит обнаруживать однократные ошибки. Назовем эти комбинации разрешенными, предназначенными для передачи информации. Все остальные 001, 010, 100, 111 – запрещенные. Любая одиночная ошибка приводит к тому, что разрешенная комбинация переходит в ближайшую, запрещенную комбинацию. Получив запрещенную комбинацию, мы обнаружим ошибку.

Комбинации для кода с $d = 3$:



Такой код может исправить одну одиночную ошибку или обнаружить две ошибки. Таким образом, увеличивая кодовое расстояние мож-

но увеличить помехоустойчивость кода. В общем случае кодовое расстояние определяется по формуле $d = (S + r) + 1$, где S – число исправляемых ошибок, r – число обнаруживаемых ошибок. Обычно $r > S$.

Большинство корректирующих кодов являются линейными кодами. Линейные коды – это такие коды, у которых контрольные символы образуются путем линейной комбинации информационных символов.

Кроме того, корректирующие коды являются групповыми кодами. Групповые коды (G_n) – это такие коды, которые имеют одну основную операцию. При этом должно соблюдаться условие замкнутости (то есть, при сложении двух элементов группы получается элемент, принадлежащий этой же группе). Число разрядов в группе не должно увеличиваться. Этому условию удовлетворяет операция поразрядного сложения по модулю 2. В группе, кроме того, должен быть нулевой элемент.

Примеры кодовых комбинаций:

- 1) 1101 1110 0111 1011 – не группа, так как нет нулевого элемента;
- 2) 0000 1101 1110 0111 – не группа, так как не соблюдается условие замкнутости (1101 + 1110 = 0011);
- 3) 000 001 010 011 100 101 110 111 – группа;
- 4) 000 001 010 111 – подгруппа.

Большинство корректирующих кодов образуются путем добавления к исходной m – комбинации k – контрольных символов. В итоге в линию передаются $n = m + k$ символов. При этом корректирующие коды называются (n, m) кодами.

Для построения кода способного обнаруживать и исправлять единичную ошибку необходимое число контрольных разрядов будет составлять $n - m \geq \log(n + 1)$.

Если необходимо исправить две ошибки, то число различных исходов будет составлять C_n^2 . Тогда $n - m \geq \log(1 + C_n^1 + C_n^2)$, в этом случае обнаруживаются однократные и двукратные ошибки. В общем случае, число контрольных символов должно быть не меньше:

$$n - m \geq \log(1 + C_n^2 + C_n^2 + \dots + C_n^t) = \log \sum_{i=0}^t C_n^i.$$

Эта формула называется неравенством Хэмминга, или нижней границей Хэмминга для числа контрольных символов.

Код Хэмминга. Код Хэмминга, являющийся групповым (n, m) кодом, с минимальным расстоянием $d=3$ позволяет обнаруживать и исправлять однократные ошибки. Построение кодов Хэмминга базируется на принципе проверки на чётность веса W (числа единичных символов) в информационной группе кодового блока.

Для каждого числа проверочных символов $k = 3, 4, 5, \dots$ существует классический код Хэмминга с маркировкой $(n, m) = (2^k - 1, 2^k - 1 - k)$, т.е. – (7,4), (15,11), (31,26) ...

Для примера рассмотрим классический код Хемминга (7,4). В простейшем варианте при заданных четырёх ($m = 4$) информационных символах (i_1, i_2, i_3, i_4) будем полагать, что они сгруппированы в начале кодового слова, хотя это и не обязательно. Дополним эти информационные символы тремя проверочными символами ($k = 3$), задавая их следующими равенствами проверки на чётность, которые определяются соответствующими алгоритмами:

$$k_1 = i_1 \oplus i_2 \oplus i_3;$$

$$k_2 = i_2 \oplus i_3 \oplus i_4;$$

$$k_3 = i_1 \oplus i_2 \oplus i_4, \text{ где знак } \oplus \text{ означает сложение по модулю 2.}$$

В соответствии с этим алгоритмом определения значений проверочных символов k_i возможны 16 кодовых слов (7,4) – кода Хемминга.

$m = 4$				$k = 3$		
i_1	i_2	i_3	i_4	k_1	k_2	k_3
0	0	0	0	0	0	0
0	0	0	1	0	1	1
0	0	1	0	1	1	0
0	0	1	1	1	0	1
0	1	0	0	1	1	1
0	1	0	1	1	0	0
0	1	1	0	0	0	1
0	1	1	1	0	1	0
1	0	0	0	1	0	1
1	0	0	1	1	1	0
1	0	1	0	0	1	1
1	0	1	1	0	0	0
1	1	0	0	0	1	0
1	1	0	1	0	0	1
1	1	1	0	1	0	0
1	1	1	1	1	1	1

На вход декодера поступает кодовое слово $V = (i_1', i_2', i_3', i_4', k_1', k_2', k_3')$, апостроф означает, что любой символ слова может быть искажён помехой в канале передачи.

В декодере в режиме исправления ошибок строится последовательность:

$$s_1 = k_1' \oplus i_1' \oplus i_2' \oplus i_3';$$

$$s_2 = k_2' \oplus i_2' \oplus i_3' \oplus i_4';$$

$$s_3 = k_3' \oplus i_1' \oplus i_2' \oplus i_4'.$$

Трёхсимвольная последовательность (s_1, s_2, s_3) называется синдромом S .

Синдром $S = (s_1, s_2, s_3)$ представляет собой сочетание результатов проверки на чётность соответствующих символов кодовой группы и характеризует определённую конфигурацию ошибок (шумовой вектор).

Число возможных синдромов определяется выражением: $S = 2^k$.

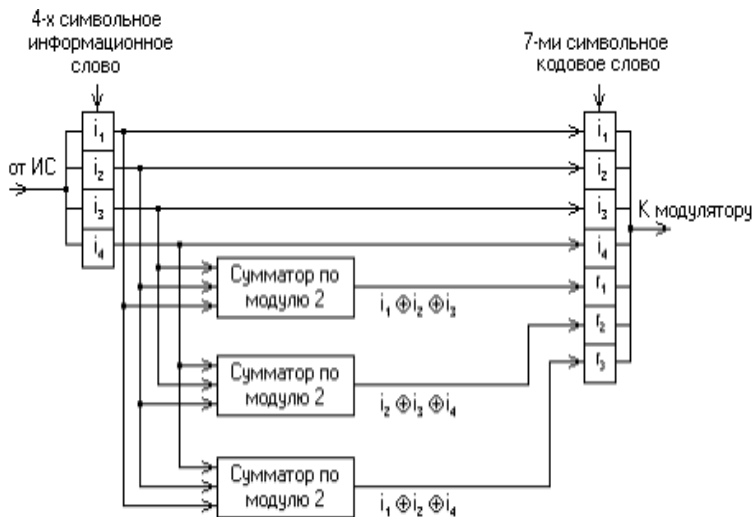


Рис. 16.1. Кодер простого (7, 4) – кода Хемминга

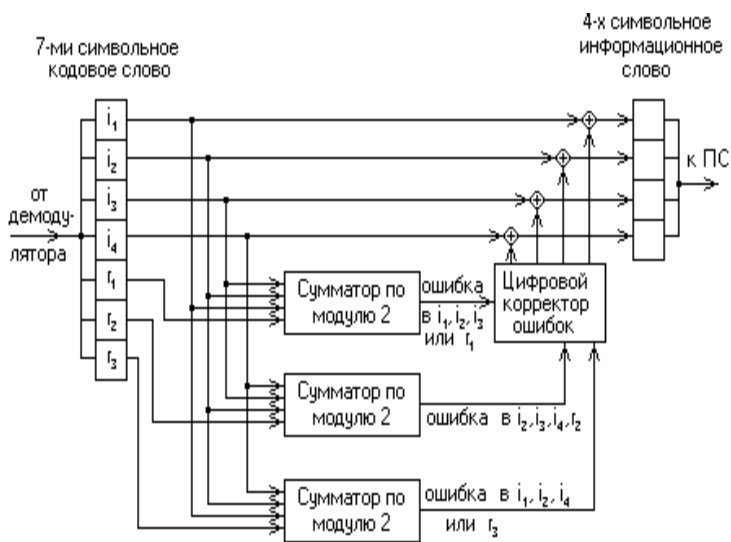


Рис. 16.2. Декодер простого (7, 4) – кода Хемминга

При числе проверочных символов $k = 3$ имеется восемь возможных синдромов ($2^3 = 8$). Нулевой синдром (000) указывает на то, что ошибки при приёме отсутствуют или не обнаружены. Всякому ненулевому синдрому соответствует определённая конфигурация ошибок, которая и исправляется. Классические коды Хемминга имеют число синдромов, точно равное их необходимому числу, позволяют исправить все однократные ошибки в любом информативном и проверочном символах и включают один нулевой синдром. Такие коды называются плотноупакованными.

Усечённые коды являются неплотноупакованными, так как число синдромов у них превышает необходимое. Так, в коде (9,5) при четырёх проверочных символах число синдромов будет равно $2^4 = 16$, в то время как необходимо всего 10. Лишние 6 синдромов свидетельствуют о неполной упаковке кода (9,5). Для рассматриваемого кода (7,4) в табл. 16.1 представлены ненулевые синдромы и соответствующие конфигурации ошибок.

Таким образом, код (7,4) позволяет исправить все одиночные ошибки. Простая проверка показывает, что каждая из ошибок имеет свой единственный синдром. При этом возможно создание такого цифрового корректора ошибок (дешифратора синдрома), который по соответствующему синдрому исправляет соответствующий символ в принятой кодовой группе.

Таблица 16.1

Соответствие синдромов конфигурациям ошибок

Синдром	001	010	011	100	101	110	111
Конфигурация ошибок	0000001	0000010	0001000	0000100	1000000	0010000	0100000
Ошибка в символе	k_3	k_2	i_4	k_1	i_1	i_3	i_2

Циклические коды. Циклическим кодом называется такой групповой код, который связан дополнительным условием цикличности. Все строки образующей матрицы такого кода могут быть получены циклическим сдвигом одной комбинации, называемой образующей для данного кода.

Циклические коды широко применяются при передаче данных в современных информационных системах благодаря ряду положительных качеств, основными из которых являются:

- высокая эффективность, так как циклические коды обладают сравнительно небольшой избыточностью, отличаются простотой реализации кодирующих и декодирующих устройств;
- высокая помехоустойчивость – за счет способности кода к обнаружению и исправлению ошибок.

Для оптимального кода количество исправляемых ошибок равно $2^k - 1$, где k – число контрольных разрядов.

Применяемые в настоящее время циклические коды, содержащие n разрядов, из которых m является информационными, а $k = n - m$ – контрольными (проверочными), расположенными в конце кодовой комбинации. Так как информационные и контрольные разряды занимают строго определенные места и длина кодовой комбинации постоянна, то циклические коды относятся к систематическим кодам.

Для описания циклических кодов обычно пользуются записью любого n -разрядного двоичного числа в виде многочлена степени $(n-1)$. Например, кодовая комбинация 1011001 записывается как многочлен $x^6 + x^4 + x^3 + 1$, т.е. коэффициенты многочлена не пишутся, а члены с коэффициентами 0 опускаются. Наивысшая степень числа с коэффициентом 1 называется степенью полинома (многочлена). Так, в примере рассмотрен многочлен 6-й степени.

Таким образом, действия над кодовыми числами можно свести к действиям над многочленами. При этом используют теорию коммутативных колец. Коммутативным кольцом называют множество, в котором особым образом определены операции сложения и умножения.

В циклическом кодировании все математические операции сложения производятся с использованием сложения по mod 2 и с приведением подобных членов.

Операцию умножения символически проводят по следующим правилам:

1) Вначале все многочлены перемножаются по обычным правилам, но с приведением подобных членов по mod 2.

2) Если старшая степень полученного в результате умножения многочлена не превышает $(n-1)$, то этот многочлен является результатом символического умножения.

3) Если старшая степень полученного в результате умножения многочлена больше $(n-1)$, то многочлен произведения делится на двучлен $x^n + 1$. В этом случае результатом символического умножения считается остаток от деления (вычет).

Пример: Имеем кодовые комбинации 001101 и 101110, где $n = 6$.

Эти комбинации соответствуют многочленам:

$$x^3 + x^2 + 1 \text{ и } x^5 + x^3 + x^2 + x.$$

Допустим, необходимо провести дважды циклический сдвиг этих кодовых комбинаций. В результате получим:

$$\begin{array}{l} 001101 \Rightarrow 110100; \\ \quad \quad \quad \uparrow \quad \uparrow \\ 101110 \Rightarrow 111010. \\ \quad \quad \quad \uparrow \quad \uparrow \end{array}$$

Для проведения этой операции с многочленами необходимо их символически умножить на $x \times x = x^2$:

1) $(x^3 + x^2 + 1) \times x^2 = x^5 + x^4 + x^2$, т.к. степень полученного многочлена не превышает $(n-1) = 5$, то этот многочлен принимается за результат умножения и действительно соответствует сдвинутой кодовой комбинации 110100.

2) $(x^5 + x^3 + x^2 + x) \times x^2 = x^7 + x^5 + x^4 + x^3$, т.к. степень полученного многочлена превышает $(n-1)$, то для получения результата символического умножения необходимо произвести деление этого многочлена на двучлен $(x^n + 1)$:

$$\begin{array}{r}
 x^7 + x^5 + x^4 + x^3 \quad | \quad x^6 + 1 \\
 x^7 + x \quad \quad \quad \quad | \quad x \\
 \hline
 \text{-----} \\
 x^5 + x^4 + x^3 + x
 \end{array}$$

Остаток от деления $x^5 + x^4 + x^3 + x$ принимается за результат символического умножения, что соответствует циклически сдвинутой кодовой комбинации 111 010.

В основе образования циклического кода лежит использование так называемого образующего (неприводимого) многочлена (полинома). Выбор образующего полинома определяет тип циклического кода и характеризует его обнаруживающие и исправляющие способности.

Степень образующего полинома равна k , т.е. числу контрольных символов. Любой многочлен циклического кода должен делиться без остатка на образующий полином.

В то же время ни один многочлен, соответствующий запрещенной кодовой комбинации, не должен делиться без остатка на образующий полином. Это свойство позволяет обнаружить ошибку, а по виду остатка и вектор ошибки, т.е. исправлять ошибки.

Для получения циклического кода, многочлен $G(x)$ (соответствующий кодовым комбинациям безызбыточного m – разрядного кода), умножают на x^k . Это соответствует приписыванию со стороны младших разрядов k нулей к кодовым комбинациям.

Затем произведение $G(x) \times x^k$ делится на образующий многочлен $P(x)$. В общем случае мы получаем в результате такого деления $Q(x)$ той же степени, что и $G(x)$ и остаток $R(x)$. Остаток $R(x)$ прибавляется к $G(x) \times x^k$. Получаем многочлен $F(x) = G(x) \times x^k + R(x)$.

Так как в комбинациях, соответствующих многочлену $G(x) \times x^k$, первые k младших разрядов – нули, а $R(x)$ – многочлен степени не выше $k-1$, то операция получения многочлена соответствует приписыванию $R(x)$ к $G(x)$ со стороны младших разрядов.

Полученный таким образом многочлен $F(x)$ будет делиться на образующий многочлен $P(x)$ без остатка.

Таким образом, циклический код можно получить, если к каждой кодовой комбинации безызбыточного кода приписывать остаток от деления многочлена, соответствующего этой кодовой комбинации на образующий многочлен.

Так как опознавателями ошибок являются остатки от деления многочленов циклического кода на образующий многочлен, то корректирующая способность кода будет тем выше, чем больше остатков от деления можно образовать. Наибольшее число остатков, равное $2^k - 1$ (исключая нулевой), может обеспечить неприводимый (простой) многочлен, т.е. такой многочлен, который делится только сам на себя. Поэтому в качестве образующего многочлена необходимо выбирать простой многочлен (или их произведение).

Пример образования циклического кода. Пусть информационный код содержит $m = 4$ разрядов. Одна из $N = 2^m$ комбинаций этого кода: 1101 в виде многочлена запишется так: $G(x) = x^3 + x^2 + 1$.

Если циклический код обнаруживает и исправляет одну ошибку, его минимальное кодовое расстояние равно: $d_{min} = S + r + 1 = 3$ (где r – число обнаруживаемых, а S – число исправляемых ошибок).

Выберем из Таблицы 16.2 значение k (для $m = 4, k = 3$).

Выберем из таблицы 16.3 полином $P(x)$ для $k = 3$: $P(x) = x^3 + x + 1$ (т.е. 1011)

Умножим $G(x)$ на x^k : $G(x) \times x^3 = (x^3 + x^2 + 1) \times x^3 = x^6 + x^5 + x^3$ (т.е. 1101 × 1000 = 1101000)

Разделим $G(x) \times x^k$ на полином $P(x)$:

$$\begin{array}{r|l}
 x^6 + x^5 + x^3 & x^3 + x + 1 \\
 \hline
 x^6 + x^4 + x^3 & \\
 \hline
 x^5 + x^4 & \\
 \hline
 x^5 + x^3 + x^2 & \\
 \hline
 x^4 + x^3 + x^2 & \\
 \hline
 x^4 + x^2 + x & \\
 \hline
 x^3 + x & \\
 \hline
 x + x + 1 & \\
 \hline
 0 + 0 + 1 &
 \end{array}
 \qquad
 \begin{array}{r|l}
 1101000 & 1011 \\
 \hline
 1011 & \\
 \hline
 1100 & 1111 \\
 \hline
 1011 & \\
 \hline
 1110 & \\
 \hline
 1011 & \\
 \hline
 1010 & \\
 \hline
 1011 & \\
 \hline
 001 &
 \end{array}$$

В результате получаем:

$$G(x) \times x^3 / P(x) = (x^3 + x^2 + x + 1) + 001 / (x^3 + x + 1) \text{ или: } 1111 + 001 / 1011$$

В соответствии с $n = m + k$: $(x^3 + x^2 + x + 1) = Q(x) \rightarrow 1111$

$1 / (x^3 + x + 1) = R(x) / P(x) \rightarrow 001 / 1011$

$R(x) = 001$

Искомый многочлен $F(x)$ равен:

$$F(x) = Q(x) \times P(x) = G(x) \times x^k + R(x) = x^6 + x^5 + x^3 + 1 \rightarrow 1101001$$

Для обнаружения и исправления ошибок принятая комбинация делится на образующий многочлен $P(x)$. Если остаток $R(x) = 0$, значит, комбинация принята без ошибок. Наличие остатка свидетельствует о том, что комбинация принята искаженной.

Таблица 16.2

Зависимость между n , m и k

n	3	5	6	7	9...15	17...31	33...63	65...127
m	1	2	3	4	5...11	12...26	27...57	28...120
k	2	3	3	3	4	5	6	7

Таблица 16.3

Неприводимые полиномы $P(x)$

Неприводимые многочлены	Их эквиваленты
$q(x) = x + 1$	11
$q(x^2) = x^2 + x + 1$	111
$q(x^3) = x^3 + x + 1$	1011
$q(x^3) = x^3 + x^2 + 1$	1101
$q(x^4) = x^4 + x + 1$	10011
$q(x^4) = x^4 + x^3 + 1$	11001
$q(x^4) = x^4 + x^3 + x^2 + x + 1$	11111
$q(x^5) = x^5 + x^2 + 1$	100101
$q(x^5) = x^5 + x^3 + 1$	101001
$q(x^5) = x^5 + x^3 + x^2 + x + 1$	101111
$q(x^5) = x^5 + x^4 + x^2 + x + 1$	110111
$q(x^5) = x^5 + x^4 + x^3 + x + 1$	111011
$q(x^5) = x^5 + x^4 + x^3 + x^2 + 1$	111101
$q(x^6) = x^6 + x + 1$	1000011
$q(x^7) = x^8$	10001001
$q(x^8) = x^8 + x^4 + x^3 + x^2 + 1$	100011101
$q(x^9) = x^9 + x^4 + 1$	1000010001
$q(x^{10}) = x^{10} + x^3 + 1$	1000001001

Если число проверочных символов меньше числа информационных символов, то проще комбинации циклического кода получить и с помощью образующей матрицы.

Образующая матрица состоит из двух подматриц – основной и дополнительной.

Основная матрица – это единичная транспортированная матрица I_m . Она содержит m столбцов и m строк (по числу информационных разрядов кода).

Дополнительная матрица – матрица остатков. Она образуется путем деления на образующий полином $P(x)$ многочлена в виде единицы с рядом нулей и выписыванием всех промежуточных остатков. Эта матрица содержит k столбцов и m строк.

$$C_{n,m} = \left\| \begin{array}{cccc|ccc} 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{array} \right\|$$

$$C_{n,m} = \tilde{I}_m, C_{(n-m),m}$$

С помощью образующей матрицы можно получить любую из $N = 2^m - 1$ разрешенных комбинаций циклического кода. Каждая из строк образующей матрицы – это уже комбинация циклического кода. Остальные $N - m$ комбинаций можно получить сложением по модулю 2 строк матрицы.

Пример:

$$\begin{array}{r} \oplus 0001011 \\ \underline{0010110} \\ 0011101 \end{array}$$

$$\begin{array}{r} \oplus 0001011 \\ \underline{0100111} \\ 0101100 \end{array}$$

Кодеры и декодеры циклических кодов в основном выполняют операции умножения и деления многочленов.

Операция деления на образующий полином $P(x)$ осуществляется также с помощью регистра сдвига, но в этом случае регистр имеет обратные связи, включенные через сумматоры по модулю 2. Деление многочленов – это операция сложения по модулю 2 делителя с разрядами делимого, начиная со старшего разряда.

Для алгоритма кодирования используется k -разрядный регистр сдвига с обратными связями через сумматоры по модулю 2. Число сумматоров равно числу членов образующего полинома, отличных от нуля, минус единица. Структурная схема кодирующего устройства приведена ниже на рис. 16.3.

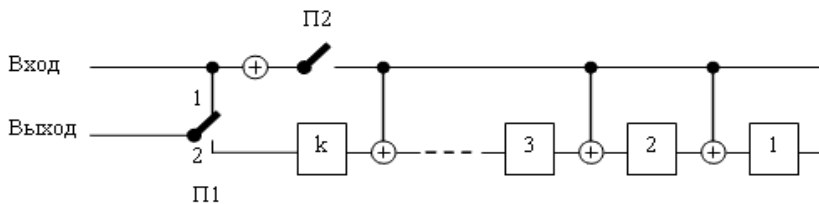


Рис. 16.3. Структурная схема кодирующего устройства

Схема работает следующим образом. В начальном состоянии ключ П2 замкнут, ключ П1 находится в положении «1». Информационная кодовая комбинация $a(x)$, имеющая m разрядов (символов) подается на вход через «1» П1 и одновременно в регистр сдвига Р1. За k тактов работы регистра Р1 в нем формируется остаток $r(x)$, который представляет собой контрольные разряды (символы) циклического кода. После этого ключ П2 размыкается, ключ П1 переводится в положение «2» и контрольные символы за k тактов работы регистра Р1 выводятся из него, следуя за информационными символами.

Необходимо иметь в виду, что информационная комбинация $a(x)$ поступает на вход кодера в обратной последовательности, т.е. начиная со старших разрядов. Комбинация циклического кода на выходе кодера имеет такую же последовательность.

Схема декодирования (образуется с помощью деления на образующий многочлен):



Рис. 16.4. Схема декодирования циклического кода. АО – анализатор ошибок

Исходная комбинация подается в буферный регистр и одновременно через ключ в декодирующий регистр. Если с приходом последнего символа, зафиксирован нулевой остаток, то ошибок нет, и, если не нулевой, то есть ошибка. Принятая комбинация подается через выходной сумматор, и искаженный сигнал исправляется анализатором ошибок (АО).

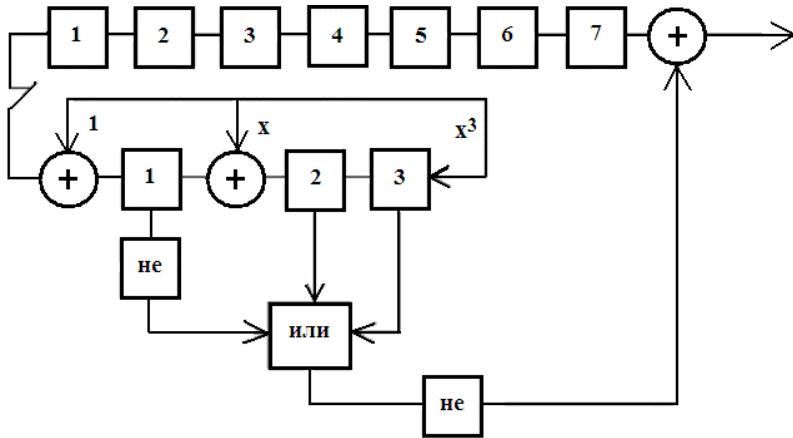


Рис. 16.5. Пример схемы декодирования методом деления на полином $P(x) = x^3 + x + 1$

Коды Рида-Соломона. Коды Рида-Соломона являются подмножеством циклических кодов и представляют собой линейные блочные коды. Код Рида-Соломона специфицируется как $RS(n, m)$ s -битных символов.

Это означает, что кодировщик воспринимает m информационных символов по s бит каждый и добавляет символы четности для формирования n символьного кодового слова. Имеется $n - m$ символов четности по s бит каждый. Декодер Рида-Соломона может корректировать до $k / 2$ символов, которые содержат ошибки в кодовом слове, где $k = n - m$.

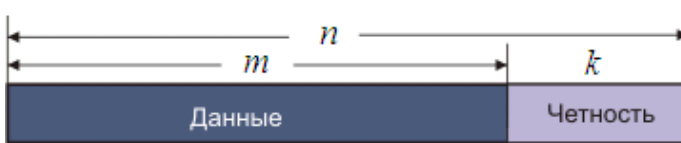


Рис. 16.6. Структура кодового слова $R-S$

Пример. Популярным кодом Рида-Соломона является $RS(255, 223)$ с 8-битными символами. Каждое кодовое слово содержит 255 байт, из которых 223 являются информационными и 32 байтами четности. Для этого кода: $n = 255$, $m = 223$, $s = 8$, $k = 32$, $k / 2 = 16$.

Декодер может исправить любые 16 символов с ошибками в кодовом слове: то есть, ошибки могут быть исправлены, если число искаженных байт не превышает 16.

Коды Рида-Соломона базируются на специальном разделе математики – полях Галуа (GF), или конечных полях. Арифметические действия (+, -, \times , / и др.) над элементами конечного поля дают результат, ко-

торый также является элементом этого поля. Для реализации этих арифметических операций требуется специальное оборудование и/или специализированное программное обеспечение.

Общая форма образующего полинома имеет вид:

$$P(x) = (x-a^i)(x-a^{i+1})\dots(x-a^{i+k}),$$

кодовое слово формируется с помощью операции:

$$F(x) = P(x).i(x),$$

где $P(x)$ – образующий полином, $i(x)$ представляет собой информационный блок, $F(x)$ – кодовое слово, называемое простым элементом поля.

VI. ПОМЕХОУСТОЙЧИВОСТЬ

17. ПОМЕХОУСТОЙЧИВОСТЬ СИСТЕМ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ

17.1. Основные понятия и термины

Основными задачами, стоящими перед техникой связи, является решение двух проблем:

- 1) эффективности связи;
- 2) помехоустойчивости связи.

Эффективность связи заключается в том, чтобы передать наибольшее количество информации наиболее экономным способом.

Скорость передачи информации по каналу связи измеряется количеством информации, передаваемой в единицу времени. Максимальная скорость передачи информации, которую может обеспечить канал связи с данными характеристиками, называется его пропускной способностью.

Помехоустойчивость связи – это способность системы сохранять свои функции неизменными или изменяющимися в допустимых пределах при действии помех.

Количественно помехоустойчивость оценивается различными показателями, использующими вероятностное описание сигналов и помех. Например, применяются такие показатели, как отношение сигнал/шум на входе и выходе приёмного устройства, вероятность правильного обнаружения сигнала, при передаче дискретных сообщений используется вероятность ошибки, а при передаче непрерывных сообщений в качестве меры различия между переданным и принятым сообщением часто используется среднеквадратическое отклонение.

В теории помехоустойчивости различают две основные задачи – анализ и синтез сигналов.

Задача анализа состоит в расчете показателей помехоустойчивости существующих (разработанных) систем. В этом случае, полагая известными вероятностное описание сигнала и помехи на входе, определяют

вероятностные характеристики выходного процесса, а по нему – показатели помехоустойчивости. Эта задача, по своей сути, сводится к анализу прохождения случайного процесса через линейные и нелинейные цепи, из которых состоит система.

Задача синтеза заключается в определении структурной схемы системы или, в более простом варианте, структурной схемы радиоприёмного устройства, которое обладало бы наилучшими, или оптимальными, показателями помехоустойчивости при заданном предназначении устройства и при известном вероятностном описании сигнала и помехи на входе.

Задача синтеза называется также задачей оптимального радиоприёма и разделяется на четыре частные подзадачи: обнаружения сигнала, различения сигналов, оценки параметров сигнала, фильтрации сигнала или сообщений.

В подзадаче обнаружения требуется по заданному критерию оптимальности на основании наблюдения процесса ответить на вопрос, содержит ли наблюдаемый процесс вместе с помехой сигнал или является только помехой?

В подзадаче различения требуется по заданному критерию оптимальности ответить на вопрос, какой именно сигнал вместе с помехой присутствует в наблюдаемом процессе, поскольку этот процесс может вместе с помехой содержать один из двух взаимно исключающих сигналов.

В подзадаче оценки параметров требуется по заданному критерию оценить неизвестные параметры сигнала. Считается, что в наблюдаемом процессе вместе с помехой существует сигнал с одним или несколькими неизвестными параметрами (параметр является случайной, но постоянной величиной на интервале наблюдения).

К задаче оценки параметров тесно примыкает задача разрешения сигнала, когда считается, что вместе с помехой в наблюдаемом процессе могут существовать один или два сигнала, неизвестные параметры которых незначительно различаются между собой. Однако сколько этих сигналов – один или два – заранее неизвестно. Требуется, увеличивая различие между параметрами сигнала, определить то наименьшее различие, при котором наступает уверенное разрешение сигналов.

В подзадаче оптимальной фильтрации требуется в каждый момент времени дать оценку меняющемуся параметру по заданному критерию оптимальности. Считается, что в соответствии со случайным законом модуляции в наблюдаемом процессе существует вместе с помехой сигнал с изменяющимся во времени параметром, т.е. параметр является случайной функцией времени.

В процессе передачи сообщений в системах связи выполняются различные преобразования, основные из которых показаны на упрощенной структурной схеме дискретной системы связи (рис. 17.1).

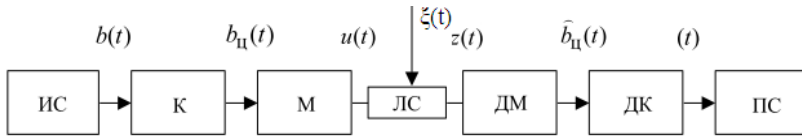


Рис. 17.1. Упрощенная структурная схема дискретной системы связи

Источник сигнала ИС включает в себя источник сообщений и преобразователь сообщения $a(t)$ в первичный сигнал $b(t)$. Первичный сигнал подвергается кодированию (экономному и/или помехоустойчивому) в кодере K , после чего сигнал $b_{ц}(t)$, называемый цифровым, поступает в модулятор M (передатчик), вырабатывающий сигнал $u(t)$, приспособленный по своим характеристикам для передачи по линии связи ЛС. В линии связи происходит искажение сигнала и его взаимодействие с помехой $\zeta(t)$ (в простейшем случае аддитивное), в результате чего на вход демодулятора ДМ (приемника) поступает наблюдаемое колебание $z(t)$. Демодулятор выполняет функцию, обратную модуляции, поэтому, в идеале, на его выходе должен быть выработан сигнал $b_{ц}(t)$. Однако в реальности вследствие воздействия помех результат демодуляции $\hat{b}_{ц}(t)$ отличается от сигнала $b_{ц}(t)$, поэтому результат декодирования $\hat{b}_{ц}(t)$ не совпадает с первичным сигналом $b(t)$.

Для облегчения восприятия в дальнейшем рассматривается идеализированный канал связи без памяти, в котором отсутствуют искажения сигнала, тогда наблюдаемое

$$z(t) = \sum_{k=-\infty}^{+\infty} b_{ц}(t) \cdot s(t - k\tau) + \xi(t), \quad (17.1)$$

где $s(t)$ – посылка длительности τ , $\xi(t)$ – помеха.

Задача демодулятора состоит в том, чтобы по наблюдаемому колебанию $z(t)$ принять такое решение $\hat{b}_{ц}(t)$ о переданном сигнале $b_{ц}(t)$, которое обеспечило бы максимальную верность. Правило (алгоритм) принятия решения – это закон преобразования $z(t)$ в $\hat{b}_{ц}(t)$. Поскольку помеха является случайной, задача построения оптимального (наилучшего) демодулятора представляет собой статистическую задачу и решается на основе методов теории вероятности и математической статистики (теории статистических решений).

Материалом для принятия решения в демодуляторе служит в анализируемом случае реализация колебания $z(t)$ на интервале длительности T . Если бы помеха отсутствовала, то эта реализация совпадала бы с элементарным сигналом (посылкой), который можно считать точкой в

гильбертовом пространстве сигналов, определенных на заданном временном интервале. Все возможные в данной системе связи посылки изображаются различными точками, и демодулятор должен вырабатывать свои решения в зависимости от того, какой именно точке соответствует принятая реализация $z(t)$.

Реализация помехи, взаимодействуя с посылкой, смещает точку, изображающую принятую реализацию, причем смещение случайно вследствие случайного характера помехи. Если смещения будут значительными, демодулятор может ошибаться. Ошибка является случайным событием, поэтому качество решения можно характеризовать вероятностью ошибки.

Задача синтеза оптимального приемника (демодулятора) ставится следующим образом: найти оптимальный алгоритм обработки и оптимальное правило, обеспечивающие максимальную вероятность безошибочного (правильного) решения.

Максимум этой вероятности академик РАН В. А. Котельников назвал потенциальной помехоустойчивостью, а приемник, реализующий этот максимум, – идеальным приемником [13].

Алгоритм работы приемника состоит в разбиении гильбертова пространства реализаций входного колебания на области так, что решение принимается в соответствии с тем, какой области принадлежит принятая реализация. Количество областей равно количеству различных кодовых символов данной системы связи. Ошибка возникает, если в результате воздействия помехи реализация попадает в «чужую» область. Оптимальный приемник разбивает пространство реализаций наилучшим образом, поэтому средняя вероятность ошибки минимальна при всех возможных разбиениях.

