МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ КАЗАХСТАН

КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ К.И. САТПАЕВА

Айтчанов Б. Х.,

Яскевич Т.В.

ТЕОРИЯ ИНФОРМАЦИИ

Рекомендовано Учебно-методическим советом
в качестве учебного пособия университета

Алматы 2009

УДК 621.3
ББК32.811я73
А 30

А 30 Айтчанов Б.Х.,Т.В. Яскевич. Теория информации: Учеб. пособие – Алматы: КазНТУ, 2009. – 110 с.

ISBN 978-601-228-080-7

В книге изложены математические модели детерминированных и случайных сигналов, принципы дискретизации и квантования сигналов. Рассмотрены различные аспекты количественной меры информации и информационные характеристики источников информации и каналов связи.

Учебное пособие предназначено для студентов специальности «Вычислительная техника и программное обеспечение». Она также полезна для студентов информационных и системотехнических специальностей.

Ил. 38. Табл. 3. Библиогр. – 25 назв.

Рецензенты:

Ш.Б. Биттеев, д-р техн. наук, проф., зав. кафедрой автоматики, телемеханики и связи КазАТК

Б.Д. Хисаров, канд. техн. наук, проф., зав. кафедрой инженерной кибернетики АИЭС

Д.З. Джурунтаев, канд. техн. наук, проф., профессор кафедры вычислительной техники КазНТУ им. К.И. Сатпаева

Печатается по плану издания Министерства образования и науки Республики Казахстан на 2009 г.

$$\text{А}\ \frac{1404000000}{00(05)-09}$$

ISBN 978-601-228-080-7

Б.Х. Айтчанов, 2009 ©

Т.В. Яскевич, 2009 ©

КазНТУ, 2009 ©

Посвящается 75-летию
Казахского национального технического
университета имени К.И. Сатпаева

ВВЕДЕНИЕ

Эффективная организация обмена информацией приобретает все большее значение как условие успешной практической деятельности людей. Объем информации, необходимый для нормального функционирования современного общества растет, примерно пропорционально квадрату развития промышленного потенциала. Доля рабочей силы, занятой вопросами обеспечения информацией, начинает превышать долю рабочей силы занятой непосредственно в производстве. Поэтому науки, изучающие структуру и закономерности протекания информационных процессов, к числу которых относится и теория информации (ТИ), в такой ситуации становятся исключительно актуальными.

Теория информации как самостоятельная научная дисциплина, связанная с восприятием, передачей и переработкой, хранением и использованием информации, была основана американским ученым К. Шенноном в конце 40-х годов XX-го века. Предложенная Шенноном теория основывалась на фундаментальном понятии количественной меры неопределенности – *энтропии* – и связанного с нею понятия – *количества информации.* Другим фактором в становлении теории информации стало осознание того, что носитель информации – *сигнал,* имеет случайную природу. На основе понятий энтропии и количества информации в теории информации введены важные характеристики сигналов и *информационных систем.*

Учебное пособие написано на основе односеместрового трехкредитного курса лекций и материалов для практических занятий, используемых авторами в учебной работе со студентами-второкурсниками в течение последних четырех лет на кафедре «Техническая кибернетика» Казахского национального технического университета имени К.И. Сатпаева.

Программой (силлабусом) курса предусмотрено изучение следующих тем: «Фундаментальные положения теории информации»;

«Количественная мера информации»; «Математические модели и формы представления детерминированных и случайных сигналов»; «Основные положения дискретизации во времени и квантования по уровню информации»; «Основные характеристики систем и каналов передачи информации», а также «Методы построения эффективных и помехоустойчивых кодов».

Учебное пособие посвящено основным положениям теории информации.

В первом разделе даны общие понятия и определения теории информации. Приведена классификация, виды и этапы обращения информации.

Во втором разделе приводится теория измерения количества информации. Подробно рассмотрены вопросы структурной, статистической и семантической мер измерения количества информации и их взаимосвязь.

Третий раздел посвящен аналитическому описанию детерминированных и случайных сигналов. Рассмотрены временная и частотная формы представления и их основные характеристики.

В четвертом разделе рассмотрена дискретизация и квантование сигналов. Приведена с доказательством теорема Котельникова, показаны аналитические методы дискретизации. Введены основные характеристики квантование информации по уровню и рассмотрены некоторые вопросы выбора оптимального уровня квантования.

В пятом разделе введены основные понятия и определения систем передачи информации. Рассмотрены методы модуляции сигналов, проведен сравнительный анализ модулированных сигналов. Рассмотрены вопросы согласования физических характеристик канала связи и сигнала.

I. ОСНОВНЫЕ ПОЛОЖЕНИЯ ТЕОРИИ ИНФОРМАЦИИ

1.1. Теория информации, и ее место среди дисциплин специальности

Широкое внедрение автоматизированного управления немыслимо без использования средств связи, телемеханики и вычислительной техники.

В процессе автоматизированного управления техникой, производственными процессами и отраслями народного хозяйства происходит интенсивный обмен информацией между отдельными звеньями систем управления, между человеком и техникой и отдельными людьми.

Комплексная автоматизация и совершенствование электронных вычислительных машин сопровождается резким возрастанием объема и скорости передачи, и обработки информации с одновременным повышением требований к ее достоверности [2, 3].

Из всего многообразия современных технических систем можно выделить особую группу систем, так называемые информационные системы, предназначенные для передачи, преобразования и хранения информации (связные, локационные, навигационные, телевизионные системы, электронно-вычислительные, и информационно-измерительные техники, автоматизированные системы управления и контроля и т. п.).

Информационные системы существенно отличаются от энергетических устройств и систем (двигателей, электрических генераторов, линий электропередач), основой функционирования которых являются процессы передачи и преобразования энергии.

Главным критерием качества работы информационной системы служит их способность передавать, накапливать или преобразовать максимальное количество информации в единицу времени при допустимых искажениях и затратах, а не к. п. д. энергетических установок [7, 9].

В информационных системах энергетические состояния играют второстепенную роль, а сама энергия является характеристикой сигнала, который используется в качестве транспортного средства.

Информационная система – система сбора, хранения, обновления, передачи, обработки и выдачи информации, необходимой для управления объектом.

Задачами информационной системы являются [2, 8, 9]:

1. Обеспечение информацией о происходящем внутри системы, а также за ее пределами.

2. Обеспечение информацией о правилах решения.

3. Регистрация, анализ и оценка данных о среде объекта.

4. Наблюдение за ходом управляемых процессов, получение, анализ и оценка данных о реальном состоянии управляемой системы.

5. Обеспечение ввода данных и обработка данных.

6. Извлечение из обработанных данных в нужной форме и необходимом объеме информации для принятия решений.

7. Получение административной и мотивационной информации для принятия решения.

8. Контроль за выполнением принятых решений.

По мере развития информационных систем происходила соответствующая дифференциация научных дисциплин. Возникли и развиваются теория автоматического управления, теория автоматического контроля, радиосвязь и проводная связь, телемеханика, программирование и т. д., занимающиеся вопросами теории и синтеза отдельных видов информационных систем. Дальнейшее развитие теории и практики этих систем сформировали общие основы в теории их построения, которые вылились в общетеоретическую дисциплину – кибернетику.

Кибернетика – наука о способах восприятия, хранения, переработки и использовании информации в машинах, живых организмах и их объединениях. Она устанавливает общие законы и принципы, согласно которым как живые организмы, так и некоторые машины выполняют целеустремленные действия на основе процессов передачи, преобразования и использования информации. Если основная задача общей кибернетики заключается в анализе информационных процессов различной природы и их алгоритмизацией, то техническая кибернетика занимается синтезом информационных устройств, способных реализовывать эти алгоритмы.

Техническая кибернетика включает в себя три направления [13, 17]:

– теорию информации, изучающую процессы получения, преобразования, накопления и передачи информации;

– теорию логических и вычислительных машин, занимающуюся методами преобразования информации в соответствии с заданными алгоритмами;

– теорию автоматического регулирования, изучающую методы использования информации для выполнения целеустремленных действий.

Возникновение и развитие теории информации связано с быстрым развитием в первой половине двадцатого столетия техники связи. Стали появляться работы, в которых предпринимались попытки научного обоснования основных характеристик систем связи и количественной оценки качества передачи сообщений.

В 1928 г. американский ученый Р. Хартли предложил логарифмическую меру оценки количества информации, впоследствии получившего название «мера Хартли». В своей знаменитой работе «О пропускной способности эфира» В.А. Котельников впервые сформулировал теорему о дискретном представлении сигнала с ограниченным спектром и ряд практических рекомендаций по оценке пропускной способности каналов связи.

Следующим крупным шагом в развитии теории информации связаны с фундаментальными трудами В.А. Котельникова по потенциальной помехоустойчивости (1946) и К. Шеннона по теории информации (1948). В этих работах сформулированы основные положения и теоремы современной теории информации.

Дальнейшее развитие теории информации связаны с трудами А.А. Харкевича, В.И. Сифорова, А.Н. Железнова, Р.М. Фано и многих других ученых. В развитие теории информации существенный вклад внесли и математики – А.Н. Колмогоров, Н. Винер, А.Я. Хинчин и др.

В настоящее время теория информации, математическим аппаратом которой является теория вероятности и математическая статистика, превратилась в строгую и достаточно универсальную науку.

1.2. Основные понятия и определения теории информации

Производственные процессы, а также процессы живой природы связаны с получением, передачей, преобразованием, накоплением и хранением информации. Существуют различные определения информации. Обычно под информацией понимают новые сведения об окружающем мире, которые получены в результате взаимодействия с ним, приспособления к нему и изменения его свойств в процессе приспособления [9].

Информация – это сведения, являющиеся объектом хранения, передачи и преобразования. Информация – это, прежде всего сведения, которые должны быть использованы.

Рассмотрим основные понятия и определения информационной системы на примере системы передачи информации (рис. 1.1).

Система передачи информации в качестве информационной системы включает следующие блоки: отправитель информации, линия связи, получатель [8,10,18].

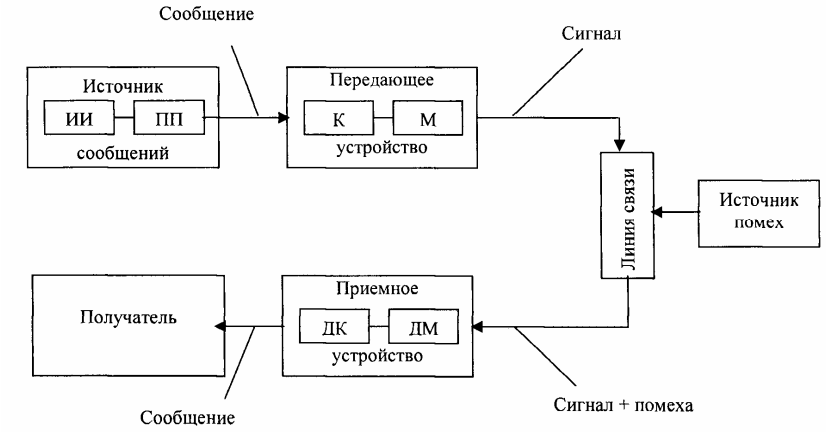


Рис.1.1. Система передачи информации

Информация поступает в систему в форме сообщений. Сообщение – информация, выраженная в определенной форме и подлежащее дальнейшему преобразованию, т. е. сообщение – форма представления информации (текст телеграмм, речь оратора, показания прибора).

Сообщение, для передачи его в соответствующий адрес, должно быть предварительно преобразовано в сигнал.

Сигнал – это материальный переносчик сообщения. Физическую среду, по которой происходит передача сигналов от передающего устройства к принимающему устройству, называют линией связи.

Возможные виды сигналов – электрические, электромагнитные, световые, механические, ультразвуковые сигналы. Для пере-

дачи сообщений необходимо применять тот переносчик, который способен эффективно распространяться по используемой линии связи. (Проводниковая линия – постоянный ток и непостоянные токи несущей частоты и т. д.).

Преобразования сообщений в сигналы удобные для прохождения по линии связи осуществляется передающим устройством связи.

Сообщение – совокупность знаков или первичных сигналов, содержащих информацию. Источник сообщений – совокупность источника информации (ИИ) и первичного преобразователя (ПП) датчика, человека – оператора и т. п.

В процессе преобразования сообщений в сигнал происходит кодирование сообщений. В широком смысле кодирование – преобразование сообщений в сигнал, удобный для передачи по данному каналу, а в узком – это отображение дискретных сообщений сигналами в виде определенных сочетаний символов. Осуществляются преобразования сообщений в сигналы удобные для передачи по конкретным линиям связи (либо для хранения в запоминающих системах).

При передаче осуществляется воздействие на один или несколько параметров выбранного носителя, изменяя его в соответствии с передаваемой информацией. Такой процесс называют модуляцией, моделируемые параметры называют информативными.

При передаче на сигналы воздействуют помехи. Под помехами понимают любые мешающие внешние воздействия возмущения (атмосферные и промышленные помехи; влияние посторонних источников), а также искажение сигналов в самой аппаратуре (аппаратные помехи), вызывающие случайные отключения принятого сигнала от передаваемого сигнала. Эффект воздействия помех стараются учесть эквивалентным изменением характеристик линий связи.

На приемнике осуществляется обратная операция – демодуляция и декодирование, т. е. восстановление по принятому сигналу переданного сообщения.

1.3. Виды информации. Этапы обращения информации

Сообщения различают по характеру измерения во времени на непрерывные и дискретные [11,14]. Непрерывные по времени сообщения отображаются непрерывной функцией времени, дискретные – дискретной функцией времени (поступают в определенный момент времени). В реальности сообщения носят случайный характер, тогда непрерывные сообщения описываются случайной функцией времени, а дискретные сообщения – как цепь случайных событий.

Сообщения различают по множеству – непрерывные и дискретные. Непрерывные по множеству сообщения характеризуются тем, что функция их описывающая может принимать множество значений в некотором интервале. Дискретные сообщения могут быть описаны с помощью конечного набора чисел или дискретных значений некоторых функций.

Дискретность по времени и множеству не связана друг с другом. Поэтому возможны следующие виды сообщений (рис. 1.2):

а) непрерывные по множеству и по времени;

б) непрерывные по множеству, дискретные во времени;

в) дискретные по множеству и непрерывные по времени;

г) дискретные по множеству и по времени.

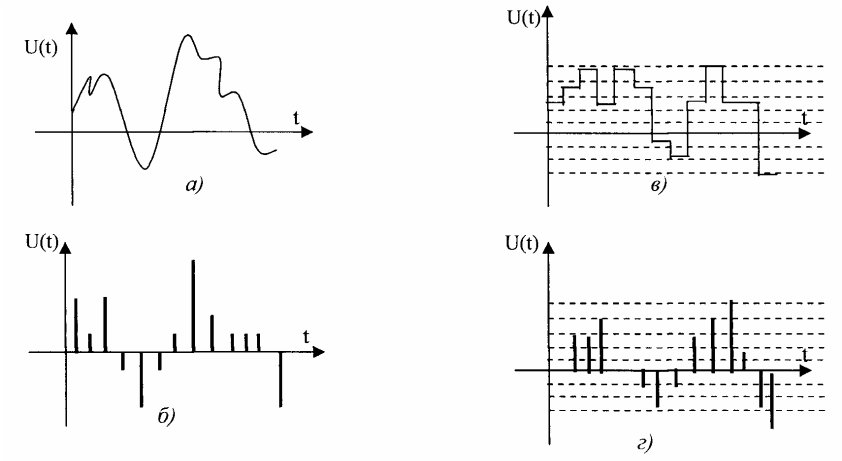


Рис.1.2. Виды сообщений

Информацию подразделяют на следующие виды:

1. По областям знаний (биологическая, техническая, экономическая и т. д.).

2. По физической природе восприятия (зрительная, слуховая, вкусовая).

3. По структурно-метрическим свойствам. Особый интерес представляет по этому виду параметрическая информация. Параметрическая информация – наборы численных оценок значений каких-либо параметров, результаты количественной оценки при исследовании.

Структуру информации можно изменять, переходя из одного вида информации к другому (табл. 1.1).

По размерности информации: информация нулевого порядка, соответствующая мощности точки; информация первого порядка – мощности линии и. т. д.

Таблица 1.1

Структура информации

Виды информации	Обозначения	Форма представления информации		
		Топологическая	Абстрактная	Лингвистическая
событие	Φ^0	точка	суждение	знак
величина	Φ^1	линия	понятие	буква
функция	Φ^2	поверхность	образ	слово
комплекс	Φ^3	объем	система	предложение
.....	.....	.....	.....	.....
поле	Φ^n	пространство	универсум	фонд

Этапы обращения информации. В информационных системах существуют следующие этапы обращения информации (рис.1.3):

1) восприятия информации;
2) подготовки информации;
3) передачи и хранения информации;
4) обработки информации;
5) отображения информации.

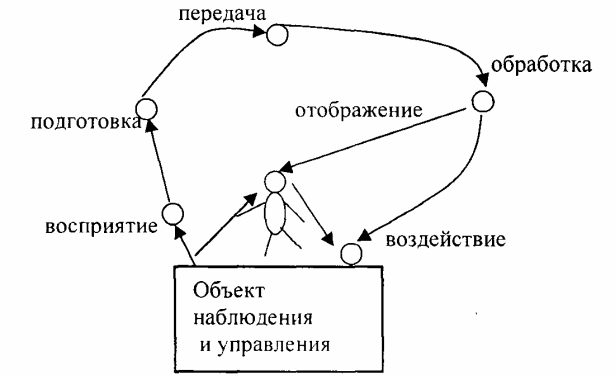


Рис.1.3. Этапы обращения информации

1. На этапе восприятия осуществляется целенаправленное извлечение и анализ информации о каком-либо объекте, в результате чего формируется образ объекта, проводятся его опознание и оценка.

2. На этапе подготовки информации проводятся такие операции, как нормализация, аналого-цифровое преобразование, шифрование. Иногда этот этап рассматривается как вспомогательный на этапе восприятия. В результате восприятия и подготовки получается сигнал в форме удобной для передачи или обработки.

3. На этапах передачи и хранения информация пересылается либо с одного места в другое, либо от одного момента времени до другого. Поскольку теоретические вопросы, возникающие на этих этапах, близки друг к другу, этап хранения часто в самостоятельный этап не выделяется, при этом передача информации получает более широкое толкование. Для передачи на расстояние используются каналы различной физической природы, самые распространенные – электрические и электромагнитные, наиболее перспективен – оптический. Для хранения информации используются в ос-

новном полупроводниковые и магнитные носители.

4. На этапах обработки информации выявляются ее общие и существенные взаимозависимости, представляющие интерес для системы. Преобразование информации на этапе обработки осуществляется средствами информационной техники, либо человеком. Если процесс обработки формализуем, то он выполняется техническими средствами (ЭВМ, микропроцессоры). Если не формализуем и требует творческого подхода, обработка информации осуществляется человеком. В системах управления целью обработки является решение задачи выбора управляющих воздействий (этап принятия решений).

5. Этап отображения информации должен предшествовать этапам, связанным с участием человека. Цель этапа отображения – предоставить человеку нужную ему информацию с помощью устройств, способных воздействовать на его органы чувств.

6. На этапе воздействия информация используется для осуществления необходимых изменений в системе. Воздействие – сигналы, производят регулирование или управление, вызывая изменения в самом объекте.

Контрольные вопросы:

1. Определите основные задачи информационной системы.
2. Дайте основные определения и понятия составляющих информационной системы.
3. Объясните сущность принципиальных различий в трактовке понятия информации.
4. Дайте определения сигнала и сообщения.
5. В чем различие между линией и каналом связи?
6. Проведите классификацию детерминированных сообщений.
7. Приведите функции основных блоков системы передачи информации.
8. Перечислите основные виды информации.
9. Определите основные этапы обращения информации.
10. Каковы основные задачи теории информации?

2. ИЗМЕРЕНИЕ ИНФОРМАЦИИ

Важным вопросом теории информации является установление меры количества и качества информации [2, 3, 6].

Информационные меры измерения отвечают трем основным направлениям в теории информации:

1. Структурная теория рассматривает дискретное строение массивов и их измерение простым подсчетом информационных элементов (квантов) или комбинаторным методом, предполагающим простейшее кодирование массивов информации.

2. Статистическая теория оперирует понятием энтропии как меры неопределенности, учитывающей вероятность появления, а, следовательно, и информативность тех или иных сообщений.

3. Семантическая теория учитывает целесообразность, полезность или существенность информации.

2.1. Структурные меры измерения информации

В структурной теории различают геометрическую меру информации, комбинаторную меру и аддитивную меру (мера Хартли).

2.1.1. Геометрическая мера

Определение количества информации геометрическим методом сводится к измерению длины линии, площади или объема геометрической модели данного информационного комплекса в виде дискретных единиц-квантов. Под информационными элементами понимаются неделимые части-кванты информации в дискретных моделях реальных информационных комплексов, а также элементы алфавита в числовых системах.

Геометрическим методом определяется потенциальное, т. е. максимально возможное количество информации в виде структурных габаритов. Это полностью будет пониматься информационной емкостью исследуемой части информационной системы. Информационная емкость вычисляется по сумме дискретных значений по всем измерениям [2, 9].

Пусть информация отражается полным комплексом XTN. Если дискретные отсчеты осуществляются по осям X, T и N, соответственно через интервалы Δx, Δt, Δn, то непрерывные координаты распадаются на элементы (кванты) количество которых составляет

$$m_x = \frac{x}{\Delta x}, \quad m_T = \frac{T}{\Delta T}, \quad m_N = \frac{N}{\Delta N}. \tag{2.1}$$

Тогда количество информации в полном комплексе XTN, определенное геометрическим методом, равно в квантах

$$M = m_X \cdot m_T \cdot m_N. \tag{2.2}$$

2.1.2. Комбинаторная мера

К комбинаторной мере целесообразно прибегать тогда, когда требуется оценить возможность передачи информации при помощи различных комбинаций. Это одна из форм кодирования информации [4, 5].

Количество информации в комбинаторной мере вычисляется как количество комбинации элементов, т. е. оценке подвергается комбинация потенциальных структурных разнообразных информационных комплексов.

В комбинаторике рассматриваются разные виды сочетаний элементов.