Каждая область соответствует предположению (гипотезе) о том, что передан был один из возможных сигналов.

Пример. Предположим, что результатом обработки в двоичной системе связи с амплитудной телеграфией является значение y , соответствующее окончанию интервала наблюдения. Если в колебании $z(t)$ присутствует только шум, имеющий гауссово распределение с нулевым математическим ожиданием, то плотность распределения величины y имеет вид:

$$w_0(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y^2}{2\sigma^2}}, \quad (17.2)$$

если кроме шума на вход приемника поступает сигнал, то результат обработки имеет ненулевое (для определенности – положительное) среднее значение a , и плотность распределения величины y имеет вид:

$$w_1(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y-a)^2}{2\sigma^2}}. \quad (17.3)$$

Гипотезы, соответствующие выражениям (17.2) и (17.3), являются простыми. Если среднеквадратическое отклонение σ неизвестно, гипотезы являются сложными.

Рассмотрим систему связи, в которой используются K различных символов. Тогда демодулятор должен различать K различных гипотез. При этом возможны ошибки: может быть принято решение D_j в пользу j -й гипотезы, в то время как справедливой является i -я гипотеза. Такая ситуация характеризуется условной вероятностью ошибки $p_{ij} = P\{D_j / H_i\}$. Различные ошибки могут наносить разный вред, поэтому вводится численная характеристика Π_{ij} , называемая потерей, или риском.

Каждая (i -я) гипотеза характеризуется некоторой вероятностью p_i осуществления, которая называется априорной вероятностью. Суммируя возможные ошибки, можно ввести усредненную характеристику (критерий) качества принятия решения, называемую средним риском:

$$R = \sum_{i=1}^K \sum_{j=1}^K p_i p_{ij} \Pi_{ij}$$

Средний риск представляет собой математическое ожидание потерь, связанных с принятием решения.

Если априорные вероятности гипотез точно известны, а потери назначены обоснованно, то приемник, обеспечивающий наименьший средний риск, будет наиболее выгодным. Критерий минимума среднего риска называют также критерием Байеса.

Иногда потери, связанные с различными ошибками, принимают равными друг другу, $\Pi_{ij} = \Pi$; $\Pi_{ii} = 0$; $i = 1, \dots, K$, тогда оптимальный байесовский приемник обеспечивает минимальную среднюю вероятность ошибки (критерий идеального наблюдателя) и называется идеальным приемником Котельникова:

$$P_{\text{ош}} = \sum_{i=1}^K \sum_{\substack{j=1 \\ i \neq j}}^K p_i p_{ij}$$

Если принять равными и априорные вероятности гипотез $p_i = 1/K$; $i = 1, \dots, K$, то критерий Байеса сводится к критерию минимума суммарной условной вероятности ошибки:

$$P_{\text{ош усл}} = \sum_{i=1}^K \sum_{\substack{j=1 \\ i \neq j}}^K p_{ij} \quad (17.4)$$

17.2. Бинарная задача проверки простых гипотез

Наиболее просто задача построения оптимального приемника решается для случая амплитудной телеграфии с пассивной паузой, что

соответствует принятию решения о том, что передавался символ 0 (сигнала нет) или символ 1 (сигнал есть).

Предполагается, что помеха в канале представляет собой гауссовский шум с нулевым средним и известной дисперсией, который взаимодействует с сигналом аддитивно (суммируется). Результатом обработки наблюдаемого колебания является случайная величина y , которая может иметь различное распределение в зависимости от того, есть ли сигнал в наблюдаемом колебании, а именно: распределение при гипотезе H_0 – «сигнала нет» – является гауссовским с нулевым средним, а распределение при гипотезе H_1 – «сигнал есть» – отличается сдвигом на величину a , зависящую от способа обработки (например, если обработка сводится к взятию отсчета в момент, когда несущее колебание достигает максимума, величина a представляет собой его амплитуду). Значение a предполагается известным.

Таким образом, проверяемые гипотезы описываются двумя условными плотностями распределения вероятностей $w(y/H_0)$ и $w(y/H_1)$, изображенными на рис. 17.2.

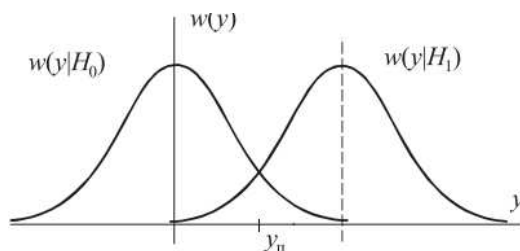


Рис. 17.2. Условные плотности распределения вероятностей величины y при простых гипотезах

Приемник в таком случае должен сравнить y с некоторым фиксированным значением (порогом) y_n и если y больше порога, принять решение о наличии сигнала, в противном случае – о его отсутствии, что можно кратко записать в следующей символической форме:

$$y \geq y_n \rightarrow \langle 1 \rangle;$$

$$y \leq y_n \rightarrow \langle 0 \rangle.$$

Каким бы ни был порог y_n , очевидно, есть некоторая ненулевая вероятность p_{01} принять решение о наличии сигнала при его фактическом отсутствии. Эта вероятность называется условной вероятностью ошибки первого рода («ложная тревога») и определяется выражением:

$$p_{01} = \int_{y_n}^{+\infty} w(y|H_0) dy.$$

Аналогично существует ненулевая вероятность принять решение об отсутствии сигнала, в то время как на самом деле он есть (условная вероятность ошибки второго рода, или пропуска сигнала):

$$p_{10} = \int_{-\infty}^{y_1} w(y | H_1) dy.$$

Анализ рис. 17.2 показывает, что сумма указанных условных вероятностей минимальна, если порог y_n находится как абсцисса точки пересечения условных плотностей $w(y|H_0)$ и $w(y|H_1)$.

Очевидно, при таком выборе порога приемник является оптимальным по критерию минимума суммарной условной вероятности ошибки (17.4) и принятие решения основывается на сравнении значений функций $w(y|H_0)$ и $w(y|H_1)$ при наблюдаемом значении y :

$$w(y|H_0) < w(y|H_1) \rightarrow "1"$$

$$w(y|H_0) \geq w(y|H_1) \rightarrow "0"$$

Это правило принятия решения можно переписать также в форме:

$$\frac{w(y | H_1)}{w(y | H_0)} > 1 \rightarrow "1"; \quad \frac{w(y | H_1)}{w(y | H_0)} \leq 1 \rightarrow "0". \quad (17.5)$$

Решение, таким образом, принимается в пользу той гипотезы, которая представляется более правдоподобной при данном значении y , поэтому отношение $\frac{w(y | H_1)}{w(y | H_0)}$ называется отношением правдоподобия и

обозначается $\Lambda(y)$. Правило (17.5) называют правилом максимального правдоподобия. Заметим, что критерий (17.4) часто называют критерием максимума правдоподобия.

Критерий идеального наблюдателя предполагает учет априорных вероятностей гипотез. Оптимальный по этому критерию приемник обеспечивает минимум средней вероятности ошибки, т.е. наименьшую сумму безусловных вероятностей ошибок первого и второго рода. Иначе говоря, сравнению подлежат функции $w(y|H_0)$ и $w(y|H_1)$, умноженные на соответствующие априорные вероятности. Правило принятия решения в таком приемнике можно записать в форме:

$$\frac{p_1 w(y | H_1)}{p_0 w(y | H_0)} > 1 \rightarrow "1"; \quad \frac{p_1 w(y | H_1)}{p_0 w(y | H_0)} \leq 1 \rightarrow "0"$$

Используя понятие отношения правдоподобия, можно записать правило в виде:

$$\Lambda(y) > \frac{p_0}{p_1} \rightarrow "1"; \quad \Lambda(y) \leq \frac{p_0}{p_1} \rightarrow "0",$$

при этом отношение правдоподобия сравнивается с пороговым значением, зависящим от априорных вероятностей.

Наконец, в случае байесовского критерия решение принимается по правилу:

$$\frac{\Pi_{10} p_1 w(y | H_1)}{\Pi_{01} p_0 w(y | H_0)} > 1 \rightarrow "1"; \quad \frac{\Pi_{10} p_1 w(y | H_1)}{\Pi_{01} p_0 w(y | H_0)} \leq 1 \rightarrow "0",$$

или

$$\Lambda(y) > \frac{p_0 \Pi_{01}}{p_1 \Pi_{10}} \rightarrow "1"; \quad \Lambda(y) \leq \frac{p_0 \Pi_{01}}{p_1 \Pi_{10}} \rightarrow "0".$$

Итак, во всех случаях оптимальный приемник «устроен одинаково»: для наблюдаемого значения y , зависящего от принятой реализации $z(t)$, вычисляется значение отношения правдоподобия, которое сравнивается с порогом; порог равен $\frac{p_0 \Pi_{01}}{p_1 \Pi_{10}}$ для оптимального приемника по

критерию минимума среднего риска, p_0 / p_1 для идеального приемника Котельникова и 1 для приемника максимального правдоподобия.

Следует отметить, что иногда удобнее вычислять не отношение правдоподобия, а его логарифм. В силу монотонности логарифмической функции это не влияет на условные вероятности ошибок, если порог также прологарифмировать.

17.3. Приём полностью известного сигнала (когерентный приём)

Рассмотрим принятие решения в системе связи при следующих условиях: синхронизация является точной и форма сигнала на интервале наблюдения точно известна, неизвестен лишь сам факт наличия либо отсутствия сигнала в наблюдаемом колебании. (Эта ситуация наиболее близка к реальности в кабельных линиях связи, где условия распространения сигналов известны и практически неизменны.)

Будем считать, что на интервале наблюдения независимо от сигнала присутствует гауссовский шум с нулевым средним значением и спектральной плотностью мощности $N_0/2$, постоянной в некоторой полосе частот $-F < f < F$ («квазибелый» шум).

Полагая, что длительность интервала наблюдения равна T , возьмем n отсчетов наблюдаемого колебания с шагом $\Delta t = 1/2F = T/n$, при этом отсчеты шума являются некоррелированными вследствие того, что корреляционная функция квазибелого шума (вида " $\sin x / x$ ") пересекает ось абсцисс при значениях времени, кратных Δt . Поэтому совместная плот-

ность распределения вероятностей взятых отсчетов (выборочных значений) в отсутствие сигнала равна:

$$w(z_1, \dots, z_n | H_0) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^n z_k^2},$$

где $\sigma^2 = N_0 / F = N_0 / 2\Delta t$.

Если сигнал присутствует и принимает в моменты взятия отсчетов значения $s_k = s(t_k)$, то совместная плотность распределения вероятностей выборочных значений

$$w(z_1, \dots, z_n | H_1) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^n (z_k - s_k)^2}.$$

Отношение правдоподобия

$$\Lambda = \frac{w(z_1, \dots, z_n | H_1)}{w(z_1, \dots, z_n | H_0)} = e^{-\frac{1}{2\sigma^2} [\sum_{k=1}^n (z_k - s_k)^2 - \sum_{k=1}^n z_k^2]}.$$

Подставляя в это выражение $2\sigma^2 = N_0 / \Delta t$, получим

$$\Lambda = e^{-\frac{1}{N_0} [\sum_{k=1}^n (z_k - s_k)^2 \Delta t - \sum_{k=1}^n z_k^2 \Delta t]}. \quad (17.6)$$

Устремляя Δt к нулю ($n \rightarrow \infty$), запишем логарифм отношения правдоподобия:

$$\begin{aligned} \ln \Lambda = & -\frac{1}{N_0} \int_0^T [z(t) - s(t)]^2 dt + \frac{1}{N_0} \int_0^T z^2(t) dt = \frac{2}{N_0} \int_0^T z(t) \cdot s(t) dt - \\ & - \frac{1}{N_0} \int_0^T s^2(t) dt. \end{aligned} \quad (17.7)$$

Поскольку логарифм является монотонной функцией, правило обнаружения сигнала известной формы на фоне гауссовского квазибелого шума, оптимальное по критерию максимума правдоподобия, основано на сравнении с нулевым порогом величины

$$\int_0^T y(t)s(t)dt - \frac{E}{2}, \quad (17.8)$$

где $E = \int_0^T s^2(t)dt$ – энергия сигнала.

Первое слагаемое в выражении (17.8) называется корреляционным интегралом, так как совпадает по форме с выражением взаимно корреля-

ляционной функции сигнала и наблюдаемого процесса при нулевом сдвиге. Энергия сигнала известна, и при обнаружении можно сравнивать значение корреляционного интеграла (случайное в силу случайности реализации $z(t)$) с порогом, равным $E/2$.

Правило различения M сигналов известной формы на фоне гауссовского квазибелого шума, оптимальное по критерию максимума правдоподобия, основано на сравнении между собой величин

$$\int_0^T z(t)s_i(t)dt - \frac{E_i}{2}, \quad i = 1, \dots, M.$$

Решение принимается в пользу того сигнала, для которого эта величина максимальна.

Структура оптимального приемника для различения M сигналов показана на рисунке 17.3.

Устройство выбора максимума УВМ выдает на выход номер k канала, в котором величина (17.8) максимальна.

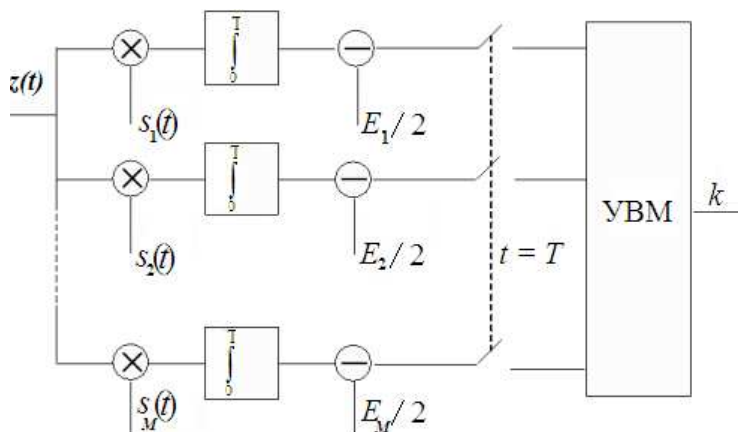


Рис. 17.3. Структура приемника максимального правдоподобия

Пример. В проводных системах связи с амплитудной телеграфией могут применяться посылки в форме прямоугольного видеоимпульса. Предположим, что сигнал, соответствующий символу «1», представляет собой прямоугольный видеоимпульс с амплитудой a и длительностью T . Тогда корреляционный интеграл имеет вид

$$\int_0^T z(t)s(t)dt = a \int_0^T z(t)dt,$$

порог равен $E/2 = a^2T/2$, тогда решающее правило имеет вид

$$\int_0^T z(t)dt > \frac{aT}{2} \rightarrow "1", \int_0^T z(t)dt \leq \frac{aT}{2} \rightarrow "0".$$

Структурная схема приемника показана на рис. 17.4. Постоянная времени интегрирующей цепи должна быть много больше длительности посылки T . В этом случае начальный участок экспоненты $a(1 - e^{-t/(RC)})$, отображающей заряд емкости, можно аппроксимировать прямой линией с тангенсом угла наклона $a/(RC)$, равным производной экспоненты в нуле. Тогда за время T напряжение на входе решающего устройства, обусловленное сигналом, составит $aT/(RC)$, а значение порога должно быть равно $aT/(2RC)$.

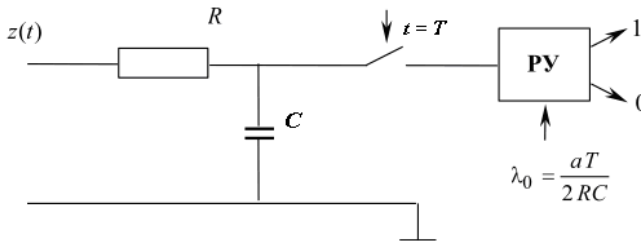


Рис. 17.4. Структурная схема приемника прямоугольного видеоимпульса

Пример. Предположим, что в двоичной системе связи с амплитудной телеграфией сигнал, соответствующий символу «1», представляет собой прямоугольный радиоимпульс с амплитудой a и длительностью T . Тогда $s(t) = \cos(\omega_0 t + \varphi)$, корреляционный интеграл имеет вид

$$\int_0^T z(t)s(t)dt = a \int_0^T z(t) \cos(\omega_0 t + \varphi)dt,$$

а порог равен $E/2 = a^2 T/4$. Сокращая на a и применяя реальный интегратор в виде RC -цепи, получаем структуру приемника, показанную на рис. 17.5.

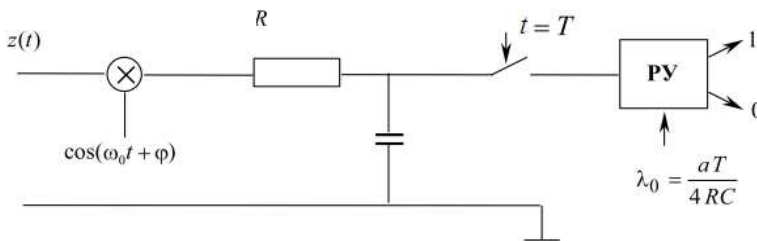


Рис. 17.5. Структурная схема приемника прямоугольного радиоимпульса

Пример. В двоичной системе связи с фазовой телеграфией сигналы $s_1(t)$ и $s_2(t)$, соответствующие символам «1» и «0», являются противоположными

$$s_1(t) = \cos(\omega_0 t + \varphi); s_2(t) = \cos(\omega_0 t + \varphi + \pi) = -s_1(t)$$

Принятие решения основано на сравнении величин

$$\int_0^T z(t)s_1(t)dt - \frac{E_1}{2} \text{ и } \int_0^T z(t)s_2(t)dt - \frac{E_2}{2}.$$

С учётом равенства энергий правило принятия решения упрощается и принимает вид

$$\int_0^T z(t) \cos(\omega_0 t + \varphi) dt > 0 \rightarrow "1", \int_0^T z(t) \cos(\omega_0 t + \varphi) dt \leq 0 \rightarrow "0".$$

17.4. Согласованная фильтрация

По существу корреляционный приемник является активным фильтром и выполняет операцию скалярного произведения

$$(\bar{x}\bar{s}_i) = \int_0^T x(t)s_i(t)dt. \quad (17.9)$$

Эту операцию можно реализовать также с помощью пассивного линейного фильтра с постоянными параметрами. Если на вход фильтра подать принимаемый сигнал, то напряжение на выходе фильтра

$$y(t) = \int_0^T x(t-\tau)g(\tau)d\tau, \quad (17.10)$$

где $g(t)$ – импульсная реакция фильтра, является зеркальным отображением $s(t)$ относительно $t = T$.

Выберем $g(t)$ такой, чтобы в момент $t = T$ получить на выходе значение $y(T)$, совпадающее со скалярным произведением (17.9).

Легко видеть, что это будет выполнено, если $g(t) = s_i(T-t)$. Действительно, при этом

$$y(T) = \int_0^T x(T-\tau)s_i(T-\tau)d\tau = \int_0^T x(t)s_i(t)dt = (\bar{x}\bar{s}). \quad (17.11)$$

Такой фильтр называют согласованным (СФ) с сигналом $s_i(t)$. Иначе говоря, фильтр является согласованным с сигналом $s(t)$, если его импульсная реакция

$$g(t) = a \cdot s(T-t), \quad (17.12)$$

где a – постоянная.

Передаточная функция СФ с импульсной реакцией (17.12) определяется преобразованием Фурье

$$\dot{K}_{\text{Ф}}(\omega) = \int_{-\infty}^{\infty} g(t)e^{-j\omega t} dt = a\dot{s}^*(\omega)e^{-j\omega T}, \quad (17.13)$$

где $\dot{S}^*(\omega) = \dot{S}(-\omega)$ – функция комплексного сопряжения со спектральной плотностью сигнала $s(t)$.

Следовательно, с точностью до коэффициента a АЧХ согласованного фильтра определяется амплитудным спектром сигнала $s(t)$. Смысл согласования проявляется в том, что СФ хорошо пропускает те частоты, которые дают большой вклад в энергию сигнала. ФЧХ СФ (без учета слагаемого $-\omega T$) обратна по знаку ФЧХ сигнала $s(t)$. Благодаря этому при $t = T$ все составляющие спектра принимаемого сигнала складываются в фазе и дают максимальный отклик.

Отметим одно важное свойство СФ, которое иногда рассматривается как его определение. Будем подавать сумму детерминированного сигнала и белого шума на вход различных линейных цепей с постоянными параметрами и измерять в момент $t = T$ отношение пиковой мощности сигнальной составляющей к средней мощности шума на выходе цепи. Оказывается, что это отношение максимально для СФ и равно

$$h_{2\text{max}}^2 = \frac{2E_s}{N_0}, \quad (17.14)$$

где E_s – энергия сигнала;

N_0 – спектральная плотность белого шума.

Иначе говоря, СФ является единственным линейным фильтром, обеспечивающим получение максимально возможного отношения сигнала к помехе на выходе.

Интересно сравнить $h_{2\text{max}}^2$ с отношением средних мощностей сигнала P_s и помехи P_ξ на входе фильтра h_1^2 :

$$h_1^2 = \frac{P_s}{P_\xi} = \frac{2T}{2T} \frac{P_s}{N_0 F_s} = \frac{2E_s/N_0}{2F_s T} = \frac{h_{2\text{max}}^2}{2F_s T}$$

Откуда

$$h_{2\text{max}}^2 = 2F_s T h_1^2 = n h_1^2, \quad (17.15)$$

где $n = 2F_s T$ – база сигнала.

Таким образом, улучшение отношения сигнала к помехе, даваемое СФ, тем больше, чем больше база сигнала n , т.е. чем сложнее форма сигнала.

Согласованный фильтр для сигнала произвольной формы может быть реализован (приближенно) на основе линии задержки с отводами (рис. 17.6).

При подаче на вход 1 линии задержки с отводами (ЛЗО) короткого импульса, на вход ФНЧ поступают (с интервалом Δt , обусловленным конструкцией линии задержки) такие же импульсы с амплитудами, определяемыми коэффициентами усиления $a_0, a_1, a_2, \dots, a_{n-1}$. Тогда на выходе ФНЧ формируется, в частности если ФНЧ является идеальным, отклик представляющий собой конечную сумму ряда Котельникова аппроксимирующую сигнал $s(t)$ требуемого вида.

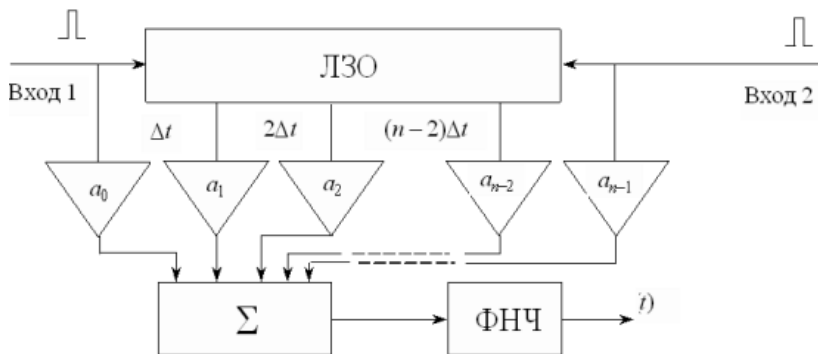


Рис. 17.6. Согласованный фильтр на основе линии задержки с отводами

Нетрудно видеть, что если короткий импульс подать на вход 2, то отклик будет зеркальной копией сигнала $s[(n-1)\Delta t - t]$. Коэффициенты $a_0, a_1, a_2, \dots, a_{n-1}$ представляют собой отсчеты сигнала $s(t)$ с шагом, определяемым верхней частотой F_v спектра сигнала.

Следует иметь в виду, что такой способ реализации согласованного фильтра является хотя и универсальным, но заведомо приближенным, так как любой сигнал конечной длительности имеет нефинитную спектральную плотность, а идеальный ФНЧ нереализуем. Тем не менее, такой фильтр применяется на практике: например, для согласованной фильтрации сигналов с линейной частотной модуляцией используют в качестве линий задержки с отводами интегральные устройства на поверхностных акустических волнах (ПАВ).

Форма сигнала на выходе такого фильтра отличается от формы входного сигнала. Но назначение согласованного фильтра состоит в вычислении корреляционного интеграла для наиболее надежного принятия решения о наличии или отсутствии сигнала на входе приемника. Иными словами, согласованный фильтр должен обеспечивать максимальное отношение сигнал/шум в момент времени t_0 .

17.5. Потенциальная помехоустойчивость когерентного приёма

По определению В.А. Котельникова потенциальной помехоустойчивостью называется максимум вероятности правильного решения, достижимый при заданных условиях приема сигналов на фоне помех (шумов) [13].

Определим потенциальную помехоустойчивость приёма двух сигналов, $s_0(t)$ и $s_1(t)$ известной формы на фоне белого гауссовского шума при равных априорных вероятностях сигналов.

Алгоритм принятия решения в приемнике, реализующем критерий максимума правдоподобия, кратко запишем в виде

$$\int_0^T z(t)s_1(t)dt - \frac{E_1}{2} \stackrel{?}{\geq} \int_0^T z(t)s_0(t)dt - \frac{E_0}{2}.$$

Это выражение можно привести к виду

$$\int_0^T z(t)[s_1(t) - s_0(t)]dt \stackrel{?}{\geq} \frac{1}{2} \frac{(E_1 - E_0)}{2}.$$

Ошибки при приеме состоят в том, что при передаче первого сигнала принимается решение о приеме второго и наоборот, поскольку гауссово распределение симметрично и априорные вероятности равны.

Найдем условную вероятность ошибки, т.е. вероятность события, заключающегося в принятии решения о наличии сигнала $s_0(t)$ при условии, что в наблюдаемом колебании присутствует сигнал $s_1(t)$. Это событие соответствует выполнению неравенства

$$\int_0^T [s_1(t) + \xi(t)] \cdot [s_1(t) - s_0(t)]dt < \frac{(E_1 - E_0)}{2},$$

которое можно переписать в виде

$$\int_0^T s_1^2(t)dt + \int_0^T \xi(t) \cdot [s_1(t) - s_0(t)]dt - \int_0^T s_1(t) \cdot s_0(t)dt < \frac{1}{2} \int_0^T s_1^2(t)dt - \frac{1}{2} \int_0^T s_0^2(t)dt$$

Проведя очевидные преобразования, получим

$$\int_0^T \xi(t) \cdot [s_1(t) - s_0(t)]dt < -\frac{1}{2} \int_0^T [s_1(t) - s_0(t)]^2 dt. \quad (17.16)$$

Левая часть неравенства представляет собой случайную величину (т.к. это интеграл по времени от случайного процесса $\xi(t)$ с весом, рав-

ным разности сигналов $s_{\Delta}(t) = [s_1(t) - s_0(t)]$, имеющую нормальное распределение (поскольку процесс $\xi(t)$ гауссов) с нулевым средним; обозначим ее v , её дисперсия (средний квадрат), равна:

$$D_v = \overline{v^2} = \int_0^T \int_0^T \overline{\xi(t_1) \cdot \xi(t_2) \cdot s_{\Delta}(t_1) \cdot s_{\Delta}(t_2)} dt_1 dt_2 = \frac{N_0 \cdot E_{\Delta}}{2}. \quad (17.17)$$

Вероятность выполнения неравенства (17.16) – это вероятность того, что нормальная случайная величина с нулевым средним и дисперсией $N_0 E_{\Delta}/2$ принимает значение меньше, чем $E_{\Delta}/2$. Эта вероятность равна

$$p_{10} = \int_{-\infty}^{-E_{\Delta}/2} \frac{1}{\sqrt{2\pi D_v}} e^{-\frac{v^2}{2D_v}} dv = \int_{\alpha}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt,$$

где $t = -\frac{v}{\sqrt{D_v}}$ – центрированная нормальная случайная величина с единичной дисперсией, а $\alpha = \frac{E_{\Delta}}{2\sqrt{D_v}}$ – положительное число.

Очевидно, p_{10} зависит только от $\alpha = \frac{E_{\Delta}}{2\sqrt{D_v}} = \sqrt{\frac{E_{\Delta}}{2N_0}}$, поэтому

можно ввести функцию

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{t^2}{2}} dt = 1 - \Phi(x),$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$ – интеграл вероятности, и записать

$$p_{10} = Q\left(\sqrt{\frac{E_{\Delta}}{2N_0}}\right).$$

(Напомним, что в силу симметрии гауссовского распределения $p_{10} = p_{01}$).

Таким образом, условная вероятность ошибки, равная средней вероятности ошибки при когерентном приеме сигналов на фоне белого шума, определяется энергией разностного сигнала $s_{\Delta}(t)$ и спектральной плотностью мощности шума N_0 .

Рассмотрим потенциальную помехоустойчивость двоичного когерентного приемника максимального правдоподобия для различных способов модуляции, считая, что энергия сигнала E фиксирована.

Представляет практический интерес сравнительная оценка потенциальной помехоустойчивости сигналов дискретной модуляции: амплитудной (ДАМ), частотной (ДЧМ), фазовой (ДФМ).

При ДАМ символу «1» соответствует сигнал $S_1(t)$, а символу «0» соответствует сигнал $S_2(t) = 0$ («пассивная пауза»), следовательно норма разностного сигнала равна

$$\|\Delta\vec{S}\|_{\text{ДАМ}} = \|\vec{S}_1 - \vec{S}_2\| = \|\vec{S}_1\| = \sqrt{\int_0^T S_1^2(t) dt} = \sqrt{E_S}.$$

При ДЧМ и ДФМ символам «1» и «0» соответствуют сигналы $S_1(t)$ и $S_2(t)$, энергия которых одинакова. При этом для нормы разностного сигнала имеем

$$\|\Delta\vec{S}\| = \sqrt{\int_0^T [S_1(t) - S_2(t)]^2 dt} = \sqrt{2E_S(1 - b_{12})};$$

$b_{12} = \frac{\int_0^T [S_1(t) - S_2(t)] dt}{E_S}$, где b_{12} – нормированный коэффициент корреляции ($-1 \leq b_{12} \leq 1$).

В частности, для ортогональных сигналов ДЧМ

$$b_{12} = 0, \|\Delta\vec{S}\|_{\text{ДЧМ}} = \sqrt{2}\|\vec{S}\| = \sqrt{2E_S},$$

а для противоположных сигналов ДФМ

$$b_{12} = -1, \|\Delta\vec{S}\|_{\text{ДФМ}} = \sqrt{2}\|\vec{S}\| = \sqrt{4E_S}.$$

Сигналы с $b_{12} = 1$ являются одинаковыми, т.е., и их невозможно различить. Для них $P_{\text{ош}} = 0,5$, что эквивалентно обрыву канала связи.

Т.о. энергия разностного сигнала $E_{\Delta S} = \|\vec{S}_1 - \vec{S}_2\|^2$ для ДЧМ в 2 раза, а для ДФМ – в 4 раза больше по сравнению с ДАМ. Соответственно возрастает и помехоустойчивость.

Заметим, что если используются три сигнала одинаковой энергии, то для достижения максимальной помехоустойчивости они должны иметь взаимный фазовый сдвиг 120° , т.е. соответствующие сигналам точки должны располагаться на окружности радиуса $E^{1/2}$ в вершинах равностороннего треугольника (рис. 17.7.з). Если сигналов четыре, то оптимальным является их размещение в вершинах правильного тетраэдра, вписанного в сферу радиуса $E^{1/2}$.

В общем случае оптимальный выбор системы из n сигналов соответствует их расположению в вершинах правильного $(n - 1)$ - мерного симплекса, вписанного в $(n - 1)$ - мерную сферу (отрезок, треугольник и тетраэдр являются одномерным, двумерным и трехмерным симплексами).

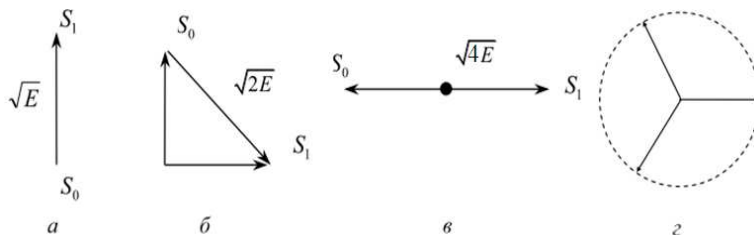


Рис. 17.7. К помехоустойчивости приема двух и трёх сигналов

Ошибки приема сигналов для различных видов манипуляций

1. Амплитудная телеграфия с пассивной паузой

В этом случае $s_0(t) = 0$ и энергия разностного сигнала равна E (норма равна $E/2$), рисунке 17.7а. Следовательно, потенциальная помехоустойчивость определяется средней вероятностью ошибки

$$P_{\text{ош}}^{\text{АТ-ПП}} = Q\left(\sqrt{\frac{E}{2N_0}}\right).$$

2. Частотная телеграфия с ортогональными сигналами

Два сигнала представляют собой радиоимпульсы одинаковой формы с различными несущими частотами, так что сигналы взаимно ортогональны (рис. 17.7б). Энергия разностного сигнала равна $2E$, а средняя вероятность ошибки

$$P_{\text{ош}}^{\text{ЧТ}} = Q\left(\sqrt{\frac{E}{N_0}}\right).$$

Повышение потенциальной помехоустойчивости при переходе от АТ-ПП к частотной телеграфии представляется естественным, так как во втором случае вдвое возрастает средняя мощность передатчика. Средняя вероятность ошибки может быть понижена без увеличения мощности передатчика, если перейти к взаимно обратным сигналам.

3. Фазовая телеграфия с манипуляцией фазы на 180°

В случае фазовой телеграфии с взаимно обратными сигналами (рис. 17.7,в) энергия разностного сигнала составляет $4E$, средняя вероятность ошибки равна

$$P_{\text{ош}}^{\text{ФТ}} = Q\left(\sqrt{\frac{2E}{N_0}}\right).$$

и дальнейшее повышение потенциальной помехоустойчивости за счет выбора сигналов при заданной энергии невозможно.

17.6. Некогерентный приём

На практике иногда не удается обеспечить условия для когерентного приема сигналов, так как один или несколько параметров принимаемого сигнала оказываются неизвестными. Такая ситуация типична, например, для систем спутниковой связи, радиосвязи с подвижными объектами, и т.п., поскольку расстояние между передатчиком и приемником изменяется случайным образом. Это приводит, в частности, к тому, что меняется начальная фаза несущего колебания.

Если изменение происходит настолько медленно, что соседние посылки имеют практически одинаковую начальную фазу, то ее можно оценить и оценку использовать вместо точного значения при организации приема. Такой прием называют квазикогерентным.

Если же начальная фаза изменяется (флюктуирует) быстро или устройство оценивания оказывается слишком сложным, тогда рассматривается задача приема сигнала со случайной начальной фазой, или некогерентного приема.

Перепишем выражение (17.7) для логарифма отношения правдоподобия при приеме сигнала $s(t)$:

$$\ln \Lambda = \frac{2}{N_0} \int_0^T z(t) \cdot s(t) dt - \frac{1}{N_0} \int_0^T s^2(t) dt. \quad (17.18)$$

Сигнал при некогерентном приеме известен с точностью до начальной фазы, поэтому обозначим его $s(t, \varphi)$ и запишем $s(t, \varphi) = \operatorname{Re}\{\hat{s}(t) \cdot e^{-j\varphi}\}$.

В этом выражении неизвестная начальная фаза сигнала представлена комплексным фазовым множителем $e^{-j\varphi}$ при аналитическом комплексном сигнале, который определяется выражением $\hat{s}(t) = s(t) + j\hat{s}(t)$, где вещественная и мнимая части связаны парой преобразований Гильберта

$$\begin{aligned} \hat{s}(t) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{s(\tau)}{t - \tau} d\tau, \\ s(t) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\hat{s}(\tau)}{t - \tau} d\tau. \end{aligned}$$

Тогда,

$$s(t, \varphi) = \operatorname{Re}\{[s(t) + j\hat{s}(t)] \cdot [\cos \varphi - j \sin \varphi]\} = s(t) \cos \varphi + j\hat{s}(t) \sin \varphi.$$

Корреляционный интеграл в выражении (17.18) в таком случае приобретает вид

$$\begin{aligned} \int_0^T z(t)s(t, \varphi) dt &= \int_0^T z(t)s(t) \cos \varphi dt + \int_0^T z(t)\hat{s}(t) \sin \varphi dt = \\ &= \operatorname{Re} \left\{ \int_0^T z(t)[s(t) + j\hat{s}(t)] \cdot [\cos \varphi - j \sin \varphi] dt \right\} = \operatorname{Re} \left\{ \int_0^T z(t)\dot{s}(t) e^{-j\varphi} dt \right\} = \\ &= \operatorname{Re} \left\{ e^{-j\varphi} \int_0^T z(t)\dot{s}(t) dt \right\} = \operatorname{Re} \{ e^{-j\varphi} \dot{V} \} = \operatorname{Re} \{ e^{-j\varphi} \cdot V \cdot e^{-j\psi} \} \end{aligned} \quad (17.19)$$

В полученном выражении фигурирует комплексная величина \dot{V} , имеющая смысл корреляционного интеграла для аналитического сигнала $s(t)$:

$$\dot{V} = \int_0^T z(t)\dot{s}(t) e^{-j\varphi} dt = \int_0^T z(t)s(t) dt + j \int_0^T z(t)s(t) dt,$$

где

$$\begin{aligned} V &= \sqrt{\left(\int_0^T z(t)s(t) dt \right)^2 + \left(\int_0^T z(t)s(t) dt \right)^2}; \\ \psi &= -\operatorname{arctg} \frac{\int_0^T z(t)s(t) dt}{\int_0^T z(t)s(t) dt}. \end{aligned}$$

Корреляционный интеграл согласно (17.19) можно переписать в виде

$$\int_0^T z(t)s(t, \varphi) dt = V \cos(\psi + \varphi),$$

тогда логарифм отношения правдоподобия

$$\ln \Lambda = \frac{2}{N_0} V \cos(\psi + \varphi) - \frac{1}{N_0} E,$$

а само отношение правдоподобия

$$\Lambda = e^{\frac{2}{N_0} V \cos(\psi + \varphi)} e^{-\frac{1}{N_0} E}.$$

Считая, что начальная фаза сигнала является случайной величиной, имеющей равномерное в интервале $(0, 2\pi)$ распределение, выполним усреднение отношения правдоподобия по ансамблю:

$$\bar{\Lambda} = e^{-\frac{E}{N_0}} \cdot \frac{1}{2\pi} \int_0^{2\pi} e^{\frac{2V}{N_0} \cos(\psi+\varphi)} d\varphi.$$

Учтём известное соотношение

$$\frac{1}{2\pi} \int_0^{2\pi} e^{a \cos(\psi+\varphi)} d\varphi = I_0(a),$$

где $I_0(a)$ – модифицированная функция Бесселя нулевого порядка, тогда

$$\bar{\Lambda} = e^{-\frac{E}{N_0}} \cdot I_0\left(\frac{2V}{N_0}\right).$$

Правило некогерентного приема сигнала со случайной равновероятной начальной фазой на фоне гауссовского шума должно быть основано на сравнении величины $\bar{\Lambda}$ с некоторым порогом, а правило различения двух сигналов – на сравнении двух отношений правдоподобия между собой. Предположим, что рассматривается прием двух сигналов $s_1(t)$ и $s_0(t)$. Сравнение усредненных отношений правдоподобия можно заменить сравнением их логарифмов

$$\ln I_0\left(\frac{2V_1}{N_0}\right) - \frac{E_1}{N_0} \stackrel{1}{\approx} \ln I_0\left(\frac{2V_0}{N_0}\right) - \frac{E_0}{N_0},$$

или сравнением с порогом разности логарифмов

$$\ln I_0\left(\frac{2V_1}{N_0}\right) - \ln I_0\left(\frac{2V_0}{N_0}\right) \stackrel{1}{\approx} \frac{E_1 - E_0}{N_0}.$$

Алгоритм сильно упрощается, если энергии сигналов равны, в этом случае в силу монотонности функции I_0 можно сравнивать между собой величины:

$$V_1 \stackrel{1}{\approx} V_0.$$

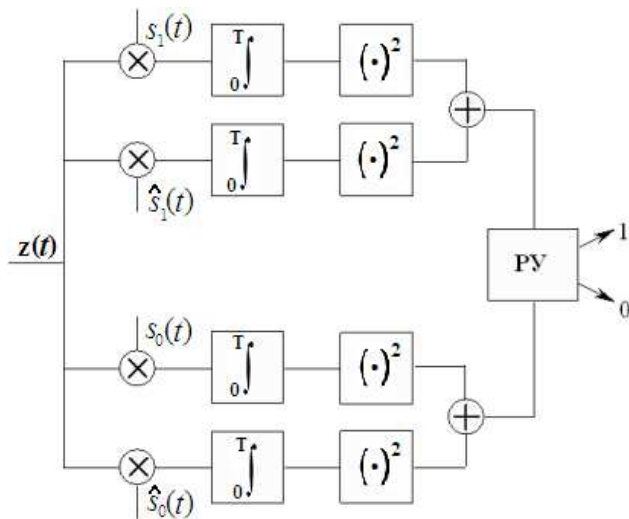


Рис. 17.8. Структура некогерентного приемника двух сигналов с равными энергиями

То же правило можно реализовать с использованием согласованных фильтров по схеме, рисунок 17.9. Здесь вычисление величин V_1 и V_0 производится устройством, называемым детектором огибающей ДО.

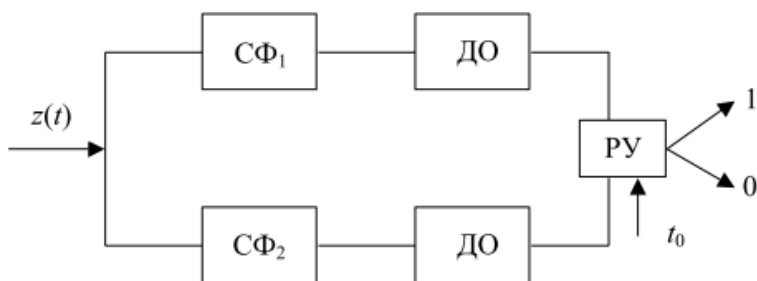


Рис. 17.9. Структура некогерентного приемника двух сигналов с использованием согласованных фильтров

17.7. Потенциальная помехоустойчивость некогерентного приёма

Определим потенциальную помехоустойчивость некогерентного приема на примере системы с пассивной паузой при равных априорных вероятностях посылок

$$s_1(t) = A \cos(\omega t + \varphi), s_0(t) = 0, p_1 = p_0 = 0,5.$$

Средняя вероятность ошибки равна

$$p_{\text{ош}} = 0,5 \cdot p_{01} + 0,5 \cdot p_{10} = 0,5 \cdot \int_0^{V_{\text{п}}} w_1(V | H_1) dV + 0,5 \cdot \int_{V_{\text{п}}}^{\infty} w_0(V | H_0) dV.$$

Здесь $w_1(V|H_1)$ и $w_0(V|H_0)$ – условные плотности распределения вероятности огибающей корреляционного интеграла при условии принятия гипотез о передаче сигналов $s_1(t)$ и $s_0(t)$ соответственно, $V_{\text{п}}$ – порог (рис. 17.10).

При гипотезе H_0 значение огибающей обусловлено только шумом, тогда квадратурные составляющие являются независимыми нормальными случайными величинами с нулевыми средними и дисперсиями $N_0E/2$ (см. выражение 17.17).

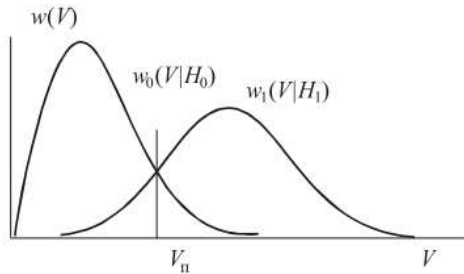


Рис. 17.10. Выбор порога при некогерентном приеме

Условная плотность распределения вероятностей огибающей имеет рэлеевский вид

$$w_0(V | H_0) = \frac{2V}{EN_0} e^{-\frac{V^2}{EN_0}}.$$

Если наблюдаемое колебание содержит сигнал $s_1(t)$, то огибающая имеет обобщенное рэлеевское распределение (распределение Рэлея-Райса)

$$w_1(V | H_1) = \frac{2V}{EN_0} e^{-\frac{V^2+E^2}{EN_0}} \cdot I_0\left(\frac{2V}{N_0}\right).$$

Средняя вероятность ошибки равна

$$p_{\text{ош}} = \frac{1}{2} \int_0^{V_{\text{п}}} \frac{2V}{EN_0} e^{-\frac{V^2+E^2}{EN_0}} \cdot I_0\left(\frac{2V}{N_0}\right) dV + \frac{1}{2} \int_{V_{\text{п}}}^{\infty} \frac{2V}{EN_0} e^{-\frac{V^2}{EN_0}} dV. \quad (17.20)$$

Второй интеграл берется по частям, при этом

$$P_{\text{ош}} = \frac{1}{2} \int_0^{V_{\text{п}}} \frac{2V}{EN_0} e^{-\frac{V^2+E^2}{EN_0}} \cdot I_0\left(\frac{2V}{N_0}\right) dV + \frac{1}{2} e^{-\frac{V^2}{EN_0}}.$$

Оптимальное значение порога, при котором достигается потенциальная помехоустойчивость некогерентного приема, является решением уравнения $dp_{\text{ош}}/dV_{\text{п}} = 0$. Точно решить полученное уравнение не удастся. Поэтому оптимальный порог определяется приближенными выражениями

$$V_{\text{п опт}} = \begin{cases} \frac{E}{2} & \text{- при больших отношениях } \frac{\text{сигнал}}{\text{шум}}, \\ \sqrt{EN_0} & \text{- при малых отношениях } \frac{\text{сигнал}}{\text{шум}}. \end{cases}$$

Подставляя в (17.20) порог $E/2$, получим среднюю вероятность ошибки при больших отношениях сигнал/шум (ОСШ):

$$P_{\text{ош}} = \frac{1}{2} \int_0^{E/2} \frac{2V}{EN_0} e^{-\frac{V^2+E^2}{EN_0}} \cdot I_0\left(\frac{2V}{N_0}\right) dV + \frac{1}{2} e^{-\frac{E}{4N_0}}$$

При больших ОСШ ($E/N_0 \geq 10$) первым слагаемым можно пренебречь, тогда

$$P_{\text{ош}} = \frac{1}{2} e^{-\frac{E}{4N_0}}.$$

Аналогично можно проанализировать помехоустойчивость приема двух ортогональных частотно-манипулированных сигналов; для этого случайная средняя вероятность ошибки

$$P_{\text{ош}} = \frac{1}{2} e^{-\frac{E}{2N_0}}.$$

Сигналы с фазовой манипуляцией при случайной начальной фазе каждой посылки применять при некогерентном приеме нельзя. Однако при медленных изменениях фазы можно использовать относительную фазовую манипуляцию, при которой начальная фаза следующей посылки совпадает с начальной фазой предыдущей посылки при передаче символа «0» и отличается от нее на 180° – при передаче символа «1». При этом средняя вероятность ошибки

$$P_{\text{ош}} = \frac{1}{2} e^{-\frac{E}{N_0}}.$$

18. ПОМЕХОУСТОЙЧИВОСТЬ СИСТЕМ ПЕРЕДАЧИ НЕПРЕРЫВНЫХ СООБЩЕНИЙ

Непрерывные сообщения (например, речь, музыка и т.п.) могут передаваться по каналу связи непосредственно (например, по местной проводной радиосети, по телефонному кабелю) или при помощи модуляции. В первом случае сигнал $s(t)$, передаваемый по каналу, может совпадать с сообщением (первичным сигналом) $b(t)$ или быть связан с ним простой пропорциональной зависимостью, во втором – передаваемый сигнал $s[t, b(t)]$ является функцией сообщения, в общем случае нелинейной (рис. 18.1).

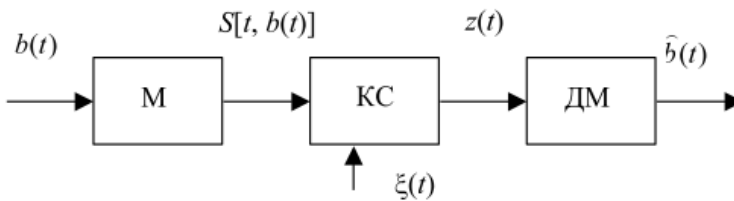


Рис. 18.1. Структура системы передачи непрерывных сообщений

Колебание на входе демодулятора $z(t) = s[t, b(t)] + \zeta(t)$ представляет собой в простейшем случае сумму передаваемого сигнала и шума $\zeta(t)$.

Задача демодулятора состоит в нахождении оценки первичного сигнала (сообщения), наилучшей в соответствии с выбранным критерием близости. В качестве критерия часто используют средний квадрат ошибки

$$\overline{\varepsilon^2} = \overline{[b(t) - \hat{b}(t)]^2}, \quad (18.1)$$

где черта означает статистическое усреднение по ансамблю. В системах телеметрии используется критерий максимальной ошибки, в радиовещании – увеличение выходного отношения сигнал/шум по сравнению с входным, критерий разборчивости речевых сообщений и т.п.

18.1. Оптимальное оценивание сигнала

Оценивание сигнала, как функции времени, – достаточно сложная задача. Во многих случаях ее можно свести к более простой задаче оценивания одного или нескольких параметров сигнала.

Простейшей задачей, связанной с оцениванием параметров сигнала, является оценка параметра, постоянного или настолько медленно

меняющегося во времени, что на интервале наблюдения его можно считать постоянным.

Рассмотрим задачу оценивания единственного параметра λ , который до опыта рассматривается как случайная величина, имеющая априорное распределение с плотностью $w(\lambda)$.

Реализация этой случайной величины представляет собой значение, постоянное на интервале $(0, T)$ наблюдения колебания $z(t) = s[t, \lambda] + \zeta(t)$.

Правило оценивания – это алгоритм обработки наблюдаемого колебания, результатом выполнения которого является значение $\tilde{\lambda}$ оценки параметра λ .

Для оценивания одного и того же параметра может существовать множество алгоритмов, вырабатывающих различные оценки. Для сравнения алгоритмов оценивания между собой и выбора наилучшего используют показатели несмещенность, состоятельность и эффективность.

Несмещенность. Оценка называется несмещенной, если выполняется условие $\overline{\tilde{\lambda} - \lambda} = 0$, означающее, что при любом значении параметра условное математическое ожидание оценки равно этому значению $\overline{\tilde{\lambda}} = \lambda$.

Другими словами, несмещенность означает отсутствие систематической ошибки оценивания. В противном случае оценка называется смещенной. Следует отметить, что смещенные оценки также находят применение, если смещение достаточно мало или стремится к нулю при увеличении времени наблюдения или мощности сигнала.

Состоятельность. Оценка называется состоятельной, если при неограниченном возрастании времени наблюдения оценка сходится по вероятности к значению параметра

$$\lim_{T \rightarrow \infty} P\{|\tilde{\lambda} - \lambda| \geq \Delta\} = 0, \text{ при любом } \Delta > 0.$$

(Здесь $P\{A\}$ обозначает вероятность события A .)

Смещенная оценка может быть состоятельной, если ее смещение стремится к нулю при $T \rightarrow \infty$. Следовательно, дисперсия ошибки для состоятельной оценки стремится к нулю

$$\lim_{T \rightarrow \infty} \overline{|\tilde{\lambda} - \lambda|^2} = 0.$$

Эффективность. Несмещенная оценка называется эффективной, если среди всех оценок, полученных при заданном времени наблюдения всевозможными алгоритмами оценивания, она обеспечивает наименьшую дисперсию ошибки

$$\overline{|\tilde{\lambda} - \lambda|^2} = \min.$$

Классический подход к оцениванию параметров сигналов основывается на формуле Байеса для апостериорной плотности распределения вероятностей оцениваемого параметра

$$w(\lambda | z) = \frac{w(\lambda)w(z | \lambda)}{w(z)}, \quad (18.2)$$

где $w(\lambda)$ – априорная ПРВ параметра λ ; $w(\lambda/z)$ – условная ПРВ наблюдаемого процесса при заданном значении λ , рассматриваемая как функция от λ при данном z (функция правдоподобия); $w(z)$ – при фиксированной реализации z постоянная величина.

Выражение (18.2) показывает, что, зная априорную плотность $w(\lambda)$ и наблюдая реализацию процесса z , можно получить уточненное представление о значении параметра λ . На рисунке 18.2 показаны примеры априорной и апостериорной ПРВ параметра (истинное значение параметра обозначено λ_0).

Апостериорное распределение обостряется по сравнению с априорным распределением, т.к., реализация z , содержит дополнительную информацию о параметре, что уменьшает исходную неопределенность априорной ПРВ.

Апостериорное распределение содержит всю информацию о параметре, которую можно получить из наблюдаемой реализации и априорных данных. Поэтому правило оценивания должно использовать апостериорную ПРВ, а способ ее использования зависит от выбранного критерия качества оценки.

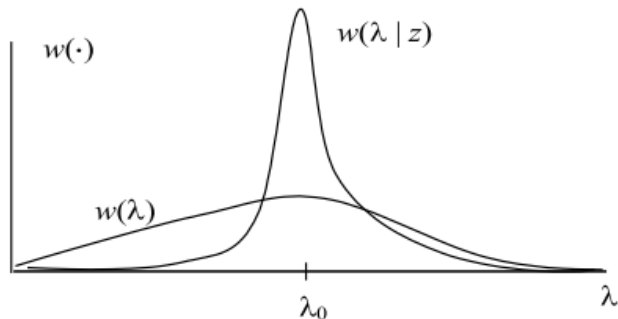


Рис. 18.2. Априорная и апостериорная ПРВ оцениваемого параметра

Ошибки оценивания параметра приводят к различным последствиям, поэтому для их учета вводится функция потерь $L(\tilde{\lambda} - \lambda)$, зависящая от разности оценки и истинного значения параметра.

Усредняя функцию потерь по апостериорному распределению параметра, получаем количественную характеристику, называемую апостериорным (условным) риском

$$r(\tilde{\lambda}, z) = \int_{(\lambda)} L(\tilde{\lambda} - \lambda) w(\lambda | z) d\lambda, \quad (18.3)$$

описывающим потери, связанные с получением оценки при наблюдении реализации z . Усреднение апостериорного риска (18.3) по всевозможным реализациям приводит к среднему риску

$$R(\tilde{\lambda}) = \int_{(z)} w(z) \left[\int_{(\lambda)} L(\tilde{\lambda} - \lambda) w(\lambda | z) d\lambda \right] dz.$$

Правило оценивания, которому соответствует наименьший средний риск, называется байесовским, а соответствующая оценка – байесовской, или оценкой по критерию минимума среднего риска. Правило, оптимальное в смысле минимума среднего риска, находится из условия минимизации условного риска (18.3).

Часто используют квадратичную функцию потерь $L(\tilde{\lambda} - \lambda) = (\tilde{\lambda} - \lambda)^2$, тогда

$$r(\tilde{\lambda}, z) = \int_{(\lambda)} L(\tilde{\lambda} - \lambda)^2 w(\lambda | z) d\lambda = \overline{(\tilde{\lambda} - \lambda)^2}, \quad (18.4)$$

Таким образом апостериорный риск равен среднему квадрату ошибки (а если оценка несмещенная, то дисперсии ошибки). Байесовская оценка в этом случае становится оценкой минимума среднеквадратической ошибки. Для нахождения правила раскроем скобки в выражении (18.4):

$$r(\tilde{\lambda}, z) = \overline{(\tilde{\lambda} - \lambda)^2} = \tilde{\lambda}^2 - 2\tilde{\lambda} \int_{(\lambda)} \lambda w(\lambda | z) d\lambda + \int_{(\lambda)} \lambda^2 w(\lambda | z) d\lambda.$$

Дифференцируя полученное выражение по λ и приравнявая результат нулю, получаем правило

$$\tilde{\lambda} = \int_{(\lambda)} \lambda w(\lambda | z) d\lambda.$$

То есть оценка, оптимальная в смысле минимума среднеквадратической ошибки, равна апостериорному среднему значению параметра.

Кроме квадратичной часто используется простая функция потерь

$$L(\tilde{\lambda} - \lambda) = \text{const} \cdot \delta(\tilde{\lambda} - \lambda), \quad (18.5)$$

Подставляя (18.5) в (18.4), получаем

$$\int_{(\lambda)} [\text{const} - \delta(\tilde{\lambda} - \lambda)] \cdot \lambda w(\lambda | z) d\lambda = \text{const} \cdot w(\lambda | z) \Big|_{\lambda=\tilde{\lambda}}.$$

Очевидно, что это выражение достигает минимума, если в качестве оценки $\tilde{\lambda}$ использовать значение параметра, доставляющее максимум апостериорной ПРВ $w(\lambda|z)$. Такая оценка называется МАВ-оценкой (оценкой максимума апостериорной вероятности).

Во многих задачах априорная ПРВ параметра неизвестна, тогда принимают ее равной константе и максимизируют функцию правдоподобия $w(\lambda|z)$. Получаемые таким образом оценки называются оценками максимального правдоподобия, или МП-оценками.

Пример. Наблюдается колебание $z(t) = \gamma s(t) + \zeta(t)$, где $s(t)$ – точно известный сигнал; γ – амплитудный множитель, подлежащий оцениванию; $\zeta(t)$ – гауссовский шум с нулевым средним и спектральной плотностью мощности $N_0/2$, постоянный в полосе частот $-F < f < F$ («квазибелый» шум). Найдем правило оценивания параметра γ , оптимальное по критерию максимального правдоподобия.

Возьмем n отсчетов наблюдаемого колебания на интервале наблюдения T с шагом $\Delta t = 1/2F = T/n$, при этом отсчеты шума являются некоррелированными. Совместная плотность распределения вероятности взятых отсчетов поэтому равна

$$w(z_1, \dots, z_n | \gamma) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^n (z_k - \gamma s_k)^2},$$

где $\sigma^2 = N_0 F = N_0 / (2\Delta t)$. Устремляя Δt к нулю ($n \rightarrow \infty$), запишем функцию правдоподобия

$$w(z | \gamma) = C \cdot \exp \left\{ -\frac{1}{N_0} \int_0^T [z(t) \cdot \gamma s(t)]^2 dt \right\},$$

где C – константа, несущественная для задачи оценивания.

Для нахождения правила оценивания следует продифференцировать функцию правдоподобия или, что проще, ее логарифм и приравнять результат нулю. Получаемое при этом уравнение правдоподобия $\frac{d[\ln w(z | \gamma)]}{d\gamma} = 0$ для данного случая имеет вид $\int_0^T [z(t) - \gamma s(t)] \cdot s(t) dt = 0$,

откуда

$$\gamma \int_0^T s^2(t) dt = \int_0^T z(t) \cdot s(t) dt.$$

Решением этого уравнения относительно γ является оценка $\tilde{\gamma}$, определяемая выражением

$$\tilde{\gamma} = \frac{1}{E} \int_0^T z(t) \cdot s(t) dt, \quad (18.6)$$

где $E = \int_0^T s^2(t) dt$ – энергия сигнала, известная по условию задачи.

Качество полученной МП-оценки можно оценить, подставив в (18.6) выражение для $z(t)$:

$$\tilde{\gamma} = \frac{\int_0^T [\gamma s(t) + \xi(t)] \cdot s(t) dt}{E} = \frac{\gamma E}{E} + \frac{\int_0^T \xi(t) \cdot s(t) dt}{E} = \gamma + \frac{1}{E} \int_0^T \xi(t) \cdot s(t) dt. \quad (18.7)$$

Второе слагаемое представляет собой ошибку оценивания, причем дисперсия интеграла равна $N_0 E/2$, поэтому дисперсия ошибки равна $N_0/2E$.

Таким образом, оценка тем точнее, чем больше энергия сигнала (для гармонического сигнала $s(t)$ увеличение энергии эквивалентно увеличению длительности интервала наблюдения) и чем меньше спектральная плотность мощности помехи. Из выражения (18.7) видно, что оценка несмещенная, так как $\xi(t)$ имеет нулевое математическое ожидание. Учитывая несмещенность и стремление дисперсии к нулю при увеличении интервала наблюдения, можно заключить, что оценка является состоятельной. Кроме того, можно показать, что оценка также эффективна.

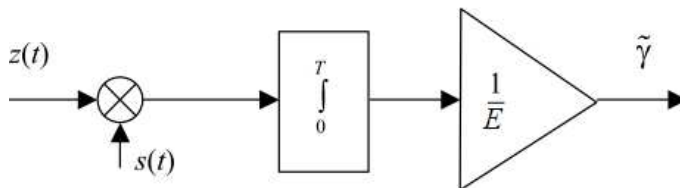


Рис. 18.3. Структура устройства оценивания амплитуды сигнала

18.2. Оптимальная фильтрация случайного сигнала

Более сложной и общей, чем задача оценивания постоянного параметра, является задача оценивания изменяющегося сообщения (первичного сигнала) на основе наблюдаемой реализации. Такое оценивание принято называть фильтрацией.

Сообщение рассматривается как реализация случайного процесса, множество всевозможных сообщений – как ансамбль реализаций с некоторым вероятностным распределением. Сообщение (первичный сигнал) модулирует несущее колебание, поэтому сигнал на выходе канала связи также случаен. Таким образом, ставится задача по наблюдаемому случайному колебанию оценить другое случайное колебание (первичный сигнал, или закон модуляции), связанное с наблюдаемым в общем случае нелинейным образом (задача нелинейной фильтрации, или демодуляции). Эта задача может быть весьма сложной.

В этом подразделе рассматривается наиболее простой случай оптимальной линейной фильтрации.

Пусть сигнал на входе линейного фильтра представляет собой сумму переданного сигнала $S(t)$ и помехи $\zeta(t)$

$$x(t) = S(t) + \zeta(t).$$

Требуется найти такую передаточную характеристику $K(\omega)$, которая минимизирует средний квадрат погрешности

$$\overline{\varepsilon^2(t)} = \overline{[S(t) - S_f(t)]^2}, \quad (18.8)$$

где $S_f(t)$ – оценка сигнала на выходе фильтра. Здесь, считаем, что время запаздывания сигнала $S(t)$ в фильтре $\tau_z = 0$, а среднее значение берется по ансамблям сигналов $S(t)$ и помех $\zeta(t)$. Считаем также, что $S(t)$ и $\zeta(t)$ стационарные взаимно-некоррелированные процессы с известными спектрами плотности мощности $G_s(\omega)$ и $G_\zeta(\omega)$. Для отклика сигнала на выходе фильтра запишем

$$y(t) = L[S(t) + \zeta(t)],$$

где $L[-]$ – оператор линейного преобразования.

Среднеквадратическую погрешность (18.8) можно представить двумя составляющими

$$\overline{\varepsilon^2(t)} = \overline{\varepsilon_s^2(t)} + \overline{\varepsilon_\zeta^2(t)},$$

где $\overline{\varepsilon_s^2(t)} = \overline{\{L[S(t) - S_x(t - \tau_z)]\}^2}$ – составляющая, обусловленная искажениями сигнала в фильтре (считаем ФЧХ линейной, т.е. $\varphi(\omega) = -\omega\tau_z$);

$\overline{\varepsilon_\zeta^2(t)} = \overline{\{L[\zeta(t)]\}^2}$ – составляющая, обусловленная действием шума.

Принимая во внимание, что

$$\varepsilon_s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(\omega) [1 - K(\omega)e^{-j\omega\tau_z}] \cdot e^{j\omega t} d\omega,$$

для спектральной плотности мощности $G_{\varepsilon_s}(\omega)$ имеем

$$G_{\varepsilon_s}(\omega) = G_s(\omega)[1-K(\omega)]^2.$$

В свою очередь, спектральная плотность мощности составляющей шума равна $G_{\varepsilon_\xi}(\omega) = G_\xi(\omega) \cdot K^2(\omega)$ а для $G_\varepsilon(\omega)$ суммарной погрешности получим

$$G_\varepsilon(\omega) = G_{\varepsilon_s}(\omega) + G_{\varepsilon_\xi}(\omega) = G_s(\omega)[1-K(\omega)]^2 + G_\xi(\omega)K^2(\omega). \quad (18.9)$$

Найдем теперь выражение $K(\omega)$ – при котором суммарная погрешность становится минимальной:

$$\frac{dG_\varepsilon(\omega)}{dK(\omega)} = -2G_s(\omega) \cdot [1-K(\omega)] + 2G_\xi(\omega) \cdot K(\omega) = 0,$$

откуда

$$K_{\text{оп}}(\omega) = \frac{G_s(\omega)}{G_s(\omega) + G_\xi(\omega)}, \quad (18.10)$$

или

$$\dot{K}_{\text{оп}}(\omega) = \frac{G_s(\omega)}{G_s(\omega) + G_\xi(\omega)} e^{-j\omega\tau_z}. \quad (18.11)$$

Фильтр (18.11), обеспечивающий минимальную среднеквадратичную погрешность случайного сигнала на фоне белого шума называется оптимальным фильтром Колмогорова-Винера. Следует отметить, что фильтр Колмогорова-Винера с коэффициентом передачи (18.11) физически не реализуем.

Конструктивные результаты в решении задачи синтеза фильтра, оптимального по критерию минимума СКП получил Р.Э. Калман, который в качестве априорных данных о сообщении использовал рекуррентную модель: состояние – наблюдение. Предполагается, что сообщение $a(t)$ порождается линейным стохастическим дифференциальным уравнением состояния

$$\frac{da}{dt} = -\alpha \cdot a(t) + n_1(t), \quad (18.12)$$

где $n_1(t)$ – нормальный белый шум (порождающий процесс). Физически это означает, что сообщение $a(t)$ моделируется как результат прохождения белого шума через интегрирующую цепь с постоянной времени $\alpha = 1/RC$.

Рассмотрим случай линейной фильтрации, когда наблюдаемый процесс на входе фильтра задан уравнением наблюдения

$$x(t) = S(t, a(t)) + \zeta(t) = a(t)f(t) + \zeta(t),$$

где сообщение $a(t)$ – задано уравнением (18.12); $S(t, a(t))$ – сигнал; $f(t)$ – носитель; $\zeta(t)$ – белый гауссовский шум.

Принимая во внимание основное условие линейной оптимальной фильтрации, состоящее в том, чтоб ошибка $\varepsilon(t) = S(t) - \hat{S}(t)$ должна быть не коррелирована со входным сигналом $x(t)$, рекуррентное соотношение, имеет вид:

$$\frac{da}{dt} = -\alpha \cdot a(t) + \frac{2k}{G_0} f(t)[x(t) - a(t) \cdot f(t)] \quad (18.13)$$

$$\frac{dk}{dt} = \frac{N_1}{2} - 2\alpha k - \frac{2}{N_0} f^2 k^2 \quad (18.14)$$

Уравнения (18.13) и (18.14) принято называть уравнениями фильтра Калмана. Уравнение (18.13) определяет алгоритм формирования оценки, а, следовательно, и структурную схему фильтра, а (18.14) – ошибку фильтрации $k = \varepsilon^2(t)$.

Структурная схема, моделирующая уравнение (18.13) приведена на рисунке 18.4. Этот фильтр представляет собой динамическую систему с переменным коэффициентом усиления, величина которого зависит от точности текущих оценок (дисперсии погрешности фильтрации) и уровня шумов наблюдения. Иначе говоря, фильтр Калмана можно рассматривать как линейную следящую систему, эффективная полоса пропускания которой, изменяется в процессе работы.

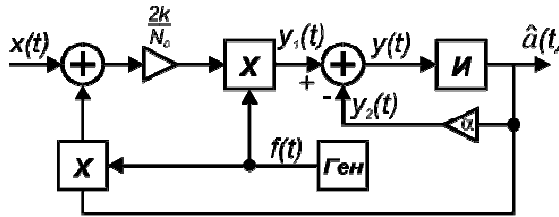


Рис. 18.4. Структурная схема, моделирующая уравнение 18.13

18.3. Потенциальная помехоустойчивость передачи непрерывных сообщений

Пусть передается непрерывное сообщение $a(t)$, причём $|a(t)| \leq 1$, а его мощность $P_a = \overline{a^2(t)} = 1$; в общем случае $a(t)$ можно представить в

виде разложения $a(t) = \sum_1^n \lambda_i \cdot \psi_i(t)$ по единичным ортогональным со-

ставляющим с коэффициентами λ_i . Тогда сигнал $S(t; a) = S(t; \lambda_1, \lambda_2, \dots, \lambda_n)$ будет получать в процессе модуляции приращения в соответствии с модуляционным вектором определяющим $D_i = \frac{dS(t)}{d\lambda_i}$ свойства модуля-

тора.

С другой стороны, принятый сигнал $S^*(t)$ за счет действующей в канале помехи $\zeta(t)$ будет отличаться от переданного $S(t)$. Как следствие этого, коэффициенты $\lambda_i = \lambda_i + \Delta \lambda_i$ также будут отличаться от переданных λ_i . В результате будем иметь

$$a^*(t) = \sum \lambda_i^* \psi_i(t) = \sum (\lambda_i + \Delta \lambda_i^*) \psi_i(t) = a(t) + \sum \Delta \lambda_i^* \psi_i(t), \quad (18.15)$$

а для погрешности передачи получим $\varepsilon(t) = a^*(t) - a(t) = \sum \Delta \lambda_i^* \psi_i(t)$.