Сочетания из h элементов по ℓ различаются составом элементов, их возможное число равно

$$Q = (h^{\ell}) = \frac{h!}{\ell!(h-\ell)!} = \frac{h(h-1)...(h-\ell+1)}{1 \cdot 2...\ell}. \tag{2.3}$$

Сочетания из h элементов по ℓ с повторениями также различаются составом элементов, но элементы в них могут повторять до ℓ раз.

$$Q = (h^{\ell})_{\text{повт}} = \frac{(h+l-1)!}{\ell!(h-1)\ !}. \tag{2.4}$$

Перестановки h элементов различаются их порядком. Число возможных перестановок h элементов равно

$$Q = h! \ . \qquad (2.5)$$

Перестановка с повторением элементов, причем один из элементов повторяется α раз, другой β раз, наконец, последний γ раз, характеризуется числом

$$Q = \frac{(\alpha + \beta + ... + \gamma)!}{\alpha!\beta!...\gamma!}. \qquad (2.6)$$

Размещения из h элементов по ℓ различаются и составом элементов, и их порядком. Возможное число размещений из h элементов по ℓ

$$Q = (h^{\ell}) = \frac{h!}{(h-\ell)!} = h(h-1)\ \ (h-2)...(h-(\ell+1))... \qquad (2.7)$$

Возможное число размещений из h элементов по ℓ с повторениями

$$Q = (h^{\ell})_{\text{повт}} = h^{\ell}. \qquad (2.8)$$

При применении комбинаторной меры возможное количество информации Q совпадает с соединениями, т. е. определение количества информации Q в комбинаторной мере заключается не в простом подсчете квантов (геометрическая мера), а в определении количества возможных или действительно существующих комбинаций, т. е. в оценке структурного разнообразия.

Пример 2.1. Вычислить количество информации для десяти элементов при их различных соединениях [1,19].

1. В случае сочетаний из 10 элементов по 0, 1, 2, 3, 4, 5, 6, 7, 8 и 9 элементам получается следующее число образований, являющихся мерой количества информации:

$$Q = \frac{10!}{0!(10-0)!} + \frac{10!}{1!(10-1)!} + ... + \frac{10!}{9!(10-9)!} + \frac{10!}{10!(10-10)!} = 1024$$

образования.

2. Перестановки этих 10 элементов дают

$$Q = h! = 1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot 9\cdot 10 = 3628800$$ образования.

3. Размещения 10 разных элементов по 10 разным позициям приведет к еще большему появлению количества информации.

$$Q = h^{\ell} = 10^{10}$$ образования.

2.1.3. Аддитивная мера (мера Хартли)

В теории информации важную роль играет комбинаторика чисел и кодов. Введем понятия глубины h и длины ℓ числа.

Глубиной h числа называется количество различных элементов, содержащихся в принятом алфавите. Глубине числа соответствует основание системы считывания и кодирования [14,15].

Длиной ℓ числа называется количество числовых гнезд, т. е. количество повторений в алфавите, необходимых и достаточных для представления чисел нужной величины. Длина числа сообщает различные системы счислений и кодирования.

Один набор из ℓ гнезд–алфавитов составляет одну числовую группу, способную предоставлять и хранить одно полное число длиной ℓ. Некоторое количество чисел N представляет числовым полем. При глубине h и длине ℓ количество чисел, которые можно представить с помощью числовых групп, выразится формулой

$$Q = h^{\ell}.$$

Как видно, из последней формулы, количество чисел Q в зависимости от длины ℓ по показательному закону. В виду этого, число Q не является удобной мерой для оценки информационной емкости, поэтому Хартли ввел аддитивную двоичную логарифмическую меру, позволяющую вычислять количество информации в двоичных единицах – битах (бит). Для этого берется не само число Q, а его двоичный логарифм:

$$J = \log_2 Q = \log_2 h^1 = \ell \log_2 h \text{ (бит)}. \qquad (2.9)$$

Здесь J – количество информации по Хартли.

Если количество разрядов (длина ℓ числа) равна единице, имеется в виду двоичная система счисления (глубина h = 2) и исполняется двоичный логарифм, то потенциальное количество информации равно 1 биту ($\log_2 2 = 1$).

Это и есть единица измерения в принятой системе. Она соответствует одному элементарному событию, которая может произойти или не произойти. Аддитивная мера удобна тем, что она

обеспечивает возможность считать пропорциональность J относительно числа ℓ.

2.2. Семантические меры измерения информации

Под семантикой понимается смысл, содержание информации. Семантика – это одна из частей более общей науки – семиотики. В семиотике различают синтаксический, прагматический аспекты информации (рис. 2.1)

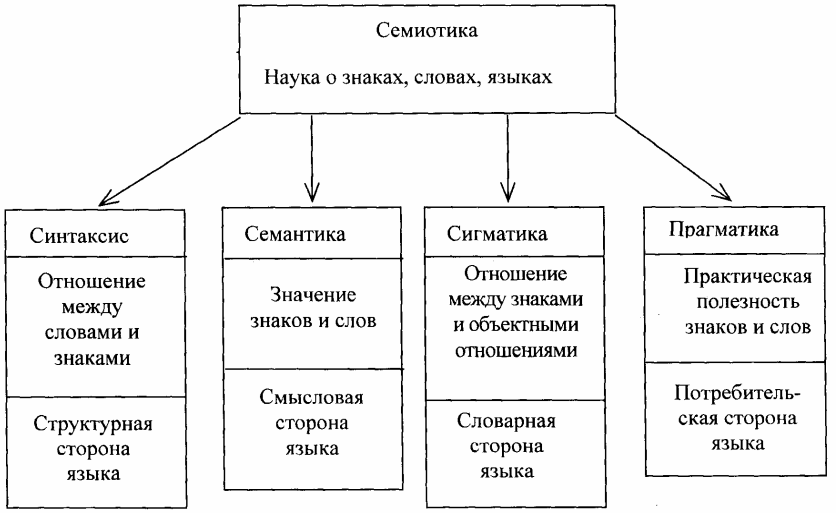


Рис. 2.1. Структура семантической меры измерения информации

В семантике имеются следующие основные понятия [2, 9]: знак, слово, язык.

Знаком называется условное изображение элемента сообщения.

Словом называется совокупность знаков, имеющих смысловое (предметное) значение.

Языком называется словарь и правила пользования им.

Структурная и статистическая мера информации относятся к синтаксическому аспекту.

Синтаксический аспект отображается теорией сигналов и кодов, рассматривающей условные обозначения элементов информации. Синтаксические оценки не имеют прямого отношения к мерам информации, и поэтому нужно опустить их рассмотрение.

Далее рассмотрим оценки, отвечающие семантическому и прагматическому аспектам теории информации. В инженерных применениях прагматические оценки сливаются с семантическими, поскольку не имеющие смысла сведения бесполезны, а бесполезные сведения бессмысленны.

Однако эффективности логического вывода, степени приближения к истине требует некоторой формализации смысла.

Один из путей формализации смысла предлагается семантической теорией информации. Карнап и Бар-Хиллел [3] предложили использовать для целей измерения смысла функции истинности и ложности логических высказываний (предложений). За основу дискретного описания объекта берется неделимое предложение. Полученная таким образом оценка получило название содержательности информации.

2.2.1. Содержательность информации

Мера содержательности обозначается cont (от английского слова content – содержание).

Содержательность события i выражается через функцию меры m(i) – содержательности его отрицания и определяется как

$$\text{cont}(i) = m(\sim i) = 1 - m(i), \qquad (2.10)$$

где i – рассматриваемое событие, m–функция меры,

$\sim$ – знак отрицания.

Логическая оценка количества информации I равна

$$\text{Inf} = \log_2\left(\frac{1}{1-\text{cont}(i)}\right) = \log_2\frac{1}{m(i)} = -\log_2 m(\sim i). \qquad (2.11)$$

2.2.2. Целесообразность информации

Если информация используется в системах управления, то ее полезность разумно оценивать по тому эффекту, который она оказывает на результат управления. В связи с этим А.А. Харкевичем

была предложена мера целесообразности информации, которая определяется как изменение вероятности достижении цели при получении дополнительной информации [9].

В общем виде мера целесообразности может быть аналитически выражена в виде соотношения

$$I_{\text{цель}} = \log_2 p_1 - \log_2 p_0 = \log_2 \frac{p_1}{p_0}, \qquad (2.12)$$

где p_0 и p_1 – начальная (до получения информации) и конечная (после получения информации) вероятность достижения цели.

Могут быть следующие случаи:

1. Полученная информация может быть пустой, т. е. не иметь вероятности достижения цели, и в этом случае ее мера равна нулю.

2. Полученная информация может изменять положение дел в нужную сторону, т. е. уменьшать вероятность достижения цели и тогда она будет дезинформацией (ложной), которая измеряется отрицательным значением количества информации.

3. Полученная информация увеличивает вероятность достижения цели и измеряется положительной величиной количества информации.

Пример 2.2. На рис. 2.2 показано исходное состояние некоторой системы, соответствующей вершине 1, из которого возможно два пути: 1–2 и 1–3. Пусть точка 3 представляет собой цель, а точка 2 некоторый промежуточный путь в стороне от цели. Если неизвестны пути к цели, то вероятности достижения цели по путям 1–2 и 1–3 будут одинаковы, а т. к. схема опыта содержит два исхода, то можно записать условие

$$p = p(1-2) = p(1-3) = \frac{1}{2}.$$

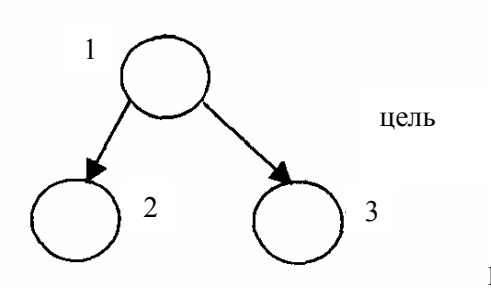


Рис. 2.2. Исходная система

1. Предположим, что достигнута точка 2 и в этой точке получается информация следующего характера рис.2.2

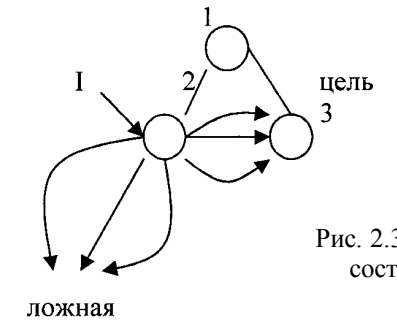


Рис. 2.3. Промежуточное состояние системы

Состояние системы описывается

$$I_{\text{цель}} = \log_2 \frac{p(2-3)}{P(1-2)} = \log_2 \frac{3/6}{3/6} = 0.$$

Эта информация оказалась пустой.

2. В точке 2 получена ложная информация, которая уменьшила вероятность достижения цели в виду преобладания ложных направлений (рис. 2.4).

В этом случае целесообразность равна

$$I_{\text{цель}} = \log_2 \frac{p(2-3)}{P(1-2)} = \log_2 \frac{1/6}{1/2} = -1,58.$$

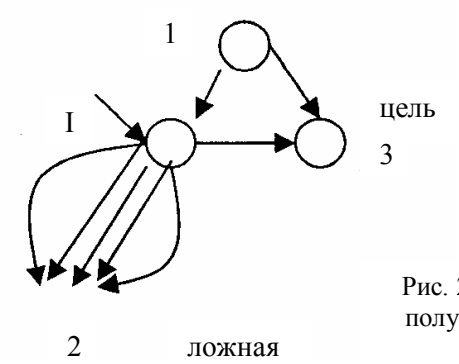


Рис. 2.4. Состояние системы при получении ложной информации

3. Благоприятный случай. В вершину 2 поступает информация, которая два исхода из места оставляет ложными, но четыре исхода из шести являются благоприятными с точки зрения достижения цели 2 (рис. 2.5).

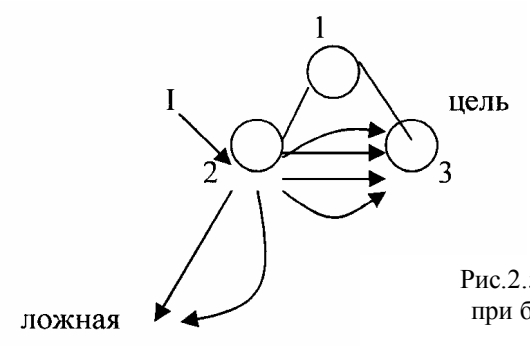


Рис.2.5. Состояние системы при благоприятном случае

В благоприятном случае целесообразность равна

$$I_{\text{цель}} = \log_2 \frac{p(2-3)}{P(1-2)} = \log_2 \frac{4/6}{1/2} = 0{,}4\,2.$$

2.2.3. Динамическая энтропия

В том случае, когда целью извлечения информации являются обращенные в 0 неопределенные ситуации, удобно пользоваться понятием динамической энтропии (в ходе распознавания образов, диагноза болезней или расследования преступлений).

Энтропия (неопределенность) ситуации изменяется во времени: $H = H(t)$.

Большинство ситуаций можно представить как множество отношений между следствиями $\{a_i\}$ и причинами $\{b_j\}$. Следствия наблюдаются, а причины раскрываются по ним. Отношения между следствиями и причинами оцениваются вероятностями $p_{ij}(t)$. При этом $p_{ij}(t) = 0$, если нет связи между следствиями и причинами, $0 \le p_{ij}(t) \le 1$, если между ними имеется небольшая связь.

Общее количество следствий в момент t обозначить N(t), а общее количество причин – M(t), тогда отдельные следствия a_i получают нумерацию i = 1,2 , N(t), а причины – j = 1,2,…, M(t)

Энтропия для данного множества отношений:

$$H(t) = -\sum_{i=1}^{N(t)} \sum_{j=1}^{M(t)} p_{ij}(t) \log p_{ij}(t). \qquad (2.13)$$

По истечении единичного интервала времени поступает дополнительная информация, которая может изменить количество следствий $\{a_i\}$ от N(t) до N(t+1), количество причин $\{b_j\}$ от M(t) до M(t+1) и вероятности отношений между ними от $p_{ij}(t)$ до $p_{ij}(t+1)$.

В результате энтропия в момент t+1 равна:

$$H(t+1) = -\sum_{i=1}^{N(t+1)} \sum_{j=1}^{M(t+1)} p_{ij}(t+1) \log p_{ij}(t+1). \qquad (2.14)$$

В качестве меры информации целесообразно принять разность:

$$I = \Delta H = H(t) - H(t+1), \qquad (2.15)$$

которая может быть как положительной, так и отрицательной величиной, в зависимости от того уменьшается или увеличивается неопределенность ситуации.

2.3. Статистические меры измерения информации

2.3.1. Вероятность и информация

При вероятностном подходе информация рассматривается как сообщение об исходе случайных событий, реализации случайных величин и функций, а количество информации ставится в зависимость от априорных вероятностей этих событий, величин, функций.

Предположим, что источник сообщений может в каждый момент времени случайным образом принять одно из конечного множества возможных состояний. Каждому состоянию источника X ставится в соответствие условное обозначение в виде знака. Совокупность знаков x_1, x_2 , …, x_i , ..., x_N, соответствующих всем N возможным состояниям источника, называют его алфавитом, а количество состояний N – объемом алфавита. Формирование таким источником сообщений сводится к выбору им некоторого состояния x_i и выдачи соответствующего знака. Таким образом, под элементарным дискретным сообщением будем понимать символ x_i , выдаваемый дискретным источником. Отдельные состояния источника могут выбираться им чаще, другие реже. Поэтому в общем случае источник X определен совокупностью состояний с вероятностями их появления, составляющими в сумме 1:

$$X = \begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots & x_n \\ p(x_1) & p(x_2) & \dots & p(x_i) & \dots & p(x_n) \end{pmatrix}, \tag{2.16}$$

$$\sum_{i=1}^{N} p(x_i) = 1, \tag{2.17}$$

Собственная информация символа x_i (количество информации, доставляемое самим символом x_i) определяется как – $\log p(x_1)$, т. е. информация о каком-либо событии измеряется логарифмом величины, обратной вероятности его *появления*. Выбор основания логарифма $\log p(x_1)$ определяет единицу количества информации. Если основание логарифма равно двум, то единица информации называется бит, если натуральному числу е – нат, а при основании логарифма 10 – дит. Переход от одной системы

логарифмов к другой равносилен простому изменению единицы измерения информации. Этот переход осуществляется по формуле.

Отсюда

$$\log_b k = \log_b a \log_a k, \qquad (2.18)$$

$$1 \text{ нат} = \log_2 e = 1{,}443 \text{ бит},$$

$$1 \text{ дит} = \log_2 10 = 3{,}32 \text{ бит}.$$

Таким образом, количество информации, содержащееся в дискретном сообщении, зависит от степени неожиданности этого сообщения, характеризуемой вероятностью его появления. Количество информации в сообщении тем больше, чем оно более неожиданно.

Источник удобнее характеризовать количеством собственной информации, содержащимся в среднем в одном символе ансамбля X, т. е. с помощью математического ожидания количества информации в отдельных событиях. Полагая $P_i = P(x_i)$, имеем

$$H(N) = -\sum_{i=1}^{N} p_i \log p_i. \qquad (2.19)$$

Это среднее количество собственной информации и названо энтропией. Мера (2.19) была предложена Клодом Шенноном в его фундаментальной работе "Математические основы теории связи", опубликованной в 1948 г., в которой были заложены основы «Современной теории информации».

Энтропию можно рассматривать и как меру количества информации, получаемой при полном устранении неопределенности относительно состояния источника.

В случае равновероятных состояний источника формула Шеннона имеет вид

$$H(N) = \log N. \qquad (2.20)$$

Понятие *энтропия* (от греческого “энтропе”–“обращение”) распространено в ряде областей знаний.

Так, Больцманом была определена энтропия физической системы следующим образом:

$$H = -\frac{1}{M_n} \sum_{i=1}^{N} m_i \ln \frac{m_i}{M_n}, \qquad (2.21)$$

где M_n – число молекул в данном пространстве,

m_i – число молекул ,обладающих скоростью $V_j + \Delta V$.

Заметим, что т. к. $\frac{m_i}{M_n}$ есть вероятность того, что молекула имеет скорость $V_j + \Delta V$, то формулу (2.20) можно записать как

$$H = -\sum_{i=1}^{N} p_i \ln p_i . \qquad (2.22)$$

Следует заметить, что энтропия характеризует среднюю неопределенность выбора одного состояния из ансамбля. При ее определении используют только вероятности состояний, полностью игнорируя их содержательную сторону. Поэтому энтропия не может служить средством решения любых задач, связанных с неопределенностью. Например, при использовании этой меры для оценки неопределенности действия лекарства, приводящего к полному выздоровлению больных, в 90 % случаев и улучшению самочувствия в остальных 10 %, она получится такой же, как и у лекарства, вызывающего в 90 % случаев смерть, а в 10 % – ухудшение состояния больных.

Энтропия как мера неопределенности согласуется с экспериментальными данными, полученными при изучении психологических реакций человека, в частности реакции выбора. Установлено, что время безошибочной реакции на последовательность беспорядочно чередующихся равновероятных раздражителей (например, загорающихся лампочек) растет с увеличением их числа так же, как энтропия. Это время характеризует неопределенность выбора одного раздражителя.

Замена равновероятных раздражителей неравновероятными приводит к снижению среднего времени реакции и ровно настолько, насколько уменьшается энтропия.

2.3.2. Свойства энтропии ансамбля

1. Энтропия $H(\vec{p})$ непрерывна на интервале $0 \le p \le 1$, т. е. при малых измерениях p величина H изменяется мало.

2. Энтропия $H(\vec{p})$ симметрична относительно p, т. е. не изменяется при любой перемене аргументов p_i.

3. Если событие X_n состоит из двух событий X'_n и X''_n с вероятностями q_1 и q_2, $q_1 + q_2 = p_N$, то

$$H(p_1, p_2, ...p_{N-1}, q_1, q_2) = H(p_1, p_2, ...p_N) + p_N \cdot H(\frac{q_1}{p_N}, \frac{q_2}{p_N}) \ . \quad (2.23)$$

4. Энтропия всегда неотрицательна.

5. Энтропия равна нулю, когда вероятность одного из событий равна единице.

6. Энтропия имеет наибольшее значение, когда все вероятности равны между собой $p_1 = p_2 = p_N$, (рис. 2.6) при этом

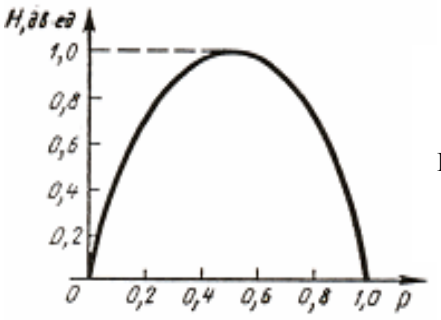


Рис. 2.6. Изменение энтропии в зависимости от вероятности элементарного события

$$H(N) = -\log \frac{1}{N} = \log N. \quad (2.24)$$

Энтропия характеризует среднюю неопределенность выбора одного состояния из ансамбля и определяется только вероятностями состояний.

Пример 2.3.

Заданы ансамбли X и Y двух дискретных случайных величин. Определить энтропию ансамбля X и Y.

$$X = \begin{vmatrix} 0.5 & 0.7 & 0.9 & 0.3 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{vmatrix}$$

$$Y = \begin{vmatrix} 5 & 10 & 15 & 8 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{vmatrix}$$

Решение: $H(X) = H(Y) = \log_2 4 = 2$.

Пример 2.4.

Определить максимум энтропии системы, состоящей из 6 элементов, каждый из которых может быть в одном из четырех состояний.

Решение:

$$H_{max} = \log_2 N = \log_2 m^n = \log_2 4^6 = 6 \cdot 2 = 12 \text{ бит.}$$.

2.3.3. Условная энтропия и энтропия объединения

Под объединением двух источников X и Y понимают обобщенный источник информации (X,Y), характеризующийся вероятностями $p(x_i, y_j)$ всех возможных комбинаций состояний x_i источника X и y_j источника Y. Аналогично трактуется и объединение ансамблей X и Y. В соответствии с определением энтропия объединения

$$H(x_i, y_j) = -\sum_{i=1}^{N}\sum_{j=1}^{k} p(x_i, y_j)\log p(x_i, y_j), \qquad (2.25)$$

здесь $p(x_i, y_j)$ – вероятности совместной реализации состояний,

x_i $(1 \le i \le N)$ и y_j $(1 \le j \le k)$.

В случае статистической независимости источников информации X и Y имеем

$$p(x_i, y_j) = p(x_i)\, p(y_j), \qquad (2.26)$$

тогда

$$H(X,Y) = -\sum_{i=1}^{N}\sum_{j=1}^{k} p(x_i)p(y_j)\log p(x_i)p(y_j) =$$

$$= -\sum_{i=1}^{N} p(x_i)\log p(x_i)\sum_{j=1}^{k} p(y_j) - \sum_{j=1}^{k} p(y_j)\log p(y_j)\sum_{i=1}^{N} p(x_i). \qquad (2.27)$$

Учитывая, что

$$\sum_{i=1}^{N} p(x_i) = 1, \qquad (2.28)$$

и

$$\sum_{j=1}^{k} p(y_j) = 1, \qquad (2.29)$$

получим

$$H(X\,Y) = H(X) + H(Y) = H(X,Y). \qquad (2.30)$$

При оценке неопределенности выбора часто необходимо учитывать статистические связи, которые в большинстве случаев имеют место:

а) между состояниями двух или нескольких источников, объединенных в рамках одной системы;

б) между состояниями, последовательно выбираемыми одним источником.