Оптимальный приемник Котельникова измеряет расстояние $R = [S^*(t) - S(t)]^2$.

Минимальному значению R соответствуют приращения $\Delta \lambda_i^*$, определяемые из условия $\frac{dR}{d\lambda_i} = -2[S^*(t) - S(t)] \frac{dS}{d\lambda_i} = 0$, т.е.

$$\Delta \lambda_i^* = \frac{\overline{\zeta(t) \cdot D_i(t)}}{D_i^2(t)}. \quad (18.16)$$

На основе (18.16) получаем выражение для средней мощности шума на выходе приемника $P_\varepsilon = \overline{(\Delta \lambda_i^*)^2}$. При слабых помехах $\zeta(t)$ спектральную мощность шума на выходе можно представить в виде

$$G_\varepsilon(\omega) = \frac{G_\zeta(\omega)}{D_i^2(t)}, \quad (18.17)$$

а P_ε — соответственно

$$P_\varepsilon = \int_0^{\omega_b} G_\varepsilon(\omega) d\omega. \quad (18.18)$$

Составим отношение средних мощностей сообщения и шума на выходе приемника

$$h_2^2 = \frac{P_a}{P_\varepsilon} = \frac{1}{\int_0^{\omega_b} G_\varepsilon(\omega) d\omega}. \quad (18.19)$$

Выражения (18.18) и (18.19) используются для сравнительной оценки потенциальной помехоустойчивости различных видов модуляции. В результате анализа получены следующие выражения $G_\epsilon(\omega)$ и h_2^2 .

Амплитудная модуляция $S(t) = A[1 + ma(t)]\cos \omega_0 t$

$$G_\epsilon(\omega) = \frac{2G_0}{m^2 A^2}; \quad h_2^2 = \frac{m^2 A^2}{2G_0 \omega_b}. \quad (18.20)$$

Фазовая модуляция $S(t) = A\cos[\omega_0 t + \Delta\varphi_m a(t)]$

$$G_\epsilon(\omega) = \frac{2G_0}{\Delta\varphi_m^2 A^2}; \quad h_2^2 = \frac{\Delta\varphi_m^2 A^2}{2G_0 \omega_b}. \quad (18.21)$$

Частотная модуляция

$$S(t) = A\cos\left[\omega_0 t + \Delta\omega_m \int_0^t a(t)dt\right]; \quad a(t) = \cos\Omega t; \quad \beta_m = \frac{\Delta\omega_m}{\nu}$$

$$G_\epsilon(\omega) = \frac{2G_0}{\Delta\omega_m^2 A^2}; \quad h_2^2 = \frac{\Delta\beta_m^2 A^2}{2G_0 \omega_b}. \quad (18.22)$$

Из (18.20 – 18.22) следует, что при АМ ($m = 1$) помехоустойчивость можно увеличить лишь за счет увеличения мощности сигнала; при ФМ и ЧМ помехоустойчивость можно увеличить также путем увеличения индексов модуляции $\Delta\varphi_m$ и β_m , что достигается за счет расширения спектра сигнала.

Все широкополосные системы модуляции обеспечивают высокую помехоустойчивость h_2^2 при условии, что отношение сигнал-помеха на входе h_1^2 больше некоторого порогового значения (рис. 18.5). При $h_1^2 < h_{1\text{пор}}^2$ широкополосные системы теряют свои преимущества (резко снижается помехоустойчивость); возникающее при этом явление называют явлением порога помехоустойчивости.

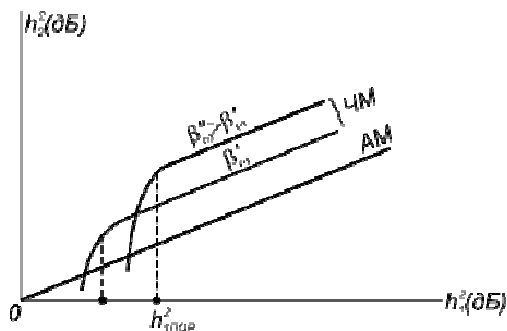


Рис. 18.5. Графики $G_\epsilon(\omega)$ для АМ, ФМ и ЧМ в соответствии с 18.20 – 18.22

19. АДАПТИВНЫЕ УСТРОЙСТВА ПОДАВЛЕНИЯ ПОМЕХ

19.1. Основы адаптивного подавления помех

Основополагающая схема подавления помех показана на рис. 19.1. Сигнал передается по каналу на приемное устройство, которое принимает смесь сигнала и не коррелированной с ним помехи p_0 . Смесь сигнала и помехи $s + p_0$ является входным сигналом устройства подавления. Другое приемное устройство принимает помеху p_1 , не коррелированную с сигналом, но некоторым неизвестным образом коррелированную с помехой p_0 . В нем формируется «эталонный сигнал» для устройства подавления. В результате фильтрации помехи p_1 , формируется сигнал y , который приблизительно представляет собой копию p_0 . Этот сигнал вычитается из входного сигнала $s + p_0$, для того, чтобы сформировать выходной сигнал системы $s + p_0 - y$.

В общем случае при известных характеристиках каналов, по которым помеха поступает на оба приемных устройства, можно синтезировать фильтр с постоянными параметрами, преобразующих p_1 в $y = p_0$. Тогда выходной сигнал фильтра можно вычесть из входного сигнала, и на выходе системы останется один сигнал.

Однако применение фильтра с постоянными параметрами не обеспечивает гибкости, поскольку считают, что характеристики трактов передачи или неизвестны, или известны только приблизительно и могут изменяться. Более того, даже если фильтр с постоянными параметрами обеспечивает гибкость, его параметры нужно настраивать с точностью, которая трудно реализуется, и малейшая ошибка может привести к увеличению на выходе системы мощности помехи.

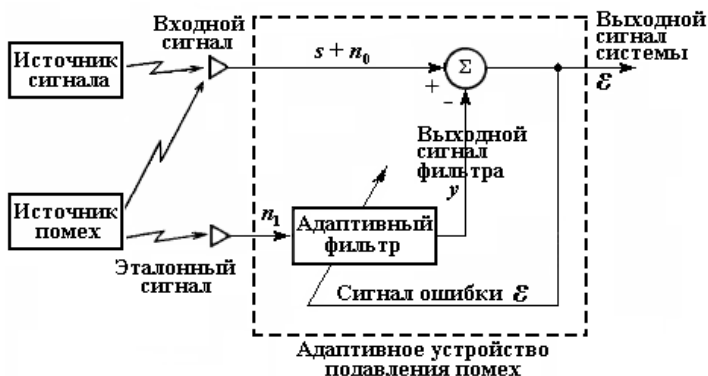


Рис. 19.1. Иллюстрация процесса адаптивного подавления помех

В приведенной на рис. 19.1 системе эталонный сигнал обрабатывается адаптивным фильтром, который автоматически перестраивает свою собственную импульсную характеристику по одному из среднеквадратических алгоритмов (например, по методу наименьших квадратов), функционирующему по сигналу ошибки, зависящему помимо всего прочего, от выходного сигнала фильтра. Таким образом, при правильном алгоритме фильтр может работать в изменяющихся условиях и перестраиваться для минимизации сигнала ошибки.

Выбор используемого для процесса адаптации сигнала ошибки зависит от конкретной области приложения. Практическое назначение систем подавления помех – формирование выходного сигнала системы $s + n_0 - y$ – у который имеет наилучшее в среднеквадратическом смысле приближение к сигналу s . Это достигается тем, что выходной сигнал системы подается на адаптивный фильтр, который перестраивается по некоторому адаптивному алгоритму так, чтобы минимизировать общую мощность выходного сигнала системы. Другими словами, в системе с адаптивным подавлением помех сигналом ошибки адаптивного процесса является выходной сигнал системы.

Можно предположить, что для синтеза фильтра или его адаптации с целью формирования сигнала y , компенсирующего помеху, необходимы некоторые априорные сведения о сигнале s или помехах n_0 и n_1 . Однако как будет показано далее, этого не требуется или почти не требуется.

Пусть s , n_0 , n_1 и y – стационарные случайные процессы с нулевыми средними значениями, и s не коррелирован с n_0 и n_1 , а n_0 и n_1 – коррелированы. Выходной сигнал

$$\varepsilon = s + n_0 - y. \quad (19.1)$$

Возведем обе части равенства в квадрат:

$$\varepsilon^2 = s^2 + (n_0 - y)^2 + 2s(n_0 - y). \quad (19.2)$$

Для обеих частей (19.2) найдем математическое ожидание и, поскольку s не коррелирован с n_0 и y , получим

$$E[\varepsilon^2] = E[s^2] + E[(n_0 - y)^2] + 2E[s(n_0 - y)] = E[s^2] + E[(n_0 - y)^2]. \quad (19.3)$$

Мощность сигнала $E[s^2]$ не изменяется при перестройке фильтра в процессе минимизации $E[\varepsilon^2]$. В соответствии с этим минимальная мощность выходного сигнала

$$E_{\min}[\varepsilon^2] = E[s^2] + E_{\min}[(n_0 - y)^2]. \quad (19.4)$$

Если фильтр построен так, что $E[(n_0 - y)^2]$ минимально, то, следовательно, минимально также и $E[\varepsilon^2]$. В этом случае выходной сигнал фильтра y является наилучшей среднеквадратической оценкой помехи

n_0 . Более того, при минимальном значении $E[(n_0 - y)^2]$ минимальное значение имеет также и $E[(\varepsilon - s)^2]$, поскольку из (19.1)

$$(\varepsilon - s) = (n_0 - y). \quad (19.5)$$

Таким образом, перестройка или адаптация фильтра для минимизации общей мощности выходного сигнала равносильна тому, что при заданных структуре адаптивного фильтра и эталонном входном сигнале выходной сигнал ε изменяется так, что он является наилучшим в среднеквадратическом смысле приближением сигнала s .

В общем случае выходной сигнал ε равен сумме сигнала s и некоторой помехи. В соответствии с (19.1) помеха на выходе равна $p_0 - y$. Поскольку при минимизации $E[\varepsilon^2]$ осуществляется минимизация $E[(n_0 - y)^2]$, то минимизация общей мощности выходного сигнала приводит к минимизации мощности помехи на выходе и, так как сигнал на выходе остается постоянным, к максимизации выходного отношения сигнал-помеха.

Из (19.3) и (19.4) видно, что наименьшая возможная мощность выходного сигнала $E_{\min}[\varepsilon^2] = E[s^2]$. Если это достижимо, то $E[(n_0 - y)^2] = 0$. Следовательно, $y = p_0$ и $\varepsilon = s$, и минимизация мощности выходного сигнала приводит к тому, что сигнал совершенно не искажен помехой.

Кроме того, если эталонный сигнал совсем не коррелирован с входным сигналом, то фильтр отключается и не увеличивает помеху на выходе. В этом случае выходной сигнал фильтра y не коррелирован с входным сигналом и мощность выходного сигнала

$$E[\varepsilon^2] = E[(s + n_0)^2] + 2E[-y(s + n_0)] + E[y^2] = -E[(s + n_0)^2] + E[y^2]. \quad (19.6)$$

Для минимизации мощности выходного сигнала необходимо, чтобы было минимальным $E[y^2]$, что достигается при равенстве всех весовых коэффициентов нулю, приводящим к тому, что $E[y^2] = 0$.

Эти рассуждения легко распространить на случай, когда входной и эталонный сигналы содержат, помимо n_0 и p_1 , составляющие аддитивного случайного шума, не коррелированные между собой и с s , n_0 и p_1 . Кроме того, эти рассуждения справедливы для случая, когда n_0 и p_1 — детерминированные, а не случайные сигналы. Отметим, что в процессе рассуждений не оговорено, что адаптивный фильтр обязательно сходится к линейному фильтру.

19.2. Подавление стационарных помех

Цель данного раздела — показать возможность повышения отношения сигнал-помеха, а также другие преимущества методов подавления помех при использовании адаптивных фильтров по сравнению с обычными методами фильтрации помех.

Как отмечалось раньше, при подавлении помех фильтры с постоянными параметрами в большинстве случаев неприменимы, так как в общем случае автокорреляционная и взаимокорреляционная функции входного и эталонного сигналов неизвестны и часто являются изменяющимися во времени. Поэтому необходимо сначала обучать адаптивные фильтры по обучающим статистикам, а затем осуществлять слежение за ними, если они медленно меняются. Однако для стационарных входных сигналов установившийся режим работы медленно адаптирующихся фильтров приближается к режиму работы винеровских фильтров, поэтому удобным математическим аппаратом для анализа статистических задач подавления помех является винеровская теория фильтрации.

На рис. 19.2 представлена несколько более подробная схема системы, показанной на рис. 19.1. Здесь показан один из способов получения входного сигнала и помех на входах системы. Входной сигнал представляет собой сумму сигнала s_k и двух помех – n_k и m_{0k} , а эталонный сигнал – сумму двух других помех m_{1k} и n_k , пришедший через тракт с импульсной характеристикой h_k и передаточной функцией $H(z)$. Обе помехи – n_k на входе системы и n_k , прошедшая через тракт с h_k имеют общий источник и являются коррелированными между собой и не коррелированными с сигналом s_k . Предполагается также, что их энергетические спектры ограничены на всех частотах. Помехи m_{0k} и m_{1k} некоррелированы между собой, с s_k , n_k на входе системы и n_k , прошедшей через тракт, с h_k . Для дальнейшего анализа будем считать, что все тракты распространения помех эквивалентны линейным фильтрам с постоянными параметрами.

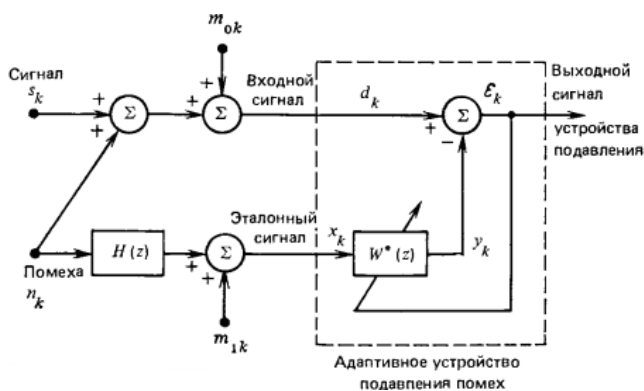


Рис. 19.2. Одноканальное адаптивное устройство подавления помех с коррелированной и некоррелированной помехами на входе устройства и на эталонном входе

Схема подавления помех на рис. 19.2. включает в себя адаптивный фильтр, входной сигнал x_k (эталонный сигнал устройства подавления) которого равен сумме m_{1k} и n_k , прошедший через тракт с h_k , а полезный отклик (входной сигнал устройства подавления) $d_k = s_k + m_{0k} + n_k$. Сигналом ошибки ε_k является выходной сигнал устройства подавления. Если считать, что адаптивный процесс завершен и найдены оптимальные в смысле минимума СКО весовые коэффициенты (оптимальное решение), то адаптивный фильтр эквивалентен винеровскому фильтру с передаточной функцией

$$W^*(z) = \Phi_{xd}(z)/\Phi_{xx}(z), \quad (19.7)$$

где $\Phi_{xx}(z)$ – спектр входного сигнала фильтра;

$\Phi_{xd}(z)$ – взаимный энергетический спектр входного сигнала фильтра и полезного отклика.

Оптимальное решение для устройства подавления на рис. 19.2 желательнее иметь по следующей причине. Выходной сигнал – это сигнал ошибки винеровского фильтра, т.е. сигнал ошибки ε_k не коррелирован с входным сигналом фильтра s_k . Следовательно, на входе системы полностью подавляются все составляющие помехи, коррелированные с составляющими помехи на эталонном входе. Однако другие составляющие не подавляются и оказываются на выходе системы.

Оптимальная передаточная функция адаптивного фильтра $W^*(z)$ есть винеровское решение (19.7), для которого теперь можно получить следующее обобщение. Спектр входного сигнала фильтра $\Phi_{xx}(z)$ можно выразить через спектры его двух некоррелированных составляющих. Спектр помехи $m_1 - \Phi_{m_1m_1}(z)$, а спектр помехи n , прошедшей через тракт с $H(z)$, – $\Phi_{nn}(z) |H(z)|^2$. Отсюда спектр входного сигнала фильтра

$$\Phi_{xx}(z) = \Phi_{m_1m_1}(z) + \Phi_{nn}(z) |H(z)|^2. \quad (19.8)$$

Взаимный энергетический спектр входного сигнала фильтра и полезного отклика зависит только от коррелированных составляющих входного и эталонного сигналов и имеет вид

$$\Phi_{xd}(z) = \Phi_{m_1}(z) H(z^{-1}), \quad (19.9)$$

Подставляя (19.9) в (19.7), получаем винеровскую передаточную функцию

$$W^*(z) = \frac{\Phi_{m_1}(z) H(z^{-1})}{\Phi_{m_1m_1}(z) + \Phi_{nn}(z) |H(z)|^2}. \quad (19.10)$$

Отметим, что $W^*(z)$ не зависит от спектра сигнала $\Phi_{ss}(z)$ и спектра некоррелированного шума $\Phi_{m_0m_0}(z)$ на входе системы.

При нулевой аддитивной помехе m_1 на эталонном входе имеем интересный частный случай, когда $\Phi_{m_1 m_1}(z) = 0$, а оптимальная передаточная функция (19.10) принимает вид

$$W(z)^* = 1/H(z). \quad (19.11)$$

Интуитивно ясно, что этот результат является правильным, так как адаптивный фильтр приводит, как и при балансе мостовой схемы, к обнулению n_k на выходе устройства подавления. Некоррелированная помеха m_{0k} не подавляется и наряду с сигналом появляется на выходе системы.

Функционирование устройства подавления помех с одним входом можно рассмотреть в более общем виде с точки зрения соотношения двух отношений – плотности мощности сигнала к плотности мощности помех на выходе системы $\rho_{\text{вых}}(z)$ и на ее входе $\rho_{\text{вх}}(z)$. Полагая, что спектр сигнала является неотрицательным на всех частотах, после соответствующих преобразований имеем

$$\frac{\rho_{\text{вых}}(z)}{\rho_{\text{вх}}(z)} = \frac{\text{плотность мощности помехи на входе}}{\text{плотность мощности помехи на выходе}} = \frac{\Phi_{nn}(z) + \Phi_{m_0 m_0}(z)}{\Phi_{\text{вых шум}}}. \quad (19.12)$$

Из рис. 19.2 следует, что энергетический спектр помехи на выходе устройства подавления равен сумме трех составляющих, одна из которых соответствует прямому прохождению на выход m_{0k} , другая – m_{1k} , прохождению через звено с передаточной функцией $W^*(z)$, а третья – n_k , прохождению через звено с передаточной функцией $1-H(z)$ $W^*(z)$. Отсюда энергетический спектр помехи на выходе

$$\Omega_{\text{вых шум}}(z) = \Phi_{m_0 m_0}(z) + \Phi_{m_1 m_1}(z) \cdot |W^*(z)|^2 + \Phi_{nn}(z) \cdot |1-H(z) \cdot W^*(z)|^2. \quad (19.13)$$

Для удобства обозначим отношения спектров коррелированных и некоррелированных помех на входе и эталонном входе соответственно через

$$A(z) \equiv \frac{\Phi_{m_0 m_0}(z)}{\Phi_{nn}(z)} \quad (19.14)$$

и

$$B(z) \equiv \frac{\Phi_{m_1 m_1}(z)}{\Phi_{nn}(z) |H(z)|^2} \quad (19.15)$$

В этих обозначениях передаточную функцию можно переписать в виде

$$W^*(z) = 1/H(z)[B(z) + 1] \quad (19.16)$$

Тогда энергетический спектр помехи на выходе

$$\begin{aligned} \Phi_{\text{вых шум}}(z) &= \Phi_{m_0 m_0}(z) + \frac{\Phi_{m_1 m_1}(z)}{|H(z)|^2 \cdot |B(z) + 1|^2} + \Phi_{nn}(z) \cdot \left| 1 - \frac{1}{B(z) + 1} \right|^2 = \\ &= \Phi_{nn}(z) \cdot A(z) + \Phi_{nn}(z) \cdot \frac{B(z)}{B(z) + 1} \end{aligned} \quad (19.17)$$

При этом отношении (19.12)

$$\frac{\rho_{\text{вых}}(z)}{\rho_{\text{вх}}(z)} = \frac{\Phi_{\text{пт}}(z) \cdot [1 + A(z)]}{\Phi_{\text{вых шум}}(z)} = \frac{1 + A(z)}{A(z) + \frac{B(z)}{B(z) + 1}} = \frac{[A(z) + 1] \cdot [B(z) + 1]}{A(z) + A(z) \cdot B(z) + B(z)} \quad (19.18)$$

Это выражение является общим представлением идеальной характеристики подавления помех в рассматриваемом случае. Оно позволяет оценить ожидаемый уровень подавления помехи при идеальной системе подавления, содержащей винеровский фильтр с двусторонней импульсной характеристикой. В такой системе сигнал проходит на выход неискаженным. В противоположность этому классические схемы Винера, Калмана и адаптивных фильтров в процессе подавления помех вносят некоторые искажения и в сигнал.

Из (19.18) видно, что возможности подавления помех ограничиваются отношениями $A(z)$ и $B(z)$. Если они относительно малы, то отношение $\rho_{\text{вых}}(z)/\rho_{\text{вх}}(z)$ велико и относительно более эффективно функционирует устройство подавления. Необходимость иметь низкий уровень некоррелированных помех на обоих входах становится ещё более очевидной при рассмотрении следующих конкретных случаев:

1) При малом $A(z)$

$$\frac{\rho_{\text{вых}}(z)}{\rho_{\text{вх}}(z)} \approx \frac{1 + B(z)}{B(z)}$$

2) При малом $B(z)$

$$\frac{\rho_{\text{вых}}(z)}{\rho_{\text{вх}}(z)} \approx \frac{1 + A(z)}{A(z)}$$

3) При малых $A(z)$ и $B(z)$

$$\frac{\rho_{\text{вых}}(z)}{\rho_{\text{вх}}(z)} \approx \frac{1}{A(z) + B(z)}$$

Из этих соотношений следует, что полное подавление возможно, когда $A(z)$ и $B(z)$ равны нулю. В этом случае на выходе системы можно полностью исключить помеху, что приводит к идеальному восстановлению сигнала. Однако при малых $A(z)$ и $B(z)$ вступают в силу другие факторы, ограничивающие характеристики системы. К этим факторам относятся конечная длина адаптивного фильтра в реальных системах и относительное среднее значение СКО, вызванное шумом оценки градиента, возникающим в процессе адаптации [14]. Влияние третьего фактора, связанного с попаданием на эталонный вход составляющих сигнала, рассматривается далее.

19.3. Адаптивный режекторный фильтр

В некоторых случаях входной сигнал представляет собой сумму составляющей сигнала и аддитивной синусоидальной помехи. Обычно для подавления такой помехи используется режекторный фильтр. В этом подразделе рассматривается реализация режекторного фильтра с помощью адаптивного устройства подавления помех. Преимущества такого режекторного фильтра заключаются в том, что он позволяет регулировать полосу частот и осуществлять адаптивное слежение за точным значением частоты и фазы помехи. Эти результаты распространяются и на случай, когда на эталонном входе имеется сигнал на многих частотах [14].

На рис. 19.3. приведена схема устройства подавления одночастотной помехи с двумя адаптивными весовыми коэффициентами. На вход устройства может подаваться сигнал любого вида – случайный, детерминированный, периодический, импульсный и т.д. – или любая комбинация этих сигналов. На эталонном входе действует чистый синусоидальный сигнал $C \cdot \cos(\Omega_0 t + \varphi)$. Отсчеты входных сигналов берутся с интервалами T секунд. Здесь x_{1k} – отсчеты эталонного сигнала, а x_{2k} – отсчеты этого сигнала, сдвинутого по фазе на 90° .

Преимущества такого режекторного фильтра заключаются в том, что он позволяет регулировать полосу частот, формировать нули и осуществлять адаптивное слежение за точным значением частоты и фазы помехи.

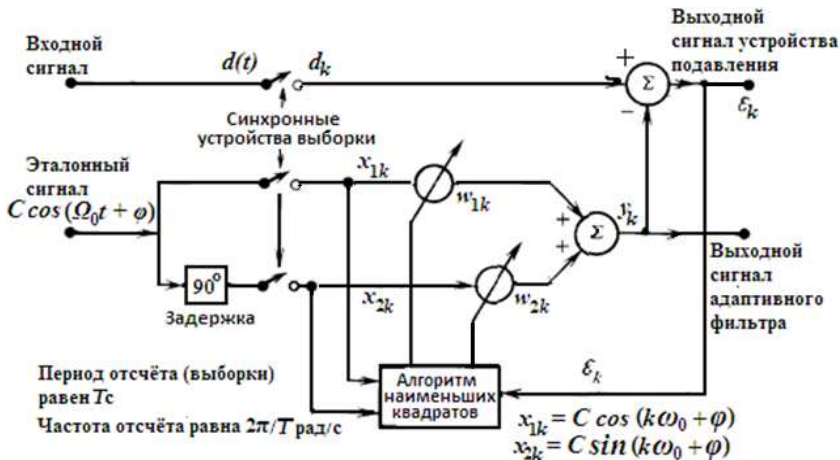


Рис. 19.3. Одночастотный адаптивный режекторный фильтр

Рассматривая прохождение сигнала от входа до выхода системы на рисунке 19.3. можно найти линейную передаточную функцию устройства подавления помех. При замкнутой петле обратной связи передаточная функция устройства подавления помех

$$H(z) = \frac{z^2 - 2z \cdot \cos \omega_0 + 1}{z^2 - 2(1 - \mu C^2)z \cdot \cos \omega_0 + 1 - 2\mu C^2}. \quad (19.19)$$

Из равенства (19.19) следует, что на частоте ω_0 эталонного сигнала устройство подавления одночастотной помехи обладает свойствами режекторного фильтра. Нули передаточной функции расположены на z -плоскости в точках

$$z = e^{\pm j\omega_0}. \quad (19.20)$$

т.е. точно на окружности единичного радиуса под углами $\pm\omega_0$ радиан.

Полуса передаточной функции расположены в точках

$$z = (1 - \mu C^2) \cdot \cos \omega_0 \pm j[(1 - 2\mu C^2) - (1 - 2\mu C^2)^2 \cdot \cos^2 \omega_0]^{1/2}, \quad (19.21)$$

т.е. внутри окружности единичного радиуса на расстоянии от начала координат $(1 - 2\mu C^2)^{1/2}$, приближенно равном $1 - 2\mu C^2$, и под углами

$$\pm \cos^{-1}[(1 - \mu C^2) (1 - 2\mu C^2)^{-1/2} \cos \omega_0]. \quad (19.22)$$

При медленной адаптации (т. е. при небольших значениях μC^2) эти углы определяются множителем, которым приблизительно равен 1.

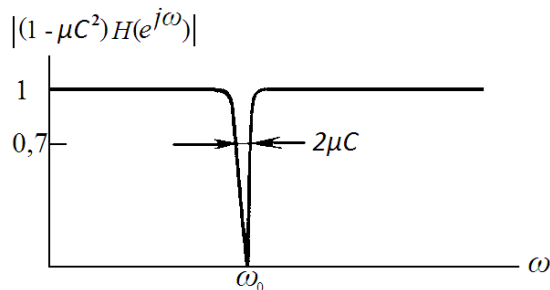


Рис. 19.4. Передаточная функция одночастотного адаптивного устройства подавления помех

Основной вывод состоит в том, что в практических случаях *углы полюсов и нулей почти равны*.

На рис. 19.4 показано расположение полюсов, нулей и точек половинной мощности передаточной функции. Поскольку нули лежат на окружности единичного радиуса, глубина режекции в децибелах для передаточной функции на частоте ω_0 равна бесконечности.

Форма провала АЧХ определяется расстоянием, которое приблизительно равно μC^2 . Длина дуги окружности единичного радиуса, заключенной между точками половинной мощности, соответствует полосе режекции фильтра и равна

$$BW = 2\mu C^2 \text{ рад} = \mu C^2 / \pi T \text{ Гц.} \quad (19.23)$$

Форма АЧХ в полосе режекции обычно определяется добротностью Q , представляющей собой отношение центральной частоты к ширине полосы режекции:

$$Q \equiv \frac{\text{центральная частота}}{\text{ширина полосы режекции}} = \frac{\omega_0}{2\mu C^2}. \quad (19.24)$$

Таким образом, устройство подавления одночастотной помехи при синусоидальном эталонном сигнале эквивалентно устойчивому режекторному фильтру. В общем случае глубина режекции адаптивного устройства выше, поскольку в результате адаптивного процесса даже при медленном изменении частоты эталонного сигнала поддерживается правильное для подавления соотношение фаз.

19.4. Адаптивный высокочастотный фильтр

Использование в адаптивном фильтре весового коэффициента для подавления низкочастотного дрейфа на входе системы является частным случаем режекции сигнала на нулевой частоте. На рис. 19.5 приведена схема только с одним весовым коэффициентом, так как не нужна подстройка фазы сигнала. Эталонным сигналом может быть постоянный сигнал единичной амплитуды.

Передающую функцию тракта прохождения сигнала от входа до выхода устройства подавления можно получить следующим образом. Из рисунка 19.5 видно, что $y_k = w_k$, поэтому для алгоритма наименьших квадратов в этом случае

$$y_{k+1} = y_k + 2\mu \varepsilon_k = y_k + 2\mu(d_k - y_k). \quad (19.25)$$

Находим z-преобразование от (19.25) и получаем для установившегося режима

$$Y(z) = \frac{2\mu \cdot D(z)}{z - (1 - 2\mu)} \quad (19.26)$$

Подставим теперь в (19.26) $Y(z) = D(z) - E(z)$ и получим

$$H(z) = \frac{E(z)}{D(z)} = \frac{z - 1}{z - (1 - 2\mu)}. \quad (19.27)$$

Из равенства (19.27) следует, что фильтр с весовым коэффициентом смещения представляет собой высокочастотный фильтр с нулем на окружности единичного радиуса на нулевой частоте и полюсом на действительной оси на расстоянии 2μ влево от нуля. Отметим, что это соответствует одночастотному режекторному фильтру (19.19) с $\omega_0 = 0$ и $C = 1$. Частота, на которой мощность режектурируемого сигнала уменьшается вдвое, равна $\omega = 2\mu$ радиан.

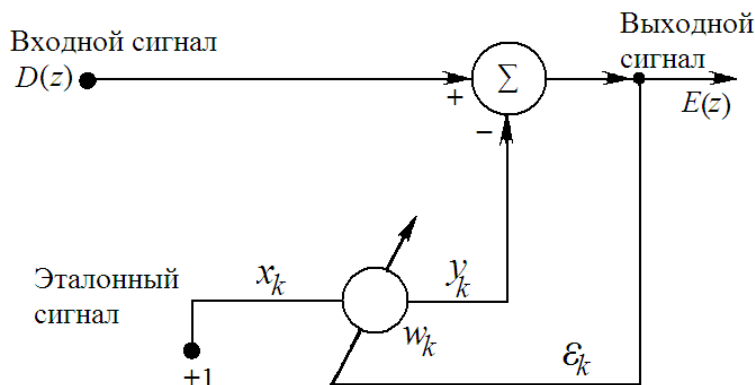


Рис. 19.5. Схема адаптивного высокочастотного фильтра, полученного путем установления на эталонном входе постоянного единичного сигнала

В адаптивном высокочастотном фильтре исключается не только постоянное смещение, но и медленно меняющийся дрейф во входном сигнале. Более того, хотя это здесь и не показано, в этом случае одновременно осуществляется подавление низкочастотной случайной помехи при постоянном эталонном сигнале.

19.5. Подавление периодической помехи с помощью адаптивного устройства предсказания

В некоторых случаях широкополосный сигнал искажается периодической помехой, при этом не имеется внешнего эталонного сигнала, не содержащего сам сигнал. В качестве примера можно назвать воспроизведение речи или музыки при наличии шумов магнитофонной пленки или диска, а также прием сейсмических сигналов в шумах автомобильного двигателя для силовых установок.

На первый взгляд может показаться, что для уменьшения или исключения такого вида помех нельзя применять адаптивное подавление. Однако при получении эталонного сигнала непосредственно из входного через заданную задержку Δ , как показано на рис. 19.6, во многих

случаях можно легко подавить периодическую помеху. Выбранная задержка должна быть достаточной для того, чтобы составляющие широкополосного сигнала на эталонном входе были не коррелированы с этим сигналом на входе. Из-за периодичности составляющие помехи остаются при этом коррелированными.

Из рисунка 19.6. следует, что система содержит адаптивное устройство предсказания. Предсказуемая составляющая входного сигнала исключается, и на выходе системы остается непредсказуемая составляющая.

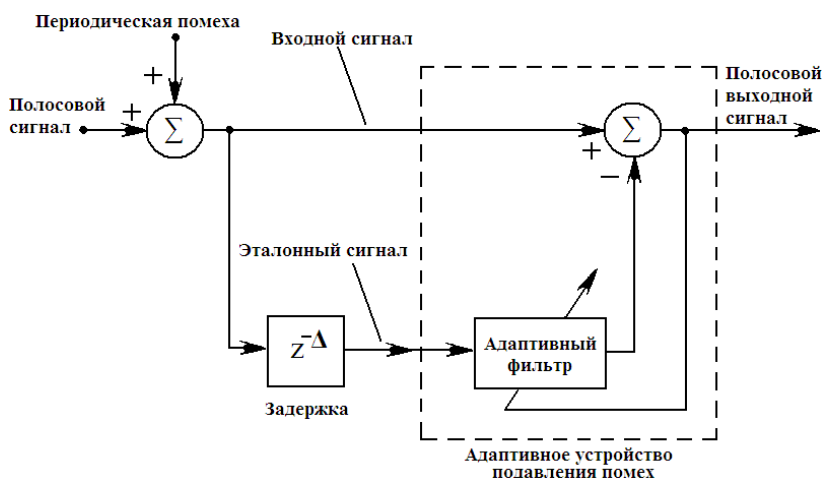


Рис. 19.6. Подавление периодической помехи в системе без внешнего эталонного сигнала

На рисунке 19.7. приведены результаты моделирования на ЭВМ входного (а) и выходного (б) сигналов устройства подавления.