Определим энтропию объединения двух статистически связанных ансамблей X и Y. Объединение таких ансамблей характеризуется матрицей p (X,Y) вероятностей $p(x_i, y_j)$ всех возможных комбинаций состояний x_i, $(1 \leq i \leq N)$ ансамбля X и состояний y_j $(1 \leq j \leq k)$ ансамбля Y:

$$p(X,Y) = \begin{vmatrix} p(x_1,y_1)... & p(x_1,y_j)... & p(x_1,y_k) \\ ..... & ..... & ..... \\ p(x_i,y_1)... & p(x_i,y_j). & p(x_i,y_k) \\ ..... & ..... & ..... \\ p(x_N,y_1).. & p(x_N,y_j).. & p(x_N,y_k) \end{vmatrix}. \qquad (2.31)$$

Свойства матрицы вероятностей p (X,Y):

$$\sum_{i=1}^{N} p(x_i, y_j) = p(y_j), \qquad (2.32)$$

$$\sum_{j=1}^{k} p(x_i, y_j) = p(x_i). \qquad (2.33)$$

Вероятность совместной реализации взаимозависимых состояний x_i и y_j можно выразить через условие вероятности $p(x_i/y_j)$ или $p(y_j/x_i)$ в соответствии с тем, какие состояния принять за причину, а какие за следствия:

$$p(x_i, y_j) = p(x_i) \cdot p(y_j / x_i) = p(y_j) \cdot p(x_i / y_j), \qquad (2.34)$$

где p (x_i / y_j) – вероятность реализации состояний x_i ансамбля X при условии, что реализовалось состояние y_j ансамбля Y; $p(y_j/x_i)$

– вероятность реализации состояния y_j ансамбля Y при условии, что реализовалось состояние x_i ансамбля X. В общем случае, если мы передаем m сигналов A и ожидаем получить m сигналов B, влияние помех в канале связи полностью описывается канальной матрицей. Со стороны источника сообщений канальная матрица, описывающая канал связи, имеет вид

$$p(Y/X) = \begin{vmatrix} p(y_1/x_1).. & p(y_j/x_1).. & p(y_k/x_1).. \\ p(y_1/x_2).. & p(y_j/x_2).. & p(y_j/x_2).. \\ ..... & .... & .... \\ p(y_1/x_N).. & p(y_j/x_N).. & p(y_k/x_N).. \end{vmatrix}. \qquad (2.35)$$

Вероятности, расположенные по диагонали, определяют правильный прием остальные – ложный.

Значение цифр, заполняющих колонки канальной матрицы, обычно уменьшаются по мере удаления от главной диагонали, и при полном отсутствии помех, все, кроме цифр, расположенных по главной диагонали, равны нулю.

Если описывается канал связи со стороны источника сообщений, то прохождение данного вида сигнала в данном канале связи описывается распределением условных вероятностей вида $p(y_j / x_i)$, так для сигнала x_1 распределением вида является

$$p(y_1/x_1), p(y_2/x_1), ..., p(y_j/x_1), ..., p(y_k/x_1). \qquad (2.36)$$

Сумма вероятностей распределения (2.36) всегда должна равняться 1. Учитывая, что $p(y_j, x_i) = p(x_i) \cdot p(y_j / x_i)$ для определения энтропии статистически связанных ансамблей, имеем

$$H(X,Y) = -\sum_{i=1}^{N}\sum_{j=1}^{k} p(x_i) \cdot p(y_j/x_i) \cdot \log p(x_i) \cdot p(y_j/x_i) =$$

$$-\sum_{i=1}^{N} p(x_i) \cdot \log p(x_i) \cdot \sum_{j=1}^{k} p(y_j/x_i) -$$

$$-\sum_{i=1}^{N} p(x_i) \cdot \sum_{j=1}^{k} p(y_j/x_i) \cdot \log p(y_j/x_i). \qquad (2.37)$$

Выражение

$$-\sum_{j=1}^{k} p(y_j / x_i) \cdot \log p(y_j / x_i) = H_{x_i}(Y). \quad (2.38)$$

называется частной условной энтропией ансамбля Y. Частная условная энтропия есть случайная величина, характеризующая неопределенность, приходящуюся на одно состояние ансамбля Y, при условии, что реализовалось конкретное состояние x_i ансамбля X. При усреднении по всем состояниям ансамбля X, получаем среднюю неопределенность, приходящуюся на одно состояние ансамбля Y, при известных состояниях ансамбля X (общую условную энтропию ансамбля Y по отношению к ансамблю X).

$$H_{x_i}(Y) = H(Y/X) = \sum_{i=1}^{N} p(x_i) \sum_{j=1}^{k} p(y_j / x_i) \cdot \log p(y_j / x_i) = H_{x_i}(Y) . \quad (2.39)$$

Таким образом, из (2.37) следует, что энтропия объединения двух статистически связанных ансамблей X и Y равна безусловной энтропии одного ансамбля плюс условная энтропия другого относительно первого.

$$H(X\,Y) = H(X) + H_y(X), \quad (2.40)$$

Распространяя правило (2.37) на объединение любого числа зависимых ансамблей, получим

$$H(XYZ...W) = H(X) + H_x(Y) + H_{xy}(Z) + ... + H_{xyz}(W). \quad (2.41)$$

Свойства:

1. $$H_x(Y) \le H(Y), \quad (2.42)$$

$$H_y(X) \le H(X). \quad (2.43)$$

Действительно, наличие сведений о результатах реализации состояний одного ансамбля никак не может увеличить неопределенность выбора состояния из другого ансамбля. Эта неопределенность может только уменьшиться, если существует взаимосвязь в реализациях состояний из обоих ансамблей.

2. В случае отсутствия статистической связи в реализациях комбинаций состояний x_i источника X и y_j источника Y сведения о результатах выбора состояний из одного ансамбля не снижают неопределенности выбора состояний из другого ансамбля, что находит отражение в равенствах

$$H_X(Y) = H(Y), \; H_Y(X) = H(X).$$

Если имеет место однозначная связь в реализациях состояний x_i ,$(1 \le i \le N)$ из ансамбля X и y_j $(1 \le j \le k)$ из ансамбля Y, то условная энтропия любого из ансамблей равна нулю.

Действительно, условные вероятности $p(x_i/y_j)$ и $p(y_j/x_i)$ в этом случае принимают значения, равные нулю или единице. Поэтому все слагаемые, входящие в выражения (2.42) и (2.45) для частных условных энтропии, равны нулю. Тогда в соответствии с (2.43) и (2.46) условные энтропии также равны нулю.

2.3.4. Количество информации и избыточность

Количество информации есть уменьшение энтропии вследствие опыта.

Если неопределенность снимается полностью, то информация равна энтропии:

$$I = H, \tag{2.44}$$

в противном случае – разнице между начальной и конечной энтропией

$$I = H_1 - H_2. \tag{2.45}$$

Наибольшее количество информации получается, когда полностью снимается максимальная неопределенность – когда вероятности всех событий были одинаковы

$$I = -\log_2 p, \tag{2.46}$$

где p – вероятность реализации в условиях равной вероятности всех событий.

Если энтропия источника сообщений не равна максимальной энтропии для алфавита с данным количеством качественных признаков (имеются в виду качественные признаки алфавита, при помощи которых составляются сообщения), то это прежде всего означает, что сообщение данного источника могли бы нести большее количество информации.

Пример 2.5

Выстрел из орудия не поражает цель с вероятностью p. Через какое число выстрелов следует поинтересоваться, уничтожена

ли цель, чтобы в результате ответа получить максимальное количество информации?

Решение:

Ансамбль интересующих нас событий А, А1–цель поражена, А2– цель не поражена. Вероятность того, что цель не поражена после *к* выстрелов равна $p^к$. Вероятность противоположного события $1–p^к$, т. к. после ответа неопределенность устраняется полностью, количество информации равно энтропии, а она максимальна при равновероятности событий, т. е. при $p^к = 1– p^к$., откуда $к = –1 / \log_2 p$.

Избыточность информации (абсолютная недогруженность на символ сообщений такого источника) есть разность между максимально возможным количеством информации и энтропией

$$D_{абс} = I – H = H_{макс} – H \frac{бит}{символ}. \qquad (2.47)$$

Информационная избыточность показывает относительную недогруженность на символ алфавита и является безразмерной величиной

$$D = \frac{H_{макс} – H}{H_{макс}} = 1 – H_{макс}, \qquad (2.48)$$

где $\frac{H}{H_{макс}}$ – коэффициент сжатия.

Так как информация есть неопределенность, снимаемая при получении сообщения, то количество информации может быть представлено как произведение общего числа сообщений k на среднюю энтропию, приходящуюся на одно сообщение:

$$I = k \cdot H.$$

Для случая равновероятных и взаимонезависимых символов первичного алфавита количество информации в *k* сообщениях алфавита *m* равно:

$$I = k \cdot \log_2 m (бит). \qquad (2.49)$$

Для неравновероятных алфавитов энтропия на символ алфавита определяется следующим образом:

$$H = \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i} = \sum_{i=1}^{m} p_i \log_2 p_i \quad \text{(бит/символ)}, \qquad (2.50)$$

а количество информации в сообщении, составленном из *k* неравновероятных символов определяется как

$$I = k \cdot \sum_{i=1}^{m} p_i \log_2 p_i. \qquad (2.51)$$

Кроме общего понятия избыточности существуют частные виды избыточности:

1. Избыточность, обусловленная неравновероятным распределением символов в сообщении:

$$D_p = 1 - (\sum p_i \cdot \frac{\log_2 p_i}{\log_2 m}). \qquad (2.52)$$

2. Избыточность, вызванная статистической связью между символами сообщения:

$$D_S = 1 \frac{-(-\sum_i \sum_j p_{a_j} \cdot p(b_j / a_i) \cdot \log_2 P(b_j / a_i)}{\sum_i p_i \log_2 p_i}. \qquad (2.53)$$

Полная информационная избыточность равна

$$D = D_s + D_o - D_s D_p. \qquad (2.54)$$

Пример 2.6

Сообщения составлены из букв алфавита a, b, c, d. Вероятности букв: p(a) = 0,2; p(b) = 0,3; p(c) = 0,4; p(d) = 0,1. Найти избыточность сообщений D.

Решение:

Избыточность сообщений D вычисляется по формуле

$$D = 1 - \frac{H}{H_{макс}}.$$

Для алфавита из четырех букв максимальная энтропия равна

$$H_{max} = \log_2 m = \log_2 4 = 2 \text{бит} / \text{символ}.$$

Средняя энтропия на символ сообщения вычисляется следующим образом:

$$H = -\sum p_i \log_2 p_i = -(0,2 \cdot \log_2 0,2 + 0,3 \cdot \log_2 0,3 + 0,4 \cdot \log_2 0,4 + 0,1 \cdot \log_2 0,1 +$$

$$+0{,}4644+0{,}5211+0{,}5288+0{,}3322=1{,}8465\,\frac{\text{бит}}{\text{символ}}$$

$$D=-\frac{1{,}8465}{2}=1-0{,}9232=0{,}070688.$$

2.3.5. Дифференциальная энтропия

На практике множество возможных состояний источника информации составляет континуум, т. е. источники – непрерывные.

Определим энтропию непрерывного источника информации следующим образом: разобьем диапазон изменения непрерывной случайной величины U, характеризующейся плотностью распределения вероятностей p(и), на конечное число n малых интервалов шириной Δu. При реализации любого значения и, принадлежащего интервалу (u_i, $u_i+\Delta u$), будем считать, что реализовалось значение u_i дискретной случайной величины U. Поскольку Δu мало, вероятность $p(u_i \le u \le u_i+\Delta u)$ реализации значения u из интервала u_i, $u_i+\Delta u$ равна

$$p(u_i \le u \le u_i+\Delta u)=\int_{u}^{u+\Delta u} p(u)du \approx p(u_i)\Delta u. \quad (2.55)$$

Тогда энтропия дискретной случайной величины U может быть записана в виде

$$H(\tilde{U})=-\sum_{i=1}^{n} p(u_i)\Delta u \log\left[p(u_i)\Delta u\right]-\log\Delta u \quad (2.56)$$

или

$$H(\tilde{U})=-\sum_{i=1}^{n} p(u_i)\Delta u \log\left[p(u_i)\right]-\sum_{i=1}^{n} p(u_i)\Delta u \log\Delta u, \quad (2.57)$$

т. к.

$$\sum_{i=1}^{n} p(u_i)\Delta u=1, \quad (2.58)$$

то

$$H(\tilde{U})=-\sum_{i=1}^{n} p(u_i)\Delta u \log\left[p(u_i)\Delta u\right]-\log\Delta u. \quad (2.59)$$

По мере уменьшения величины Δu вероятность p ($u_i < u \le u_i+\Delta u$) все больше приближается к вероятности $p(u_i)$, равной нулю, а свойства дискретной величины U – к свойствам непрерывной

случайной величины U. Переходя к пределу при $\Delta u \to 0$, получаем следующее выражение для энтропии H(U) непрерывного источника:

$$\lim_{\Delta u \to 0} \tilde{H}(U) = \lim_{\Delta u \to 0} \left\{ -\sum_{i=1}^{n} p(u_i) \Delta u \log p(u_i) \right\} - \lim_{\Delta u \to 0} \log \Delta u \qquad (2.60)$$

или

$$H(U) = -\int_{-\infty}^{\infty} p(u) \log p(u) du - \lim_{\Delta u \to 0} \log \Delta u. \qquad (2.61)$$

Эта величина при $\Delta u \to 0$ стремится к бесконечности, что полностью соответствует интуитивному представлению о том, что неопределенность выбора из бесконечно большого числа возможных состояний (значений) бесконечно велика.

Первый член в правой части соотношения (2.61) имеет конечное значение, которое зависит только от закона распределения непрерывной случайной величины U и не зависит от шага квантования Δu . Он имеет точно такую же структуру, как энтропия дискретного источника.

Второй член того же соотношения, наоборот, зависит лишь от шага квантования случайной величины U, из-за него H(U) обращается в бесконечность.

Один из подходов для получения конечной характеристики информационных свойств непрерывного источника состоит в том, что в качестве неопределенности непрерывного источника принимают первый член соотношения (2.61)

$$h(U) = -\int_{-\infty}^{\infty} p(u) \log p(u) du . \qquad (2.62)$$

Она получила название относительной дифференциальной энтропии или просто *дифференциальной энтропии* непрерывного источника информации (непрерывного распределения случайной величины U).

Ее можно трактовать как среднюю неопределенность выбора случайной величины U с произвольным законом распределения по сравнению со средней неопределенностью выбора случайной величины U', изменяющейся в диапазоне, равном единице, и имеющей равномерное распределение.

Основные свойства дифференциальной энтропии:

а) если единственным ограничением для случайной величины U является область ее возможных значений [α, β], то максимальной дифференциальной энтропией обладает равномерное распределение вероятностей в этой области;

б) если ограничения на область значений непрерывной случайной величины U отсутствуют, но известно, что дисперсия ее ограничена, то максимальной дифференциальной энтропией обладает нормальное распределение величины U.

Пример 2.7. Дифференциальная энтропия гауссовского источника. Дифференциальная энтропия гауссовского источника равна:

$$\frac{H(X)}{\text{нат}} = -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \log_2 \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)\right] dx =$$

$$= -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(-\frac{x^2}{2\delta^2}\right) \log_2 \left[\frac{1}{\sqrt{2\pi\delta^2}} - \frac{x^2}{2\delta^2}\right] dx. \quad (2.63)$$

Выражение в квадратных скобках может быть разложено на два интеграла. Таким образом, окончательно имеем

$$\frac{H(X)}{\text{нат}} = \ln\sqrt{2\pi\sigma^2} + \frac{1}{2\sigma^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x^2 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \frac{1}{2}\ln\left[2\pi\sigma^2\right] + \frac{1}{2} =$$

$$= \frac{1}{2}\ln\left[2\pi\delta^2 e\right] \quad . \quad (2.64)$$

Численные примеры для трех, наиболее употребительных распределений приведены в табл. 2.1.

Таблица 2.1

Пример дифференциальной энтропии

Гауссовское	$f(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$	$\frac{\ln(2\pi\sigma^2 e)}{2\ln 2} = 2{,}04$
Равномерное	$f(x) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{для} \lvert x\rvert \le \sqrt{3} \\ 0 = \text{const} \end{cases}$	$\frac{\ln(2\sqrt{3}\sigma)}{\ln 2} = 1{,}79$
Лапласа	$f(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left(-\sqrt{2}\frac{\lvert x\rvert}{\sigma}\right.$	$\frac{\ln(\sqrt{2}\sigma e)}{\ln 2} = 1{,}94$

Пример 2.8. Телефония

Практическая польза приведенных выше результатов может быть наглядно показана при помощи оценки скорости передачи информации в битах в цифровых телефонных линиях. Современные стандартные методы цифровой передачи речи (логарифмические PCM) требуют затраты 8 бит на кодирование одного отсчета, при частоте отсчетов 8 кГц. Таким образом, скорость передачи речи составляет 64 кбит/с.

Исходя из равномерного распределения вероятностей в интервале [–1,1], опытным путем получим $\sigma^2 = 1/3$. Таким образом, дифференциальная энтропия на один отсчет составляет

$$\frac{H'(X)}{\text{бит}} = \frac{\ln(2\sqrt{3}\delta)}{\ln 2}\Big|_{\delta^2=\frac{1}{3}} = 1.$$

Так как отсчеты производятся с частотой 8 кГц, получаем, что необходимая скорость передачи речи составляет 8 кбит/с. При оценке энтропии мы не принимали во внимание связи между соседними отсчетами (память источника) и, поэтому, реальная дифференциальная энтропия источника речи будет еще меньше. Известно, что современные алгоритмы кодирования речи позволяют осуществлять передачу речевого сигнала со скоростью около 8 кбит/с при качестве, сравнимом со стандартным PCM.

2.3.6. Эпсилон-энтропия

Пусть:

1. Отдельные состояния источника информации есть независимые реализации случайной величины U. Ансамбль реализаций случайной величины U описывается плотностью распределения вероятностей p(и).

2. О значениях случайной величины U можно судить по значениям другой случайной величины Z, если мера их различия не превышает заданной точности воспроизведения.

3. Требование к точности задается с использованием критерия

$$\widetilde{V}(ZU) = \iint p(u)\, p_u(z)\,(z-u)^2\, du\, dz \le \varepsilon^2, \qquad (2.65)$$

где $p_u(z)$ – условная плотность распределения – функция правдоподобия того, что конкретный сигнал и будет воспроизведен как сигнал z; ε – заданное значение точности.

Так как плотность p(u) определена, то для выполнения условия (2.65) варьировать можно только условной плотностью распределения $p_и$ (z).

Если случайная величина Z воспроизводит случайную величину U с некоторой точностью ε, то количество информации, содержащееся в воспроизводящей величине Z относительно U, конечно и определяется по формуле

$$I(ZU) = \iint p(u) p_u(z) \log \frac{p_u(z)}{p(z)}, \qquad (2.66)$$

где p(z) = p(u) p_u(z) du – плотность воспроизводящей величины Z.

Желательно обеспечить заданную верность воспроизведения при минимальном количестве получаемой информации. Поэтому среди множества функций $p_u(z)$, удовлетворяющих условию (2.65), целесообразно выбрать такую, которая обеспечивает наименьшее I (ZU).

Минимальное количество информации в одной случайной величине Z относительно другой U, при котором удовлетворяется заданное требование к точности воспроизведения величины U, называется ε (эпсилон) – энтропией величины U и обозначается $H_ε(U)$:

$$H_\varepsilon(U) = \min_{(p_u(z))} I(ZU), \qquad (2.67)$$

при

$$V(ZU) \le \varepsilon^2. \qquad (2.68)$$

Большая практическая важность этой величины обусловлена тем, что *ε* (эпсилон) – энтропия есть минимальное количество двоичных цифр, которое нужно передать, чтобы воспроизвести сообщение с заданной верностью (точностью).

2.3.7. Информационные потери при передаче по дискретным каналам связи с шумами

Потери информации в каналах связи с шумами обычно описывают при помощи условной энтропии и энтропии объединения.

Если помех нет или их уровень настолько низок, что они не в состоянии уничтожить сигнал или имитировать полезный сигнал в отсутствие передачи, то при передаче x_i мы будем твердо уверены, что получим y_j – сигнал, соответствующий переданному a_i – му сигналу. События X и Y статистически жестко связаны, условная вероятность максимальна

$$p(y_1 / x_i) = 1, \tag{2.69}$$

а частная условная энтропия

$$H_{a_i}(Y) = -\sum_{i=1}^{m} p(y_j / x_i) \log p(y_j / x_i) = 0, \tag{2.70}$$

т. к. $\log p(y_j/x_i) = \log 1 = 0$.

В этом случае количество информации, содержащееся в принятом ансамбле сообщений Y равно энтропии передаваемых сообщений ансамбля X, т. е.

$$I(Y, X) = H(X).$$

(2.71)

При высоком уровне помех любой из принятых сигналов y_j может соответствовать любому переданному сигналу x_i, статистическая связь между переданными и принятыми сигналами отсутствует.

В этом случае вероятности $p(x_i)$ и $p(y_i)$ есть вероятности независимых событий и $p(y_j / x_i) = p(y_j)$, $p(x_i / y_j) = p(x_i)$.

$$H(X/Y) = -\sum_i \sum_j p(y_j) p(x_i / y_j) \log p(x_i / y_j) =$$

$$= -\sum_i \sum_j p(y_j) p(x_i) \log p(x_i) = \sum_i p(y_j) H(X) = H(X), \tag{2.72}$$

т. к $\sum p(y_i) = 1$, т. е. условная энтропия равна безусловной, а количество информации, содержащееся в Y относительно X равно нулю:

$$I(X,Y) = H(X) - H(X/Y) = 0. \quad (2.73)$$

Информационные характеристики реальных каналов связи лежат между этими двумя предельными случаями. При этом потери информации при передаче k символов по данному каналу связи вычисляются по формуле

$$I = kH(X/Y). \quad (2.74)$$

Так как имеет место

$$H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y), \quad (2.75)$$

то потери в канале связи могут быть учтены при помощи энтропии объединения следующим образом:

$$I(X,Y) = H(X) + H(X) - H(Y,X), \quad (2.76)$$

а с использованием условной энтропии определяется выражением

$$I(Y,X) = H(X) - H(X/Y) = H(Y) - H(Y/X). \quad (2.77)$$

Найденные зависимости можно пояснить при помощи диаграммы потоков информации (рис. 2.8).

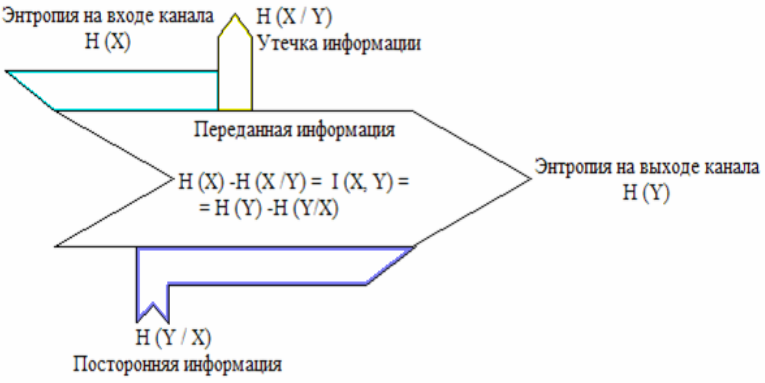


Рис. 2.8. Диаграмма потоков информации

Условная энтропия H(X/Y) – определяет среднюю меру неопределенности посланного символа источника X в том случае, когда символы приемника источника Y известны, т. е. оставшуюся неопределенность. Величину H (X/Y) часто называют также и утечкой информации, т. к. как энтропию H (X) можно интерпретировать как собственную информацию источника X, H (X) = I (X, X).

В бесшумном канале H (X/Y) = 0, т. е. информация передается от источника к приемнику и обратно без потерь.

Если канал полностью зашумлен, то H (X/Y) = H (X) и передачи информации не происходит.

H (Y/X) определяет среднюю неопределенность принятого символа при известных посланных символах, поэтому ее называют «посторонней» шумовой информацией.

2.3.8. Скорость передачи информации по дискретному каналу связи

Характеризуя дискретный канал связи, используют два понятия скорости передачи: технической и информационной.

Под технической скоростью передачи V_τ, называемой также скоростью манипуляции, подразумевает число элементарных сигналов (символов), передаваемых по каналу в единицу времени. Она зависит от свойств линии связи и быстродействия аппаратуры канала, с учетом возможных различий в длительностях символов

$$V_\tau = \frac{1}{\tau_{ср}}, \qquad (2.78)$$

где $\tau_{ср}$ – среднее значение длительности символа. При одинаковой продолжительности τ, всех передаваемых символов, $\tau_{ср} = \tau$. Единицей измерения технической скорости служит бод-скорость, при которой за одну секунду передается один символ.

Информационная скорость или скорость передачи информации, определяется средним количеством информации, которое передается по каналу в единицу времени. Она зависит как от характеристики данного канала связи, таких как объем алфавита исполь-

зуемых символов, техническая скорость их передачи, статистические свойства помех в линии, так и от вероятностей, поступающих на вход символов и их статистической взаимосвязи.

При известной скорости манипуляции V_τ, скорость передачи информации $V_u(X, Y)$ задается соотношением

$$V_u(X, Y) = V_\tau I(X, Y), \qquad (2.79)$$

где $I(X,Y)$ – среднее количество информации, переносимое одним символом.

Пример 2.9. Сообщения передаются в двоичном коде ($m = 2$), время передачи 0 $t_0 = 1$ сек., длительность импульса, соответствующего $t_1 = 5$ сек. Определить скорость передачи информации, если вероятность появления символа 0 $p_0 = 0{,}37$, вероятность появления символа 1 $p_1 = 0{,}63$.

Решение:

$$C = \frac{H_1}{\tau_{ср}} = \frac{-\sum_{i=1}^{m} p_1 \cdot \log_2 p_1}{\sum_{i=1}^{m} \tau_1 \cdot p_1} =$$

$$= \frac{-(0{,}37 \cdot \log_2 0{,}37 + 0{,}63 \cdot \log_2 0{,}63)}{0{,}63 \cdot 5 + 0{,}37 \cdot 1} = 0{,}27 .$$

2.3.9. Пропускная способность дискретного канала

Для теории и практики важно иметь представление, до какого предела и каким путем можно повысить скорость передачи информации по конкретному каналу связи. Предельные возможности канала по передаче информации, характеризуются его пропускной способностью.

Пропускная способность канала $C_П$ равна той максимальной скорости передачи информации по данному каналу, которой можно достичь при самых совершенных способах передачи и приема

$$C_п = \max V_u(X,Y) = \max V_\tau I(X,Y). \qquad (2.80)$$

При заданном алфавите символов и фиксированных основных характеристиках канала (например, полосе частот, средней и пиковой мощности передатчика) остальные характеристики должны

быть выбраны такими, чтобы обеспечить наибольшую скорость передачи по нему элементарных сигналов, т. е. обеспечить максимальное значение V_τ.

Максимум среднего количества информации, приходящейся на один символ принятого сигнала I(X,Y), определяется на множестве распределений вероятностей между символами $x_1, \ldots, x_i, \ldots, x_m$.

Пропускная способность канала измеряется числом двоичных единиц информации в секунду (бит/с).

При отсутствии помех существует взаимно - однозначное соответствие между множеством символов {Y} на выходе канала и {X} на его входе

$$I(X, Y) = I(Y, X) = H(X). \qquad (2.81)$$

Максимум возможного количества информации на символ равен log m, где m – объем алфавита символов, откуда пропускная способность дискретного канала без помех

$$C_п = V_u \log m. \qquad (2.82)$$

Следовательно, для увеличения скорости передачи информации по дискретному каналу без помех и приближения ее к пропускной способности канала, последовательность букв сообщения должна подвергнуться такому преобразованию в кодере, при котором различные символы в его выходной последовательности появлялись бы по возможности равновероятно, а статистические связи между ними отсутствовали бы.

При наличии помех соответствие между множествами символов на входе канала связи перестает быть однозначным.

Среднее количество информации, передаваемое в этом случае, определяется соотношением

$$I(X,Y) = H(X) - H(X/Y) = H(Y) - H(Y/X). \qquad (2.83)$$

Если объем алфавита входных символов X равен m_1, а выходных Y – m_2, то

$$I(X,Y) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} p(y_j, x_i) \log \frac{p(y_j, x_i)}{p(y_j) \cdot p(x_i)}. \qquad (2.84)$$

Соответственно скорость передачи информации по каналу с помехами вычисляется по формуле

$$V_u(X,Y) = V_\tau \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} p(y_j, x_i) \log \frac{p(y_j, x_i)}{p(y_j) \cdot p(x_i)}. \qquad (2.85)$$

Считая скорость манипуляции V_τ предельно допустимой при заданных технических характеристиках канала, величину $V_u(X,Y)$ можно максимизировать, изменяя статистические свойства последовательностей символов на входе канала посредством преобразователя (кодера канала).

Получаемое при этом предельное значение $C_П$ скорости передачи информации по каналу называют пропускной способностью дискретного канала связи с помехами

$$C_П = \max_{p(X)} V_\tau I(X,Y), \qquad (2.86)$$

где $p\{X\}$ – множество возможных распределений вероятностей входных сигналов. Важно подчеркнуть, что при наличии помех пропускная способность канала определяет наибольшее количество информации в единицу времени, которое может быть передано со сколь угодно малой вероятностью ошибки.

Контрольные вопросы:

1. Назовите основные меры информации.
2. Определите количество информации геометрическим методом.
3. Определите и сравните количество информации, полученные различными методами комбинаторики.
4. Вычислите количество информации мерой Хартли.
5. Проведите сравнительный анализ количества информации, вычисленного комбинаторным методом и мерой Хартли
6. Покажите связь вероятности и информации.
7. Определите понятие энтропии.
8. Перечислите основные свойства энтропии ансамбля.
9. Где используется понятие условной энтропии в теории информации?
10. Определите вероятность совместного появления взаимозависимых символов.

12. Определите состав канальной матрицы и ее свойства.
13. Запишите выражение для условной энтропии и поясните ее смысл.
14. Как определить энтропию нескольких взаимозависимых ансамблей?
15. Как связаны между собой понятия количества информации и энтропии?
16. Определите понятие избыточности.
17. Определите количество информации в сообщении, составленном из k неравновероятных символов?
18. Определите среднее количество информации в условиях помех, содержащееся в принятом ансамбле сообщений B относительно переданного с помощью энтропии объединения.
19. Определите среднее количество информации в условиях помех, содержащееся в принятом ансамбле сообщений B относительно переданного с помощью условной энтропии.
20. Как определяются информационные потери при передаче сообщений по каналам связи с шумами?
21. Приведите основные виды измерения информации в семиотике.
22. В чем смысл измерения информации при семантическом методе?
23. В чем смысл содержательности информации?
24. Как определяется содержательность информации?
25. Понятие целесообразности информации, и ее физический смысл.
26. Определите понятие технической скорости передачи.
27. Определите понятие скорости передачи информации.
28. Поясните сущность понятия пропускной способности канала.
29. Запишите выражения для пропускной способности дискретного канала с помехами и без помех.
30. Дайте определение дифференциальной энтропии и сформулируйте ее основные свойства.
31. Какие распределения обладают максимальной дифференциальной энтропией:

а) при ограничении на диапазон изменения случайной величины;
б) при ограничении на дисперсию случайной величины.

3. МАТЕМАТИЧЕСКИЕ МОДЕЛИ СИГНАЛОВ

3.1. Сигнал и его модели

Понятие «сигнал» имеет неоднозначное толкование. Под сигналом в широком толковании понимают материальный носитель информации, он может быть естественным или специально созданным для определенных целей. В книге будем понимать сигнал в узком смысле, т. е. *сигнал* – материальный носитель информации, специально создаваемый для передачи сообщения в информационной системе.

В качестве носителей информации используются *колебания различной природы* [3,9]. Колебания принято подразделять на *детерминированные и случайные.*

Детерминированные колебания точно определены в любые моменты времени.

Случайные колебания могут иметь параметры, значения которых предсказать невозможно.

Сигнал принципиально представляет собой случайное колебание и его аналитической моделью может быть случайный процесс. Но изучение моделей детерминированных сигналов необходимо по многим причинам:

1) случайный процесс может быть представлен набором детерминированных функций;

2) детерминированные сигналы специально создаются для целей измерения, наладки, регулирования объектов информационной техники.

Сигналы (как правило, электрические) являются носителями информации лишь при определенном воздействии (модуляции) на один или несколько их параметров [8,10,11]. Такие параметры принято называть информативными или кодовыми. Сигналы по своей структуре разделяются на непрерывные и дискретные. Непрерывные сигналы принимают непрерывные множества значений на некотором отрезке времени и в диапазоне ограничивающем его максимальную и минимальную величины (примеры непрерывных сигналов: постоянный ток и напряжение, гармонические колебания тока). В первом случае информативным параметром сигнала может

быть только величина тока или напряжение, во втором случае – может быть амплитуда, частота или фаза колебаний. Непрерывный сигнал служит, как правило, для передачи непрерывного сообщения (речь, музыка, телевизионное изображение). Непрерывные сообщения преобразуют в электрические непрерывные сигналы. Например, если в качестве сообщения служит речь, то микрофон преобразует звуковые колебания воздушной среды в электрические колебания. Электрические сигналы передаются в место приема или их используют для модуляции высокочастотного колебания.

Дискретным называется сигнал, принимающий конечное множество значений в определенном интервале времени и диапазоне величин. С помощью дискретного сигнала передается как дискретное, так и непрерывное сообщение. Дискретные сообщения – это конечная последовательность отдельных символов (букв). Длительность последовательности ограничены. Пример: телеграмма

Используются дискретные сообщения в телеграфии, при передаче данных и телеметрии. Для передачи дискретного сообщения, т. е. преобразования его в сигнал необходимо выполнить операцию кодирования сообщения.

Для передачи непрерывных сообщений дискретными методами непрерывные сигналы с помощью операций дискретизации во времени квантования по уровню преобразуют в дискретные, которые затем передают дискретными методами. Дискретные сигналы могут представляться в виде отдельных простейших элементов (посылок) или совокупности таких элементов. Такими элементами могут быть:

1) однополярные импульсы напряжения;

2) синусоидальные изменяющиеся напряжения (ток определенной длительности);

3) периодическая последовательность импульсов напряжения (токов определенной длительности);

4) пауза (отсутствие напряжения).

Каждому элементу (посылке) свойственны свои качественные признаки (признаки посылок). Один или несколько из этих признаков могут быть использованы для передачи сообщения.

Для однополярных импульсов качественными признаками являются величина, длительность и временное положение импульсов,

для гармонических колебаний – амплитуда, частота и фаза колебаний, длительность и временное положение посылки. Для периодической последовательности импульсов качественными признаками являются амплитуда, длительность, полярность, частота следования и фаза импульсов, длительность и временное положение посылки. Применяют следующие способы разделения посылок:

1) полярное разделение;

2) амплитудное разделение: смежные посылки имеют различные значения амплитуд;

3) фазовое разделение, используемое для посылок в виде синусоидального колебания;

4) временное разделение, при котором посылки разделены друг от друга временным интервалом;

5) частотное разделение, при котором смежные посылки различаются друг от друга частотой;

6) простое разделение, при котором посылки сигнала передаются одновременно по разным каналам связи.

Посылки сигнала сообщения, обладающие информативными кодовыми признаками, называют активными, а не обладающие такими признаками и служащие для разделения посылок называют разделяющими.

Необходимо найти такое представление сигнала, которое облегчает задачи исследования прохождения реальных сигналов через системы связи.

С этой целью сложные сигналы представляются совокупностью базисных элементарных сигналов. Тогда сложный сигнал $u(t)$ можно представить в виде взвешенной суммы базисных функций $\Psi_k(t)$:

$$u(t) = \sum_k C_k \Psi_k(t). \qquad (3.1)$$

Если в качестве базисных выбираются ортогональные функции, т. е. такие $\Psi_0(t), \Psi_1(t), ..., \Psi_n(t)$ на отрезке $[t_a, t_b]$, что для всех $k = \overline{0, n}$, $j = \overline{0, n}$, кроме $k = j$, имеет место

$$\int_{t_a}^{t_b} \Psi_k(t)\Psi_j(t)dt = 0. \qquad (3.2)$$

Эта система функций будет ортонормированной, если для всех $j=\overline{0,n}$ справедливо соотношение

$$\int_{t_a}^{t_b}\Psi_j^{\,2}(t)dt=1\cdot \qquad (3.3)$$

Определим коэффициенты C_k при представлении сигнала $u(t)$ совокупностью ортонормированных функций в виде (3.1). Правую и левую части уравнения (3.1) умножаем на $\Psi_j(t)$ и интегрируем на интервале $[t_1,t_2]$, где $[t_1,t_2]\in[t_a,t_b]$

$$\int_{t_1}^{t_2}u(t)\cdot\Psi_j(t)dt=\sum_k C_k\cdot\int_{t_1}^{t_2}\Psi_j(t)\cdot\Psi_k(t)dt. \qquad (3.4)$$

В силу справедливости (3.2) все интегралы в правой части выражения (3.4) при $k\neq j$ будут равны нулю. При $k=j$ в соответствии с (3.3) интеграл равен единице

$$C_k=\int_{t_2}^{t_1}u(t)\Psi_j(t)dt\cdot$$

3.2. Временная форма представления сигнала

Временным представлением сигнала называется такое разложение сигнала $u(t)$ [12,16], при котором в качестве базисных функций используются единичные импульсные функции – *дельта - функции*:

$$\delta(t)=\begin{cases}\infty,\text{при} & t=\xi_1\\ 0,\text{при} & t\neq\xi_1\end{cases},\qquad \int_{-\infty}^{\infty}\delta(t-\xi_1)dt=1\cdot$$

Единственным параметром, правильно выражающим реальный сигнал, является время его действия. С помощью δ – функции значение реального сигнала можно выразить как

$$u(t)=\int_{-\infty}^{\infty}U(\xi)\delta(t-\xi)d\xi. \qquad (3.5)$$

Символическое изображение дельта - функции приведено на рис. 3.1.

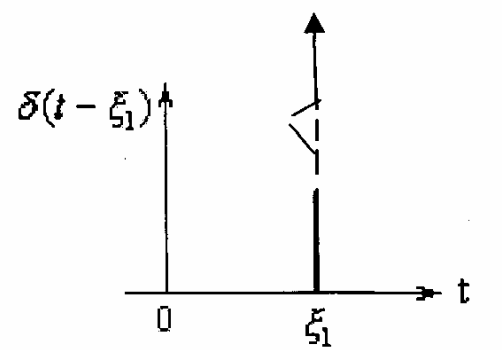


Рис. 3.1. Символическое изображение дельта-функции

Таким образом, функция $u(t)$ выражена в виде совокупности примыкающих друг к другу импульсов бесконечно малой длительности. Разложение (3.5) имеет большое значение в теории линейных систем [12,16], т. к. установив реакцию системы на элементарный входной сигнал в виде дельта-функции, можно определить реакцию системы на произвольный входной сигнал как суперпозицию реакций на бесконечную последовательность смещенных дельта-импульсов с «площадями», равными соответствующим значениям входного сигнала.

3.3. Частотная форма представления сигнала

В этом случае в качестве базисных функций выбраны

$$\Psi_j(t) = e^{jk\omega_1 t}. \qquad (3.6)$$

Для периодического сигнала u(t) коэффициенты C_K для базисных функций называются спектром и определяются

$$\underline{C}_k = \underline{A}_k = \frac{2}{T}\int_{-T/2}^{T/2} u(t)e^{-jk\omega_1 t}dt\ , \qquad (3.7)$$

где k = 0, $\pm 1,\pm 2,\pm 3,...$; T– период сигнала, $\omega_1 = \frac{2\pi}{T}$.

Представление периодического сигнала u(t) с помощью базисных функций (3.6) называется разложением в ряд Фурье [6] в комплексной форме и имеет вид

$$u(t) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \underline{A}_k e^{jk\omega_1 t} .$$

Тригонометрическая форма разложения в ряд Фурье имеет вид

$$u(t) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \underline{A}_k \cos(k\omega_1 t - \varphi_k)$$

или

$$u(t) = a_0 + \sum_{k=1}^{\infty} (a_k \cos k\omega_1 t + b_k \sin k\omega_1 t),$$

где $A_k \cos(k\omega_1 t - \varphi_k)$ – *k*-я гармоническая составляющая, $A_k, k\omega_1, \varphi_k$ – амплитуда, частота и начальная фаза *k*-й гармонической составляющей; a_0 –постоянная составляющая и выражает среднее значение сигнала за период

$$a_0 = \frac{1}{T} \int_{-T/2}^{T/2} u(t)dt ; \; a_k = A_k \cos\varphi_k ; \; b_k = A_k \sin\varphi_k ; \; A_k = \sqrt{a_k^2 + b_k^2} ;$$

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

где $\underline{A}_k$ – комплексная амплитуда гармонической составляющей частоты $\omega_k = k\omega_1$ связана с коэффициентами a_k и b_k следующим образом: $\underline{A}_k = a_k - jb_k$.

Характерной особенностью спектра периодического сигнала является его прерывность (дискретность). Расстояния между соседними спектральными линиями одинаковы и равны частоте основной гармоники [11, 18].

Непериодический сигнал можно рассматривать как периодический, период изменения которого равен бесконечности. При увеличении периода T интервал между смежными частотами в спектре сигнала и амплитуды спектральных составляющих уменьшаются и в пределах $T \to \infty$ становятся бесконечно малыми величинами. При этом ряд Фурье, отображающий спектральное разложение периодического сигнала, преобразуется в интеграл Фурье, отобра-

жающий спектральное разложение непериодического сигнала:

$$u(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(j\omega)\exp\{j\omega t\}d\omega, \qquad (3.8)$$

где $S(j\omega) = |S(j\omega)|\exp\{j\varphi(\omega)\}$ – спектральная плотность;

$|S(j\omega)| = S(\omega)$ – амплитудно-частотная характеристика сигнала;

$\varphi(\omega)$ – фазочастотная характеристика сигнала.

Выражение (3.8) называется формулой обратного преобразования Фурье.

Спектральная плотность связана с временной функцией сигнала через прямое преобразование Фурье

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

Спектральная плотность однозначно отображает непериодический сигнал и удовлетворяет условиям:

1. $\lim\limits_{\omega \to \infty} S(\omega) = 0.$

2. Модуль спектральной плотности является четной, а аргумент – нечетной функцией частоты, т. е. $S(\omega) = S(-\omega)$ и $\varphi(\omega) = -\varphi(-\omega)$.

Спектральное представление сигнала. Совокупность гармонических составляющих, образующих сигнал несинусоидальной формы называют спектром этого негармонического сигнала.

Спектр наглядно можно представить в виде спектральной диаграммы – так принято называть графическое изображение коэффициентов ряда Фурье для конкретного сигнала. Различают амплитудные и фазовые спектральные диаграммы. При этом по горизонтальной оси в некотором масштабе откладывают частоты гармоник, а по вертикальной оси представлены их амплитуды и начальные фазы.

Например, если некоторый периодический несинусоидальный сигнал описывается рядом

$$u(t) = 100\sin(\omega_1 t + 40^0) + 50\sin(2\omega_1 t + 60^0) + 25\sin(3\omega_1 t + 20^0),$$

то его амплитудная спектральная диаграмма представлена на рис.3.2, *а* , а частотная – на рис. 3.2, *б*.

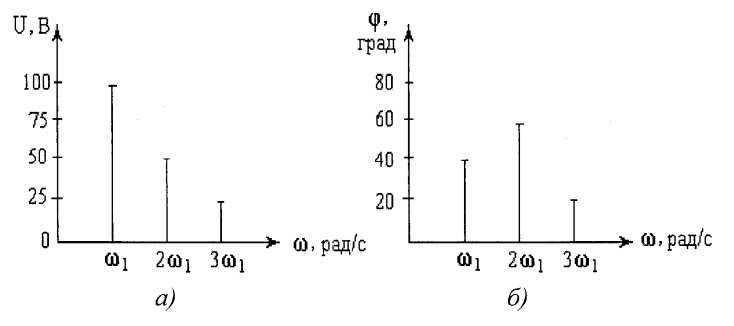


Рис.3.2. Спектральная диаграмма сигнала

Пример 3.1. Вычислить комплексные амплитуды гармонических составляющих и записать ряд Фурье для колебания, изображенного на рис. 3.3.

Аналитическое выражение этого сигнала:

$$u(t)=\begin{cases} E \\ -E \end{cases} \quad при \quad \begin{matrix} 2m\dfrac{T}{2} < t < (2m+1)\dfrac{T}{2} \\ (2m+1)\dfrac{T}{2} < t < 2(m+1)\dfrac{T}{2} \end{matrix},$$

где m – целое число.