Входной сигнал представляет собой сумму сигнала в виде небелого гауссовского шума и помехи в виде синусоидального сигнала. Поскольку задача решалась методом моделирования и входной широкополосный сигнал известен точно, он показан наряду с выходным сигналом на графике, из которого видно хорошее совпадение этих сигналов. Это совпадение не является полным из-за того, что фильтр имеет ограниченную длину и скорость адаптации.



Рис. 19.7. Результат эксперимента по подавлению периодической помехи

19.6. Адаптивный следящий фильтр

Описанный в предыдущем подразделе эксперимент можно, кроме того, использовать для иллюстрации еще одного важного приложения адаптивного устройства подавления помех. Во многих случаях, когда имеется входной сигнал, состоящий из смеси периодического и полосового сигналов, представляют интерес периодические, а не полосовые составляющие. Если выходной сигнал системы на рис. 19.6. снимается с

адаптивного фильтра, то в результате имеем адаптивный следящий фильтр, который может выделять из широкополосного шума периодический сигнал.

На рис. 19.8. показана схема адаптивного устройства подавления в качестве следящего фильтра.

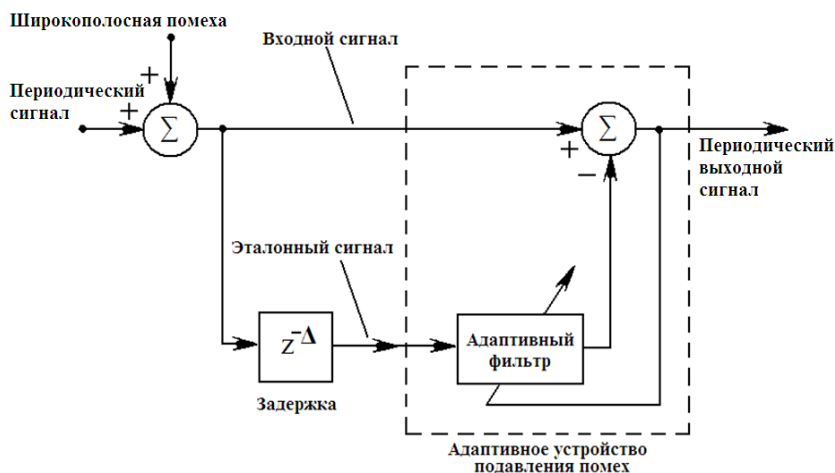


Рис. 19.8. Схема адаптивного устройства подавления помех в качестве следящего фильтра

Выходной сигнал этой системы получен с помощью моделирования на ЭВМ для входного сигнала из предыдущего примера (рис. 19.7а). Действительный входной синусоидальный сигнал и полученное в результате адаптации его приближение показаны на рис. 19.9. Здесь сигнал ошибки представляет собой случайный процесс с малой амплитудой.

На рис. 19.10 приведены импульсная характеристика и передаточная функция адаптивного фильтра после сходимости.

Импульсная характеристика (рис. 19.10а) близка к синусоидальному сигналу. При полосовой входной составляющей оптимальное устройство приема является согласованным фильтром, а его импульсная характеристика имеет синусоидальную форму.

Приведенная на рис. 19.10б передаточная функция равна дискретному преобразованию Фурье импульсной характеристики. На частоте помехи ее амплитуда близка к единице – значению, требуемому для полного подавления. Сдвиг фазы на этой частоте не равен нулю, но при сложении с фазовым сдвигом, вносимым задержкой Δ , приводит к суммарному сдвигу на 360° .

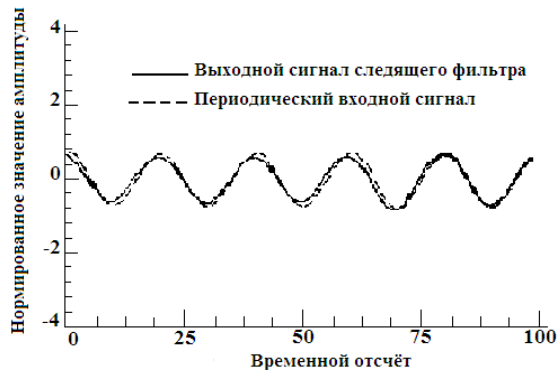


Рис. 19.9. Входные сигналы системы подавления помех в качестве следящего фильтра

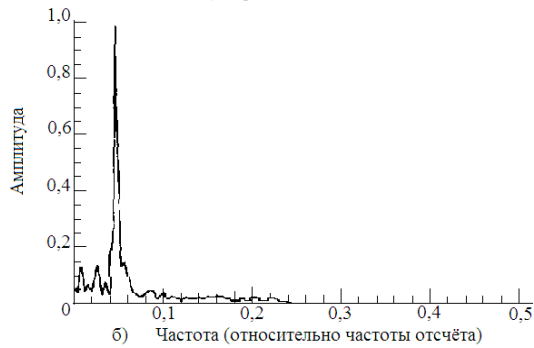
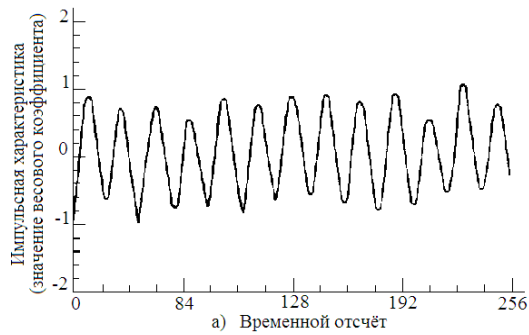


Рис. 19.10. Экспериментальные характеристики адаптивного следящего фильтра:
 а) импульсная характеристика адаптивного фильтра после адаптации;
 б) АЧХ адаптивного фильтра после адаптации

Аналогичные эксперименты проведены для сумм синусоидальных сигналов и широкополосной случайной помехи. В этих экспериментах

адаптивный фильтр имеет резонансы на всех частотах, на которых во входном сигнале расположены периодические спектральные составляющие. Таким образом, эту систему можно использовать в качестве автоматического устройства поиска.

19.7. Адаптивный накопитель

Адаптивный следящий фильтр можно использовать в качестве накопителя при обнаружении в шуме синусоидальных сигналов с очень низким уровнем.

Адаптивный следящий фильтр, который, как показано выше, обладает свойством разделять в сигнале периодические и случайные составляющие (когда их уровни сравнимы), при обнаружении в шуме синусоидального сигнала с очень низким уровнем может работать как адаптивный накопитель. Это адаптивное устройство соперничает с прецизионным обнаружителем, использующим алгоритм быстрого преобразования Фурье, и обладает возможностями, которые могут превосходить возможности обычных анализаторов спектра, когда неизвестный синусоидальный сигнал имеет конечную полосу частот или является модулированным по частоте. Этот способ иллюстрируется схемой на рис. 19.11.

Здесь входной сигнал представляет собой сумму синусоидального сигнала и шума. На выходе системы формируется дискретное преобразование Фурье (ДПФ) от импульсной характеристики фильтра. Сигнал обнаруживается тогда, когда спектральный пик превышает фоновый шум.

Отметим, что в схеме на рис. 19.11, кроме того, имеется выходной сигнал фильтра y , который можно использовать непосредственно или в качестве входного сигнала анализатора спектра или устройства фазовой автоподстройки. Более того, способ, иллюстрируемый на рис. 19.11, можно применять при одновременном обнаружении множества синусоидальных сигналов.

На входе действует белый шум с общей мощностью v^2 , а входной сигнал имеет на нормированной частоте ω_0 мощность $C^2/2$. Идеальная импульсная характеристика, эквивалентная отклику согласованного фильтра, представляет собой отсчеты синусоиды, имеющей частоту ω_0 . Общий сдвиг фазы отклика на частоте ω_0 и задержки кратен 360° . Если значение пика передаточной функции равно a , то значение пика весовых коэффициентов составляет примерно $2a/(L+1)$, где $L+1$ – число весовых коэффициентов.

В результате адаптивного процесса минимизируется среднеквадратическая ошибка $E[\varepsilon_k^2]$. В рассматриваемом случае она является суммой трех составляющих: мощности шума на входе, мощности шума на вы-

ходе адаптивного фильтра и мощности синусоидального сигнала, т.е. СКО можно записать в виде

$$\text{СКО} = v^2 + v^2 \frac{2a^2}{L+1} + \frac{C^2}{2} (1-a)^2. \quad (19.28)$$

В этом выражении v^2 – мощность шума на входе, во втором слагаемом v^2 умножено на сумму квадратов весовых коэффициентов фильтра, а третье слагаемое представляет собой разность мощностей входного и отфильтрованного синусоидального сигналов; при этом полагаем, что сигналы вычитаются когерентно.

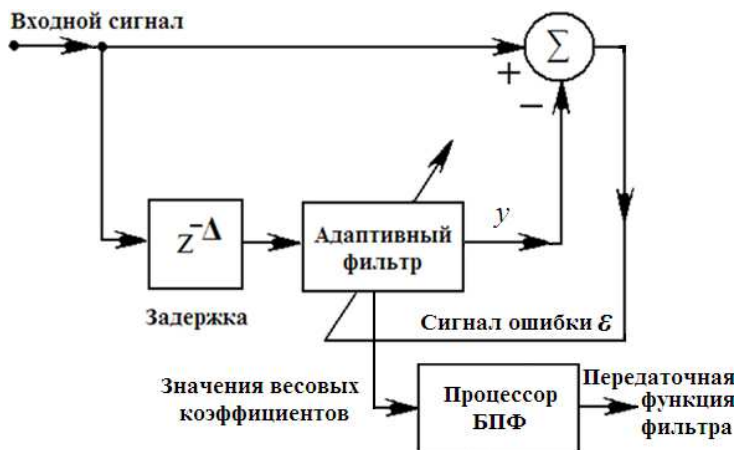


Рис. 19.11. Схема адаптивного накопителя

Приравняв нулю производную выражения (19.14) по a , получаем оптимальное значение a^* , при котором минимальная мощность сигнала ошибки:

$$a^* = \frac{\frac{C^2/2}{v^2} \cdot \frac{L+1}{2}}{1 + \frac{C^2/2}{v^2} \cdot \frac{L+1}{2}} = \frac{\text{ОСШ}(L+1)/2}{1 + \text{ОСШ}(L+1)/2}. \quad (19.15)$$

При больших значениях сигнал-шум a^* близко к 1, при малых – меньше 1. Для того чтобы a^* приближалось к 1 при малых отношениях сигнал-шум, можно использовать большое число адаптивных весовых коэффициентов.

На рисунке 19.12. для сравнения приведены некоторые экспериментальные результаты моделирования на ЭВМ процесса измерения

спектральной плотности мощности с помощью ДПФ и с применением адаптивного накопителя. В каждом из трех случаев одни и те же данные анализируются методом ДПФ, а затем подаются на адаптивный накопитель.

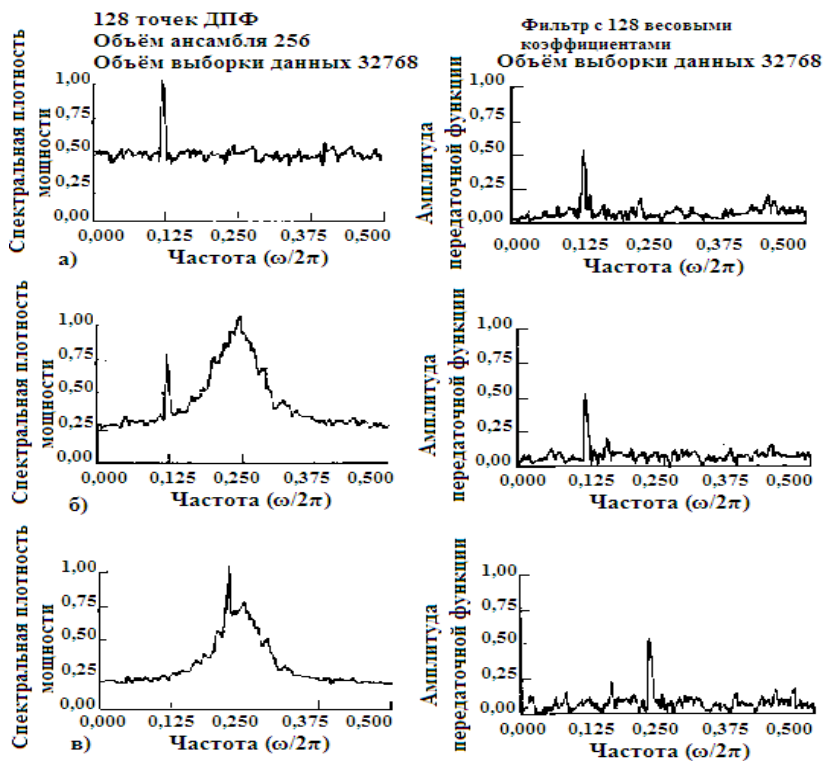


Рис. 19.12. Классический спектральный анализ (ДПФ) – слева и адаптивное накопление – справа: а – одночастотный сигнал в белом шуме; б – одночастотный сигнал в смеси из белого (50%) и окрашенного (50%) шумов; в – одночастотный сигнал с другой частотой в смеси белого (50%) и окрашенного (50%) шумов

Анализ всех кривых, приведенных на рисунке 19.12. в одном и том же масштабе, показывает, что амплитуды составляющих сигнала приблизительно одинаковы и, кроме того, аналогичны уровни фонового шума как для алгоритма ДПФ, так и для накопителя. В каждом случае накопитель реализуется с задержкой в 256 отсчетов, выбранной так, чтобы составляющие окрашенного шума были некоррелированными и подавлялись.

Очевидно, что адаптивный накопитель можно применять в качестве альтернативы к алгоритму ДПФ при обнаружении и оценке слабых сигналов на фоне шума. Кроме того, его можно использовать в качестве следящего фильтра, который автоматически настраивается на сигналы или отстраивается от них. При этом реализуется методология спектрального анализа, связанная с методом максимальной энтропии. По своей структуре накопитель полностью отличен от устройства вычисления ДПФ, и в некоторых случаях его легче реализовать.

Адаптивное подавление помех представляет собой способ оптимальной фильтрации, который можно применять всегда, когда имеется подходящий эталонный входной сигнал. Принципиальными достоинствами этого способа являются его адаптивность, низкий уровень помех на выходе и малые вносимые искажения сигнала. Адаптивность позволяет обрабатывать входные сигналы с неизвестными свойствами и в некоторых случаях нестационарные сигналы. Этот способ приводит к устойчивой системе, которая автоматически отключается, если не достигается улучшения отношения сигнал-шум. В общем случае уровень помех на выходе и степень искажения сигнала ниже достигаемых в обычных схемах оптимальных фильтров.

VII. МНОГОКАНАЛЬНАЯ СВЯЗЬ И РАСПРЕДЕЛЕНИЕ ИНФОРМАЦИИ

20. ПРИНЦИПЫ МНОГОКАНАЛЬНОЙ СВЯЗИ И РАСПРЕДЕЛЕНИЯ ИНФОРМАЦИИ

20.1. Общие положения

Наиболее дорогостоящей частью системы передачи информации (СПИ) являются линии связи – кабельные, волноводные, световодные, радиорелейные и т.п. Естественно возникает задача совместного использования этого оборудования различными пользователями (абонентами), т.е. многоканальной связи или уплотнения. Тем самым повышается эффективность использования ресурсов линии.

Многоканальная связь возможна лишь тогда, когда пропускная способность совместно используемого оборудования больше суммарной информационной производительности всех источников. При этом ресурсы линии связи должны быть некоторым образом распределены между пользователями.

На рисунке 20.1 изображена упрощенная схема передающей и приемной частей многоканальной СПИ, предназначенной для передачи в одном направлении.

Генератор поднесущих выдает n периодических сигналов (гармонических, последовательностей импульсов и т.п.), имеющих разную форму в разных каналах. Вид поднесущих и, следовательно, возможные способы их модуляции определяются методом уплотнения каналов. Модулированные поднесущие $S_1(t), \dots, S_n(t)$ называются канальными сигналами.

Операция уплотнения каналов – это обычное суммирование канальных сигналов

$$S(t) = \sum_{j=1}^n s_j(t) \cdot \quad (20.1)$$

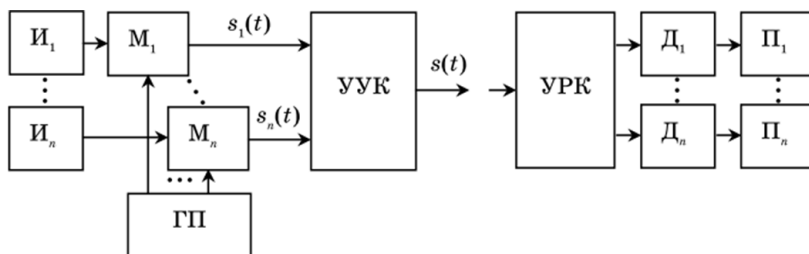


Рис. 20.1. Упрощенная схема многоканальной СПИ:

I_1, \dots, I_n – источники информации (сообщений); ГП – генераторы поднесущих;
 M_1, \dots, M_n – каналные модуляторы; УУК – устройство уплотнения каналов;
 УРК – устройство разделения каналов; D_1, \dots, D_n – каналные демодуляторы;
 Π_1, \dots, Π_n – получатели сообщений

Для разделения каналов, как правило, применяются также линейные устройства. Считают, что в приемнике происходит разделение канальных сигналов без возникновения междуканальных помех, если выполняется следующее условие: при любых значениях сообщений, передаваемых в каналах, напряжение на j -м выходе линейного устройства разделения каналов равно нулю тогда и только тогда, когда $S_j(t) = 0$.

Известно, что для выполнения этого условия необходимо и достаточно, чтобы система n функций, описывающих канальные сигналы, была линейно независимой при любых значениях передаваемых сообщений.

Если выполняется лишь это условие, устройство разделения каналов может оказаться достаточно сложным.

Поэтому обычно выдвигают еще одно, дополнительное требование: канальные сигналы должны быть к тому же попарно ортогональны

$$\int_{-\infty}^{+\infty} s_j(t) \cdot s_k(t) dt = \begin{cases} E_j, & j = k, \\ 0, & j \neq k. \end{cases} \quad (20.2)$$

где E_j – энергия j -го канального сигнала.

Приведем примеры типичных СПИ, использующих ортогональные канальные сигналы.

1. Канальные сигналы не перекрываются во времени (рис. 20.2). Условие ортогональности (20.2) здесь выполняется, так как при любом t подынтегральное выражение равно нулю при $j \neq k$. Это метод временного разделения каналов (ВРК). Используются также термины: временное мультиплексирование, многостанционный доступ с временным разделением каналов (МДВП), Time Division Multiple Access (TDMA).

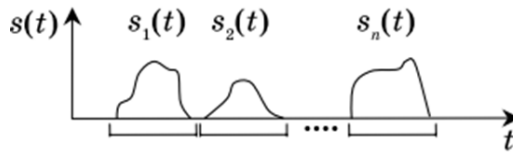


Рис. 20.2. Канальные сигналы, не перекрывающиеся во времени

Фактически линия связи поочередно подключается на короткое время к каждому из источников. Поэтому принципиальными элементами такой СПИ являются мультиплексор (устройство уплотнения каналов) и демультиплексор (устройство разделения каналов), всегда работающие синхронно (рис. 20.3). Поэтому неотъемлемой составной частью любой СПИ с ВРК является система синхронизации приемного устройства.

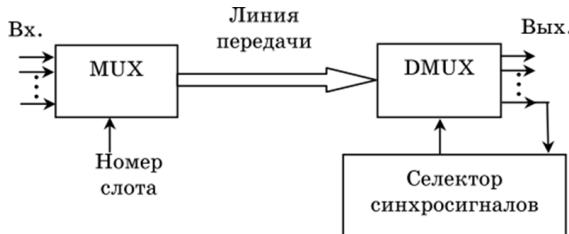


Рис 20.3. Принципиальные элементы СПИ с ВРК

Важно отметить, что длительность канального интервала (слота) мала (в противном случае у пользователей исчезнет эффект «одновременности» передачи), поэтому за время одного подключения к линии удастся передать лишь часть сообщения и необходимо неоднократно повторять этот процесс, а все перечисленные элементы сигналов источников должны быть сжаты во времени.

2. Спектры канальных сигналов не перекрываются на оси частот. Спектр группового сигнала (сумма спектров канальных сигналов) показан на ри. 20.4. Условие ортогональности (20.2) здесь выполняется для спектров, следовательно, и для функций времени. Это метод частотного разделения каналов (ЧРК). Используются также термины: частотное мультиплексирование, многостанционный доступ с частотным разделением каналов (МДЧР), Frequency Division Multiple Access (FDMA).

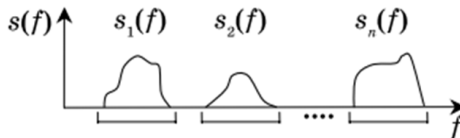


Рис. 20.4. Спектры канальных сигналов не перекрывающиеся на оси частот

Каждой паре «источник – получатель» выделена своя полоса частот. Поэтому принципиальными элементами такой СПИ являются блок n генераторов гармонических колебаний разных частот поднесущих в передатчике и набор полосовых фильтров для разделения канальных сигналов в приемнике.

3. Все канальные сигналы передаются одновременно (перекрываются во времени) в одной общей полосе частот (их спектры также перекрываются), но взаимная ортогональность сигналов обеспечена тем, что все они имеют разную форму. Тогда устройство разделения каналов – это набор n параллельно включенных корреляционных приемников или согласованных фильтров, на входы которых подается принимаемый групповой сигнал. В каждом приемнике весовая функция $w(t)$ совпадает по форме с соответствующим канальным сигналом, поэтому он реагирует лишь на сигнал данного канала. Последнее требование удастся выполнить лишь в цифровой СПИ.

Такой метод называется многоканальной передачей с разделением по форме сигналов. В качестве набора канальных сигналов можно использовать любую известную из курса математики систему ортогональных функций (функции Эрмита, Лагерра, Чебышева и т.п.), но все эти функции имеют довольно сложную форму. Поэтому для практического использования предложен более простой вариант метода, когда в качестве канальных сигналов применяются импульсы с внутриимпульсной фазовой модуляцией, причем закон модуляции каждого канального сигнала задается индивидуальной двоичной кодовой комбинацией.

В таком виде метод получил название кодового разделения каналов (КРК). Используются также термины: кодовое мультиплексирование, многостанционный доступ с кодовым разделением каналов (МДКР), Code Division Multiple Access (CDMA).

20.2. Частотное разделение каналов

Сообщения с выходов источников (рис. 20.1) поступают на каналные модуляторы, где происходит модуляция гармонических колебаний с различными частотами, называемых поднесущими. Частоты поднесущих колебаний должны различаться настолько, чтобы спектры модулированных сигналов не накладывались друг на друга во избежание взаимных помех. После модуляции информационные сигналы занимают ограниченные полосы частот, которые могут отличаться по ширине от спектров исходных колебаний (например, при частотной или фазовой модуляции), или совпадать с ними (при ОБП-модуляции). Важно, чтобы полосы частот, занимаемые различными сигналами, не только не перекрывались, но и отстояли друг от друга на ширину некоторого защитного интервала, что облегчает их последующее разделение при помощи

реальных фильтров, имеющих конечную крутизну АЧХ в переходной полосе.

Индивидуальные каналные (модулированные) сигналы суммируются и поступают на групповой передатчик, где происходит модуляция несущего колебания групповым сигналом, после чего модулированный линейный сигнал передается в линию связи. Групповой приемник производит демодуляцию линейного сигнала, после чего каждый каналный приемник выделяет при помощи полосового фильтра «свой» каналный сигнал, демодулирует его и выделяет сообщение.

Как видно, частотное разделение каналов основано на распределении одного из ресурсов – полосы пропускания группового канала – между различными индивидуальными каналами.

Частотному разделению каналов свойственны следующие недостатки.

Во-первых, из-за неидеальности полосовых фильтров необходимы защитные интервалы, которые суммарно составляют около 20% полосы пропускания группового канала связи. Например, в многоканальных телефонных системах для передачи речевых сигналов установлена полоса частот 3100 Гц (считается, что при передаче речи для обеспечения разборчивости с сохранением индивидуальных голосовых признаков достаточен диапазон от 300 до 3400 Гц), а ширина защитного интервала составляет 900 Гц; таким образом, при объединении N телефонных каналов общая ширина полосы частот группового канала составляет $4N$ кГц.

Во-вторых, предъявляются очень жесткие требования к линейности канала (нелинейность приводит к появлению кратных и комбинационных составляющих, а поскольку спектры каналных сигналов имеют ширину значительно больше защитного интервала, эти составляющие попадают в «чужие» каналы и разделить их путем фильтрации или каким-либо другим способом невозможно).

20.3. Временное разделение каналов

Временное уплотнение (временное разделение каналов, ВРК) основано на распределении временного ресурса группового канала между различными индивидуальными каналами. Каждый пользователь передающий информацию многократно получает канал в свое пользование на короткое время. Но так можно передавать лишь отсчеты сигнала, взятые с шагом, равным периоду следования тактовых интервалов. Таким образом, в основе ВРК лежит использование теоремы отсчетов, и передавать можно лишь сигналы с финитным спектром.

Для формирования каналных сигналов используются различные виды импульсной модуляции (АИМ, ВИМ, ШИМ). Групповой сигнал

может передаваться непосредственно по линии или модулировать гармоническую несущую.

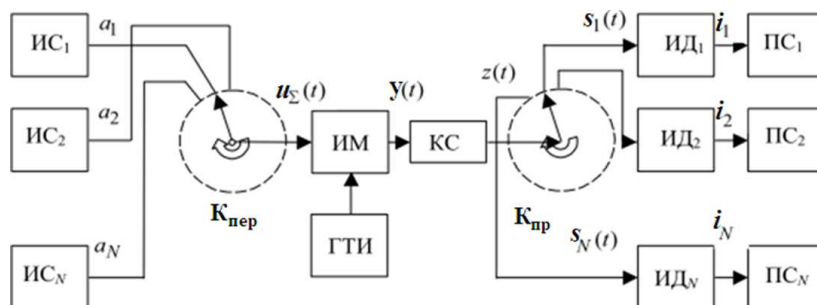


Рис. 20.5. Упрощенная структурная схема СПИ с ВРК

Источники первичных сигналов $ИС$ соединены с коммутатором передатчика $K_{пер}$. Коммутатор поочередно подключает источники к импульсному модулятору $ИМ$, который модулирует сигнал-переносчик – периодическую последовательность импульсов. В результате этого получается групповой сигнал, поступающий в канал связи $КС$, который может включать модулятор гармонической несущей, линию связи и общий демодулятор. После общей демодуляции групповой сигнал разделяется коммутатором приемника $K_{пр}$ на каналные сигналы $s_1(t), \dots, s_N(t)$, которые после демодуляции в импульсных демодуляторах $ИД$ поступают к получателям сигналов $ПР$.

Следует отметить, что в системах связи с ВРК во избежание межканальных помех необходима синхронизация приемной и передающей станций, поэтому в линейный сигнал, кроме каналных сигналов, добавляется периодическая последовательность синхроимпульсов, которые должны достаточно сильно отличаться от каналных импульсов, чтобы их можно было легко выделить.

Для снижения уровня взаимных помех приходится вводить защитные временные интервалы между соседними импульсами, что приводит или к уменьшению числа каналов, или к укорочению каналных импульсов, что ведет к расширению требуемой полосы частот группового канала.

Например, при передаче речевых сигналов (ширина спектра 3100 Гц), минимальная частота дискретизации должна быть равна 6200 Гц. Однако в действительности частоту дискретизации принимают равной 8 кГц, что требует для передачи каналной импульсной последовательности полосы частот около 4 кГц.

В результате общая полоса частот группового канала практически совпадает с полосой, требуемой при ЧРК. Если учесть необходимость

дополнительной передачи синхроимпульсов, то сравнение получается не в пользу ВРК.

Вместе с тем роль временного разделения возрастает в связи с широчайшим распространением цифровых систем связи.

Режим асинхронной передачи применяется только в цифровых СПИ. При подключении источника к линии передается пакет, то есть небольшая часть сообщения, например файла, содержащая обычно от сотен до нескольких тысяч бит. Порядок подключения источников к линии – произвольный, по мере возникновения необходимости, поэтому циклы, как таковые, здесь отсутствуют. В начальной части каждого пакета имеется заголовок, где, кроме прочего обязательно указан номер получателя.

Параметры цифровых СПИ с временным разделением каналов стандартизованы. Первый из наиболее важных стандартов носит название плезиохронной цифровой иерархии (ПЦИ) – Plesyochronous Digital Hierarchy (PDH) – и определяет способ мультиплексирования стандартных цифровых телефонных сигналов для их передачи по медному кабелю.

Так стандартный цифровой телефонный сигнал (DSO) получается в результате АЦП стандартного аналогового телефонного сигнала и характеризуется скоростью 64 кбит/с.

Первая ступень уплотнения PDH (аппаратура ИКМ-30) рассчитана на организацию 32 каналов методом временного побайтного мультиплексирования, в итоге на выходе формируется цифровой поток E1 со скоростью $64 \times 32 = 2048$ кбит/с. Из них 2 канала являются служебными (передача сигналов цикловой синхронизации, передача команд для управления коммутирующими приборами на АТС и др.), а остальные 30 – пользовательскими.

Каждая из ступеней более высокого уровня (аппаратура ИКМ-120, ИКМ-480, ИКМ-1920) рассчитана на уплотнение четырех цифровых потоков низшего уровня методом временного побитного мультиплексирования, а пятый уровень пока не имеет аппаратурной реализации (и, по-видимому, не будет иметь из-за конкуренции с методами передачи по оптическому кабелю).

Другой важный стандарт носит название синхронной цифровой иерархии (СЦИ) – Synchronous Digital Hierarchy (SDH) – и определяет способ мультиплексирования цифровых потоков для их передачи по оптоволоконному кабелю. Аппаратура SDH не предназначена для мультиплексирования мелких цифровых потоков, поступающих от пользователей (для этого используется аппаратура PDH, например ИКМ-30), она применяется для транспортировки более крупных потоков между узлами сети. Синхронная цифровая иерархия создавалась для работы в синхронных сетях связи, где существует синхронизация всех узлов сети из единого центра. Благодаря этому скорости передачи всех однотипных по-

токов строго одинаковы. Передача цифровых сигналов на каждом уровне мультиплексирования осуществляется блоками длительностью 125 мкс.

20.4. Кодовое разделение каналов

Этот метод применяется только в двоичных цифровых СПИ с ФМ. В одноканальной СПИ с ФМ для передачи каждого бита от источника используется радиоимпульс прямоугольной формы, начальная фаза которого определяется значением этого бита, например, 0° при передаче единицы и 180° при передаче нуля. Длительность этого информационного импульса $\tau_{\text{и}} = 1/V$, где V – скорость передачи, бит/с. Ширина спектра такого сигнала примерно равна $\Delta f \approx 2/\tau_{\text{и}} = 2V$. При малой скорости передачи сигнал занимает узкую полосу частот.

Для организации кодового разделения каналов в многоканальной СПИ каждый каналный информационный импульс (рис. 20.6а) подвергается дополнительной внутриимпульсной фазовой модуляции (рис. 20.6б).

Для этого импульс разбивается на B частей равной длины (на B более коротких импульсов-чипов).

Для данного канала раз и навсегда задана B -разрядная двоичная последовательность, определяющая закон внутриимпульсной фазовой модуляции, причем элементы этой последовательности удобно обозначить как $+1$ и -1 вместо традиционных 1 и 0. Далее каждый элемент сигнала умножается на соответствующий элемент последовательности, то есть остается без изменений, если множитель равен $+1$ и инвертируется в случае -1 .

В итоге ширина спектра каналного сигнала увеличивается в B раз по сравнению с традиционным методом передачи без внутриимпульсной модуляции, так как теперь сигнал фактически состоит из более коротких импульсов. Кстати, метод внутриимпульсной ФМ является самым распространенным методом искусственного расширения спектра сигнала и часто применяется в других областях радиоэлектроники, например в радиолокации.

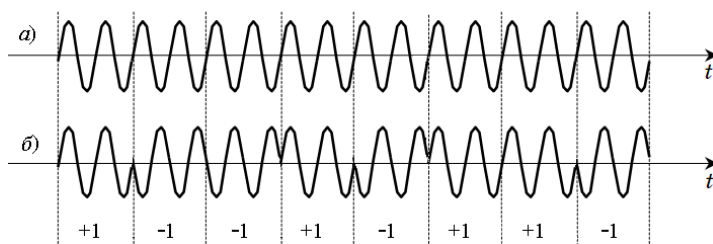


Рис. 20.6. Преобразование информационного импульса в сигнал с внутриимпульсной ФМ: а – до преобразования; б – после преобразования

Все каналные сигналы передаются одновременно в одной и той же широкой полосе частот. Каждому каналному сигналу присваивается своя двоичная последовательность, определяющая закон внутриимпульсной ФМ, так, чтобы сигналы всех каналов были взаимно ортогональны. Два каналных ФМ сигнала ортогональны, если в двоичных последовательностях, определяющих законы их внутриимпульсной ФМ, число совпадающих элементов равно числу несовпадений. Например, если одному из каналов присвоена последовательность, показанная на рис. 20.6, то другим каналам можно присвоить, допустим, такие последовательности:

$$+1, -1, +1, -1, -1, +1, -1, +1, \dots$$

$$+1, -1, -1, +1, +1, -1, -1, +1, \dots$$

В системах МДКР операции выделения нужного каналного сигнала из суммарного, группового сигнала и демодуляции можно совместить. Для этого можно использовать корреляционный приемник (или заменяющий его согласованный фильтр) и в качестве опорного сигнала использовать целиком импульс с внутриимпульсной ФМ. Таким образом, в системах МДКР устройство разделения каналов и каналные демодуляторы – это набор, состоящий из n корреляционных приемников.

Максимальное число ортогональных последовательностей равно B , следовательно, именно таково количество каналов, работающих без взаимных помех.

Для практического применения подбирают последовательности, обладающие лучшими спектральными свойствами (широкий спектр). Хорошие результаты дает использование отрезков псевдослучайной двоичной последовательности (ПСП).

Так как в системах МДКР не предусматривается взаимная синхронизация каналов, т.е. модулирующие последовательности в разных каналах произвольно сдвинуты во времени, это может привести к нарушению их ортогональности и явиться причиной междуканальных помех.

Метод МДКР используется пока лишь в гражданских и военных системах радиосвязи с подвижными объектами. Следует упомянуть стандарт IS-95, в соответствии с которым построены сотовые системы CDMA и спутниковая система связи с подвижными объектами Globalstar ($B = 64$, ширина спектра группового и каждого из каналных сигналов равна 1,25 МГц).