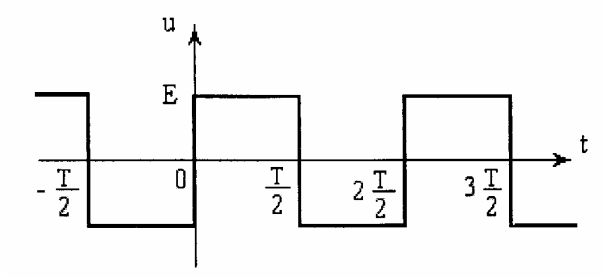


Рис. 3.3. Форма разлагаемого сигнала

В соответствии с (3.7) и учитывая формулу Эйлера:

$\cos\alpha = \dfrac{e^{j\alpha} + e^{-j\alpha}}{2}$, а также $\omega_1 = \dfrac{2\pi}{T}$, имеем

$$\underline{A}_k = \frac{2}{T}\int_{-T/2}^{T/2} U(t)e^{-jk\omega_1 t}dt = \frac{2}{T}\int_{-T/2}^{0}(-E)e^{-jk\omega_1 t} + \frac{2}{T}\int_{0}^{T/2} Ee^{-jk\omega_1 t}dt =$$

$$= -\frac{2E}{T}\frac{e^{-jk\omega_1 t}}{-jk\omega_1 t}\Bigg|_{-T/2}^{0} + \frac{2E}{T}\frac{e^{-jk\omega_1 t}}{-jk\omega_1 t}\Bigg|_{0}^{T/2} = \frac{2E}{jk\pi}(1-\cos k\pi)\cdot$$

Так как значение cosk π при k = 0, 2, 4, 6 равно единице, то нулевой составляющей и четных гармоник данное колебание не содержит.

При $k = 1$

$$\underline{A}_1 = \frac{2E}{j\pi}(1+1) = \frac{4E}{\pi}e^{-j\frac{\pi}{2}},\ ,(-j = e^{-j\frac{\pi}{2}})\,..$$

Аналогично при $k = 3, 5,\ldots$

$$\underline{A}_3 = \frac{2E}{j3\pi}(1+1) = \frac{4E}{3\pi}e^{-j\frac{\pi}{2}},\quad \underline{A}_5 = \frac{2E}{j5\pi}(1+1) = \frac{4E}{5\pi}e^{-j\frac{\pi}{2}}\,..$$

При переходе от комплексной формы представления к тригонометрической получим

$$u(t) = \frac{4E}{\pi}\cos(\omega_1 t - \frac{\pi}{2}) + \frac{4}{3}\frac{E}{\pi}\cos(3\omega_1 t - \frac{\pi}{2}) + \frac{4}{5}\frac{E}{\pi}\cos(5\omega_1 t - \frac{\pi}{2}) =$$

$$= \frac{4E}{\pi}\sin\omega_1 t + \frac{4E}{3\pi}\sin 3\omega_1 t + \frac{4E}{5\pi}\sin 5\omega_1 t.$$

Спектральная диаграмма представлена на рис. 3.4.

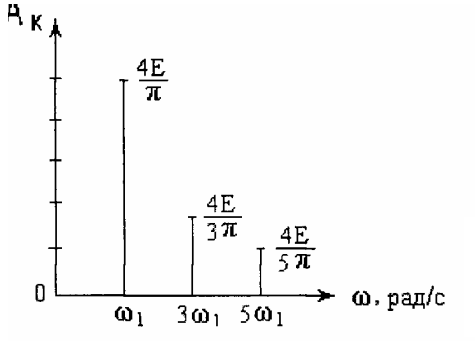


Рис. 3.4. Спектральная диаграмма последовательности прямоугольных импульсов

Пример 3.2. Задан одиночный прямоугольный импульс (рис.3.5).

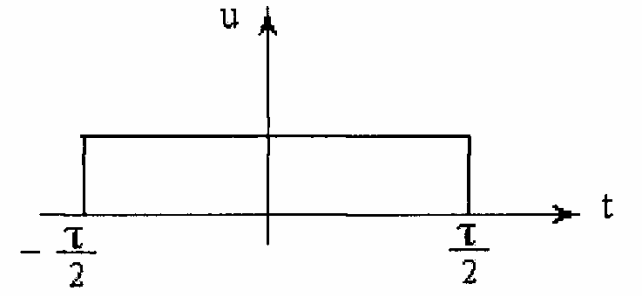


Рис. 3.5. Одиночный прямоугольный импульс

Уравнение одиночного прямоугольного импульса имеет вид

$$u(t)=\begin{cases} E & \text{при } -\dfrac{\tau}{2}<t<\dfrac{\tau}{2},\\ 0 & \text{в противном случае.}\end{cases}$$

Необходимо определить спектральную функцию (рис. 3.6).

$$S(j\omega)=\int_{-\infty}^{\infty}u(t)e^{-j\omega t}dt=\int_{-\tau/2}^{\tau/2}Ee^{-j\omega t}dt=\frac{Ee^{-j\omega t}}{-j\omega}\Bigg|_{-\tau/2}^{\tau/2}=$$

$$=\frac{E\,e^{-j\omega\tau/2}}{-j\omega}-\frac{E\,e^{j\omega\tau/2}}{-j\omega}=E\tau\frac{\sin\dfrac{\omega\tau}{2}}{\dfrac{\omega\tau}{2}}=g\frac{\sin\dfrac{\omega\tau}{2}}{\dfrac{\omega\tau}{2}},$$

где $g = E \cdot \tau -$ площадь импульса.

3.4. Распределение энергии сигнала в спектре

Рассмотрим непериодический сигнал u(t), физическим представлением которого, будем считать электрическое напряжение на резисторе с сопротивлением в 1 Ом.

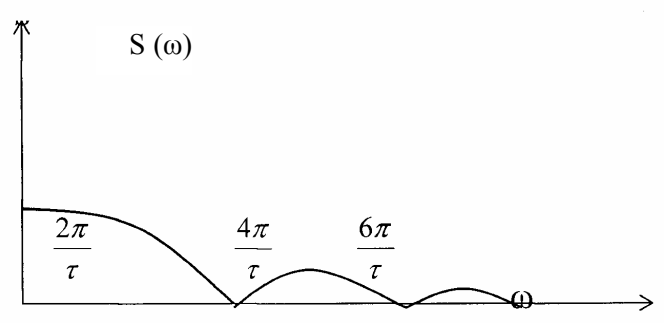


Рис. 3.6. Спектральная диаграмма одиночного прямоугольного импульса

Тогда энергия, выделяемая на этом резисторе, определяется

$$E = \int_{-\infty}^{\infty} u^2(t) d t . \qquad (3.9)$$

В предположении, что интеграл (3.9) сходится, выразим энергию через модуль спектральной характеристики S(ω) сигнала u(t). Квадрат этого модуля запишем в виде [8,17]

$$|S(\omega)|^2 = S(j\omega) S(-j\omega), \qquad (3.10)$$

где $S(-j\omega) = \int_{-\infty}^{\infty} u(t) e^{j\omega t} d t$ – функция, комплексно -сопряженная спектральной характеристике S(jω) сигнала u(t).

Тогда

$$\int_{-\infty}^{\infty} |S(\omega)|^2 = \int_{-\infty}^{\infty} S(j\omega) S(-j\omega) d\omega = \int_{-\infty}^{\infty} S(j\omega) \int_{-\infty}^{\infty} u(t) e^{j\omega t} d t d\omega .$$

После изменения последовательности интегрирования и использования обратного преобразования Фурье (3.8) получим

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

Окончательно имеем

$$\int_{-\infty}^{\infty} |u(t)|^2\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |S(\omega)|^2\, d\omega = \frac{1}{\pi}\int_{0}^{\infty} |S(\omega)|^2\, d\omega. \qquad (3.11)$$

Соотношение (3.11) известно как *равенство Парсеваля.* Оказывается, что энергию, выделяемую непериодическим сигналом за время его существования, можно определить, интегрируя квадрат модуля его спектральной характеристики в интервале частот.

Каждое из бесконечно малых слагаемых $(1/\pi)\,|S(\omega)|^2 d\omega$, соответствующих бесконечно малым участкам спектра, характеризует энергию, приходящуюся на спектральные составляющие сигнала, сосредоточенные в полосе частот от ω до ω + *dω.*

Анализируя спектр одиночного прямоугольного импульса (рис. 3.6) можно установить, что при увеличении его длительности от 0 до ∞ спектр сокращается от безграничного (у дельта-функции) до одной спектральной линии в начале координат, соответствующей постоянному значению сигнала. Это свойство сокращения ширины спектра сигнала при увеличении его длительности и наоборот справедливо для сигналов любой формы. Оно вытекает непосредственно из особенностей прямого и обратного интегрального преобразования Фурье, у которых показатель степени экспоненциальной функции в подынтегральных выражениях имеет переменные t и ω в виде произведения.

Рассмотрим функцию u(t) определенной продолжительности и функцию u(λt), длительность которой при λ>1 будет в λ раз меньше. Считая, что u(t) имеет спектральную характеристику S(jω), найдем соответствующую характеристику $S_\lambda(j\omega)$ для *u(λt):*

$$S_\lambda(j\omega) = \int_{-\infty}^{\infty} u(\lambda t)\, e^{-j\omega t} dt = = \frac{1}{\lambda}\int_{-\infty}^{\infty} u(t')\, e^{-j\frac{\omega t'}{\lambda}} dt' = \frac{1}{\lambda} S\left(j\frac{\omega}{\lambda}\right), \qquad (3.12)$$

где t' =λt.

Следовательно, спектр укороченного в λ раз сигнала ровно, в λ раз шире. Коэффициент $\frac{1}{\lambda}$ перед $S(j\frac{\omega}{\lambda})$ изменяет только амплитуду гармонических составляющих и на ширину спектра не

влияет.

Реальные сигналы ограничены во времени и не могут содержать гармонические составляющие сколь угодно высоких частот.

В этой связи, используя энергетический критерий [6,11], определим практическую длительность T и практическую ширину спектра ω так, чтобы в них была сосредоточена подавляющая часть энергии сигнала.

Для сигналов, начинающихся в момент времени $t_0 = 0$, практическая длительность определяется из соотношения

$$\int_0^{T_п} |u(t)|^2 \, dt = \eta \int_0^{\infty} |u(t)|^2 \, dt, \qquad (3.13)$$

где η – коэффициент, достаточно близкий к 1 (от 0,9 до 0,99 в зависимости от требований к качеству воспроизведения сигнала).

Принимая во внимание равенство Парсеваля (3.11), для практической ширины спектра сигнала соответственно имеем

$$\frac{1}{\pi} \int_0^{\omega_п} |S(\omega)|^2 \, d\omega = \frac{\eta}{\pi} \int_0^{\infty} |S(\omega)|^2 \, d\omega. \qquad (3.14)$$

Для каждого импульса определена база:

$$B = \tau \cdot \Delta f, \qquad (3.15)$$

где τ – длительность импульса (с); Δf – ширина спектра импульса в Гц.

База импульса для каждого импульса постоянная, $1.13 \geq B \geq 0.2\,2$`.

База характеризует экономичность сигнала, наиболее экономичным с этой точки зрения является колокольный импульс, который требует наименьшей полосы частот $\Delta\omega$ при заданной длительности τ.

При грубых оценках, достаточных для практики, считают $\Delta f \approx \frac{1}{\tau}$.

Свойства спектров сигналов:

1. Спектры сигналов обладают бесконечной протяженностью, имеют тенденцию к затуханию с увеличением частоты.

2. Форма спектра, степень и характер затухания зависит от формы и длительности сигнала.

3. Амплитуда импульса не влияет на ширину импульса, а влияет на амплитуду спектра.

4. Увеличение длительности сигнала в α раз приводит к сужению его спектра во столько же раз.

3.5. Случайный процесс как модель сигнала

Наряду с полезными информационными составляющими в реальных сигналах присутствуют помехи и шумы. К помехам обычно относят сигналы от других посторонних источников, "наводки" аппаратуры, влияние дестабилизирующих факторов на основной сигнал и т. п., физическая природа которых не случайна, после соответствующего изучения может переводиться в разряд детерминированной помехи и исключаться из сигнала. К шумам – случайные флуктуации сигнала, обусловленные природой его источника или устройств детектирования и формирования сигнала. При неизвестной природе помех они также могут относиться к числу случайных.

Теория вероятностей рассматривает случайные величины и их характеристики в "статике". Задачи описания и изучения случайных сигналов "в динамике", как отображения случайных явлений, развивающихся во времени или по любой другой переменной, решает теория случайных процессов.

В качестве универсальной координаты для распределения случайных величин по независимой переменной будем использовать, как правило, переменную "t" и трактовать ее, чисто для удобства, как временную координату. Распределения случайных величин во времени равно и сигналов их отображающих в любой математической форме, обычно называют случайными (или стохастическими) процессами. В технической литературе термины "случайный сигнал" и "случайный процесс" используются как синонимы.

В отличие от детерминированных сигналов значения случайных сигналов в произвольные моменты времени не могут быть вычислены. Они могут быть только предсказаны в определенном диапазоне значений с определенной вероятностью, меньшей единицы.

Количественные характеристики случайных сигналов, позволяющие производить их оценку и сравнение, называют статистическими.

В процессе обработки и анализа физико-технических данных обычно приходиться иметь дело с тремя типами сигналов, описываемых методами статистики. Во-первых, это информационные сигналы, отображающие физические процессы, вероятностные по своей природе, как, например, акты регистрации частиц ионизирующих излучения при распаде радионуклидов.

Во-вторых, информационные сигналы, зависимые от определенных параметров физических процессов или объектов, значения которых заранее неизвестны, и которые обычно подлежат определению по данным информационным сигналам. И, в-третьих, шумы и помехи, хаотически изменяющиеся во времени, которые сопутствуют информационным сигналам, но, как правило, статистически независимы от них как по своим значениям, так и по изменениям во времени. При обработке таких сигналов обычно ставятся задачи:

- обнаружение полезного сигнала;
- оценка параметров сигнала;
- выделение информационной части сигнала (очистка сигнала от шумов и помех);
- предсказание поведения сигнала на некотором последующем интервале (экстраполяция).

Таким образом, реальные сигналы (помехи) являются случайными процессами [12]. Случайный процесс описывается случайной функцией x(t), значение которой при любом значении аргумента является случайной величиной. При неизменяемых условиях опыта случайный процесс x(t) может принимать ту или иную конкретную форму $x_n(t)$.

Конкретный вид случайного процесса, зарегистрированный в определенном опыте, называют его реализацией.

Совокупность всех возможных реализаций $\{x_n(t)\}$ случайного процесса называется ансамблем.

Основные признаки, по которым классифицируются случайные процессы: пространство состояний, временной параметр, ста-

тистические зависимости между случайными величинами x (t_i) в разные моменты времени t_i.

Пространством состояний называют множество возможных значений случайной величины x(t_i):

- Случайный процесс, у которого множество состояний континуум, а изменение состояний возможны в любые моменты времени – непрерывный случайный процесс.
- Если изменения состояний возможны в конечном или счетном числе моментов времени – непрерывная случайная последовательность.
- Случайные процессы с конечным множеством состояний, которые могут изменяться в произвольные моменты времени – дискретные случайные процессы.
- Если изменение состояния возможны только в конечном или счетном числе моментов времени – дискретная случайная последовательность.

Среди случайных процессов с дискретным множеством состояний, есть такие, у которых статистические зависимости распространяются на ограниченное число k, следующих друг за другом значений. Они – обобщенные процессы k-го порядка,

В соответствии с определением, случайный процесс x(t) может быть описан системой N, обычно зависимых случайных величин $x_1 = x(t_1), ..., x_i = x(t_i), ..., x_N = x(t_N)$, , взятых в различные моменты $t_1, ..., t_i, ..., t_N$. При неограниченном увеличении N такая система эквивалентна рассматриваемому случайному процессу x(t).

Исчерпывающей характеристикой указанной системы является N - мерная плотность вероятности $W_N(x_1, ..., x_i, ..., x_N, t_1, ..., t_i, ..., t_N)$. Она позволяет вычислить вероятность p_N реализации, значения которой в моменты времени $t_1, ..., t_i, ..., t_N$ будут находиться соответственно в интервалах $(x_1, x_1 + \Delta x_1), ..., (x_i, x_i + \Delta x_i), ..., (x_N, x_N + \Delta x_N)$, где $x_i \ (1 \le i \le N)$ – значение, принимаемое случайной величиной x_i.

3.5.1. Вероятностные характеристики случайных процессов

Пусть x(t) случайный процесс (рис. 3.7). В некоторый фиксированный момент t_2 различные реализации его имеют различные значения $x_1(t_2), x_2(t_2),..., x_n(t_2)$. Значения $x_n(t_k)$ является случайной величиной $k = 1,...,n$.

Полной характеристикой случайного процесса является n-мерная функция распределения $P_n(x_1, t_1,..., x_n, t_n)$.

Одномерная функция распределения $P_1(x_1, t_1)$ определяется как вероятность того, что случайная величина $x(t_1)$ не превышает некоторого значения x_1 $P_1(x_1, t_1) = P[x(t_1) \le x_1]$.

Частная производная $\frac{\partial P_1(\xi_1, t_1)}{\partial x_1} = w_1(\varsigma_1, t_1)$ – одномерная плотность вероятности случайного процесса $\{x_n(t)\}$. Для времени $t \ge t_1$ характеризует распределение одной случайной величины x_1, взятой в производный момент t_1, и не отражает зависимость случайной величины в различные моменты времени.

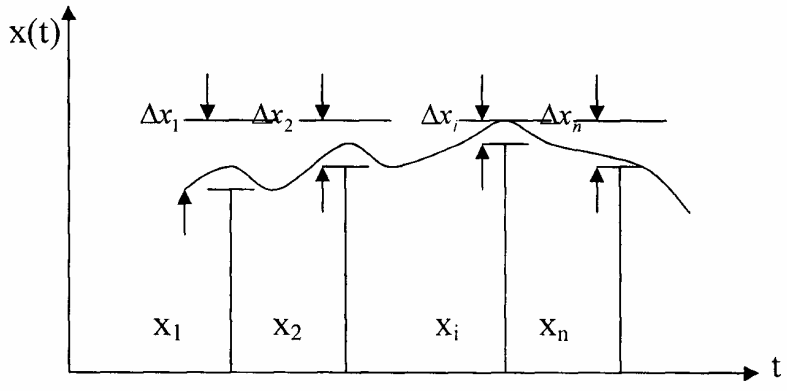


Рис.3.7. Случайный процесс

Для установления связи между возможными значениями случайного процесса x(t) в последующий момент времени со значе-

ниями в предыдущие моменты, вводится понятие двухмерные плотности вероятности $W_2(x_1, t_1; x_2, t_2)$.

Это означает вероятность того, что в момент времени t_1 величина x находится в интервале $(x_1, x_1 + d\,x_1)$, а в момент t_2 – в интервале $(x_2, x_2 + d\,x_2)$ будет $W_2(x_1, t_1; x_2, t_2)d\,x_1 d\,x_2$. Это вероятность того что кривая x(t) пройдет возле двух точек (x_1,t_1) и (x_2,t_2).

Вводится n – мерная плотность вероятности $W_n(x_1, t_1; x_2, t_2; ...; x_n, t_n)$..

Тогда $W_n(x_1, t_1; x_2, t_2; ...; x_n, t_n)d\,x_1, ..., d\,x_n$ – вероятность того, что x(t) пройдет вблизи n точек.

Таким образом, случайный процесс определяется функциями $W_1, W_2, ... W_n$ и связью между ними.

Двумерная плотность вероятности $P_2 = P_2(x_1, x_2, t_1, t_2)$ позволяет определить вероятность совместной реализаций двух значений случайных величин x_1 и x_2 в производные моменты времени t_1 и t_2 (т. е. оценить динамику процесса). Двумерная плотность вероятности случайного процесса x(t):

$$W_1(x_1; t_1) = \int_{-\infty}^{\infty} W_2(x_1, x_2; t_1, t_2)d\,x_2 .$$

Использование плотности вероятности даже низших порядков в практических приложениях часто приводит к неоправданным усложнениям. Достаточно знание простейших характеристик случайных процессов, аналогично числовым характеристикам случайных величин (математическое ожидание и дисперсия), а также корреляционной функции [2,12].

1. Математическое ожидание случайного процесса x (t) – неслучайная функция времени $m_x(t_i)$, которая при любом аргументе t_i, равна среднему значению случайной величины $x(t_i)$ по всему множеству возможных реализаций

$$m_x(t_1) = M_{-\infty}\{x(t_1)\} = \int_{-\infty}^{\infty} x_1 w_1(x_1 t_1)d\,x_1 ..$$

2. Дисперсия случайного процесса x(t) характеризует степень разброса случайных значений процесса $x(t_1)$ от своего среднего значения $m_x(t_i)$ для момента времени t_1:

$$D_x(t_1) = M\{[x(t_1) - m_x(t_1)]^2\} = \int_{-\infty}^{\infty} [x(t_1) - m_x(t_1)]^2 W_1(x_1; t_1) dx_1,$$

$$D_x(t_1) = \sigma_x^2(t_1),$$

где $\sigma_x(t_1)$ – среднее квадратичное отклонение.

Для оценки статистической зависимости мгновенных значений случайного процесса x(t) в производные моменты t_1 и t_2 используется неслучайная функция аргументов $R_x(t_1, t_2)$ – автокорреляционная функция и она равна корреляционному моменту значений процесса $x(t_1)$ и $x(t_2)$

$$R_x(t_1, t_2) = M[\overset{0}{x}(t_1)\overset{0}{x}(t_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overset{0}{x}(t_2)\overset{0}{x}(t_2) w_2(x_1, x_2; t_1, t_2) dx_1 dx_2, \quad (3.16)$$

где $\overset{0}{x}(t) = x(t) - m_x$ – центрированный случайный процесс.

Взаимные моменты случайных процессов второго порядка дают возможность оценить совместные свойства двух случайных процессов X(t) и Y(t) путем анализа произвольной пары выборочных функций $x_k(t)$ и $y_k(t)$.

Мера связи между двумя случайными процессами X(t) и Y(t) также устанавливается корреляционными функциями, а именно – функциями взаимной корреляции и взаимной ковариации. В общем случае, для произвольных фиксированных моментов времени $t_1 = t$ и $t_2 = t + \tau$:

$$R_{XY}(t, t+\tau) = M\{X(\tau)\, Y(t+\tau)\}. \quad (3.17)$$

$$K_{XY}(t, t+\tau) = M\{(X(t) - m_x(t))(Y(t+\tau) - m_y(t+\tau))\}. \quad (3.18)$$

Взаимные функции являются произвольными функциями (не обладают свойствами четности или нечетности) и удовлетворяют следующим соотношениям:

$$R_{xy}(-\tau) = R_{yx}(\tau), \quad (3.19)$$

$$|R_{xy}(\tau)|^2 \le R_x(0)\, R_y(0).$$

Нормированная взаимная ковариационная функция (коэффициент корреляции двух процессов), которая характеризует степень линейной зависимости между случайными процессами при данном сдвиге τ одного процесса по отношению ко второму, определяется выражением

$$\rho_{xy}(\tau) = \frac{R_{yx}(\tau)}{\sigma_x(\tau)\sigma_y(\tau)}. \quad (3.20)$$

Статистическая независимость случайных процессов определяет отсутствие связи между значениями двух случайных величин X и Y. Это означает, что плотность вероятности одной случайной величины не зависит от того, какие значения принимает вторая случайная величина. Двумерная плотность вероятностей при этом должна представлять собой произведения одномерных плотностей вероятностей этих двух величин:

$$p(x, y) = p(x)\, p(y).$$

Это условие является обязательным условием статистической независимости случайных величин. В противном случае между случайными величинами может существовать определенная статистическая связь, как линейная, так и нелинейная. Мерой линейной статистической связи является коэффициент корреляции

$$r_{xy} = \frac{[\, M\{X \cdot Y\} - M\{X\} \cdot M\{Y\}\,]}{\sqrt{D(X)\, D(Y)}}.$$

Значения r_{xy} могут изменяться в пределах от –1 до +1. В частном случае, если случайные величины связаны линейным соотношением $x = ay + b$, коэффициент корреляции равен ± 1 в зависимости от знака константы a. Случайные величины некоррелированы при $r_{xy} = 0$, при этом из выражения для r_{xy} следует

$$M\{X \cdot Y\} = M\{X\} \cdot M\{Y\}.$$

Из статистической независимости величин следует их некоррелированность. Обратное не очевидно. Так, например, случайные величины $x = \cos\varphi$ и $y = \sin\varphi$, где φ – случайная величина с равномерным распределением в интервале $0 \ldots 2\pi$, имеют нулевой коэффициент корреляции, и вместе с тем их зависимость очевидна.

3.5.2. Стационарные случайные процессы

Случайные процессы различают по степени однородности их протекания во времени (по аргументу). Кроме моментов первого и второго порядка случайные процессы имеют моменты и более высоких порядков. По мере повышения порядка моментов вероятностная структура случайных процессов и их выборочных реализаций описывается все более детально. Однако практическая оценка этих

моментов по выборкам ограничена в основном только стационарными случайными процессами.

В информационных системах часто встречаются случайные процессы, протекающие во времени приблизительно однородно. Они имеют вид непрерывных случайных колебаний около некоторых средних значений, причем ни средние значения, ни характер этих колебаний не претерпевают существенного изменения во времени. Такие процессы – стационарные (пример, шум электрических устройств).

В любой динамической системе случайный процесс начинается с так называемого «переходного» процесса, затем переходит в установившийся режим, который с некоторым приближением можно считать стационарным. Строго говоря, стационарные случайные процессы бесконечны во времени, т. е. не имеют ни начала, ни конца. Однако на практике такие процессы не существуют, но многие случайные процессы на определенных отрезках времени с достаточной приближенностью можно считать стационарным.

Под стационарностью в узком смысле понимаются случайные процессы, для которых функции распределения плотности вероятности $w_n(x_1,t_1;x_2,t_2;...;x_n,t_n)$ произвольного порядка n не меняется при любом сдвиге всей группы точек $t_1,t_2,...t_n$ вдоль оси t, т. е. для любых n и τ [12]

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

Из определения стационарности в узком смысле следует следующие свойства:

1. Одномерная функция распределения плотности вероятности не зависит от времени:

$$w_1(x_1,t_1) = w_1(x_1t_1+\tau) = w_1(x).$$

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

$$w_2(x_1t_1;x_2t_2) = w_2(x_1x_2;t_2-t_1) = w_2(x_1,x_2\tau).$$

3. Трехмерная функция распределения плотности вероятности зависит только от двух разностей времени

$$t_2 - t_1 = \tau_1 \text{ и } t_3 - t_1 = \tau_2$$

$$w_3(x_1, t_1, ..., x_3, t_3) = w_3(x_1, x_2, x_3, \tau_1, \tau_2).$$

Поскольку математическое ожидание и дисперсия выражаются через одномерную функцию распределения плотности вероятности, то на основе свойства 1 можно утверждать, что для стационарного случайного процесса математическое ожидание и дисперсия не зависят от времени.

На основании свойства 2, корреляционная функция случайного процесса зависит только от разности времен τ, т. е.

$$\begin{aligned} m_x(t_1) &= m_x = \text{const}, \\ D_x(t_1) &= D_x = \text{const}, \\ R_x(t, t_1 + \tau) &= R_x(\tau). \end{aligned} \qquad (3.21)$$

На практике часто встречаются случайные процессы, для которых при выполнении условия (3.21) моменты высшего порядка зависят от времени. Поэтому понятия стационарности целесообразно расширить, приняв за основу определения стационарности условия (3.21).

Стационарность в широком смысле (по Хинчину). Согласно стационарности в широком смысле стационарными процессами являются случайные процессы, у которых математическое ожидание и дисперсия не зависят от времени, а корреляционная функция зависит только от разности времени $\tau = t_2 - t_1$.

Случайные процессы стационарные в узком смысле и будут стационарными в широком смысле, но не наоборот.

Стационарные процессы по своей природе проще, чем нестационарные и описываются более простыми характеристиками.

Среди стационарных процессов многие удовлетворяют условию эргодичности. Оно проявляется в том, что каждая реализация случайного процесса достаточной продолжительности несет практически полную информацию о свойствах всего ансамбля реализаций.

Характеристики случайных процессов оцениваются путем усреднения по ансамблю реализаций в определенные моменты времени (по сечениям процессов). Но большинство стационарных случайных процессов обладают эргодическим свойством. Сущность его заключается в том, что по одной достаточно длинной реализации процесса можно судить обо всех его статистических

свойствах так же, как по любому количеству реализаций. Другими словами, закон распределения случайных величин в таком процессе может быть одним и тем же как по сечению для ансамбля реализаций, так и по координате развития. Такие процессы получили название эргодические (ergodic). Для эргодических процессов имеет место

$$m_x = \lim_{T\to\infty}\frac{1}{T}\int_0^t x(t)\,dt = x_0, \quad D_x = \lim_{T\to\infty}\frac{1}{T}\int_0^T [x(t)-x_0]^2\,dt,$$

$$R_x(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T [x(t)-x_0][x(t+\tau)-x_0]\,dt.$$

Эргодичность является очень важным свойством случайных стационарных, и только стационарных процессов. Математическое ожидание эргодического случайного процесса равно постоянной составляющей любой его реализации, а дисперсия является мощностью его флюктуационной составляющей, т. к. определение функций производится по ограниченным статистическим данным одной реализации и является только определенным приближением к соответствующим фактическим функциям процессов, целесообразно называть эти функции статистическими.

Свойства эргодичности могут проявляться только по отношению к двум первым моментам случайного процесса, что вполне достаточно для использования соответствующих методик исследования процессов. Практическая проверка эргодичности процесса обычно производится проверкой выполнения условия:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T R(\tau)\,d\tau = 0.$$

Если корреляционная функция процесса стремится к нулю при возрастании значения аргумента τ, то процесс относится к числу эргодических, по крайней мере, относительно моментов первого и второго порядков.

Эти соотношения лежат в основе корреляционной теории сигналов.

Теперь рассмотрим свойства корреляционной функции.

Свойства корреляционной функции случайного процесса**:**

I. По мере увеличения τ зависимость между $x(t)$ и $x(t+\tau)$ ослабевает и при $t \to \infty$ становится незначительной, т. е.

$$R_x = \lim_{\tau\to\infty} R_x(\tau) = \lim_{\tau\to\infty} m_1[\,x(t)\,x(t+\tau)\,] = m_x^2(x) = R_x(\infty).$$

При $\tau \to \infty$ значения функции автокорреляции (ФАК) для сигналов, конечных по энергии, стремятся к нулю, что прямо следует из физического смысла ФАК. Это позволяет ограничивать длину ФАК, определенным максимальным значением τ_{max} – временем корреляции, за пределами которого отсчеты можно считать независимыми. Интегральной характеристикой времени корреляции случайных величин обычно считают *эффективный интервал корреляции*, определяемый по формуле:

$$T_k = \int_{-\infty}^{\infty} |\rho(\tau)|\, d\tau = \frac{1}{R_x(0)} \int_{-\infty}^{\infty} |R(\tau)|\, d\tau. \qquad (3.22)$$

Отсчеты (сечения) случайных функций, отстоящие друг от друга на расстояние большее T_k, при инженерных расчетах считают некоррелированными.

2. При уменьшении τ зависимость между $x(t)$ и $x(t+\tau)$ усиливается и в пределе при $\tau \to 0$

$$R_x(0) = \lim_{\tau\to 0} R_x(\tau) = \lim_{\tau\to 0} m_1[x(t), x(t+\tau)] = m_1[\,x^2(t)\,] = m_2[\,x(t)\,].$$

Таким образом, при $\tau = 0$ корреляционная функция стремится к начальному моменту второго порядка функции x(t). Физический смысл: корреляционная функция при $\tau = 0$ характеризует полную среднюю мощность стационарного случайного сигнала.

Следовательно, дисперсия случайного стационарного сигнала равна

$$D(x) = R_x(0) - R_x(\infty).$$

3. Автокорреляционная функция является четной функцией

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

Последнее также очевидно

$$X(t)\,X(t+\tau) = X(t-\tau)\,X(t) \quad \text{при } t = t - \tau.$$

Говоря иначе, моменты двух случайных величин $X(t_1)$ и $X(t_2)$ не зависят от последовательности, в которой эти величины рассматриваются, и соответственно симметричны относительно своих аргументов

$$R_x(t_1, t_2) = R_x(t_2, t_1).$$

4. Максимум корреляционной функций наблюдается при $\tau = 0$. Это очевидно, т. к. при $\tau = 0$ вычисляется степень связи отсчетов с собой же, которая не может быть меньше связи разных отсчетов. Значение максимума функции корреляции равно средней мощности сигнала.

Доказательство этого утверждения. Представим момент первого порядка

$$m_1\{[\,x(t)+x(t+\tau)\,]^2\} = m_1\{x^2\} + 2m_1[\,x(t)x(t+\tau)\,] + \\ + m_1[\,x^2(t+\tau)\,] = 2R_{xx}(0) + 2R_{xx}(\tau),$$

т. к., среднее значение квадрата суммы не может быть отрицательным, то

$$2R_x(0) + 2R_x(\tau) \ge 0, \quad |R_x(\tau| \le |R_x(0)|.$$

5. Если к случайной функции $X(t)$ прибавить неслучайную функцию $f(t)$, то ковариационная функция не изменяется. Обозначим новую случайную функцию как $Y(t) = X(t) + f(t)$. Функция математического ожидания новой величины: $\overline{y}(t) = \overline{x}(t) + f(t)$. Отсюда следует, что $Y(t) - \overline{y}(t) = X(t) - \overline{x}(t)$, и соответственно

$$R_y(t_1, t_2) = R_x(t_1, t_2).$$

6. Если случайную функцию $X(t)$ умножить на неслучайную функцию $f(t)$, то ее корреляционная функция $R_x(t_1, t_2)$ умножается на $f(t_1) \cdot f(t_2)$. Обоснование данного свойства проводится по методике, аналогичной предыдущему пункту.

7. При умножении функции случайного процесса на постоянное значение С значения ФАК увеличиваются в C^2 раз.

Пример 3.3. Приведем примеры типовых корреляционных функций (КРФ) стационарного случайного сигнала (рис. 3.8).

1. Монотонная корреляционная функция $R(\tau)$ (рис. 3.8, *а*).
2. Немонотонная корреляционная функция (рис. 3.8, *б*).

На практике вместо стационарного случайного сигнала x(t) рассматривается его отклонение от математического ожидания, т. е. центрированный стационарный случайный сигнал (пульсации или флюктуаций процесса)

$$x_0(t) = x(t) - m_x(x).$$

Корреляционная функция пульсации стационарного случайного сигнала определяется следующим образом:

$$R_x^0(\tau) = m_1[x_0(t)\, x_0(t+\tau)] = m_1\{[x(t) - a(x)] \times$$
$$\times [x(t+\tau) - a(x)]\} = R_x(\tau) - a^2(x). \qquad (3.23)$$

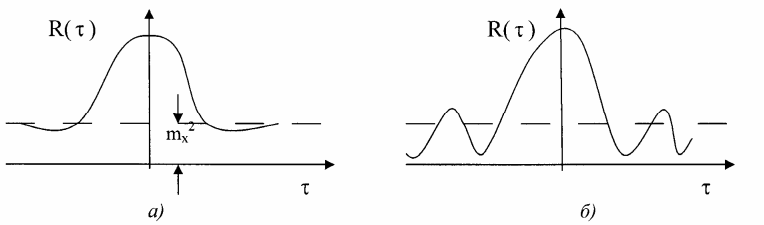


Рис. 3.8. Типовые корреляционные функции

Математическое ожидание центрированного стационарного случайного сигнала равно нулю, а дисперсия

$$D_0(x) = R_x^0(0).$$

Отношение корреляционной функции центрированного стационарного случайного сигнала к его дисперсии называется нормированной корреляционной функцией КРФ (коэффициент корреляции) пульсаций стационарного процесса

$$r_x^0(\tau) = \frac{R^0{}_x(\tau)}{D_0(x)}.$$

Пример 3.4. На рис. 3.9 приведены типовые нормированные корреляционные функции случайных сигналов.

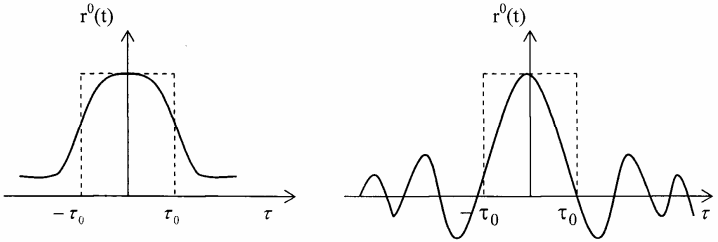


Рис. 3.9. Типовые нормированные корреляционные функции

Для чисто стационарного случайного процесса всегда можно указать значение времени корреляции $\tau = \tau_0$, что при значении $\tau > \tau_0$ величины $x(t)$ и $x(t+\tau)$ можно считать независимыми, причем практически независимость понимается в сигнале, что при $\tau > \tau_0, |r_x^0(\tau)| < \sigma = 0.05$. Интервал τ_0 называют временем корреляции стационарного процесса. Время корреляции определяется как половина ширины основания прямоугольного импульса единичной высоты, площадь которого равна площади под графиком нормированной корреляционной функции

$$\tau_0 = \frac{1}{2}\int_{-\infty}^{\infty} r_x^0(\tau)\,d\tau = \frac{1}{R_{xx}^0(0)}\int_0^{\infty} R^0{}_x(\tau)\,d\tau.$$

Стационарные случайные процессы наиболее часто встречаются при решении физических и технических задач. Теория стационарных случайных функций разработана наиболее полно. Случайные процессы, удовлетворяющие условиям стационарности на ограниченных, интересующих нас интервалах, также обычно рассматривают в классе стационарных и называют квазистационарными.

В общем случае значения математического ожидания, дисперсии и корреляции могут быть зависимыми от момента времени t, т. е. изменяться во времени. Такие процессы составляют класс нестационарных процессов.

Пример 3.5. Случайная функция задана выражением $Z(t) = X(t) + Y$, Y – случайная величина, некоррелированная с X(t). X(t) – стационарная эргодичная функция. Эргодична ли функция Z (t)?

Функция Z(t) стационарна, но не эргодична, т. к. при $\tau \to \infty$ имеет место

$$R_z(\tau) \to D_y.$$

3.5.3. Спектральная плотность стационарного случайного процесса

Для детерминированных сигналов удобным оказался гармонический анализ. В связи с этим желательно использовать аппарат преобразований Фурье к случайному процессу. Однако непосредственное приложение гармонического анализа к случайному процессу невозможно, т. к.:

1. Реализации случайного процесса $x_k(t)$ не удовлетворяют условию абсолютной интегрируемости

$$\int_{\infty}^{\infty} |x_k(t)|\, dt < \infty .$$

2. Для случайного процесса $x(t)$ частотный спектр также случайная функция.

Можно обобщить гармонический анализ, усредняя спектральные разложения, полученные для отдельных реализаций [2].

Для этого вводим функцию (рис. 3.10)

$$x_k'(t) = \begin{cases} x_k(t), \text{при } t \le |T/2|, \\ 0, \quad \text{при } t > |T/2|. \end{cases} \qquad (3.24)$$

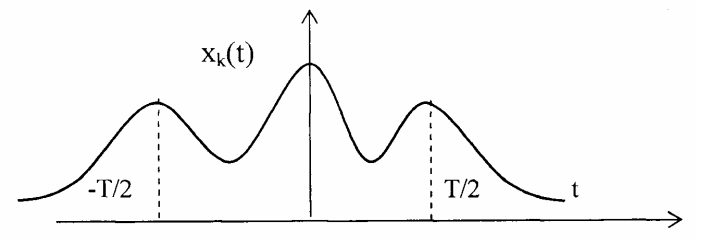


Рис. 3.10. Функция определяемая (3.24)

Для такой функции применим справедливо преобразование Фурье

$$S_k(j\omega) = \int_{-T/2}^{T/2} x_k(t)\, e^{-j\omega t}\, dt$$

и

$$P_k = \frac{1}{T}\int_{-T/2}^{T/2} [x_k(t)]^2\, dt \; - \text{средняя мощность сигнала.}$$

С другой стороны, средняя мощность сигнала определяется через частотный спектр

$$S_k(\omega) = \frac{|S_k(j\omega)|^2}{\pi T} \text{ и } P_T = \int_0^{\infty} \frac{|S_k(j\omega)|^2}{\pi T} d\omega .$$

Спектральная функция $S_k(\omega)$ характеризует распределение

мощности реализации по спектру частот (спектральная плотность мощности).

Спектральная плотность совокупности реализации характеризует усреднение по ансамблю возможных значений $S_k(\omega)$

$$m_1[S_k(\omega)] = \frac{1}{\pi T} m_1\{|S_k(j\omega)|^2] = \frac{1}{\pi T} m_1\{S_k(j\omega)S_k(-j\omega)\} =$$

$$= \frac{1}{\pi T} m_1\{\int_{-T/2}^{T/2} x_n(t_1)\, e^{-j\omega t_1} d\, t_1 \int_{-T/2}^{T/2} x_k(t_2)\, e^{j\omega t_2} d\, t_2\} =$$

$$= \frac{1}{\pi T} \int_{-T/2}^{T/2} m_1\{x_n(t_1)\, x_n(t_2)\} \exp(-j\omega(t_1 - t_2)) d\, t_1 d\, t_2.$$

Так как для стационарного случайного сигнала, корреляционная функция является функцией одного аргумента $R_x(t_1 - t_2) = R_x(\tau)$, то

$$m_1[G_k(\omega)] = \frac{1}{\pi T} \int_{-T/2}^{T/2} R_x(\tau) e^{-j\omega\tau} d\tau \int_{-T/2}^{T/2} d\, t_2 .$$

Переходя к пределу при $T \to \infty$

$$G(\omega) = \lim_{T\to\infty} G_k(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} R(\tau)\, e^{-j\omega\tau} d\tau = \frac{2}{\pi} \int_0^{\infty} R(\tau) \cos\omega\tau\, d\tau . \qquad (3.25)$$

Спектральная плотность $G(\omega)$ является усредненной характеристикой совокупности реализации случайного процесса и определяется прямым преобразованием Фурье от корреляционной функции (3.25). С другой стороны по обратному преобразованию Фурье можно определить корреляционную функцию

$$R(\tau) = \int_0^{\infty} G(\omega) \cos\tau d\omega . \qquad (3.26)$$

Преобразования (3.25) и (3.26) называют преобразованиями Хинчина – Винера. Введя вместо круговой частоты ω обычную частоту f, получим

$$R(\tau) = \int_0^{\infty} G(f) \cos 2\pi f\, \tau\, d\, f ;$$

$$G(f) = 4\int_0^{\infty} R(\tau) \cos 2\pi f\, \tau\, d\, \tau .$$

Согласно эргодической гипотезе для стационарного случайного процесса усреднение по множеству может быть замечено усреднением по времени, то

$$G(\omega) = \lim_{T\to\infty} \frac{|S_k(j\omega)|^2}{\pi T}. \qquad (3.27)$$

Таким образом, спектр стационарного случайного сигнала можно определить двумя способами [3,12]:

а) непосредственным наблюдением одной реализаций x(t) и нахождением предела (3.27);

б) нахождением преобразования Фурье от корреляционной функции (3.25).

Функция $G(\omega)$ играет большую роль в теории информации. Для выяснения физического смысла $G(\omega)$ в корреляционной функции примем $\tau = 0$ и получим

$$R(0) = \int_0^\infty G(\omega)d\omega \cdot$$

Так как $R(0)$ – мощность сигнала, то $G(\omega)$ характеризует усредненную энергетическую картину распределения мощности сигнала по частотному спектру.

Из преобразований Фурье следует, что при растяжении $R(\tau)$ в α раз ее частотный спектр $G(\omega)$ во столько же раз сужается (рис. 3.11).

В связи с этим свойством рассматриваем граничный и наиболее интересный случай, когда $G(\omega) = a_0 = \text{const}$, т.е. когда спектральная плотность мощности равномерна во всех частотах, так называемый белый шум (рис. 3.12).

Корреляционная функция такого сигнала определятся как

$$R(\tau) = \int_0^\infty S_0 \cos\omega\tau\, d\omega = \frac{1}{2}\int_{-\infty}^{\infty} S_0(\omega) e^{j\omega\tau} d\omega = \pi a_0 \delta(\tau), \qquad (3.28)$$

поскольку интеграл $\frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(j\omega\tau)\, d\omega$ равно обратному преобразованию Фурье от функции $\delta(\tau)$.

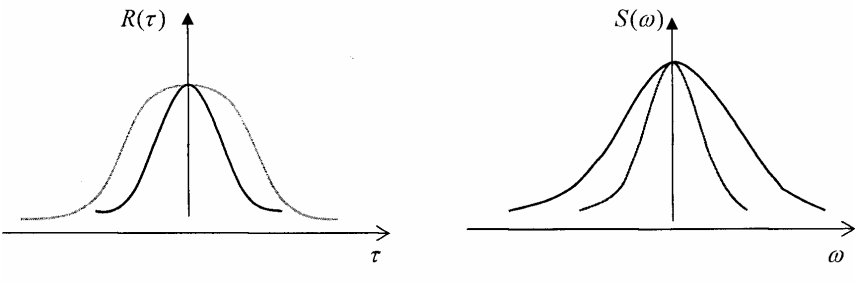


Рис. 3.11. Корреляционная функция и спектр

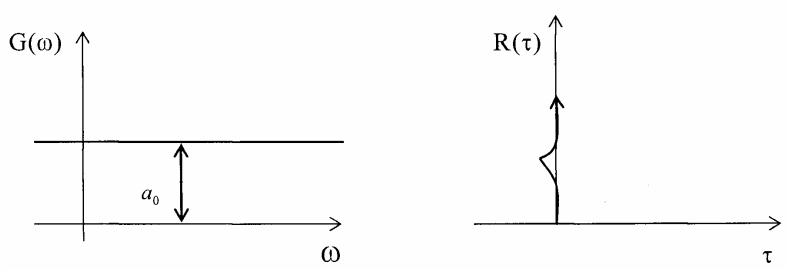


Рис. 3.12. Кореляционная функция и спектр белого шума

Таким образом, корреляционная функция белого шума выражается дельта-функцией

$$R(\tau) = \begin{cases} \infty, \tau = 0 \\ 0, \tau \neq 0 \end{cases}.$$

Характерная особенность такого сигнала в том, что значения $x(t)$ в любые сколь угодно близкие моменты времени некоррелированы [16]. Такой процесс – абсолютно чистый процесс.