20.5. Синхронизация в СПИ с многостанционным доступом

Проблема синхронизации не возникает лишь в аналоговых СПИ с ЧРК. Все элементы, предназначенные для преобразования импульсных

сигналов, в остальных СПИ работают под управлением тактовых импульсов, задающих ритм их работы.

Работой передающей части многоканальной СПИ в синхронном режиме передачи управляет достаточно стабильный первичный генератор тактовых импульсов. Период следования этих импульсов равен длительности самого короткого элемента группового сигнала. В цифровой СПИ с ВРК это один бит, в СПИ с КРК это один из B чипов информационного бита. Например, в аппаратуре ИКМ-30 тактовые импульсы имеют частоту 2048 кГц.

Точно такой же комплект генераторного оборудования используется для управления работой всех устройств в приемной части СПИ. Система синхронизации приемного устройства обеспечивает условия, при которых тактовые импульсы всех видов в приемнике формируются одновременно с соответствующими тактовыми импульсами в передатчике (точнее, с одной и той же задержкой, равной текущему времени распространения сигнала от передатчика к приемнику).

Чтобы система синхронизации могла выполнять свои функции, кроме сигналов, несущих информацию о передаваемых сообщениях, необходимо еще передавать специальные синхросигналы разных видов. Поэтому в передающей части СПИ имеются формователи синхросигналов запускаемые соответствующими тактовыми импульсами от генераторного оборудования.

В приемной части СПИ имеются селекторы синхросигналов, которые постоянно сравнивают временное положение принимаемых информационных и синхронизирующих сигналов, с одной стороны, и тактовых импульсов от местного генераторного оборудования, с другой стороны, и по мере необходимости подают команды на коррекцию последних.

Специальные синхросигналы в момент начала очередного мельчайшего элемента группового сигнала, например бита и каналные синхросигналы, как правило, не передают. Достаточно того, что передается кадровый синхросигнал в начале каждого кадра (как сигналы точного времени достаточно передавать раз в час, а не каждую секунду). Для передачи кадрового синхросигнала выделяется отдельный канал, и в нем в цифровой СПИ передается специальная кодовая комбинация той же длины, что и в любом другом канале.

Например, в аппаратуре ИКМ-30 в начале каждого четного цикла на 2–8 битовых позициях передается синхросигнал вида 0011011. Единственное отличие циклового синхросигнала от комбинаций в информационных каналах состоит в том, что он всегда один и тот же, а комбинации в информационных каналах изменяются. При сбое синхронизации в селекторе он переходит в режим поиска, то есть проверяет подряд все слоты до тех пор, пока не встретится комбинация 0011011. После этого переходит в нормальный режим, то есть режим слежения.

Естественно, что требования к помехоустойчивости селектора синхросигналов существенно выше, чем требования к помехоустойчивости других решающих устройств.

Решение проблемы синхронизации в системах радиосвязи с много-станционным доступом имеет ряд особенностей. Если в многоканальной СПИ формирование цифрового сигнала, кодирование, модуляция и уплотнение каналов осуществляются в пункте передачи под управлением общего генераторного оборудования, то в системах радиосвязи передатчик, устройство уплотнения канальных сигналов (базовая станция в системах сотовой связи, ретранслятор на ИСЗ в системах спутниковой связи и т.п.) и приемник находятся в разных пунктах. Здесь приходится налаживать совместную синхронную работу не двух, а трех объектов в условиях, когда расстояния между ними меняются во времени (следовательно, меняются и задержки сигнала).

В таких системах, как правило, осуществляется двусторонняя передача информации, то есть на обоих участках имеются два равноценных канала: в прямом и обратном направлениях. Здесь ведущая роль принадлежит ретранслятору, поэтому в нем находится генераторное оборудование, задающее временной график работы всей системы, и формируются синхросигналы. Приемники всех абонентов синхронизируют свое генераторное оборудование по сигналам, передаваемым ретранслятором, и это же оборудование управляет работой своего передатчика.

Остается лишь один вопрос: с каким опережением должен работать передатчик, чтобы его сигнал достиг ретранслятора к тому моменту, когда начнется нужный временной слот? Для оценки текущего времени запаздывания на пути от станции до ретранслятора каждая станция периодически принимает свой сигнал, возвратившийся от ретранслятора, и измеряет его запаздывание относительно переданного сигнала.

Работа СПИ в асинхронном режиме передачи (передача пакетов по мере возникновения требований) имеет ряд особенностей. Все современные протоколы требуют, чтобы пакет содержал целое число байтов.

Опишем два типичных способа обеспечения синхронизации, где система должна указать приемнику байты, соответствующие началу и концу пакета.

1. В начале и в конце каждого пакета передается флаг, то есть специальная восьмибитовая комбинация, допустим, 01111110. Такая комбинация не может встретиться нигде больше внутри пакета. Селектор синхросигнала проверяет все восьмибитовые последовательности и первую обнаруженную комбинацию 01111110 идентифицирует как начало пакета, а вторую – как его конец.

2. Номер байта, соответствующий началу пакета, и количество байтов в пакете передаются по отдельным, служебным каналам.

Кроме того существуют свободные интервалы, не занятые для передачи пакетов. В действительности в это время передается «пустая» последовательность символов для того, чтобы обеспечить устойчивую работу селекторов синхросигналов в приемной части СПИ.

20.6. Коммутация в сетях связи

Для обмена информацией между многими абонентами создаются сети связи, в которых производится распределение информации в соответствии с заданными адресами.

Сети связи подразделяются на некомутируемые, в которых связь абонентов осуществляется по принципу «каждый с каждым» по закрепленным каналам, и коммутируемые, в которых связь осуществляется по временно выделяемым каналам. Выделение каналов парам абонентов производится узлами коммутации. Таким образом, сеть связи состоит из оконечных (абонентских) устройств, каналов связи и узлов коммутации.

Сети могут иметь различную структуру: линейную, радиальную, кольцевую, радиально-узловую и т.д. Построение и оптимизация сетей связи осуществляются на основе теории графов и теории массового обслуживания.

Наиболее широко известными узлами коммутации можно считать автоматические телефонные станции (АТС). Основной составной частью АТС как узла коммутации является коммутационное поле. Коммутационное поле может быть пространственным, характерным для аналоговых систем, или пространственно-временным. В первом случае коммутационное поле электрически соединяет отдельные линии на все время соединения. Во втором случае линии соединяются на короткие временные интервалы в соответствии с методом временного уплотнения. В цифровых сетях связи применяется цифровая коммутация канальных интервалов, которая осуществляется записью принятых сегментов сообщений в память и считыванием их в определенном порядке.

Для эффективного использования имеющихся каналов связи и узлов коммутации форма представления информации должна быть стандартизована. Существующий стандарт имеет иерархическую (многоуровневую) структуру, основанную на модели взаимодействия открытых систем [15].

Эталонная модель содержит семь уровней. Каждый уровень позволяет рассматривать некоторый аспект функционирования системы (сети связи), абстрагируясь от содержания остальных уровней (рис. 20.7).

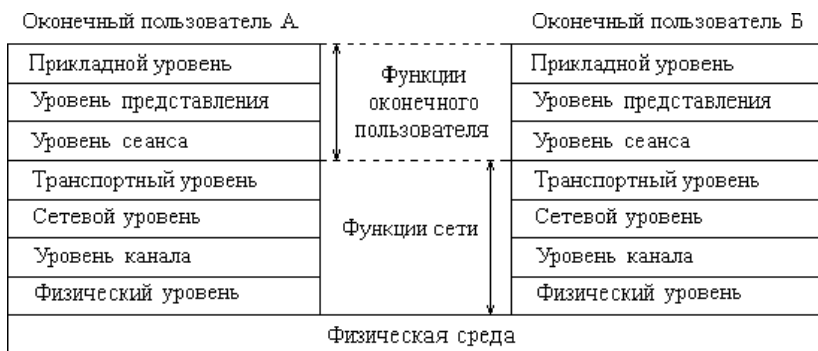


Рис. 20.7. Эталонная модель взаимодействия

Физический уровень модели непосредственно взаимодействует с физической средой распространения сигналов и обеспечивает передачу сигналов между двумя узлами.

Сетевой уровень обеспечивает установление адреса и маршрута для передачи пакета данных от узла передачи до узла назначения.

Транспортный уровень соответствует передаче данных между абонентами сети и характеризуется максимальным временем установления соединения, пропускной способностью, временем задержки при передаче сообщений и т.п.

Четыре нижних уровня реализуют функции сети, оставшиеся три уровня ориентированы на услуги, предоставляемые конечным пользователям.

Уровень сеанса обеспечивает организацию диалога, очередность передачи данных, приоритеты и т.д.

Уровень представления определяет коды, форматы данных, способы сжатия и т.п.

Прикладной уровень служит для реализации услуг, предоставляемых сетью пользователям (электронная почта, телетекст, факс, электронные переводы, пакетная передача речи и т.п.).

Уровни модели взаимодействуют со смежными (верхним и нижним соседними) уровнями; кроме того, возможно взаимодействие различных пользователей на одинаковых уровнях. Правила взаимодействия одного уровня называются протоколом. Взаимодействие на некотором уровне обеспечивается предоставлением ему услуг смежным нижележащим уровнем.

VIII. ЭФФЕКТИВНОСТЬ СИСТЕМ СВЯЗИ

21. ОЦЕНКА ЭФФЕКТИВНОСТИ И ОПТИМИЗАЦИЯ ПАРАМЕТРОВ ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ (ТКС)

21.1. Критерии эффективности

Решение вопросов выбора наиболее целесообразных вариантов ТКС в конечном итоге сводится к решению задач оптимизации этих систем по выбранным критериям качества.

Обобщенной характеристикой эффективности систем связи является коэффициент использования канала по пропускной способности (информационная эффективность) который характеризует реальную скорость передачи информации R по отношению к пропускной способности C канала связи [5]:

$$\eta = R/C. \quad (21.1)$$

Информационная эффективность η всегда меньше единицы; чем ближе η к единице, тем совершеннее система.

Для оценки эффективности систем связи вводятся также коэффициент использования канала по мощности (энергетическая эффективность)

$$\beta = R/(P_c/N_0), \quad (21.2)$$

и коэффициент использования канала по полосе частот (частотная эффективность)

$$\gamma = R/\Delta F. \quad (21.3)$$

В этих формулах P_c – мощность сигнала; N_0 – спектральная плотность шума; ΔF – ширина полосы частот, занимаемой сигналом.

Предельные возможности системы передачи информации можно оценить с помощью выражения для пропускной способности гауссовского непрерывного канала связи с полосой частот ΔF :

$$C = \Delta F \log_2(1 + P_c/P_m). \quad (21.4)$$

Здесь $P_c = E_b B$ – средняя мощность сигнала: E_b – энергия, затрачиваемая на передачу одного бита информации; $B = 1/T_b$ – скорость передачи информации источника; T_b – время передачи источником одного бита информации; $P_{ш} = N_0 \Delta F$ – средняя мощность шума в полосе частот.

В реальных ТКС скорость передачи информации B [Бит/с], меньше пропускной способности непрерывного канала: $B \leq C$. Можно показать, что после элементарных преобразований это неравенство приводится к виду [5]:

$$\beta \leq \gamma / (2^\gamma - 1), \quad (21.5)$$

$$\text{где } \beta = 1/h_2^2 = N_0/E_b \quad (21.6)$$

Тогда информационная эффективность для гауссовского непрерывного канала может быть найдена по формуле:

$$\eta = \gamma / \log_2(\gamma / (\beta + 1)). \quad (21.7)$$

Согласно теореме Шеннона, при соответствующих способах передачи и приема величина η может быть сколь угодно близкой к единице. При $\eta = 1$ получаем предельную зависимость между β и γ :

$$\beta = \gamma / (2^\gamma - 1). \quad (21.8)$$

Наглядно данная зависимость представляется в виде кривой на β γ на плоскости (рис. 21.1).

Эта зависимость, часто называется границей (пределом) Шеннона: она отражает наилучший обмен между β и γ в непрерывном канале.

Анализ соотношения (21.6) и предела Шеннона показывает, что повышение частотной эффективности (т.е. снижение затрат полосы $1/\gamma$) требует увеличения энергетических затрат (снижения энергетической эффективности). Для непрерывного канала частотная эффективность изменяется в пределах от 0 до ∞ , в то время как энергетическая эффективность ограничена сверху:

$$\beta_{\max} = \lim_{\gamma \rightarrow 0} \beta = \lim_{\gamma \rightarrow 0} \frac{\gamma}{2^\gamma - 1} = \frac{1}{\ln 2} \approx 1,443.$$

Аналогичные предельные зависимости $\beta = f(\gamma)$ можно получить и для других моделей канала, если в (21.2) и (21.3) вместо скорости R подставить выражение для пропускной способности соответствующего канала. Предельные зависимости β - γ -номограммы позволяют определить системы, удовлетворяющие заданным требованиям по энергетической и частотной эффективности, и установить, насколько эти показатели близки к предельным.

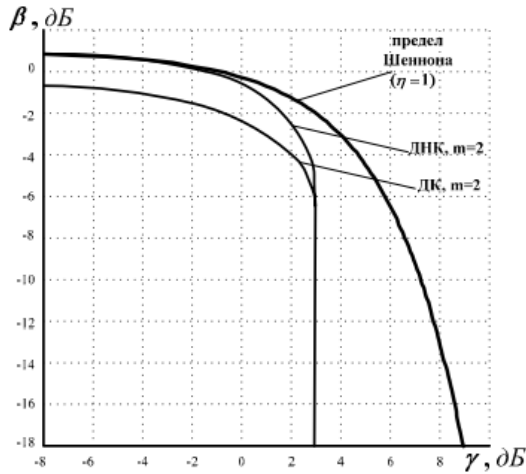


Рис. 21.1. Связь частотной (γ) и энергетической (β) эффективности

21.2. Эффективность аналоговых и цифровых систем

В системах передачи дискретных сообщений сигнал формируется с помощью кодирования и модуляции. При этом кодирование осуществляется обычно в два этапа: кодирование источника с целью сокращения его избыточности и кодирование канала с целью уменьшения вероятности ошибки за счет введения избыточности кода. При этом выражение (21.1) для информационной эффективности системы передачи дискретных сообщений можно представить в виде произведения:

$$\eta = R/C = \eta_{\text{ки}} \cdot \eta_{\text{кк}} \cdot \eta_{\text{м}}, \quad (21.9)$$

где $\eta_{\text{ки}}$ – эффективность кодера источника; $\eta_{\text{кк}}$ – эффективность кодера канала; $\eta_{\text{м}}$ – эффективность модема, зависящая от вида модуляции и способа обработки сигнала в канале.

Средняя скорость передачи информации в системе при использовании многопозиционных сигналов длительностью T равна $R = R_{\text{кк}} \cdot (\log_2 m)/T$ (бит/с), где $R_{\text{кк}} = k/n$ – скорость помехоустойчивого кода. Тогда энергетическая эффективность

$$\beta = \frac{R}{P_c/N_0} = \frac{R_{\text{кк}} \log_2 m}{E_0/N_0} = \frac{1}{E_b/N_0}, \quad (21.10)$$

частотная эффективность может быть найдена по формуле

$$\gamma = R/\Delta F = \log_2 m / T\Delta F, \quad (21.11)$$

где $E_0 = P_c T = E_b R_{\text{кк}} \log_2 m$ – энергия сигнала; $E_b = E_0 / R_{\text{кк}} \log_2 m$ – энергия, затрачиваемая на передачу одного бита информации.

Для определения β и γ могут использоваться приближенные формулы:

$$\gamma \approx (\log_2 m)/n; \beta \approx 1/(E_b / N_0), \quad (21.12)$$

где n – размерность сигнала, в m -позиционной системе. В табл. 21.1 приведены значения m и формулы для приближенных расчетов γ некоторых ансамблей сигналов.

Таблица 21.1

Формулы для приближенных расчетов частотной эффективности некоторых ансамблей сигналов

Ансамбль сигналов	Ортогональный	Биортогональный	Симплексный
m	n	$2n$	$n + 1$
γ	$(\log_2 m)/m$	$2(\log_2 m)/m$	$(\log_2 m)/(m - 1)$

В реальных системах вероятность ошибки всегда имеет ненулевое значение и $\eta < 1$. В этих случаях при заданном значении $p_{\text{ош}} = \text{const}$ можно определить отдельно β и γ и построить кривые $\beta = f(\gamma)$.

В координатах β и γ каждому варианту реальной системы будет соответствовать точка на плоскости (рис. 21.2) [5]. Все эти точки располагаются ниже предельной кривой Шеннона и ниже предельной кривой соответствующего канала. Ход этих кривых зависит от вида модуляции, метода кодирования и способа обработки сигналов. Около графиков на рис. 21.2. указано число позиций дискретного сигнала m . Кривые рассчитаны на основании формул оценки помехоустойчивости различных методов модуляции для оптимального приема сигналов при вероятности ошибки на бит $p_{\text{ош}} = 10^{-5}$.

Занимаемая полоса частот для ЧМн $\Delta F = m/(T \cdot \log_2 m)$, а для ФМн (АМн) $\Delta F = 1/(T \cdot \log_2 m)$.

Анализ рисунка 21.2. показывает, что в системах с ЧМн при увеличении числа позиций m энергетическая эффективность β увеличивается, а частотная эффективность γ уменьшается. В системах с ФМн и ОФМн, наоборот, с увеличением m коэффициент β уменьшается, а γ – увеличивается. Таким образом, условия обмена β на γ за счет изменения числа позиций сигналов в системах связи с ЧМн и ФМн различны.

Представленные на рисунке 21.2. результаты позволяют определить системы, удовлетворяющие заданным требованиям по энергетиче-

ской и частотной эффективности, и установить, насколько эти показатели близки к предельным.

После выбора системы по показателям β и γ , информационная эффективность вычисляется с использованием формулы (21.7).

Например, для сигналов АМн-2 показатель информационной эффективности составляет $\eta \approx 0,228$, а для ЧМн-2 $\eta \approx 0,145$; для ФМн-2 $\eta \approx 0,25$, а для ФМн-4 $\eta \approx 0,47$.

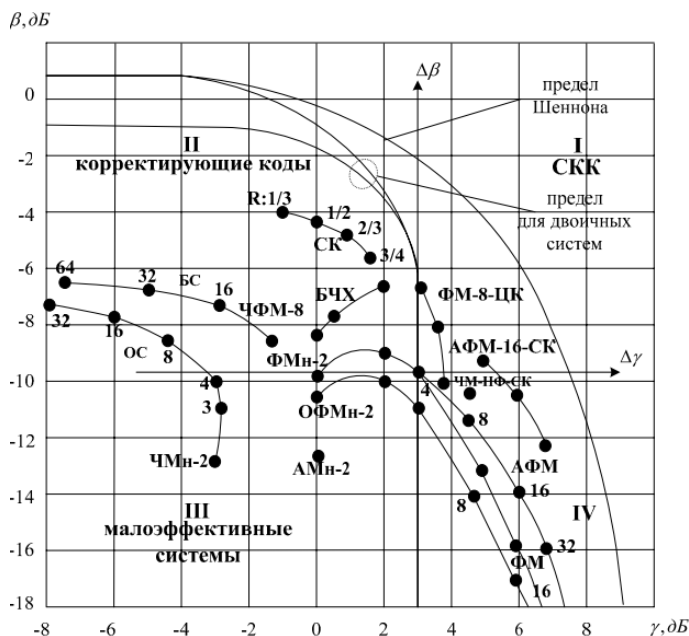


Рис. 21.2. Кривые энергетической и частотной эффективности цифровых систем связи

Анализ предельных кривых показывает, что эффективность дискретных систем передачи можно существенно повысить, если вместо двоичных применять многопозиционные сигналы ($m > 2$).

Эффективность передачи непрерывных сообщений в значительной степени зависит от вида модуляции. Для сравнительного анализа различных видов модуляции обычно используют выигрыш по отношению сигнал/шум ($h_{\text{вых}}$) и коэффициент использования пропускной способности каналов связи (ν)[5]:

$$\eta = \frac{\log h_{\text{вых}}}{\nu \cdot \log(h_{\text{вых}}/g + 1)}. \quad (21.13)$$

В таблице 21.2 приведены данные сравнительного анализа эффективности различных видов модуляции, полученные при $h_{\text{вых}} = 40$ дБ и пик-факторе $\Pi = 3$ для гауссовского канала при оптимальной обработке сигналов [5].

Таблица 21.2

Значения выигрыша и информационной эффективности некоторых систем передачи непрерывных сообщений

Вид модуляции	$\nu = F/F_c$	$g = h_{\text{вых}} / h_{\text{вх}}$	$g' = g/\nu$	$\eta = R/C$
АМ	2	0,2	0,1	0,42
БМ	2	2	1	0,50
ОМ	1	1	1	1
ФМ	20	222	11,1	0,12
ЧМ	20	666	33,3	0,17
ФИМ-АМ	20	666	33,3	0,17
ИКМ-АМ	20	250	12,5	0,23
ИКМ-ЧМ	20	500	25	0,32
ИКМ-ФМ	20	1000	50	0,48
ИС	20	6310	315	1

Анализ показывает, что наибольшая информационная эффективность достигается при однополосной модуляции, однако значение обобщенного выигрыша для этого вида модуляции ($g'=1$) свидетельствует о том, что в системе отсутствует выигрыш по помехоустойчивости.

Одноканальные системы ЧМ и ФИМ примерно равноценны. В этих системах, а также в цифровых системах с ИКМ, высокая помехоустойчивость может быть достигнута с помощью увеличения ширины спектра сигнала, т.е. за счет частотной избыточности. При больших индексах ФМ и ЧМ приближаются по помехоустойчивости к идеальной системе (выигрыш составляет десятки и сотни раз), но информационная эффективность таких систем мала (0,12 ÷ 0,17) из-за большой частотной избыточности. Основными способами повышения эффективности передачи непрерывных сообщений являются устранение избыточности, статистическое уплотнение и применение цифровых видов модуляции.

Аналоговые системы ОМ, АМ и узкополосная ЧМ обеспечивают высокую частотную эффективность при сравнительно низкой энергетической эффективности. Применение этих систем целесообразно в каналах с хорошей энергетикой (при больших значениях $h_{\text{вх}}$) или в тех случаях, когда требуемое значение $h_{\text{вых}}$ мало. Цифровые системы обеспечи-

вают высокую β -эффективность при достаточно хорошей γ -эффективности. В каналах с ограниченной энергетикой (при малых значениях $h_{\text{вх}}$) преимущества цифровых систем особенно заметны. При высоком качестве передачи, когда требуемые значения $h_{\text{вых}}$ велики, широкополосная ЧМ и цифровые системы обеспечивают примерно одинаковую эффективность.

В многоканальных системах эффективность связи снижается за счет несовершенства системы разделения сигналов. Показатели частотной, энергетической и информационной эффективности для систем с множественным доступом определяются на основании суммарной скорости передачи, зависящей от методов формирования и обработки информационных сигналов в парциальных каналах и методов доступа.

21.3. Выбор сигналов и помехоустойчивых кодов

Эффективность систем передачи дискретных сообщений можно существенно повысить путем применения многопозиционных сигналов и корректирующих кодов.

На рисунке 21.2. приведены $\beta\gamma$ -диаграммы для некоторых ансамблей многопозиционных сигналов. Центральное место на рис. 21.2. занимают кривые для систем с сигналами ФМн-4, которые относятся к классу многопозиционных при $m = 4$. В цифровых сетях система ФМн-4 является наиболее распространенной и принята в качестве стандарта, поэтому при сравнительной оценке эффективности систем она принята за эталон. Если начало координат перенести в точку, соответствующую ФМн-4, то в новой системе координат по вертикальной оси будет отсчитываться энергетический выигрыш $\Delta\beta$ рассматриваемых систем по сравнению с ФМн-4, а по горизонтальной оси – выигрыш $\Delta\gamma$ по удельной скорости. В этой системе координат все возможные системы связи можно условно разделить на четыре группы, соответствующие четырем квадрантам на плоскости:

- малоэффективные системы (III квадрант), имеющие относительно ФМн-4 проигрыш по β и γ , например, АМн-2, ЧМн-2;

- системы с высокой энергетической эффективностью (II квадрант), обеспечивающие выигрыш по β и проигрыш по γ (системы с корректирующими кодами);

- системы с высокой частотной эффективностью (IV квадрант), обеспечивающие выигрыш по γ и проигрыш по β (системы с многопозиционными ФМн и АФМ сигналами);

- высокоэффективные системы (I квадрант), позволяющие получить одновременно выигрыш по обоим показателям β и γ на основе применения сложных сигнально-кодовых конструкций.

Наряду с многопозиционными сигналами для повышения эффективности широко используются помехоустойчивые коды. Применение корректирующих кодов позволяет повысить верность передачи сообще-

ний или при заданной верности повысить энергетическую эффективность системы. Это особенно важно для систем с малой энергетикой, например, систем спутниковой связи.

На практике используются как блочные, так и непрерывные коды. На рис. 20.2. приведены кривые эффективности для циклического кода Боуза-Чоудхури-Хоквингема (БЧХ) и для сверточного кода (СК) с декодированием по алгоритму Витерби.

Применение циклического кода позволяет получить энергетический выигрыш $\Delta\beta = 2 \dots 4$ дБ, а сверточного кода $\Delta\beta = 4 \dots 6$ дБ в обмен на снижение частотной эффективности примерно в 2 раза (3 дБ).

Энергетический выигрыш $\Delta\beta$ от применения помехоустойчивого кодирования тем больше, чем выше требуемая верность передачи. Для непрерывного канала с белым гауссовским шумом при требуемой вероятности ошибки 10^{-5} предельный энергетический выигрыш кодирования по сравнению с ФМн-2 без кодирования при оптимальном когерентном приеме составляет примерно 10 дБ.

Применение циклического кода в канале с ФМн или сверточного кода в канале с АФМ позволяет повысить одновременно энергетическую, так и частотную эффективность.

Построение высокоэффективных систем на основе сигнально-кодовых конструкций ведет к неизбежному увеличению сложности системы. Не пропускная способность, а сложность является ограничивающим фактором при построении высокоэффективных систем. Задача состоит в том, чтобы построить систему, удовлетворяющую высоким показателям эффективности, при допустимой сложности.

При современной элементной базе затраты на реализацию кодирующих и декодирующих устройств значительно сократились. В то же время стоимость энергетики канала практически не изменилась. Таким образом, «цена» выигрыша $\Delta\beta$ за счет кодирования может быть существенно меньше цены того же выигрыша, полученного за счет увеличения энергетики канала (мощности сигнала или размеров антенн).

Отметим, что выбор способов кодирования и модуляции зависит от характеристик канала. Улучшение этих характеристик, например, путем адаптации к помехам и оценивания искажений сигнала и их последующей компенсации, снижает потери в канале и создает лучшие условия для применения корректирующих кодов.

Сравнение между собой различных ТКС осуществляется по степени использования ими основных ресурсов канала связи (пропускной способности, мощности, занимаемой полосы частот), выражаемой через показатели информационной, энергетической и частотной эффективности. Создание ТКС, в которых достигаются близкие к предельным показатели эффективности, требует согласования методов модуляции и кодирования с учетом статистических свойств непрерывного канала.

22. ОЦЕНКА ЭФФЕКТИВНОСТИ РАДИОТЕХНИЧЕСКОЙ СИСТЕМЫ СВЯЗИ

22.1. Тактико-технические параметры радиотехнической системы связи

Для радиотехнической системы морской связи, максимальная протяженность которой $r_B = 8000$ км, рабочая частота (в дневное время) составляет $f_p = 15$ МГц, мощность передатчика $P_1 = 400$ Вт, действующая длина приемной антенны – 30 м, коэффициент усиления передающей антенны $G_1 = 50$. При радиотелеграфной работе используются амплитудная, частотная и фазовая манипуляции со скоростью $B = 50$ бит/с; при радиотелефонной работе – амплитудная и частотная модуляции с первичным сигналом звукового вещания, имеющие следующие параметры: $K_A^2 = 2$ дБ – коэффициент амплитуды модулирующего сигнала; $M = 0,6$ – коэффициент модуляции при АМ; $300 < F < 3400$ Гц – частота модуляции; $\Delta f = 5$ кГц – девиация частоты при частотной модуляции.

Требуется определить:

- отношение сигнал/помеха на входе приемного устройства и обосновать применяемый метод обработки сигнала в приемнике;
- количество информации при передаче дискретных и непрерывных сигналов;
- реальную и потенциальную помехоустойчивость при оптимальном и неоптимальном приеме сигналов радиотелеграфии;
- реальную и потенциальную помехоустойчивость при оптимальном и неоптимальном приеме непрерывных сигналов;
- выигрыш демодулятора при передаче сигналов с амплитудной и частотной модуляциями;
- производительность источника сообщений;
- пропускную способность при приеме сигналов радиотелеграфии и радиотелефонии.

22.2. Оценка отношения сигнал/помеха на входе радиоприемника радиотехнической системы связи

Для рассматриваемой радиотехнической системы морской связи определим отношение сигнал/помеха на входе радиоприемника декаметровых волн для указанных выше исходных данных. Расчет проведем для значения критической частоты слоя E , равного 3 МГц, и действующей высоты отражения волны 350 км. Для расчета воспользуемся методикой, изложенной в работе [16]. Считаем, что число скачков $n = 1$.

Медианное значение напряженности поля в точке приема определяется по формуле

$$E_{\text{д}} = \frac{100\sqrt{P_1 G_1}}{r_B} 0,8^{n-1} \exp(-\Gamma_{\text{н}}), \quad (22.1)$$

где интегральный коэффициент поглощения $\Gamma_{\text{н}}$ находят с помощью графиков и формулы (22.2) [16]. Подставляя численные значения, получаем следующие величины:

$$\Gamma_{\text{н}} = \frac{A_{\Sigma}}{(f_{\text{р}} + f_{\text{пл}})^2} \cdot B_{F_2} \cdot f_{\text{р}}^2 = \frac{260}{(15 + 0,7)^2} + 8 \cdot 10^{-4} \cdot 15^2 = 1,23, \quad (22.2)$$

где A_{Σ} – суммарный коэффициент не отклоняющего поглощения; $B_{F_2} \approx 0,02 \cdot \cos^2 \varphi_{F_2}$ – отклоняющее поглощение (φ_{F_2} – угол зависящий от длины трассы и высоты отражения волны); $f_{\text{р}}$ – рабочая частота в МГц; $f_{\text{пл}}$ – продольная составляющая гиромагнитной частоты (в средних широтах обычно принимают $0,7 \div 0,8$ МГц), и соответственно

$$E_{\text{д}} = \frac{100\sqrt{15 \cdot 50}}{3000} \exp(-1,23) = 0,269 \text{ мВ/м} \cdot$$

Напряжение сигнала на выходе приёмника равно произведению напряженности поля на действующую длину приёмной антенны: $U_{\text{с}} = 0,269 \cdot 30 = 8,07$ мВ, или в децибелах относительно 1 мкВ $U_{\text{с}} = 20 \cdot \lg 8070 = 78,1$ дБ.

Медианное значение напряжения помех станций (преобладающих в диапазоне ВЧ), измеренное в дневное время в полосе частот 1 кГц не превышает 10 дБ над 1 мкВ [16]. В полосе передачи 3 кГц напряжение помех составит: $U_{\text{н}} = 10 + 20 \cdot \lg \sqrt{3} = 14,8$ дБ или 5,5 мкВ.

Таким образом, для рассматриваемой радиотехнической системы отношение сигнал/помеха по напряжению вычисляется по формуле

$$U_{\text{с}} / U_{\text{н}} = (8,07 \cdot 10^{-6}) / (5,5 \cdot 10^{-6}) = 1,47$$

или

$$U_{\text{с}} / U_{\text{н}} = 78,1 - 14,8 = 63,3 \text{ дБ},$$

т.о.

$$\rho_{\text{вх}} = 1,47 \text{ или } \rho_{\text{вх}} = 63,3 \text{ дБ}.$$

22.3. Оптимальная фильтрация непрерывных сигналов

Отфильтровать сигнал от помехи можно только при наличии отличий между ними. При оптимальной фильтрации сигнала минимизируется среднеквадратическая разность:

$$\varepsilon^2(t) = \frac{1}{T} \int_0^T [S_{\text{неп}}(t) - S_{\text{пр}}(t)]^2 dt \cdot$$

где T – длительность сообщения; $S_{\text{пер}}(t)$ – переданное сообщение; $S_{\text{пр}}(t)$ – принятое сообщение.

Оптимальный фильтр, удовлетворяющий этому выражению (фильтр Колмогорова – Винера), имеет амплитудно-частотную характеристику:

$$H(\omega) = \frac{G_c(\omega)}{G_c(\omega) + G_n(\omega)},$$

где $G_c(\omega)$, $G_n(\omega)$ – соответственно спектральные плотности мощности сигнала и помехи. Фазо-частотная характеристика должна быть линейной при любых сигналах и помехах, так как только в этом случае не будет линейных искажений.

Коэффициент передачи оптимального фильтра уменьшается с увеличением спектра помехи. Тем самым в оптимальном фильтре создаются условия, при которых подавление спектра помехи сопровождается меньшим подавлением спектра сигнала.

В радиотехнических системах связи типичным является случай, когда спектральная плотность мощности помехи намного меньше сигнала, а отношение сигнал/помеха много больше единицы. В таких случаях помеха представляет собой случайный процесс типа белого шума. Тогда квазиоптимальным оказывается или идеальный полосовой фильтр, или фильтр низких частот с амплитудно-частотной характеристикой $H_{\text{опт}}(\omega) = 1$ при ω , принадлежащей полосе, занимаемой сигналом, и $H_{\text{опт}}(\omega) = 0$ при ω вне полосы частот сигнала.

В другом случае возможна ситуация, когда спектры сигнала и помехи перекрываются, но помеха является более узкополосной по сравнению с сигналом и ее спектральная плотность мощности намного превышает спектральную плотность мощности сигнала (воздействие на сигнал мощных сосредоточенных помех). Тогда оптимальным оказывается фильтр с амплитудно-частотными характеристиками:

$H_{\text{опт}}(\omega) = 0$ при ω , принадлежащей полосе частот помехи, и

$H_{\text{опт}}(\omega) = 1$ при ω , принадлежащей полосе частот сигнала.

В этом случае в тракт приемника включают, кроме идеального полосового фильтра, идеальный заграждающий фильтр, обеспечивающий подавление помехи в ее полосе.

Для рассматриваемой радиотехнической системы морской связи приемник должен содержать фильтр, близкий по своим характеристикам к идеальному полосовому фильтру. Кроме указанного фильтра входная цепь приемника должна содержать фильтры, близкие по своим характеристикам к идеальным заграждающим фильтрам.

22.4. Количество информации при приёме дискретных сигналов радиотехнической системы связи

Для рассматриваемой радиотехнической системы морской связи найдем вероятность ошибки при оптимальных когерентном и некогерентном приемах двоичных АМ, ЧМ и ФМ-сигналов, если отношение сигнал/шум по напряжению $h = 1,47$. В таблице 22.1 приведены значения вероятностей ошибок, вычисленные для различных видов сигналов и приёма.

Количество информации, получаемой при приеме двух равновероятных сигналов с вероятностями ошибки, соответствующими табл. 22.1, можно вычислить по формуле

$$I = H_0 - H_1,$$

где H_0 – априорная энтропия; H_1 – апостериорная энтропия.

Для двух равновероятных сигналов априорная энтропия вычисляется по формуле

$$H_0 = -\sum_{i=1}^n p_{1i} \log p_{1i} = -0,5 \log 0,5 = 1 \text{ бит}.$$

Для сигнала с амплитудной модуляцией апостериорная энтропия составляет:

$$H_1 = -\sum_{i=1}^n p_{2i} \log p_{2i} = -0,017 \log 0,017 - 0,983 \log 0,983 = 1,124 \text{ бит}.$$

Количество информации определяется как

$$I = H_0 - H_1 = 1 - 0,124 = 0,876 \text{ бит}.$$

Аналогично вычислено количество информации для сигналов с частотной и фазовой манипуляциями.

Таблица 22.1

Вероятность ошибок для различных видов сигналов и приёма

Вид дискретных сигналов	Вероятность ошибки при когерентном приеме	Вероятность ошибки при некогерентном приеме
АМ	0,017	0,056
ЧМ	0,001	0,006
ОФМ	0,000003	0,000083

Результаты вычислений сведены в табл. 22.2. Результаты расчетов показывают, что предпочтительной является относительная фазовая манипуляция, позволяющая получить наибольшее количество информации при передаче сообщений.

Таблица 22.2

Количество информации для различных видов сигналов и приёма

Вид дискретных сигналов	Количество информации при когерентном приеме, бит	Количество информации при некогерентном приеме, бит
АМ	0,876	0,69
ЧМ	0,982	0,94
ОФМ	0,999	0,999

Таким образом, сравнение помехоустойчивости различных видов манипуляции можно произвести по вероятности ошибки при передаче определенных сигналов, по количеству передаваемой информации на два переданных сигнала. Кроме этого, очень часто сравнение помехоустойчивости производится по энергетическому выигрышу, под которым понимают различие в энергиях сигналов, обеспечивающих одинаковое значение вероятности ошибки при различных видах манипуляции, способах приема, кодирования и т.д. Энергетический выигрыш определяется по формуле

$$g_s = 10 \cdot \lg(E_{S1} / E_{S2}) \text{ при } P_{ош} = \text{const}, N_0 = \text{const},$$

где E_{S1} , E_{S2} – соответственно энергии сигналов $S_1(t)$ и $S_2(t)$.

Если $g_s < 0$ дБ, то имеются энергетические потери. Обычно значения энергий для расчета g_s находят по графическим зависимостям. Такие зависимости приведены на рисунке 22.1. Вероятность ошибки по оси ординат для удобства отложена в логарифмическом масштабе. Из графиков следует, что наибольшую помехоустойчивость имеет система с ФМ-сигналами, наименьшую – с АМ-сигналами. Проведя вычисления g_s , получим, что система с противоположными ФМ-сигналами дает энергетический выигрыш на 3 дБ (или в 2 раза) по сравнению с ортогональными ЧМ-сигналами и на 6 дБ (или в 4 раза) – по сравнению с АМ-сигналами. Физически это объясняется тем, что сигналы максимально отличаются друг от друга.

Для рассматриваемой радиотехнической системы морской связи определим энергетический выигрыш оптимального когерентного приема двоичных АМ-сигналов по сравнению с оптимальным некогерентным приемом этих же сигналов. Пусть $P_{ош} = 1,79 \cdot 10^{-2}$. Из точек пересече-

чения прямой $P_{ош} = 1,79 \cdot 10^{-2}$ зависимостей 4 (когерентный прием) и 5 (некогерентный прием) следует, что для обеспечения $P_{ош} = 1,79 \cdot 10^{-2}$ при когерентном приеме необходимо $h_2^2 = 10$, а при некогерентном приеме $h_2^2 = 14$. Энергия сигнала определится как $E = h^2 N_0$, энергетический выигрыш составит:

$$g_s = 10 \cdot \lg(14/10) = 1,46 \text{ дБ.}$$

Для рассматриваемой радиотехнической системы морской связи определим энергетический выигрыш оптимального когерентного приема двоичных ЧМ сигналов по сравнению с оптимальным некогерентным приемом этих же сигналов. Пусть $P_{ош} = 1,79 \cdot 10^{-2}$. Из точек пересечения прямой $P_{ош} = 1,79 \cdot 10^{-2}$ зависимостей 2 (когерентный прием) и 3 (некогерентный прием) следует, что для обеспечения $P_{ош} = 1,79 \cdot 10^{-2}$ при когерентном приеме необходимо $h_2^2 = 4,1$, а при некогерентном приеме $h_2^2 = 6,1$. Энергия сигнала определяется по формуле $E = h_2^2 \cdot N_0$, энергетический выигрыш составит

$$g_s = 10 \cdot \lg(6,1/4,1) = 1,725 \text{ дБ.}$$

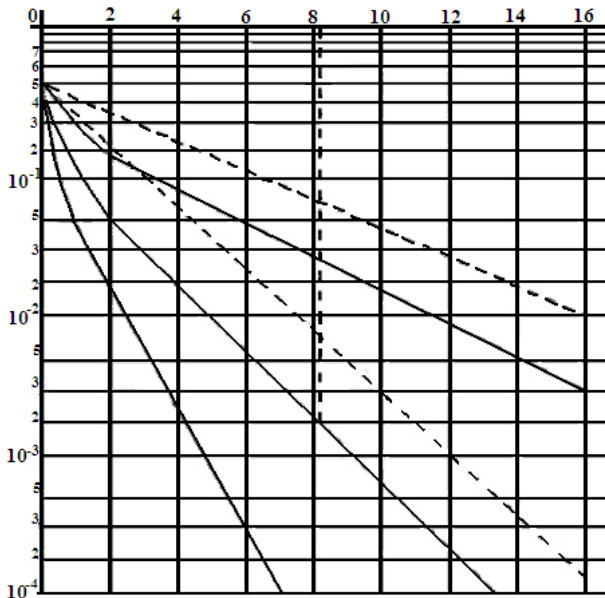


Рис. 22.1. Зависимость вероятности ошибок от отношения сигнал/помеха для различных видов модуляции (манипуляции)

Таким образом, для рассматриваемой радиотехнической системы морской связи энергетический выигрыш g_s оптимального когерентного

приема двоичных ЧМ-сигналов по сравнению с оптимальным некогерентным приемом этих же сигналов равен 1,725 дБ.

Из примера хорошо видно, что энергетические потери некогерентного приема незначительны (1,46 дБ для АМ и 1,725 для ЧМ), поэтому при практической реализации приемников АМ и ЧМ-сигналов целесообразно использовать более простой некогерентный приемник.

22.5. Количество информации при оптимальном приеме непрерывных сигналов

В качестве критерия оптимальности при приеме непрерывных сигналов принимают минимум среднеквадратического отклонения между переданным $u(t)$ и принятым $u_{\text{пр}}(t)$ сигналами [13]:

$$\min \varepsilon_{u(t)}^2 = \min [u_{\text{пр}}(t) - u(t)]^2.$$

Этот критерий учитывает не только помехи, но и искажения принимаемых сигналов. Минимально возможное значение среднеквадратической ошибки $\min \varepsilon_{u(t)}^2$ при заданных условиях передачи определяет потенциальную помехоустойчивость приема непрерывных сигналов. Физически $\min \varepsilon_{u(t)}^2$ означает мощность помехи, поэтому расчет потенциальной помехоустойчивости сводится к вычислению минимально возможной мощности помехи на демодуляторе. Абсолютное значение мощности помехи не может быть объективной характеристикой ее влияния на сигнал, так как надо учитывать еще и уровень сигнала. Поэтому оценку помехоустойчивости приема непрерывных сигналов можно произвести количеством информации, получаемой при приеме этих сигналов.

Для рассматриваемой радиотехнической системы морской связи непрерывный сигнал может быть частотно-модулированным, его можно представить в виде ряда Котельникова. Для этого сигнала ранее найдено отношение сигнал/помеха: $h = \sqrt{P_c/P_{\text{п}}} = 1,47$. В качестве помехи примем белый шум со спектральной плотностью $N_0 = 10^{-5} \text{ В}^2/\text{Гц}$; ширина полосы частот сигнала $\Delta f = 3,0 \text{ кГц}$. Требуется определить количество информации при передаче сигнала по каналу связи.

Определим мощность помехи на входе приемного устройства:

$$P_{\text{п}} = N_0 \Delta f = 10^{-5} \cdot 3,0 \cdot 10^3 = 3,0 \cdot 10^{-2} \text{ Вт}.$$

Среднеквадратическое отклонение помехи составит:

$$\sigma_{\text{п}} = \sqrt{P_{\text{п}}} = \sqrt{3,0 \cdot 10^{-2}} = 0,173 \text{ В}.$$

Найдем математическое ожидание напряжения сигнала:

$$U_c = m_c = \sqrt{E_c} = \sqrt{h^2 P_{\text{п}}} = \sqrt{1,47^2 \cdot 3,0 \cdot 10^{-2}} = 0,255 \text{ В}.$$

Вероятность ошибки при появлении одного отсчета на входе приемника составит:

$$P_{\text{ош}} = 0,5 \exp(-0,5h^2) = 0,5 \exp(-0,5 \cdot 1,47^2) = 0,017.$$

Это же можно записать и в виде

$$P_{\text{ош}} = \Phi^*((\infty - m_n)/\sigma_n) - \Phi^*((l - m_n)/\sigma_n) = 1 - \Phi^*(l/\sigma_n) = 0,017,$$

где l – порог обнаружения; $m_n = 0$ – математическое ожидание помехи.

Следовательно,

$$\Phi^*(l/\sigma_n) = 1 - 0,017 \approx 0,83.$$

По таблице нормального распределения [3] находим $\Phi^*(l/\sigma_n) \approx 0,83$. Тогда

$$l/\sigma_n = 0,96; l = 0,96 \cdot \sigma_n = 0,96 \cdot 0,173 = 0,166.$$

Вероятность отсутствия ошибки при приеме составит:

$$P_{\text{пр}} = 1 - P_{\text{ош}} = 1 - 0,017 \approx 0,83,$$

$$P_{\text{пр}} = \Phi^*((\infty - m_c)/\sigma_c) - \Phi^*((l - m_c)/\sigma_c) = 1 - \Phi^*((0,166 - 0,255)/\sigma_c).$$

После преобразования получим:

$$\Phi^*(-0,089/\sigma_c) = 1 - P_{\text{пр}} = 0,83.$$

По таблице нормального распределения находим:

$$\Phi^*(-0,089/\sigma_c) = \Phi^*(-0,96) \approx 0,83,$$

$$-0,089/\sigma_c = -0,96, \sigma_c = 0,093.$$

Количество информации на один отсчет определяется по формуле

$$I_{Y \leftrightarrow X} = \log \frac{\sqrt{\sigma_c^2 + \sigma_n^2}}{\sigma_n} = \log \frac{\sqrt{0,093^2 + 0,173^2}}{0,173} = \log 1,14 = 0,198 \text{ бит}.$$

22.6. Выигрыш в отношении сигнал/помеха

Другим способом определения помехоустойчивости приема непрерывных сигналов является вычисление отношения средних мощностей сигнала P_c и помехи P_n на выходе демодулятора

$$h_{\text{вых}} = P_c / P_n.$$

В любом демодуляторе отношение сигнал/помеха на выходе $h_{\text{вых}}$ зависит не только от качественных показателей демодулятора, но и от отношения сигнал/помеха на его входе $h_{\text{вх}}$. Помехоустойчивость систем передачи непрерывных сигналов оценивают выигрышем в отношении сигнал/помеха:

$$g = h_{\text{вых}} / h_{\text{вх}} = (P_{c \text{ вых}} / P_{n \text{ вых}}) / (P_{c \text{ вх}} / P_{n \text{ вх}}),$$

причем средние мощности помех на входе и выходе демодулятора определяются в полосе частот сигналов.

Выигрыш g показывает изменение отношения сигнал/помеха демодулятором. При $g > 1$ демодулятор улучшает отношение сигнал-помеха, при $g < 1$ получается не «выигрыш», а «проигрыш».

Расчетные формулы выигрыша для оптимального демодулятора при различных видах модуляции и помехе в виде аддитивного белого гауссовского шума приведены в табл. 22.3 где обозначены: $K_A^2 = 10 \lg P_{\max}/P$; $\alpha = \Delta f_c/F_m$ – коэффициент расширения полосы, показывающий, во сколько раз ширина спектра модулированного сигнала Δf_c превышает максимальную частоту модулирующего сигнала F_m ; M – коэффициент модуляции; t – индекс модуляции; K_A – коэффициент амплитуды модулирующего сигнала, представляющий собой отношение его максимальной мощности к средней и определяемый в логарифмических единицах по формуле (табл. 22.3).

Таблица 22.3

Расчетные формулы выигрыша оптимального демодулятора при различных видах модуляции

Вид	Выигрыш	Вил	Выигрыш
АМ	$\frac{M^2}{M^2 + K_A^2} \alpha_{AM}$	ЧМ	$\frac{3}{K_A^2} m_{ЧМ}^2 \alpha_{AM} \approx \frac{3}{4K_A^2} \alpha_{ЧМ}$
БМ	$\alpha_{БМ} = 2$	ФМ	$\frac{m_{ФМ}^2}{K_A^2} \alpha_{ФМ} \approx \frac{\alpha_{ФМ}^3}{4K_A^2}$
ОМ	$\alpha_{ОМ} = 1$	АИМ-АМ	$\frac{0,6}{4K_A^2} \alpha_{АИМ-АМ}^3$
АИМ	$\alpha_{АИМ}$	ФИМ-АМ	$\frac{3}{4K_A^2} \alpha_{ФИМ-АМ}^3$

Анализ формул таблицы 22.3 показывает, что для АМ максимальный выигрыш $g_{AM} = 0,666$ достигается при $M = 1$ и $K_A^2 = 2$. Практически всегда $M < 1$ и $K_A^2 = 2$, поэтому $g_{AM} < 0,666$, т.е. система с АМ не обеспечивает «выигрыш». Физически малый выигрыш для АМ объясняется тем, что большая часть мощности модулированного сигнала сосредоточена в несущей частоте, а полезная информация, создающая сигнал на выходе детектора, содержится в маломощных боковых коле-

баниях. Поэтому устранение несущей в АМ-сигнале (переход к БМ и ОМ) увеличивает выигрыш до значения $g = \alpha$.

В широкополосных видах модуляции (ЧМ, ФМ, ФИМ и др.) выигрыш может быть намного больше единицы и резко возрастает при расширении спектра модулирующего сигнала (кубическая зависимость от коэффициента расширения полосы α). В связи с этим для увеличения выигрыша следует повышать девиацию частоты угловых модуляций или уменьшать длительность импульса несущей импульсных модуляций.

Формулы выигрыша являются исходными как для определения качества приема непрерывных сигналов, так и для сравнения различных систем передачи по помехоустойчивости.

Для рассматриваемой радиотехнической системы морской связи определим отношение сигнал/помеха на выходе демодулятора АМ-сигналов, если отношение сигнал/помеха на его входе $h_{\text{вх}} = 1,47$ дБ. При этом параметры модуляции следующие: $M = 0,6$; коэффициент амплитуды $K_A = 14$ дБ. При приеме используется фильтр, согласованный со спектром сигнала.

Подставив в формулу выигрыша АМ параметры модуляции ($m_{\text{АМ}} = 0,6$ и $K_A^2 = 2$, $\alpha = 2$), получим:

$$g = \frac{0,6^2}{0,6^2 + 4} \cdot 2 = 0,165$$

Тогда $h_{\text{вых}} = g \cdot h_{\text{вх}} = 0,165 \cdot 1,47 = 0,243$.

В логарифмических единицах $h_{\text{вых}} = 10 \cdot \lg 0,243 \approx -6,1$ дБ, т.к. $g < 1$ — это свидетельствует о проигрыше демодулятора.

Для рассматриваемой радиотехнической системы морской связи определим отношение сигнал/помеха на выходе демодулятора ЧМ-сигналов, если отношение сигнал/помеха на его входе $h_{\text{вх}} = 1,47$ дБ. При этом параметры модуляции следующие: $m_{\text{ЧМ}} = 10$, коэффициент амплитуды $K_A = 14$ дБ, прием оптимальный.

Подставив в формулу выигрыша ЧМ параметры модуляции $m_{\text{ЧМ}} = 10$ и $K_A^2 = 10^{1,4} = 25,1$, $\alpha \approx 2 \cdot \Delta f / F_m = 2 \cdot 3 \cdot 10^3 / (3,4 \cdot 10^3) = 1,77$, получим:

$$g_{\text{ЧМ}} = \frac{3}{K_A^2} m_{\text{ЧМ}}^2 \alpha_{\text{АМ}} = \frac{3}{25,1} 10^2 \cdot 1,77 = 21,16 \cdot$$

Тогда $h_{\text{вых}} = g \cdot h_{\text{вх}} = 21,16 \cdot 1,47 = 31,1$. В логарифмических единицах $h_{\text{вых}} = 10 \cdot \lg 31,1 = 14,93$ дБ.

В общем случае при оптимальном приеме также имеет место проигрыш при применении АМ-сигналов по сравнению с ЧМ-сигналами для одинаковых условиях приема, т.е. равенстве мощностей модулированных сигналов и спектральной плотности мощности помех на входах приемников. При одних и тех же условиях отношение сигнал/помеха в

системе с ЧМ не менее чем в $(4,5 \cdot m_{\text{ЧМ}}^2)$ раз больше, чем в системе с АМ. На практике в системах с ЧМ применяют, как правило, индекс модуляции $m_{\text{ЧМ}}^2 \geq 5$, и тогда преимущество ЧМ по сравнению с АМ весьма значительное. Это преимущество получается за счет расширения полосы занимаемых частот.

Выигрыш при применении модулированных сигналов объясняется когерентным сложением в демодуляторе спектральных составляющих сигнала. Сложение составляющих помех осуществляется некогерентно.

Однако из изложенного не следует, что для различных видов модуляции и больших значений коэффициента a достигаются огромные значения выигрыша. Так, для ЧМ при $m_{\text{ЧМ}} = 60$ и $K_A^2 = 2$ можно получить выигрыш $g = 6,6 \cdot 10^5$. Казалось бы, задача обеспечения высокого качества передачи непрерывных сигналов решается достаточно просто увеличением ширины спектра модулированного сигнала. Но с расширением спектра растет мощность помехи на входе демодулятора и соответственно снижается отношение сигнал/помеха $h_{\text{вх}}$. При некотором пороговом значении $h_{\text{вх пор}}$ резко увеличивается уровень помех на выходе демодулятора, а отношение сигнал/помеха на выходе демодулятора скачкообразно уменьшается.

На рисунке 22.2 приведены кривые помехоустойчивости оптимальных демодуляторов при различных видах модуляции. Порогом помехоустойчивости демодулятора является минимальное отношение сигнал /помеха на его входе, ниже которого система передачи информации с заданной модуляцией теряет преимущество по помехоустойчивости. Пороговый эффект ограничивает возможность применения модуляций для повышения качества передачи непрерывных сигналов. Появление порога можно объяснить эффектом подавления сильным сигналом слабого в детекторе.

В надпороговой области сигнал превышает помеху, и в детекторе подавляется более слабая помеха. В подпороговой области помеха превышает сигнал, и в детекторе подавляется уже более слабый сигнал более сильной помехой. Пороговые явления начинают наблюдаться при равенстве пиковых значений сигнала и помехи. Обычно коэффициент амплитуды помехи $K_A \approx 3$, порог помехоустойчивости $h_{\text{вх пор}} \approx 10$ дБ (рис. 22.2).

При синхронном детектировании АМ, БМ, ОМ-сигналов пороговый эффект не наблюдается.

В настоящее время разработаны и внедрены методы снижения порога помехоустойчивости для систем передачи информации с ЧМ как наиболее распространенной. С этой целью используются следящие фильтры додетекторной обработки сигнала. Следящий фильтр имеет полосу пропускания меньше, чем ширина спектра модулированного сигнала, и следит за мгновенной частотой ЧМ-сигнала, которая изменяется

сравнительно медленно по закону модулирующего сигнала. Это позволяет уменьшить мощность помехи на выходе следящего фильтра примерно в $m_{\text{ЧМ}}$ раз, что ведет к понижению порога на 5–7 дБ.

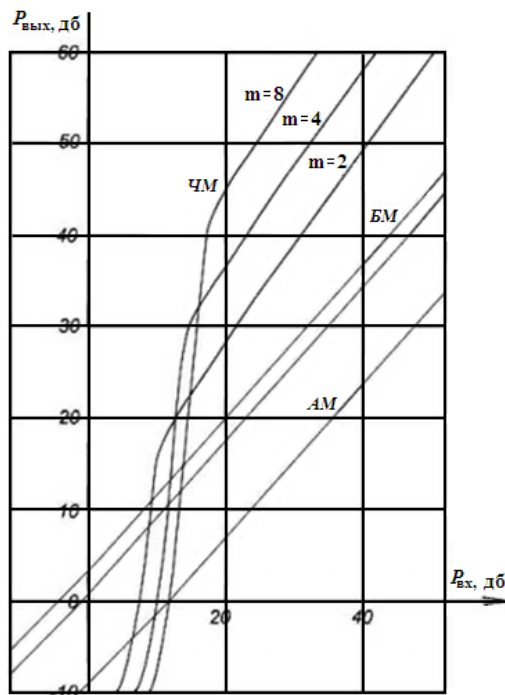


Рис. 22.2. Помехоустойчивость оптимальных демодуляторов при различных видах модуляции

Вместо следящего фильтра часто используют следящий гетеродин, частота которого изменяется синхронно с частотой принимаемого сигнала. При этом полоса пропускания фильтра промежуточной частоты остается неизменной: $\Pi \approx 2F_m$, где F_m – максимальная частота модулирующего сигнала.

Для рассматриваемой радиотехнической системы морской связи определим производительность источника дискретных сообщений при скорости передачи информации $B = 50$ Бод пятиэлементным двоичным кодом.

Для равновероятных букв вероятность одной буквы русского алфавита $P(a_i) = 1/32$. В одной букве содержится $I(a_i) = -\log 1/32 = 5$ бит информации. При коэффициенте избыточности русского текста $\nu = 0,5$ энтропия текста определится как $H(A) = 2,5$ бит на букву. Длительность

передачи одного символа находится по формуле $t_1 = 1/B$, а длительность передачи пяти символов и расстояния между буквами, т.е. одной буквы, по формуле $t_5 = 7,5/B = 7,5/50 = 0,15$ с. Тогда производительность дискретного источника составит:

$$V(A) = H(A)/t_{cp} = 2,5/0,15 = 16,67 \text{ бит/с.}$$

22.7. Пропускная способность каналов радиотехнической системы связи

Скорость передачи информации определяется средним количеством информации, которое получают на выходе канала связи в единицу времени. При вычислении скорости передачи информации в канале связи с помехами необходимо учитывать потери информации из-за действия помех.

Для дискретного канала связи скорость передачи информации находится по формуле

$$R_{дк} = [H(U) - H_{пот}(U)] / t_{cp};$$

для непрерывного канала связи скорость передачи информации определяется как

$$R_{нк} = 2F_m[h(U) - h_{пот}(U)],$$

где $H(U)$ – энтропия передаваемого дискретного первичного сигнала; $h(U)$ – дифференциальная энтропия передаваемого непрерывного сигнала; $h_{пот}(U)$ – соответственно энтропии потерь в канале для дискретного и непрерывного первичных сигналов; t_{cp} – средняя длительность дискретного первичного сигнала; F_m – максимальная частота спектра непрерывного первичного сигнала.

Наибольшее значение скорости R передачи информации по каналу связи при заданных ограничениях называют пропускной способностью канала связи. Под заданными ограничениями понимают тип канала связи (дискретный или непрерывный), характеристики сигналов и помех.

Пропускная способность двоичного дискретного канала определяется так:

$$C_{дк} = B[1 + p \cdot \log p + (1 - p) \log(1 - p)].$$

Для рассматриваемой радиотехнической системы морской связи определим пропускную способность при $B = 50$ Бод и вероятностях ошибок:

для АМ: $P_{AM} = 2 \cdot 10^{-2}$,

$$C_{AM} = 50[1 + 2 \cdot 10^{-2} \log 2 \cdot 10^{-2} + 0,98 \log 0,98] = 47,12 \text{ бит/с;}$$

для ЧМ: $P_{ЧМ} = 1,57 \cdot 10^{-3}$,

$$C_{AM} = 50[1 + 1,57 \cdot 10^{-3} \log 1,57 \cdot 10^{-3} + 0,99843 \log 0,99843] = 49,15 \text{ бит/с;}$$

для ФМ: $P_{\text{ФМ}} = 1,62 \cdot 10^{-5}$,

$$C_{\text{АМ}} = 50[1 + 1,62 \cdot 10^{-5} \log 1,62 \cdot 10^{-5} + 0,999 \log 0,999] \approx 50 \text{ бит/с.}$$

Для непрерывного канала максимальная скорость передачи информации достигается для гауссовского канала с постоянными параметрами, при условии, что для сигнала распределение вероятностей мгновенных значений при ограниченной средней мощности близко к нормальному распределению.

Расчетная формула пропускной способности гауссовского канала, которая выведена Шенноном и носит его имя, имеет вид:

$$C_{\text{НК}} = F \cdot \log(1 + P_{\text{с}}/P_{\text{п}}),$$

где F – ширина полосы пропускания канала; $P_{\text{с}}/P_{\text{п}}$ – отношение сигнал/помеха по мощности.

Для рассматриваемой радиотехнической системы морской связи определим пропускную способность канала передачи информации с учетом ширины полосы пропускания для сигнала с частотной модуляцией $F \approx 2\Delta f = 6,2$ кГц и отношения сигнал/помеха по мощности $P_{\text{с}}/P_{\text{п}} = 1,47^2 = 2,16$. Тогда

$$C_{\text{НК}} = F \cdot \log(1 + P_{\text{с}}/P_{\text{п}}) = 6200 \cdot \log(1 + 2,16) = 10290 \text{ бит/с.}$$

Для рассматриваемой радиотехнической системы морской связи величина пропускной способности должна быть намного больше производительности источника сообщений и удовлетворять основной теореме Шеннона. Эта теорема формулируется так: если производительность источника меньше пропускной способности канала, то существует способ кодирования (преобразования сообщения в сигнал на входе) и декодирования (преобразования сигнала в сообщение на выходе канала), при котором вероятность ошибочного декодирования может быть сколь угодно мала.

IX. ТЕОРЕТИКО-ИНФОРМАЦИОННАЯ КОНЦЕПЦИЯ КРИПТОЗАЩИТЫ СООБЩЕНИЙ В ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

23. ОСНОВЫ КРИПТОЗАЩИТЫ СООБЩЕНИЙ В СИСТЕМАХ СВЯЗИ

23.1. Основные понятия криптографии

Информация, которая нуждается в защите, называется защищаемой (приватной, конфиденциальной, секретной). Принято говорить, что такая информация содержит тайну (государственную, коммерческую, врачебную, личную, тайну следствия и т.п.).

Во всех случаях имеется круг законных пользователей информации. Кроме того, могут существовать люди или группы людей, которые заинтересованы в том, чтобы использовать информацию в своих целях, а значит – во вред законным пользователям. Для краткости эти незаконные пользователи называются противниками. Действия противника представляют угрозу для законного пользователя (угрозу разглашения, подмены, имитации, уничтожения информации и т.п.).

Возможны три подхода к защите информации [17]:

- создание абсолютно надежного канала связи, совершенно недоступного для незаконных пользователей;
- передача информации по общедоступному каналу, сохраняя в тайне сам факт передачи;
- передача информации по общедоступному каналу в такой форме, чтобы воспользоваться ею мог только адресат.

Первый способ при современном развитии техники практически нереализуем или сопряжен с чрезвычайно большими затратами.

Реализацией второго способа занимается стеганография. Известно немало примеров стеганографического сокрытия информации: от нанесения сообщения на обритуемую голову раба, которого посылали к адресату

ту после того, как волосы отрастали, до использования симпатических чернил, которыми секретное сообщение вписывается между строк другого сообщения.

В настоящее время используется, например, прием, состоящий в выборочной записи символов сообщения в младшие разряды цифрового изображения – на качество его визуального восприятия эти искажения практически не влияют, а найти «спрятанные» данные в больших массивах, которыми представляются изображения, не так легко.

Третий способ – криптографическая защита информации путем шифрования – в настоящее время является одним из самых эффективных. Следует различать понятия кода и шифра.

Задачей криптографии является разработка методов преобразования сообщений, затрудняющих (в идеале – исключаящих) извлечение противником информации из перехватываемых сообщений. По открытому каналу связи при этом передается криптограмма (шифртекст) – результат преобразования сообщения с помощью шифра (криптографического алгоритма). Наблюдение криптограммы для противника, не имеющего ключа, является бесполезным с точки зрения получения информации.

Зашифрованием называется процесс преобразования открытого сообщения в зашифрованное (криптограмму) с помощью шифра, а расшифрованием – обратный процесс преобразования зашифрованных (закрытых) данных в открытые с помощью шифра (шифрование – общее наименование для зашифрования и расшифрования).

Дешифрованием (вскрытием, взломом) называют процесс преобразования зашифрованных данных в открытые при неизвестном (частично или полностью) шифре. Решением всех перечисленных задач занимается наука, называемая криптологией. Отрасль криптологии, занимающаяся разработкой методов шифрования, называется криптографией, а отрасль, занимающаяся разработкой методов взлома – криптоанализом.

Помимо получения информации из перехваченного сообщения противник может преследовать и другие цели, например, он может попытаться изменить содержание сообщения. В защищенной системе, таким образом, получатель сообщения должен иметь возможность проверить его подлинность и целостность. Защита от навязывания ложных данных называется имитозащитой. Для этого к криптограмме добавляется имитовставка, представляющая собой последовательность данных фиксированной длины, полученную по определенному алгоритму из открытых данных и ключа. Получатель зашифрованного сообщения может проверить, соответствует ли имитовставка содержанию расшифрованного сообщения.

Криптостойкостью называется характеристика шифра, определяющая его способность противостоять дешифрованию. Обычно количест-

венной мерой криптостойкости считается время, необходимое для дешифрования [18].

Проблема криптозащиты существует многие сотни лет, что естественно, если вспомнить, что у любого человека есть свои секреты. Математические основы криптографии были разработаны К. Шенноном [19]. В общем случае секретная система связи может быть представлена структурной схемой, показанной на рис. 23.1.

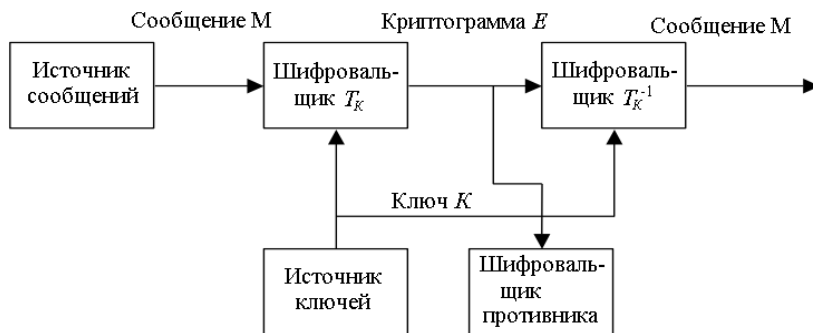


Рис. 23.1. Структура секретной системы

Сообщение M , вырабатываемое источником сообщений, поступает к шифровальщику. Шифровальщиком может быть как человек, так и устройство. От источника ключей к шифровальщику поступает ключ K . Шифровальщик на основе ключа осуществляет преобразование сообщения в криптограмму $E = f(M, K)$.

Удобнее понимать это выражение не как функцию двух переменных, а как отображение $E = T_K\{M\}$, при этом ключ играет роль параметра этого отображения. Выбранный ключ должен быть каким-то образом передан (сообщен) шифровальщику на приемной стороне системы, который при помощи ключа расшифровывает сообщение. Чтобы секретная система была пригодной для передачи информации, необходимо иметь возможность восстановить по криптограмме и ключу исходное сообщение, т.е. должно существовать (единственное) обратное отображение, так что

$$M = T_K^{-1}\{E\} = T_K^{-1}\{T_K[M]\},$$

$$T_K^{-1}T_K = I,$$

где I – тождественное отображение.

Предполагается, что множество всех сообщений конечно и сообщения M_1, \dots, M_n вырабатываются источником с вероятностями q_1, \dots, q_n . Аналогично источник ключей вырабатывает ключи K_1, \dots, K_m с вероятностями p_1, \dots, p_m .

В современной криптографии принято считать, что противнику известно множество отображений $T_{K_i}, i = \overline{1, m}$ и набор априорных вероятностей p_1, \dots, p_m .

Это предположение является пессимистическим, но, в конечном счете, реалистическим, так как рано или поздно всякая секретная система может быть раскрыта [19]. Если противник не сможет взломать алгоритм шифрования, даже зная, как он устроен, то тем более это окажется ему не по силам без такого знания.

В настоящее время используются следующие криптографические методы защиты [18]:

- методы замены (подстановки);
- методы на основе датчика псевдослучайных чисел;
- методы перемешивания (алгоритмические);
- системы с открытым ключом.

23.2. Метод замены

Для примера используем буквы русского алфавита (без буквы «ё») плюс пробел, обозначаемый символом $_$. Для удобства приведем русский алфавит в виде таблицы, содержащей также номера букв.

Буква	А	Б	В	Г	Д	Е	Ж	З	И	Й	К
№	1	2	3	4	5	6	7	8	9	10	11
Буква	Л	М	Н	О	П	Р	С	Т	У	Ф	Х
№	12	13	14	15	16	17	18	19	20	21	22
Буква	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	$_$
№	23	24	25	26	27	28	29	30	31	32	33

В качестве примера рассмотрим шифр простой подстановки. В основе шифрования этим методом лежит алгебраическая операция над алфавитом сообщения, называемая подстановкой.