Понятие «белый шум» основано на спектральном свойстве случайного процесса и совсем не связано с законами распределения плотности вероятности. Белый шум – идеализация, никогда не встречается в природе, потому что достаточно близкие значения случайного процесса практически коррелированы, а также реальные процессы имеют конечную мощность. Подобная идеализация очень важна и

значительно упрощает математический анализ и не вносит сколько-нибудь существенной погрешности.

Спектры реальных процессов практически ограничены полосой $\Omega_Э = \omega_b - \omega_h$, вследствие ограниченности полосы пропускания реальных систем.

На рис. 3.13 показан спектр и корреляционная функция реального сигнала. Для приведенного на рис. 3.13 спектра реального сигнала определим корреляционную функцию по формуле

$$R(\tau) = \int_0^{\omega_в} G_0 \cos\omega\tau \, d\omega = G_0\omega_0 \frac{\sin\omega_в\tau}{\omega_в\tau} = P_0 \frac{\sin\omega_в\tau}{\omega_в\tau},$$

где $P_0 = G_0\omega_в$ – средняя мощность случайного сигнала.

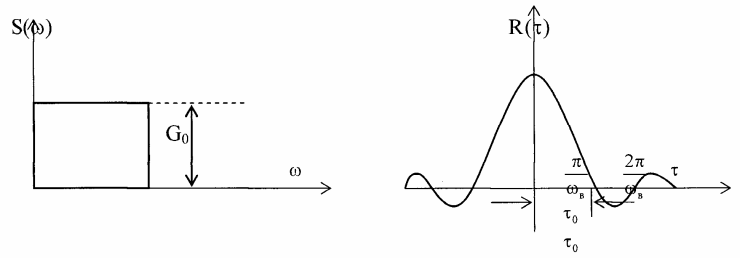


Рис. 3.13. Спектры реального сигнала

Для такого процесса время корреляции τ_0 имеет конечную величину, например, как интервал между точкой $\tau = 0$ и точкой, где корреляционная функция $R(\tau)$ первый раз обращается в нуль, т. е. $\tau_0 = \frac{\pi}{\omega_в}$ (рис. 3.13). Таким образом, ограничение спектра влечет за собой появление корреляции, причем по мере сокращения полосы частот $\Omega_э$ интервал корреляции τ_0 сигнала возрастает.

3.5.4. Эффективная ширина спектра случайного процесса

На практике часто распространены сигналы с неравномерным спектром (рис. 3.14).

При анализе случайного процесса с неравномерным спектром

используется понятие эквивалентной или эффективной ширины спектра [9]

$$\Omega_э = \frac{\int_0^\infty S(\omega)d\omega}{S_{max}}. \qquad (3.29)$$

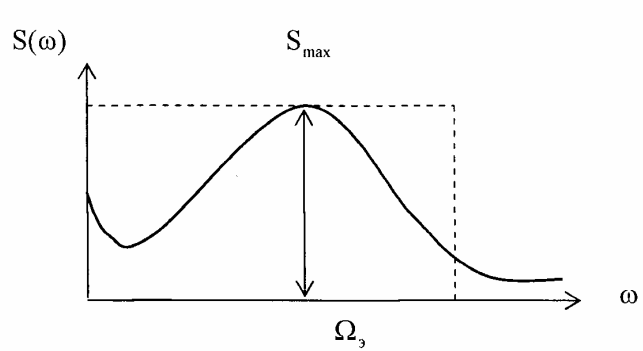


Рис. 3.14. Сигнал с неравномерным спектром

При этом средняя мощность сигнала равна

$$\int_0^\infty S(\omega)d\omega = R(0) = S(0)\Omega_э. \qquad (3.30)$$

И так интервал корреляции сигнала с учетом $S(\omega) = \frac{2}{\pi}\int_0^\infty R(\tau)\cos\omega\tau\, d\tau$ определяется выражением

$$\tau_0 = \frac{\int_0^\tau |R(\tau)|\, d\tau}{R(0)} = \frac{\pi S(0)}{2R(0)} = \frac{\frac{2}{\pi}\int_0^\infty R(\tau)\cos\omega\tau\, d\tau}{R(0)}. \qquad (3.31)$$

Отсюда интервал корреляции τ_0 равен

$$\tau_0 = \frac{\int_0^\tau |R(\tau)|d\tau}{R(0)} = \frac{\int_0^\tau |R(\tau)|d\tau}{S(0)\Omega_э} = \frac{\pi\int_0^\tau |R(\tau)|d\tau}{2\cdot\int_0^\tau R(\tau)d\tau\Omega_э} = \frac{\pi}{2\Omega_э} = \frac{1}{4F_э}. \qquad (3.32)$$

Контрольные вопросы:

1. Что понимают под детерминированным сигналом?
2. Перечислите разновидности математических представлений детерминированных сигналов.
3. Запишите условия ортонормированности и ортогональности системы функций.
4. Почему необходимо изучение моделей детерминированных сигналов?
5. Опишите временную форму представления сигнала.
6. В чем сущность спектрального представления сигнала?
7. Назовите преимущества частотного представления сигнала.
8. Дайте определения спектра амплитуды и спектра фаз.
9. В чем различие спектров периодического и непериодического сигналов?
10. Дайте определение практической ширины спектра.
11. Какая зависимость между длительностью сигнала и шириной его спектра?
12. Дайте определение случайного процесса.
13. Приведите и поясните смысл n-мерной функции и плотности вероятности случайного сигнала.
14. Поясните, как можно задать случайный сигнал.
15. Перечислите числовые характеристики случайного сигнала.
16. Поясните понятие и смысл корреляционной функции случайного сигнала.
17. Дайте определение стационарного случайного сигнала.
18. Какая взаимосвязь между корреляционной функцией и спектральной плотностью
    стационарного случайного сигнала?
19. Как определяется эффективная ширина спектра случайного процесса?

4. КВАНТОВАНИЕ ИНФОРМАЦИИ

В ряде случаев значительное преимущество при передаче, хранении и обработке информации дает переход от аналогового представления сигналов к цифровому (квантование по времени и уровню).

4.1. Общая постановка задачи дискретизации информации

Дискретизация во времени заключается в замене непрерывного сообщения последовательностью его мгновенных значений (отсчетов), взятых в дискретные моменты времени (рис. 4.1). При такой замене исключаются все множества значений непрерывных функций, находящиеся внутри интервала дискретизации Δt. Полученная при этом функция имеет вид последовательности отчетов, взятых в дискретные моменты времени [9,11].

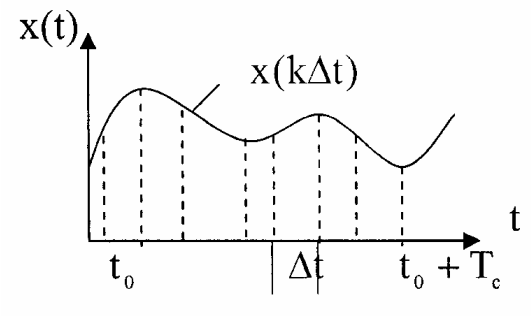


Рис. 4.1. Дискретизация сообщения во времени

Дискретизация по времени может быть равномерная (принудительная), когда интервал дискретизации Δt неизменен и неравномерной, когда Δt меняется в соответствии с изменением тех или иных характеристик сообщения, например в соответствии скоростного изменения сообщения во времени (адаптивная).

Теоретическое расстояние и практическая реализация адаптивной дискретизации представляет значительные трудности, т. к. для такой дискретизации необходимо располагать большим объемом ап-

риорных сведений о дискретизируемых сообщениях, что существенно усложняет устройство дискретизации.

Наиболее распространена равномерная дискретизация. Теория и техника такой дискретизации достаточно разработана.

В основе математического описания дискретизации непрерывных функций во времени может так называемая импульсная функция дискретизации $a_д(t)$, представляя собой периодическую последовательность δ-функций, следующих через интервалы времени Δt

$$a_д(t) = \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t). \qquad (4.1)$$

Используя преобразование Фурье, получим спектр импульсной функции дискретизации

$$A_д(f) = \frac{1}{\Delta t} \sum_{k=-\infty}^{\infty} \delta(f - \frac{k}{\Delta t}). \qquad (4.2)$$

Из выражения (4.2) видно, спектр импульсной функции дискретизации представляет собой периодическую последовательность δ-функции, следующих через частотные интервалы $\Delta f = \frac{1}{\Delta t}$.

В соответствии со свойством δ- функции $\int_{-\infty}^{\infty} \delta(\xi - \xi_0) d\xi = 1$, площадь каждой составляющей в выражении (4.1) равна единице, а (4.2) стремится к $\frac{1}{\Delta t}$. Эти значения площадей определяют “веса” δ- функции.

Дискретизация непрерывных функций времени $x(t)$ с математической точки зрения представляет собой умножение этой функции на функцию $a_д(t)$

$$x_д = x(t)a_д(t) = \sum_{k=-\infty}^{\infty} x(t)\delta(t - k\Delta t). \qquad (4.3)$$

В соответствии с фильтрующим свойством δ-функции

$\int_{-\infty}^{\infty} x(\xi)\delta(\xi-\xi_0)d\xi = x(\xi_0)$, получим

$$x(t)\delta(t-k\Delta t) = x(k\Delta t)\delta(t-k\Delta t). \qquad (4.4)$$

Выражение (4.4) означает, что умножение $x(t)$ на единичный δ - импульс приводит к тому, что значение этого импульса становится равной вместо единицы значению $x(t)$ в момент времени $t = k\Delta t$. Эту площадь обычно называют весовым коэффициентом δ - импульса. Иными словами, умножение $x(t)$ на единичный δ - импульс соответствует получению отчета функции $x(t)$ в момент $t = k\Delta t$.

С учетом выражения (4.4) соотношение (4.3) примет вид

$$x_д = \sum_{k=-\infty}^{\infty} x(k\Delta t)\delta(t-k\Delta t). \qquad (4.5)$$

Таким образом, умножение $x(t)$ на импульсную функцию дискретизации $a_д(t)$ приводит к преобразованию периодической последовательности δ-импульсов, веса которых равны мгновенным значениям сообщений в моменты $t = k\Delta t$.

На практике реализовать импульсную функцию дискретизации $a_д(t)$ в виде δ - импульсов невозможно. Технически дискретизация непрерывных сообщений реализуется ключевыми устройствами, управляемыми периодической последовательностью коротких импульсов. При этом длительность отсчетов конечно, поскольку отсчет берется не в одной точке, а в некотором интервале времени, равном длительности импульса δ. Если выбрать длительность τ так, чтобы выполнялось $\frac{\tau}{\Delta t} << 1$, образуется последовательность коротких импульсов, амплитуды которых пропорциональны мгновенным значениям сообщения.

Практическая реализация дискретизации по времени часто называют импульсным преобразованием непрерывного сообщения (амплитудно-импульсная модуляция АИМ) [7].

При решении задач дискретизации по времени возникает ряд вопросов:

1. Из каких соображений надо исходить при выборе интервала Δt.

2. Какова точность замены непрерывного сообщения последовательностью его отсчетов, взятых в дискретные моменты, и от чего она зависит.

3. Каков максимально допустимый интервал Δt, при котором еще принципиально возможно восстановление непрерывных сообщений.

Решить эти задачи можно только в случае, если проблему дискретизации рассматривать в неразрывной связи с обратной проблемой – восстановлением непрерывных функций времени по его мгновенным значениям, известным только в дискретные моменты времени. Сложность этой проблемы заключается в том, что нужно восстановить утраченные при дискретизации сведения о поведении непрерывных функций в промежутке между отчетами.

Очевидно, чем меньшим количеством отсчетов в сообщении заменяется сообщение длительностью T, тем продолжительнее Δt, и сложнее выполнить восстановление исходных функций. И наоборот, чем больше отсчетов, тем короче Δt и тем проще восстановление.

Таким образом, рассматриваемая проблема дискретизации не имеет однозначные решения. Осуществление более экономного импульсного преобразования непрерывного сообщения приводит к усложнению задачи восстановления, а неэкономичное импульсное преобразование – к упрощению процедуры восстановления.

4.2. Предельная дискретизация непрерывных сообщений

В теории дискретизации непрерывных сообщений особую важность приобретает вопрос о максимальном (предельном) интервале $\Delta t = \max \Delta t$, при котором возможно восстановление непрерывной функции с заданной точности по ее отсчетам. При этом необходимое число отсчетов будет минимальной

$$N = \min N = [\frac{T}{\max \Delta t}].$$

Дискретизация, соответствующая вышеприведенному условию (берется целая часть соотношения) называется предельной (оптимальной). При этом обеспечивается представление непрерывного сообщения с заданной точностью минимальным количеством дискретных отсчетов.

Проблема дискретизации сигналов сложна, и несмотря на значительное количество исследований, далека до завершения.

В настоящее время наиболее разработанной и широко применяемой предельной дискретизацией непрерывных сообщений является дискретизация, основанная на теореме В.А. Котельникова [2,9].

Остановимся на исходных положениях этой теоремы. При построении теории дискретизации необходимо опираться на некоторую модель сообщения. Допустим, что известна одна из реализаций x(t) квазистационарного случайного процесса {x(t)}, соответствующего совокупности возможных непрерывных сообщений.

Как бы сложна не была эта реализация, она представляет собой некоторую неслучайную (детерминируемую) функцию времени. На основе преобразования Фурье комплексный спектр имеет вид [18]

$$G_x(j\omega) = \int_0^{T_c} x(t)\, e^{-j\omega t} d\, t\, . . \qquad (4.6)$$

Исходя из требуемой точности допустимого уровня помех, ограничим этот спектр некоторой частотой $\omega_c = 2\pi F_c$, т. е. введем модель сообщения с ограниченным амплитудным спектром. Особенности этой модели следующие:

– относится к одной реализации случайного процесса, т. е. соответствует детерминированной функции;

– имеет ограниченный спектр.

Для такой модели верна следующая теорема. Если непрерывная функция времени x (t) имеет спектр, ограниченный полосой частот от нуля до F_c, то эта функция полностью определяется последовательностью своих мгновенных значений, взятых в моменты

времени, отсчитываемые через интервалы $\Delta t = \frac{1}{2F_c}$ (теорема Котельникова).

Доказательство теоремы Котельникова. Покажем, что непрерывная функция x(t), заданная на интервале $-\infty < t < \infty$ и не содержащая составляющих с частотой выше $\omega_c = 2\pi F_c$, вполне определяется значениями $x(\frac{n}{2F_c})$ в точках отсчета $(n = 1,2\ldots)$, расположенных на расстоянии $\Delta t = \frac{1}{2F_c}$ друг от друга.

На рис. 4.2 приведен дискретизируемый сигнал и его спектр.

По условию теоремы спектр $S(j\omega) = 0$ при $|\omega| > 2\pi F_c$. На интервале $(-2\pi F_c\, 2\pi F_c)$ спектр $S(j\omega)$ можно разложить как периодическую функцию с периодом $4\pi F_c$ в ряд Фурье

$$S(j\omega) = \sum_{n=-\infty}^{\infty} C_n e^{j2\pi\frac{\omega}{2\omega_c}} \qquad (4.6)$$

или

$$S(j\omega) = \sum_{n=-\infty}^{\infty} C_n e^{j\omega\frac{n}{2F_c}}, \qquad (4.7)$$

где

$$C_n = \frac{1}{4\pi F_c}\int_{-2\pi F_c}^{2\pi F_c} S(j\omega)\, e^{-j\omega\frac{n}{2F_c}} d\omega. \qquad (4.8)$$

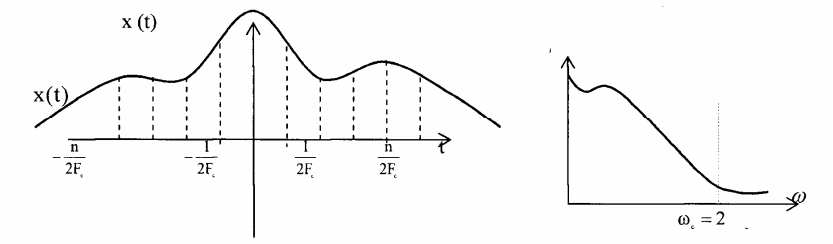


Рис. 4.2. Дискретизируемый сигнал и его спектр

С другой стороны по обратному преобразованию Фурье

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(j\omega)e^{j\omega t}d\omega, \qquad (4.9)$$

и с учетом соотношения $S(j\omega) = 0$ при частотах $|\omega| > 2\pi F_c$

$$x(t) = \frac{1}{2\pi}\int_{-2\pi F_c}^{2\pi F_c} S(j\omega)\, e^{j\omega t} d\omega. \qquad (4.10)$$

Определим x(t) для дискретных моментов времени $t = \frac{n}{2F_c}$,

где $n = -\infty, 0, .., \infty$.

Для этого в выражение (4.10) подставим $t = \frac{n}{2F_c}$

$$x(\frac{n}{2F_c}) = \frac{1}{2\pi}\int_{-2\pi F_c}^{2\pi F_c} S(j\omega)\, e^{j\omega\frac{n}{2F_c}} d\omega. \qquad (4.11)$$

Из сравнения выражений (4.11) и (4.8) следует, что коэффициенты C_n разложения $S(j\omega)$ в соотношении (4.7) пропорциональны отсчетам непрерывного сигнала x(t) в дискретные моменты времени $t = \frac{n}{2F_c}$, т. е.

$$C_n = \frac{1}{2F_c} x(-\frac{n}{2F_c}). \qquad (4.12)$$

Выражение (4.12) доказывает теорему В.А. Котельникова.

Действительно, если функция времени x(t) известна в точках отсчета $\frac{n}{2F_c}..., \; n = -\infty, 0, .., \infty$, то коэффициенты C_n определены.

Эти коэффициенты определяют спектр $S(j\omega)$, а спектр $S(j\omega)$ по обратному преобразованию Фурье – исходный непрерывный сигнал x(t). Это означает, что существует единственная функция, не содержащая частот выше $\omega_c = 2\pi F_c$ и проходящая через

заданные мгновенные значения в точках отсчета, отстающих друг от друга на интервал $\Delta t = \frac{1}{2F_c}$.

Связь между сигналом x(t) и его спектром $S(j\omega)$ определяем подставив (4.12) в (4.7):

$$S(j\omega) = \sum_{n=-\infty}^{\infty} \frac{1}{2F_c} x(-\frac{n}{2F_c}) e^{j\omega \frac{n}{2F_c}} = \frac{1}{2F_c} \sum_{n=-\infty}^{\infty} x(\frac{n}{2F_c}) e^{-j\omega \frac{n}{2F_c}}. \quad (4.13)$$

Так между сигналом x(t) и его спектром $S(j\omega)$ существует однозначная связь, то x(t) как и $S(j\omega)$ однозначно определяется дискретными значениями $\left\{ x\left(\frac{n}{2F} \right) \right\}$.

Таким образом, методика доказательства теоремы Котельникова В.А. сводится к следующим этапам:

1. Спектральная функция $S(j\omega)$ представляется рядом Фурье с коэффициентом C_n (4.8).

2. По обратному преобразованию Фурье определяется функция времени $x(\frac{n}{2F})$ в точках $t = \frac{n}{2F_c}$ (4.11).

3. Сравнивая (4.8) и (4.11) устанавливается связь коэффициентов ряда Фурье C_n с отчетами функции $x(-\frac{n}{2F})$, что и доказывает теорему.

4.3. Методы восстановления непрерывного сигнала

Для восстановления функции x(t) по ее отсчетам $x(\frac{n}{2F_c})$ необходимо подставить полученное выражение спектра (4.13) в выражение обратного преобразования Фурье (4.10)

$$x(t)=\frac{1}{2\pi}\int_{-2\pi F_c}^{2\pi F_c}S(j\omega)e^{j\omega t}d\omega=\frac{1}{4\pi F_c}\int_{-2\pi F_c}^{2\pi F_c}\sum_{n=-\infty}^{\infty}x(\frac{n}{2F_c})e^{-j\omega\frac{n}{2F_c}+j\omega t}d\omega. \quad (4.14)$$

Окончательно получим [17]

$$x(t)=\frac{1}{4\pi F_c}\sum_{n=-\infty}^{\infty}x(\frac{n}{2F_c})\int_{-2\pi F_c}^{2\pi F_c}e^{-j\omega(\frac{n}{2}-t)}d\omega=\frac{1}{4\pi F_c}\sum_{n=-\infty}^{\infty}x(\frac{n}{2F_c})\frac{-1}{j(\frac{n}{2F_c}-t)}\times$$

$$\times\left[e^{-j2\pi F_c(\frac{n}{2F_c}-t)}-e^{j2\pi F_c(\frac{n}{2F_c}-t)}\right]=\frac{1}{4\pi F_c}\sum_{n=-\infty}^{\infty}x(\frac{n}{2F_c})\frac{1}{2j(2\pi F_c t-\pi n)}\cdot\left[e^{j(2\pi F_c t-\pi n)}-e^{-j(2\pi t-\pi n)}\right].$$

Учитывая, что $\frac{e^{j(2\pi F_c t-\pi n)}-e^{-j(2\pi F_c t-\pi n)}}{2j}=\sin(2\pi F_c t-\pi n)$, сигнал $x(t)$ определим через его отсчеты

$$x(t)=\sum_{n=-\infty}^{\infty}x(\frac{n}{2F_c})\frac{\sin(2\pi F_c t-\pi n)}{2\pi F_c t-\pi n}=\sum_{n=-\infty}^{\infty}x(n\Delta t)\frac{\sin 2\pi F_c(t-n\Delta t)}{2\pi F_c(t-n\Delta t)}. \quad (4.15)$$

Ряд (4.15) называется рядом В.А. Котельникова.

Таким образом, получено разложение функции x(t) в ряд по координатным функциям вида $\frac{\sin z}{z}$, а коэффициентами разложения являются значения самой функции x(t), отстающее друг от друга на расстоянии $\frac{1}{2F_c}$.

Тем самым не только доказана сформулированная теорема, но и указан способ построения (восстановление функции) x(t), если имеется в наличие лишь ее отсчеты.