Подстановкой называется взаимно однозначное отображение конечного множества M на себя. Это означает, что каждому элементу множества (например, символу алфавита) ставится в соответствие один и только один элемент этого же множества. Очевидно, для задания конкретной подстановки природа элементов не имеет значения, важно лишь их количество N (называемое степенью подстановки). Можно поэтому в качестве множества рассматривать множество целых чисел $M = \{1, 2, 3, \dots, N\}$.

Любая подстановка может быть описана матрицей $2 \times N$, в которой первая строка содержит числа $1, 2, \dots, N$, а вторая строка состоит из тех же чисел, расположенных в порядке, определяемом подстановкой, например, если число i после подстановки получает значение k_i , то можно записать

$$S = \begin{pmatrix} 1 & 2 & \dots & N \\ k_1 & k_2 & \dots & k_N \end{pmatrix},$$

где $k_i \in M \forall i$.

Очевидно, если последовательно применить две подстановки S_1 и S_2 , то их результат снова будет подстановкой, равной композиции (произведению) исходных подстановок $S = S_1 \times S_2$. Множество подстановок одинаковой степени образует группу относительно «умножения», определенного таким образом. Отсюда следует, что для каждой подстановки существует обратная операция. Нейтральным элементом группы подстановок служит тривиальная подстановка, оставляющая на месте все элементы множества.

Пусть открытый текст представляет собой фразу «открытый текст», а подстановка задана таблицей

$$S = \begin{pmatrix} 1 & 2 & 3 & \dots & 32 & 33 \\ 33 & 32 & 31 & \dots & 2 & 1 \end{pmatrix}.$$

Шифртекст имеет вид «тоцреоечаоыщпо». Очевидно, что ключом к этому шифру является сама подстановка S . Количество всевозможных подстановок для алфавита из N символов равно $N!$, что делает раскрытие ключа методом полного перебора проблематичным. На первый взгляд, криптостойкость такого шифра должна быть очень высокой. Однако недостатком шифра простой подстановки является то, что статистические свойства текста (частоты появления букв) при шифровании сохраняются, благодаря чему криптостойкость шифра на самом деле низка. Если в распоряжении дешифровальщика окажется достаточно длинный шифртекст (несколько десятков знаков), то шифр может быть взломан за несколько минут путем подсчета частот букв и сравнения с известными статистическими характеристиками русского (английского и т.д.) текста.

23.3. Методы шифрования на основе датчика псевдослучайных чисел

Принцип шифрования заключается в генерировании псевдослучайной последовательности, называемой гаммой, и наложении гаммы на открытое сообщение некоторым обратимым образом.

Например, если открытые данные и гамма представлены в двоичном коде, то подходящей является операция сложения по модулю 2;

если исходным является русский текст, то в качестве гаммы можно использовать последовательность независимых псевдослучайных целых чисел от 1 до 33 с равными вероятностями, а операция наложения будет заключаться в сложении номеров букв с числами гаммы по модулю 33 и т.д. Расшифрование требует повторного генерирования гаммы и применения обратной операции (в рассмотренных случаях вычитания по соответствующему модулю). Ключом в данном случае может служить некоторый параметр датчика псевдослучайных чисел.

Линейный датчик псевдослучайных чисел реализуется на цифровой вычислительной машине и описывается рекуррентной формулой

$$g_{i+1} = [Ag_i + C] \bmod M,$$

где g_i , $i = 1, 2, 3, \dots$ – последовательность псевдослучайных чисел,

g_0 – стартовое значение,

A и C – некоторые константы. Обычно принимается $M = 2^b$, где b – разрядность машины.

Если гамма имеет период больший, чем длина зашифрованного сообщения, и если не известна никакая часть исходного текста, то этот шифр можно раскрыть только прямым перебором возможных ключей, поэтому криптостойкость определяется размером ключа [18].

Шифрование на основе датчиков псевдослучайных чисел наиболее часто применяется в программной реализации криптозащиты данных, так как, с одной стороны, он достаточно прост для программирования, а с другой стороны, обладает высокой криптостойкостью.

Одной из форм реализации описанного метода является метод «одноразового блокнота». Суть этого метода состоит в том, что у шифровальщиков на передающей и приемной сторонах имеются два идентичных блокнота с одинаковой гаммой, содержащей случайные независимые равновероятные числа. Эта гамма, являющаяся ключом, используется только один раз, после чего соответствующий лист блокнота вырывается и уничтожается. Если эти условия соблюдены, шифр является абсолютно криптостойким, т.е. не может быть взломан в принципе [20]. Однако этот метод очень сложно реализовать, так как трудно снабдить всех возможных получателей шифртекстов идентичными шифровальными блокнотами и обеспечить их однократное использование. Абсолютно стойкие шифры применяются только в системах секретной связи с небольшим объемом передаваемой информации, обычно для передачи особо важной государственной информации [21].

23.4. Методы перемешивания

Методы перемешивания основаны на работе К. Шеннона [19], в которой задачи криптологии рассматривались с информационной точки

зрения. Прежде чем перейти к конкретным системам шифрования на основе перемешивания, приведем основные положения работы Шеннона.

Основные вопросы, которые представляют интерес с теоретической точки зрения, состоят в следующем. Насколько устойчива система, если шифровальщик противника располагает всеми необходимыми средствами для криптоанализа и неограниченным временем? Имеет ли криптограмма единственное решение (даже если оно может быть найдено за практически неприемлемое время), а если нет, то сколько решений она имеет? Какой объем шифрованного текста нужно перехватить, чтобы решение стало единственным? Существуют ли секретные системы, для которых нельзя найти единственное решение независимо от объема перехваченной криптограммы? Существуют ли системы, в которых противник вообще не получает никакой информации независимо от объема перехваченного шифртекста? Рассмотрение этих вопросов основывается на понятиях теории информации.

Пусть имеется конечное множество сообщений M_1, \dots, M_n с априорными вероятностями $P(M_1), P(M_2), \dots, P(M_n)$ и эти сообщения преобразуются в криптограммы E_1, E_2, \dots, E_n , так что $E = T_i\{M\}$.

После перехвата криптограммы E шифровальщик противника может вычислить условные (апостериорные) вероятности различных сообщений $P_E(M) = P(M | E)$. Тогда совершенная секретная система удовлетворяет следующему условию: для всех E все апостериорные вероятности равны априорным

$$P_E(M_i) = P(M_i) \quad \forall i \quad \forall E.$$

В этом и только в этом случае перехват шифртекста не дает противнику никакой информации.

По теореме Байеса

$$P_E(M) = \frac{P(M)P_M(E)}{P(E)}, \quad (23.1)$$

где $P(M)$ – априорная вероятность сообщения M ;

$P_M(E)$ – условная вероятность криптограммы E при условии, что выбрано сообщение M ;

$P(E)$ – вероятность получения криптограммы E ;

$P_E(M)$ – апостериорная вероятность сообщения M при условии, что перехвачена криптограмма E .

Из выражения (23.1) следует, что для совершенной секретности необходимо и достаточно, чтобы выполнялось равенство $P_M(E) = P(E)$ для любых сообщения M и криптограммы E .

Очевидно, различных криптограмм должно быть не меньше, чем различных сообщений, так как при любом ключе сообщению M должна однозначно соответствовать криптограмма E , которой, в свою очередь,

однозначно соответствует сообщение M . Кроме того, число различных ключей должно быть также не меньше, чем различных сообщений. Например, совершенная система получается, если число сообщений равно числу ключей и числу криптограмм, причем все ключи равновероятны [19].

В терминах теории информации секретная система содержит два источника неопределенности: источник сообщений с энтропией

$$H(M) = -\sum_{i=1}^n P(M_i) \log P(M_i),$$

и источник ключей с энтропией

$$H(K) = -\sum_{j=1}^m P(K_j) \log P(K_j).$$

Количество информации источника не может быть больше, чем $\log n$ (это количество соответствует равновероятным сообщениям). Эта информация может быть полностью скрыта, только если неопределенность ключа не меньше $\log n$. Таким образом, неопределенность источника ключей устанавливает предел – максимальное количество информации, которое может быть скрыто при помощи ключей данного источника.

Если источник порождает последовательности бесконечной длины, то никакой конечный ключ не дает совершенной секретности. Пусть длина сообщения равна N , а длина ключа M . Тогда для совершенной секретности требуется

$$N \log n \leq M \log m,$$

где n и m – объемы алфавитов для сообщений и ключей соответственно.

Совершенные системы применяются на практике для засекречивания сравнительно коротких сообщений и в тех случаях, когда полная секретности придается чрезвычайное значение.

В качестве теоретической меры секретности Шеннон предложил использовать ненадежность. Имеются две ненадежности:

ненадежность ключа

$$H_E(K) = H(K | E) = -\sum_{E,K} P(E, K) \log P(K | E), \quad (23.2)$$

и ненадежность сообщения

$$H_E(M) = H(M | E) = -\sum_{E,K} P(E, M) \log P(M | E), \quad (23.3)$$

где $P(E, M)$ – совместная вероятность криптограммы E и сообщения M , $P(E, K)$ – совместная вероятность криптограммы E и ключа K , $P(M | E)$ и $P(K | E)$ – апостериорные вероятности для сообщения и ключа при усло-

вии перехвата криптограммы E . Суммирование в (23.2) и (23.3) проводится по всем возможным криптограммам и ключам (соответственно по всем возможным криптограммам и сообщениям) определенной длины, поэтому ненадежности зависят от числа N перехваченных букв криптограммы. Постепенное убывание ненадежностей с ростом N соответствует увеличению сведений о ключе и сообщении, имеющихся у шифровальщика противника. Если ненадежность становится нулевой, это означает, что единственное сообщение (или единственный ключ) имеет апостериорную вероятность 1, а все остальные – нулевые апостериорные вероятности. Шеннон доказал, что в принципе можно построить систему, в которой ненадежности не стремятся к нулю при $N \rightarrow \infty$, и даже «строго идеальную систему», в которой $H_E(K) = H(K)$. Для сообщений на естественных языках характерны неравновероятность и зависимость букв исходного текста, поэтому для них построение идеальной системы хотя и возможно, но осуществить его практически очень сложно [19].

Статистическая зависимость символов (букв) естественного языка позволяет раскрывать многие типы шифров. Например, шифры на основе подстановок раскрываются на основе простого подсчета частот встречаемости различных букв в криптограмме и сравнения их с известными для данного языка частотами. Для некоторых шифров могут потребоваться более тонкие методы статистического анализа. В любом случае статистический анализ криптограмм строится на основе некоторых статистик. Под статистикой понимается функция наблюдаемой криптограммы, значение которой позволяет сделать какие-либо более или менее правдоподобные выводы относительно использованного ключа. Удачный выбор подходящих статистик определяет успех криптоаналитика при взломе шифра, многократно сужая возможный круг ключей до того предела, когда уже можно установить правильный ключ путем перебора.

Для того чтобы затруднить раскрытие шифров статистическими методами, Шеннон предложил использовать две идеи, которые он назвал «распылением» и «запутыванием». Смысл «распыления» состоит в таком преобразовании текста, при котором статистические связи между его элементами становятся трудно наблюдаемыми, хотя избыточность сообщения при этом и не изменяется. Идея «запутывания» (перемешивания) заключается в том, чтобы сделать соотношения между простыми статистиками в пространстве криптограмм и простыми подмножествами в пространстве ключей весьма сложными и беспорядочными [19]. На этих идеях основано использование составных шифров, состоящих часто в последовательном применении простых подстановок и перестановок.

В шифрах подстановки буквы сообщения заменяются другими буквами, поэтому статистические свойства текста сохраняются, но относятся

ся теперь к буквам другого алфавита. В шифрах перестановки буквы сообщения остаются теми же, но изменяется их расположение в тексте. Комбинирование поочередно применяемых подстановок и перестановок дает возможность создавать весьма стойкие шифры.

Один из примеров реализации такого подхода – государственный стандарт шифрования США, известный под названием DES (Data Encryption Standard). Алгоритм шифрования является открытым; секретным при каждом использовании алгоритма является только ключ. Ключ длиной в 56 бит обеспечивает высокий уровень стойкости шифра, так как общее количество ключей составляет около $7,2 \cdot 10^{16}$. Открытый текст и криптограмма при этом являются двоичными 64-значными последовательностями. Криптоалгоритм DES представляет собой суперпозицию из 16 последовательных шифроциклов, в каждом из которых происходят подстановки и перестановки в четырехбитовых группах, при этом в каждом цикле используются 48 бит ключа, которые выбираются из полного ключа длиной в 56 бит [18]. По утверждению Национального бюро стандартов США, метод DES имеет высокую криптостойкость, которая делает его раскрытие дороже получаемой при этом прибыли, экономичен в реализации и эффективен в смысле быстрейшего действия.

Наиболее существенным недостатком считается слишком короткий ключ: для дешифрования требуется перебрать 2^{56} (или $7,2 \cdot 10^{16}$) ключей, что достаточно много для современной аппаратуры, но может оказаться преодолемым в ближайшем будущем. Впрочем, метод допускает простую модификацию в виде последовательного применения к сообщению нескольких циклов шифрования с различными ключами: например, уже при трех ключах для дешифрования требуется выполнение около 2^{168} (т.е. $3,7 \cdot 10^{50}$) операций.

Отечественный стандарт шифрования данных, предусмотренный ГОСТ 28147-89, предназначен для шифрования данных аппаратным или программным путем [18]. Стандарт предусматривает различные режимы шифрования, но в любом случае используется 256-разрядный двоичный ключ. Двоичные данные, подлежащие зашифрованию, разбиваются на блоки по 64 разряда, над которыми выполняются преобразования, включающие суммирование по модулям 2, 2^{32} и $2^{32}-1$, подстановки и циклические сдвиги. Стандарт обладает всеми достоинствами алгоритма DES и в то же время свободен от его недостатков, однако имеет собственный существенный недостаток – очень низкое быстрейшее действие при программной реализации [18].

23.5. Криптосистемы с открытым ключом

Системы с открытым ключом названы так потому, что ключ, используемый при зашифровании данных, не является секретным и может

быть, например, опубликован в средствах массовой информации. Также несекретным является алгоритм зашифрования. Защита данных обеспечивается тем, что для расшифрования необходим другой (секретный) ключ, причем он не может быть определен по открытому ключу зашифрования. Алгоритмы шифрования с открытым ключом называют поэтому несимметричными алгоритмами [20]. Наиболее известен метод шифрования с открытым ключом RSA.

Согласно методу RSA, для генерирования ключей необходимо выполнить следующие действия [18].

1. Выбрать два больших простых числа p и q .
2. Определить их произведение $n = pq$.
3. Выбрать число d , взаимно простое с числом $(p - 1)(q - 1)$.
4. Определить число e , для которого выполняется условие

$$ed \bmod [(p - 1)(q - 1)] = 1.$$

5. Назвать открытым ключом числа e и n , а секретным ключом числа d и n .

Для применения полученных ключей открытое сообщение необходимо закодировать числами от 0 до $n - 1$. Каждое такое число $M(i)$ зашифровывается при помощи открытого ключа по формуле

$$C(i) = [M(i)]^e \bmod(n). \quad (23.4)$$

Для расшифрования используется формула с секретным ключом

$$M(i) = [C(i)]^d \bmod(n). \quad (23.5)$$

Пример. Рассмотрим в качестве простого примера алгоритм RSA, основанный на очень малых числах p и q . Предположим, что шифрованию подлежит сообщение на русском языке [18]. Буквы сообщения можно представить числами от 0 до 32 (см. разд. 23.2).

Тогда за n можно принять число 33, а за простые числа p и q соответственно 3 и 11. Итак, согласно описанному алгоритму:

- 1) выберем два простых числа $p = 3$ и $q = 11$;
- 2) найдем $n = 3 \cdot 11 = 33$;
- 3) за число d взаимно простое с числом $(p - 1)(q - 1) = 20$, примем число 3;
- 4) соотношению $e \cdot 3 \bmod(20) = 1$ удовлетворяют числа 7, 27, 47, ..., выберем $e = 7$.

Итак, открытым ключом для зашифрования является пара чисел $e = 7$ и $n = 33$, а закрытым (секретным) ключом для расшифрования – пара чисел $d = 3$ и $n = 33$.

Зашифруем слово ДОМ. Буквам Д, О и М соответствуют числа 5, 15 и 13. Используя открытый ключ, получим на основании (23.4) криптограмму, состоящую из чисел:

$$C_1 = 5^7 \bmod(33) = 78125 \bmod(33) = 14;$$

$$C_2 = 15^7 \bmod(33) = 170859375 \bmod(33) = 27;$$

$$C_3 = 13^7 \bmod(33) = 62748517 \bmod(33) = 7.$$

Для расшифрования криптограммы {14, 27, 7} воспользуемся формулой (23.5) и секретным ключом:

$$M_1 = 14^3 \bmod(33) = 2744 \bmod(33) = 5;$$

$$M_2 = 27^3 \bmod(33) = 19683 \bmod(33) = 15;$$

$$M_3 = 7^3 \bmod(33) = 343 \bmod(33) = 13.$$

Легко видеть, что в результате расшифрования получилось исходное открытое сообщение ДОМ. Следует отметить, что на практике применяются настолько большие числа p и q , что, зная e и n (открытый ключ), невозможно найти d за приемлемое время, так как в настоящее время не только не известен достаточно эффективный (полиномиальный) алгоритм разложения большого числа на простые множители, но и сам вопрос о существовании таких алгоритмов (а следовательно, о возможности взлома систем с открытым ключом в будущем) остается открытым [18, 21]. Тем не менее нельзя исключить открытие в будущем эффективных алгоритмов определения делителей целых чисел (факторизации), вследствие чего метод шифрования с открытым ключом станет абсолютно бесполезным. Пока этого не произошло, метод RSA имеет важные преимущества перед другими криптосистемами, такие как очень высокая криптостойкость и простота аппаратной и программной реализации.

23.6. Цифровая подпись

Одной из важных задач, связанных с передачей документов по каналам связи или с пересылкой их на машинных носителях в электронной форме, является задача аутентификации (подтверждения подлинности). Для документа в обычной (бумажной) форме эта проблема решается за счет жесткой связи информации с носителем (бумагой). Документ в электронной форме такой связи не имеет и иметь не может.

При передаче секретных документов (военного или дипломатического содержания) весьма вероятным является перехват документа противником, который может либо создать и переслать вместо него подложный документ, либо изменить содержимое документа законного источника. Чтобы обнаружить подмену, в зашифрованное сообщение встраивается так называемая имитовставка, которая представляет собой последовательность, полученную по определенному алгоритму на основе всего текста открытого сообщения. Получатель после расшифрования криптограммы подвергает полученный открытый текст повторной обработке тем же алгоритмом и сравнивает полученную имитовставку с

принятой. Если совпадения нет, принятое сообщение считается ложным.

Совершенно другая ситуация может иметь место при обмене информацией коммерческого характера. Здесь возможна ситуация, когда партнеры по информационному обмену не доверяют друг другу и являются в каком-то смысле противниками. Один из партнеров может изготовить документ, зашифровать его и заявить, что получил его от партнера.

Для такого случая необходимо применять схему шифрования следующего вида. Передающий абонент зашифровывает подпись с помощью своего секретного ключа, а получатель расшифровывает ее своим несекретным ключом. Несекретный ключ может представлять собой набор проверочных соотношений, позволяющих установить подлинность подписи, но не восстановить секретный ключ. Таким образом, никто, кроме законного автора документа, не в состоянии сгенерировать правильную подпись.

ЗАКЛЮЧЕНИЕ

Современные системы связи становятся все более совершенными, предоставляя пользователю широкий выбор различных телекоммуникационных услуг. С точки зрения специалиста в области телекоммуникации это означает, что системы связи все более усложняются и их эксплуатация требует не только некоторых практических навыков, но и глубоких знаний в области теории связи.

Настоящее пособие содержит основополагающие знания по современным технологиям преобразования и передачи сообщений и сигналов по каналам связи. Эти знания, являясь базовыми, позволят видеть и понимать состояние и тенденции развития современных телекоммуникационных систем, ориентироваться во всем их многообразии.

Следует отметить, что даже при использовании самых современных технологий базирующихся на высокопроизводительных вычислительных системах, повышение эффективности современных и перспективных систем связи ставят перед теорией новые задачи решения постоянно возникающих проблем. Поэтому теорию электрической связи ни в коей мере нельзя назвать устоявшейся и тем более завершённой, она постоянно развивается и совершенствуется.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ*

1. Котельников, В.А. О пропускной способности «эфира» и проволоки в электросвязи // **Материалы по радиосвязи к I Всесоюзному съезду по вопросам технической реконструкции связи / В.А. Котельников.** – М.: Всесоюзный энергетический комитет, 1933.
2. Зюко, А.Г. **Теория электрической связи / А.Г. Зюко, Д.Д. Кловский, В.И. Коржик, М.В. Назаров; под ред. Д.Д. Кловского.** – М.: Радио и связь, 1999. – 432 с.
3. Бронштейн, И.Н. Справочник по математике для инженеров и учащихся втузов / И.Н. Бронштейн, К.А. Семендяев. – 13-е изд., испр. – М.: Наука, Гл. ред. физ.-мат. лит., 1986. – 544 с.
4. Борисов, В.А. Радиотехнические системы передачи информации / В.А. Борисов, В.В. Калмыков, Я.М. Ковальчук; под ред. В.В. Калмыкова. – М.: Радио и связи, 1990. – 304 с.
5. Панфилов, И.П. Теория электрической связи / И.П. Панфилов, В.Е. Дырда. – М.: Радио и связь, 1991. – 344 с.
6. Теплов, Н.Л. Теория передачи сигналов по электрическим каналам связи / Н.Л. Теплов. – М.: МО СССР, 1976. – 424 с.
7. **Оппенгейм, А.В. Цифровая обработка сигналов / А.В. Оппенгейм, Р.В. Шафер.** – М.: Связь, 1979. – 416 с.
8. Баскаков, С.И. Радиотехнические цепи и сигналы / С.И. Баскаков. – М.: Высш. шк., 2002. – 448 с.
9. Кловский, Д.Д. Модели непрерывных каналов связи на основе стохастических дифференциальных уравнений / Д.Д. Кловский, В.Я. Конторович, С.М. Широков. – М.: Радио и связь, 1984. – 248 с.
10. Волков, Л.Н. Системы цифровой радиосвязи: базовые методы и характеристики: учеб. пособие / В.Л. Волков, М.С. Немировский, Ю.С. Шинаков. – М.: Эко-Трендз, 2005. – 392 с.
11. Голяницкий, И.А. Математические модели и методы в радиосвязи / И.А. Голяницкий, Ю.А. Громакова; под ред. Ю.А. Громакова. – М.: Эко-Трендз, 2005. – 440 с.

* Литература, выделенная в списке жирно, является обязательной к изучению.

12. Тихонов, В.И. Марковские процессы / В.И. Тихонов, М.А. Миронов. – М.: Советское радио, 1977. – 488 с.
13. Котельников, В.А. Теория потенциальной помехоустойчивости / В.А. Котельников. – М.: Госэнергоиздат, 1956. – 152 с.
14. Уидроу, Б. Адаптивная обработка сигналов / Б. Уиндрои, С. Стринз; пер. с англ. – М.: Радио и связь, 1989. – 440 с.
15. Васюков, В.Н. Теория электрической связи: учебник / В.Н. Васюков. – Новосибирск.: Изд-во НГТУ, 2005. – 392 с.
16. Черенкова, Е.Л. Распространение радиоволн / Е.Л. Черенкова, О.В. Чернышев. – М.: Радио и связь, 1984. – 271 с.
17. Ященко, В.В. Основные понятия криптографии / В.В. Ященко // Математическое просвещение. – 1998. – Сер. 3. Вып. 2. – С. 53–70.
18. Петраков, А.В. Основы практической защиты информации / А.В. Петраков. – М.: Радио и связь, 2000. – 368 с.
19. Шеннон, К. Теория связи в секретных системах // Работы по теории информации и кибернетике / К. Шенон. – М.: ИЛ, 1963. – 830 с.
20. Брассар Ж. Современная криптология [Текст]. – М.: Изд. полиграф. фирма ПОЛИМЕД, 1999. – 176 с.
21. Введение в криптографию / под общ. ред. В.В. Ященко. – М.: МЦНМО, 2000. – 272 с.

ОГЛАВЛЕНИЕ

СПИСОК СОКРАЩЕНИЙ	3
ОСНОВНЫЕ ОБОЗНАЧЕНИЯ	5
ВВЕДЕНИЕ	9
I. СООБЩЕНИЯ, СИГНАЛЫ И ПОМЕХИ	11
1. ОБЩИЕ СВЕДЕНИЯ О СИСТЕМАХ ЭЛЕКТРИЧЕСКОЙ СВЯЗИ	11
1.1. Информация, сообщения, сигналы и помехи	11
1.2. Общие принципы построения систем связи	16
1.3. Классификация систем связи	20
2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СИГНАЛОВ	21
2.1. Математическое описание сигнала.....	21
2.2. Математическое представление сигналов	22
2.3. Геометрическое представление сигналов	24
2.4. Представление сигналов в виде рядов ортогональных функций	25
3. СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ СИГНАЛОВ.....	29
3.1. Спектральное представление периодических сигналов	29
3.2. Спектральное представление непериодических сигналов.....	30
3.3. Основные свойства преобразования Фурье:.....	32
4. СИГНАЛЫ С ОГРАНИЧЕННЫМ СПЕКТРОМ. ТЕОРЕМА КОТЕЛЬНИКОВА.....	41
4.1. Разложение непрерывных сигналов в ряд Котельникова.....	41
4.2. Спектр дискретизированного сигнала.....	46
4.3. Спектр сигнала дискретизированного импульсами конечной длительности (амплитудно-импульсно модулированный (АИМ) сигнал)	47
4.4. Восстановление непрерывного сигнала из отсчётов.....	49
4.5. Погрешности дискретизации и восстановления непрерывных сигналов.....	50
5. СЛУЧАЙНЫЕ ПРОЦЕССЫ	52
5.1. Характеристики случайных процессов	52
5.2. Нормальный случайный процесс (гауссов процесс).....	54

5.3. ФПВ и ФРВ для гармонического колебания со случайной начальной фазой	56
5.4. ФПВ для суммы нормального случайного процесса и гармонического колебания со случайной начальной фазой	58
5.5. Огибающая и фаза узкополосного случайного процесса	58
5.6. Флуктуационный шум	59
6. КОМПЛЕКСНОЕ ПРЕДСТАВЛЕНИЕ СИГНАЛОВ И ПОМЕХ	62
6.1. Понятие аналитического сигнала	62
6.2. Огибающая, мгновенная фаза и мгновенная частота узкополосного случайного процесса	67
7. КОРРЕЛЯЦИОННАЯ ФУНКЦИЯ ДЕТЕРМИНИРОВАННЫХ СИГНАЛОВ	69
7.1. Автокорреляция вещественного сигнала	69
7.2. Автокорреляция дискретного сигнала	71
7.3. Связь корреляционной функции с энергетическим спектром.....	73
7.4. Практическое применение корреляционной функции.....	74
 II. МЕТОДЫ ФОРМИРОВАНИЯ И ПРЕОБРАЗОВАНИЯ СИГНАЛОВ.....	
8. МОДУЛЯЦИЯ СИГНАЛОВ.....	76
8.1. Общие положения.....	76
8.2. Амплитудная модуляция гармонического колебания.....	77
8.3. Балансная и однополосная модуляция гармонической несущей.....	80
9. МЕТОДЫ УГЛОВОЙ МОДУЛЯЦИИ.....	81
9.1. Принципы частотной и фазовой (угловой) модуляции	81
9.2. Спектр сигналов угловой модуляции	84
9.3. Формирование и детектирование сигналов амплитудной и однополосной амплитудной модуляции.....	86
9.4. Формирование и детектирование сигналов угловой модуляции.....	88
10. МАНИПУЛЯЦИЯ СИГНАЛОВ.....	90
10.1. Временные и спектральные характеристики амплитудно-манипулированных сигналов	90
10.2. Временные и спектральные характеристики частотно-манипулированных сигналов	92
10.3. Фазовая (относительно-фазовая) манипуляция сигналов.....	96
 III. АЛГОРИТМЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ.....	
11. ОСНОВЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ	103
11.1. Общие понятия о цифровой обработке	103
11.2. Квантование сигнала.....	105
11.3. Кодирование сигнала	106
11.4. Декодирование сигнала	110
12. ОБРАБОТКА ДИСКРЕТНЫХ СИГНАЛОВ	113
12.1. Алгоритмы дискретного и быстрого преобразований Фурье	113

12.2. Стационарные линейные дискретные цепи	118
12.3. Цепи с конечной импульсной характеристикой (КИХ-цепи)	120
12.4. Рекурсивные цепи	121
12.5. Устойчивость ЛИС-цепей	122
13. ЦИФРОВЫЕ ФИЛЬТРЫ	124
13.1. Методы синтеза КИХ-фильтров	124
13.2. Синтез БИХ-фильтров на основе аналого-цифровой трансформации	130
IV. КАНАЛЫ СВЯЗИ	137
14. КАНАЛЫ СВЯЗИ	137
14.1. Основные определения	137
14.2. Модели непрерывных каналов	141
14.3. Модели дискретных каналов	143
V. ТЕОРИЯ ПЕРЕДАЧИ И КОДИРОВАНИЯ СООБЩЕНИЙ	147
15. ТЕОРИЯ ПЕРЕДАЧИ ИНФОРМАЦИИ	147
15.1. Количество информации переданной по дискретному каналу	147
15.2. Пропускная способность дискретного канала	148
15.3. Пропускная способность симметричного дискретного канала без памяти	149
15.4. Методы сжатия дискретных сообщений	150
15.5. Количество информации, переданной по непрерывному каналу	155
15.6. Пропускная способность непрерывного канала	156
16. ТЕОРИЯ КОДИРОВАНИЯ СООБЩЕНИЙ	159
16.1. Основные понятия	159
16.2. Коды с обнаружением ошибок	159
16.3. Корректирующие коды	161
VI. ПОМЕХОУСТОЙЧИВОСТЬ	174
17. ПОМЕХОУСТОЙЧИВОСТЬ СИСТЕМ ПЕРЕДАЧИ ДИСКРЕТНЫХ СООБЩЕНИЙ	174
17.1. Основные понятия и термины	174
17.2. Бинарная задача проверки простых гипотез	178
17.3. Приём полностью известного сигнала (когерентный приём)	181
17.4. Согласованная фильтрация	185
17.5. Потенциальная помехоустойчивость когерентного приёма	188
17.6. Некогерентный приём	192
17.7. Потенциальная помехоустойчивость некогерентного приёма	195

18. ПОМЕХОУСТОЙЧИВОСТЬ СИСТЕМ ПЕРЕДАЧИ НЕПРЕРЫВНЫХ СООБЩЕНИЙ	198
18.1. Оптимальное оценивание сигнала.....	198
18.2. Оптимальная фильтрация случайного сигнала	203
18.3. Потенциальная помехоустойчивость передачи непрерывных сообщений.....	206
19. АДАПТИВНЫЕ УСТРОЙСТВА ПОДАВЛЕНИЯ ПОМЕХ.....	209
19.1. Основы адаптивного подавления помех	209
19.2. Подавление стационарных помех	211
19.3. Адаптивный режекторный фильтр	216
19.4. Адаптивный высокочастотный фильтр	218
19.5. Подавление периодической помехи с помощью адаптивного устройства предсказания	219
19.6. Адаптивный следящий фильтр	221
19.7. Адаптивный накопитель.....	224
VII. МНОГОКАНАЛЬНАЯ СВЯЗЬ И РАСПРЕДЕЛЕНИЕ ИНФОРМАЦИИ.....	228
20. ПРИНЦИПЫ МНОГОКАНАЛЬНОЙ СВЯЗИ И РАСПРЕДЕЛЕНИЯ ИНФОРМАЦИИ	228
20.1. Общие положения	228
20.2. Частотное разделение каналов.....	231
20.3. Временное разделение каналов.....	232
20.4. Кодовое разделение каналов	235
20.5. Синхронизация в СПИ с многостанционным доступом.....	236
20.6. Коммутация в сетях связи	239
VIII. ЭФФЕКТИВНОСТЬ СИСТЕМ СВЯЗИ	241
21. ОЦЕНКА ЭФФЕКТИВНОСТИ И ОПТИМИЗАЦИЯ ПАРАМЕТРОВ ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ (ТКС).....	241
21.1. Критерии эффективности	241
21.2. Эффективность аналоговых и цифровых систем	243
21.3. Выбор сигналов и помехоустойчивых кодов.....	247
22. ОЦЕНКА ЭФФЕКТИВНОСТИ РАДИОТЕХНИЧЕСКОЙ СИСТЕМЫ СВЯЗИ.....	249
22.1. Тактико-технические параметры радиотехнической системы связи.....	249
22.2. Оценка отношения сигнал/помеха на входе радиоприемники радиотехнической системы связи	249
22.3. Оптимальная фильтрация непрерывных сигналов.....	250
22.4. Количество информации при приёме дискретных сигналов радиотехнической системы связи	252
22.5. Количество информации при оптимальном приёме непрерывных сигналов.....	255
22.6. Выигрыш в отношении сигнал/помеха	256
22.7. Пропускная способность каналов радиотехнической системы связи	261

IX. ТЕОРЕТИКО-ИНФОРМАЦИОННАЯ КОНЦЕПЦИЯ КРИПТОЗАЩИТЫ СООБЩЕНИЙ В ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ	263
23. ОСНОВЫ КРИПТОЗАЩИТЫ СООБЩЕНИЙ В СИСТЕМАХ СВЯЗИ	263
23.1. Основные понятия криптографии.....	263
23.2. Метод замены	266
23.3. Методы шифрования на основе датчика псевдослучайных чисел	267
23.4. Методы перемешивания	268
23.5. Криптосистемы с открытым ключом	272
23.6. Цифровая подпись.....	274
ЗАКЛЮЧЕНИЕ	276
СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ.....	277

Учебное издание

Павликов Сергей Николаевич
Убанкин Евгений Иванович
Левашов Юрий Александрович

ОБЩАЯ ТЕОРИЯ СВЯЗИ

Учебное пособие

В авторской редакции
Компьютерная верстка М.А. Портновой

Подписано в печать 01.11.2016. Формат 60×84/16.
Бумага писчая. Печать офсетная. Усл. печ. л. 16,5.
Уч.-изд. л. 12,5. Тираж 200 экз. Заказ

Издательство Владивостокского государственного университета
экономики и сервиса
690014, Владивосток, ул. Гоголя, 41
Отпечатано во множительном участке ВГУЭС
690014, Владивосток, ул. Гоголя, 41