Разложение непрерывной функции $f(t)$ в ряд в виде (4.15) очень важно и приобрело в теории информации и других областях, не меньшее значение, чем разложение Фурье

$$f(t)=\sum_{k=-\infty}^{\infty}f_k\varphi_k(t)\ .$$

Множитель $\frac{\sin 2\pi F_c(t-n\Delta t)}{2\pi F_c(t-n\Delta t)}$ носит название функции отсчета. Рассмотрим основные свойства функции отсчета.

1. В момент времени $t = n\Delta t$, она достигает максимального значения (оно равно единице).

2. В моменты времени кратные Δt ($t = (k \pm l)\Delta t$, где 1 – любое целое число), функция отсчета обращается в нуль.

3. Амплитуда функции отсчета быстро затухающая, в силу быстрого возрастания знаменателя и ограниченного значения амплитуды синуса.

4. Функции отсчетов на временном интервале $[-\infty, \infty]$ ортогональны.

Покажем свойство ортогональности функции отсчета. Для этого рассмотрим последовательности функции отсчета, отличающиеся между собой величиной сдвига на величину кратную Δt.

Из этой последовательности выберем две функции отсчета

$$\frac{\sin \omega_c(t-n\Delta t)}{\omega_c(t-n\Delta t)} \text{ и } \frac{\sin \omega_c(t-m\Delta t)}{\omega_c(t-m\Delta t)}.$$

Интеграл от произведения этих функций

$$\int \frac{\sin \omega_c(t-n\Delta t)}{\omega_c(t-n\Delta t)} \cdot \frac{\sin \omega_c(t-m\Delta t)}{\omega_c(t-m\Delta t)} dt = \begin{cases} \frac{1}{2F_c} = \frac{\pi}{\omega_c}, & \text{если } m = n \\ 0, & \text{если } m \neq n \end{cases},$$ т.

е. рассматриваемые нами функции отсчета ортогональны.

Для получения ортонормированного ряда нужно ввести нормирующий множитель. Тогда ортогональная последовательность функции отсчета примет вид

$$\varphi_n(t) = \sqrt{\frac{\omega_c}{\pi}} \frac{\sin \omega_c(t-n\Delta t)}{\omega_c(t-n\Delta t)}.$$

Теорема Котельникова может быть обобщена и на случай не симметричной относительно частоты $\omega_ń$ спектра. Временной ин-

тервал в этом случае определяется как $\Delta t = \frac{1}{\omega_1 - \omega_2}$.

Если у сообщения длительностью T_c ограничить спектр на частоте F_c, то в соответствии с теорией Котельникова для количества отсчетов получим

$$m = \frac{T_c}{\Delta t} = 2F_c T_c. \qquad (4.16)$$

В этом случае ряд В.А. Котельникова примет вид

$$x(t) = \hat{x}(t) = \sum_{k=-m/2}^{m/2} x_k \varphi_k(t). \qquad (4.17)$$

Ряд (4.17) имеет конечное число членов, следовательно, представление непрерывного сигнала таким рядом будет неточным. Приближенность этого выражения проявляется в том, что при конечном числе членов ряда их сумма точно совпадает с мгновенными значениями x(t) не на всем интервале T_c, а только в точках отсчета [8]. В интервалах между точками отсчета x(t) и $\hat{x}(t)$ различаются, и появляется погрешность. Уменьшить эту погрешность можно путем увеличения числа членов ряда. Из анализа (4.16) видно, при конечном значении T_c для этого необходимо уменьшить Δt, т. е. увеличивая значение F_c.

Среднеквадратическая ошибка, вызванная погрешностью, связанную с конечным числом членов ряда В.А. Котельникова определяется следующей формулой:

$$\delta_T^2 = \frac{\int_0^{T_c} [x(t) - \hat{x}(t)]^2 dt}{\int_0^{T_c} x^2(t) dt} = \frac{\int_0^{T_c} E_T^2(t) dt}{E}, \qquad (4.18)$$

где $E_T(t)$ – погрешность между x(t) и $\hat{x}(t)$, E – энергия непрерывных сообщений.

Таким образом, дискретизация непрерывных сообщений конечной длительности в соответствии с теоремой В.А. Котельникова

связана с ошибкой, одна составляющая которой обусловлена введением модели с ограниченным спектром:

$$\delta_F^{\,2} = \frac{\int_{\omega_c}^{\infty} G_x(\omega)d\omega}{\int_{0}^{\infty} G_x(\omega)d\omega} = \frac{\Delta P_x}{P_x} = \frac{\Delta E_x}{E_x}, \qquad (4.19)$$

а вторая – учетом конечного числа членов ряда разложения (4.18).

Общая среднеквадратическая ошибка определяется как

$$\delta_g^2 = \delta_T^2 + \delta_F^2 \geq 2\delta_F = \frac{2}{E_x}\int_{\Omega_b}^{\infty} G(\omega)d\omega = \frac{2\Delta E_x}{E_x}. \qquad (4.20)$$

Методы восстановления непрерывного сигнала. Восстановление непрерывной функции времени с конечной длительностью по ее отсчетам в соответствии (4.13) может выполняться двумя способами:

– фильтрационный метод с применением аналогового фильтра;

– интерполяционный метод с применением специальных интерполяторов или средств вычислительной техники.

При фильтрационном способе восстановления последовательность отсчетов определяется выражением (4.13), в котором $\Delta t = \frac{1}{2F_c}$, подается на фильтр нижней частоты ФНЧ. Напряжение на выходе фильтра определяется суперпозицией откликов фильтра на каждый поступающий отсчет.

Применяемый ФНЧ в этом случае должен иметь импульсную переходную функцию вида $h(t) = 2F_c \frac{\sin \omega_c t}{\omega_c t}$. Такая переходная функция соответствует фильтру нижней частоты ФНЧ с прямоугольной передаточной функцией и, поэтому физически нереализуем.

При интерполяционном способе (4.13) алгоритм восстановления непрерывной функции сводится к созданию $m = 2F_cT_c$ функ-

ции отсчета $\varphi_k(t) = \dfrac{\sin 2\pi F_c(t - k\Delta t)}{2\pi F_c(t - k\Delta t)}$ и суммированию их с учетом весовых коэффициентов, равных дискретным отсчетам $x_k = x(k\Delta t)$.

Технически это сложная операция, требующая запоминания отсчетов генерируемых функций $\varphi_к$ и их суммирования с учетом их весовых коэффициентов.

По указанным причинам предельная дискретизация непрерывных сообщений в соответствии с теоремой В.А. Котельникова, а также последующее восстановление сообщений рассмотренными способами не находят практического применения.

Дискретизация случайного сигнала. Дальнейшее развитие теории дискретизации сигналов связано с работами Н.А. Железнова, где решен вопрос максимального интервала Δt между отсчетами для случайного сигнала.

Отличительные особенности модели Н.А. Железнова:

1. Спектр сигнала сплошной и отличен от нуля на всей частотной области.

2. Функция корреляции сигналов равна нулю вне интервала корреляции τ_0.

3. Сигнал имеет конечную длительность и длительность сигнала намного больше интервала корреляции τ_0 ($T_c >> \tau_0$).

4. Рассматриваются стационарные и нестационарные процессы.

Неограниченность спектра сигнала и его конечная длительность является большими преимуществами этой модели: она большей степени отвечает реальным сигналам. Это случай модели сигнала с ограниченным интервалом корреляции.

Единственное ограничение – ограничение функции корреляции:

$$R_{xx}(\tau) = \begin{cases} R_{xx}(\tau), & \tau \le \tau_0, \\ 0, & \tau > \tau_0. \end{cases}$$

Для стационарных случайных сигналов, отвечающих перечисленным выше свойствам, Н.А. Железновым показано, что они могут быть предсказаны системой линейного прогнозирования со среднеквадратической ошибкой $\overline{\varepsilon}^2$, как угодно, мало отличающейся от нуля лишь в промежутке времени, равным интервалу корреляции τ_0. Тогда количество дискретных отсчетов равно

$$N = \frac{T_c}{\tau_0}.$$

Рассмотрим некоторые примеры по определению шага дискретизации непрерывного сообщения.

Пример 4.1. Определить по теореме Котельникова шаг дискретизации Δt для детерминированной функции

$$x(t) = \begin{cases} 2e^{-t} & \text{при} \quad t \geq 0, \\ 0 & \text{при} \quad t \prec 0, \end{cases}$$

с практической шириной спектра (3.11) и коэффициента $\eta = 0{,}95$.

Спектральная характеристика детерминированной функции x(t) вычисляется по формуле

$$S(j\omega) = 2\int_0^\infty e^{-t} e^{-j\omega t} dt = \frac{2}{1 + j\omega},$$

тогда модуль спектральной характеристики детерминированной функции x(t)

$$S(\omega) = 2\int_0^\infty e^{-t} e^{-j\omega t} dt = \frac{2}{\sqrt{1 + \omega^2}}.$$

Пользуясь формулой (4.19) определяем практическую ширину спектра

$$\int_0^\infty \frac{4}{1 + \omega^2} d\omega = \frac{0{,}95}{\pi} \int_0^\infty \frac{4}{1 + \omega^2} d\omega,$$

т. к. $\int_0^\infty \frac{1}{1 + \omega^2} d\omega = \operatorname{arctg}\omega \Big|_0^\infty = \frac{\pi}{2}$, имеем

$$\operatorname{arctg}\omega_{п} = 0{,}95\frac{\pi}{2} = 1{,}49 .$$

По таблице значений тангенсов получаем

$$\omega_{п} = 13{,}1 \ \frac{1}{с}.$$

Следовательно, шаг дискретизации Δt для детерминированной функции x(t)

$$\Delta t = \frac{\pi}{\omega_{п}} = 0{,}24 \ \text{с}.$$

4.4. Дискретизация по критерию наибольшего отклонения

В настоящее время в технических приложениях выбор интервала дискретизации непрерывных сообщений определяется в значительной мере требованиями удобства и простоты восстановления этих сообщений по их отсчетам.

В процессе дискретизации непрерывная функция $x(t)$, имеющая (n+1) ограниченных производных, аппроксимируется многочленом n-й степени.

В зависимости от выбранного способа восстановления он может быть интерполирующим или экстраполирующим. Задача обеспечения минимальной погрешности при восстановлении сигнала обычно не ставится, а указывается ее допустимая погрешность ε_0.

Погрешность восстановления $\varepsilon(t)$ функции $x(t)$ многочленом $\varphi_{п}(t)$ на каждом участке аппроксимации определяется остаточным членом $L_{п}(t)$

$$\varepsilon(t) = L_{п}(t) = x(t) - \varphi_{п}(t). \qquad (4.21)$$

Следовательно, шаг дискретизации выбирается из условия $L_{п}(t) \le \varepsilon_0$.

Выбор аппроксимирующего многочлена более высокой степени при малой допустимой погрешности ε_0 обеспечивает меньше число отсчетов, но при этом существенно возрастает сложность технической реализации метода.

Поэтому обычно ограничиваются многочленами нулевой, первой и второй степеней (ступенчатая, линейная и параболическая соответственно аппроксимации).

На практике широко применяется метод восстановления с помощью интерполяционных полиномов Лагранжа [2,9].

Многочлен Лагранжа для равносторонних узлов t_i на отрезке $[t_0, t_n]$ может быть записан в виде

$$L_n(t) = L_n[t_0 + \chi\Delta t] = (-1)^n \frac{\chi(\chi-1)..(\chi-n)}{n!} \sum_{i=}^{n} (-1)^i \frac{C_n^i x(t_i)}{\chi - i}, \quad (4.22)$$

где $t_1 - t_0 = t_2 - t_1 = ... = t_i - t_{i-1} = \Delta t;\ \chi = (\frac{t - t_0}{\Delta t}), n = 0,1,2,...,i = 0,1..\cdot$

Если все отсчеты для равноотстоящих точек функции $x(t)$ точны, то интерполяционный полином Лагранжа совпадает с исходной функцией в точках t_i. В остальных точках отрезка интерполяции $[t_0, t_n]$ совпадения не будет, кроме случая, когда исходная функция $x(t)$ является полиномом степени не выше n.

Погрешность воспроизведения исходной непрерывной функции x(t) интерполяционными полиномами Лагранжа $L_n(t)$, имеющими ограниченные производные порядка n, определяется остаточным членом:

$$R_n(t) = \left|x(t) - L_n(t)\right| \le \frac{M_{(n+1)}}{(n+1)!} \Delta t^{n+1} \left|\chi(\chi-1)...\chi - n\right|, \quad (4.23)$$

где $M_{(n+1)} = \max\left|x^{n+1}(t)\right|$ – модуль максимального значения (n+1) производных.

Шаг дискретизации Δt выбирают таким, чтобы по отсчетам функции $x(t_i)$ можно было восстановить $x(t)$ полиномом Лагранжа $L_n(t)$ с погрешностью

$$\varepsilon(t) = R_n(t) \le \varepsilon_0, \quad (4.24)$$

где ε_0 – заданная допустимая погрешность дискретизации.

Формулы (4.23) и (4.24) являются исходными для определения шага равномерной дискретизации Δt.

Рассмотрим наиболее часто применяемые способы дискретизации.

Пример 4.2. Определим шаг равномерной дискретизации на основе интерполирующих полиномов Лагранжа нулевой степени.

Значение восстанавливающей функции y(t) любой момент времени t на каждом j-ом интервале времени $t_{j-1} \le t \le t_j$ принимается равным отсчету $x(t_j)$ (рис. 4.3).

В этом случае остаточный член (4.24) имеет вид

$$R_0(t) \le M_1 \chi \Delta t . \qquad (4.25)$$

Очевидно, что максимальное значение $R_0(t)$ принимает при значении $\chi = 1$. С учетом (4.24) и (4.25) шаг равномерной дискретизации

$$\Delta t = \frac{\varepsilon_0}{M_1} . \qquad (4.26)$$

Частота отсчетов при использовании полиномов нулевой степени

$$F_0 = \frac{M_1}{\varepsilon_0} .$$

Восстановление исходного сигнала x(t) по отсчетным значениям $x(t_i)$ может проводиться как с экстраполяцией, так и интерполяцией.

Воспроизводящая функция $y(t) = L_0(t)$ для j-того интервала дискретизации $[t_{0j}, t_{0j+1}]$, при восстановлении с экстраполяцией равна

$$y_j(t) = x(t_{0j}) ,$$

при восстановлении с интерполяцией

$$y_j(t) = x(t_{0j+1}) . \qquad (4.27)$$

Показатель качества приближения на каждом из интервалов дискретизации Δt, выбранных в соответствии (4.23) может достигать величины $\lambda_{max} = 1$.

Если $y_i(t) = \dfrac{[x(t_{0j}) + x(t_{0j+1})]}{2}$, то шаг равномерной дискретизации может быть увеличен в два раза без увеличения погрешности.

Пример 4.3. Определим шаг равномерной дискретизации на основе интерполирующих полиномов Лагранжа первой степени.

При восстановлении исходного сигнала x(t) на каждом интервале времени $[t_{j-1}, t_j]$ используются два отсчета $x(t_j)$ и $x(t_{j-1})$. Они соединяются прямой линией (рис. 4.4). Максимальное значение для остаточного члена R_{1max} найдем приравняв нулю его производную

$$R_{1max} = \frac{M_2}{2}\left|\frac{(t_1 - t_0)(t_0 - t_1)}{4}\right| = \frac{M_2(\Delta t)^2}{8}. \qquad (4.28)$$

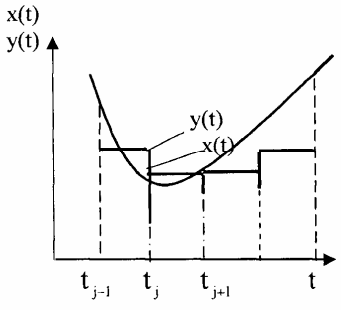


Рис. 4.3. Ступенчатая аппроксимация

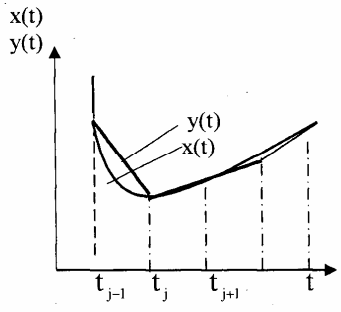


Рис. 4.4. Линейная аппроксимация

Отсюда находим допустимый шаг дискретизации

$$\Delta t = 2\sqrt{\frac{2\varepsilon_0}{M_2}}. \qquad (4.29)$$

4.5. Квантование сигналов по уровню

Квантование по уровню – это преобразование непрерывных

(по уровню) сигналов $x(t_i)$, полученных в результате квантования по времени, в моменты отсчета t_i в дискретные. В результате непрерывное множество значений сигнала $x(t_i)$ в диапазоне от x_{min} до x_{max} преобразуется в дискретное множество значений x_k – уровней квантования (рис. 4.5).

Шаг квантования Δx определяется по формуле

$$\Delta x = x_j - x_{j-1}.$$

Можно сказать, что квантование по уровню – это измерение сигнала.

В самом деле, по рис. 4.5 видно, что сигнал $x(t_1)$ составляет 0 уровней квантования (k = 0), а сигнал $x(t_4)$ – 2 уровня квантования (k = 2).

При квантовании по уровню не всегда сигнал $x(t_i)$ совпадает с уровнем квантования (см. сигнал $x(t_2)$ на рис.4.5). В таком случае поступают одним из следующих способов:

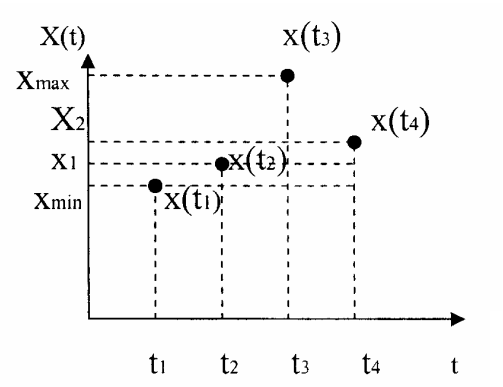


Рис. 4.5. Квантование сигналов по уровню

– $x(t_i)$ отождествляют с ближайшим значением (в нашем случае – с x_2);

– $x(t_i)$ отождествляют с ближайшим меньшим (или большим) значением. Тогда при отождествлении с ближайшим большим зна-

чением сигнал $x(t_2)$ отождествится с x_2 независимо от того, насколько близко он находится к этому уровню квантования. При отождествлении с ближайшим меньшим значением сигнал $x(t_2)$ отождествится с x_1 также независимо от того, насколько близко он к этому уровню квантования находится.

Очевидно, и при квантовании по уровню возникает погрешность квантования $\xi(x_k)$

$$\xi(x_k) = x(t_i) - x_k .$$

Погрешность квантования по уровню тем меньше, чем меньше шаг квантования.

Различают следующие виды квантования по уровню:

– равномерное, когда диапазон изменения сигнала разбивается на n одинаковых частей. Тогда, зная размер шага квантования, для представления x_k достаточно знать число k;

– неравномерное, когда диапазон изменения сигнала разбивается на n различных частей.

При равномерном квантовании шаг квантования Δx равен

$$\Delta x = \Delta x_i = x_i - x_{i-1} = \frac{x_{max} - x_{min}}{n} .$$

Определим ошибку квантования по критерию наибольшего отклонения

$$\sup_x \xi(x) = \sup_x \left| x(t) - \bar{x}_i \right| .$$

При заданном значении n наибольшая ошибка квантования $\sup|\xi(x)|$ минимальна, если $\Delta_{x_i} = \Delta_x = \text{const}$ и уровень квантования x_i, выбирается в середине интервала $\bar{x}_i = \frac{1}{2}(x_i + x_{i-1})$.

Рассматривая сигнал $x(t_i) = x$ как реализацию случайной величины X с плотностью распределения p(x), можно найти характеристики случайной величины Δx – погрешности квантования, также зависящей от x. Погрешность квантования

$$\Delta x_k = x - x_k \qquad (4.30)$$

изменяется в пределах $-0{,}5\delta_k \le x_k \le 0{,}5\delta_k$,

где $\max \Delta x_k = \max\left|x(t_i) - x_k\right| = \delta_k$.

Для k-го уровня квантования, $k = 0,\ldots,n-1$, математическое ожидание ошибки Δx_k:

$$M(\Delta x_k) = \int_{x_к - \frac{\delta_k}{2}}^{x_к + \frac{\delta_k}{2}} (x - x_k) p(x) dx, \qquad (4.31)$$

а дисперсия

$$D(\Delta x_k) = \int_{x_к - \frac{\delta_k}{2}}^{x_к + \frac{\delta_k}{2}} (x - x_k)^2 p(x) dx. \qquad (4.32)$$

Полагая, что интервал δ_k является малым по сравнению с диапазоном изменения сигнала x(t), можно принять значение $p(x)$ на этом интервале постоянной и равной своему значению для некоторого среднего уровня $p(x_k)$, т. е.

$$p(x_k)\delta_k = \int_{x_к - \frac{\delta_k}{2}}^{x_к + \frac{\delta_k}{2}} p(x) dx.$$

Тогда $M(\Delta x_k) = 0$, а

$$D(\Delta x_k) = \frac{1}{12} p(x_k)\delta_k^2 = \frac{1}{12}\delta_k^2 [p(x_k)\delta_k]. \qquad (4.33)$$

Выражение (4.33) показывает, что дисперсия погрешности квантования по уровню k-го уровня представляет собой дисперсию равномерно распределенных в этом интервале сигнала, умноженную на вероятность попадания x в этот интервал.

Дисперсию полной погрешности квантования по уровню для диапазона изменения сигнала (x_{min}, x_{max}) можно получить, суммируя (4.33) по всем уровням квантования:

$$D[\Delta x] = \sum_{k=0}^{m-1} D\left(\Delta x_k\right) = \frac{1}{12}\sum_{k=0}^{m-1} p(x_k)\delta_k^3. \qquad (4.34)$$

В случае равномерного шага квантования ($\delta_k = \delta$) дисперсия погрешности квантования по уровню равна

$$D[\Delta x] = \frac{\delta^2}{12}\sum_{k=0}^{m-1} p(x_k)\delta . \qquad (4.35)$$

Полагая, вероятность появления значения сигнала x(t), превосходящий максимальный уровень x_{min} пренебрежительно малым, запишем

$$\sum_{k=0}^{m-1} p(x_k)\delta = \sum_{k=0}^{m-1} p(x_k) = 1 .$$

Тогда для равномерного квантования сигнала по уровню получим

$$D[\Delta x] = \sum_{k=0}^{m-1} D(\Delta x_k) = \frac{1}{12}\sum_{k=0}^{m-1} p(x_k)\delta_k^3 = \frac{\delta^2}{12} . \qquad (4.36)$$

Отсюда среднеквадратичная погрешность квантования сигнала по уровню для равномерного квантования

$$\sigma[\Delta x] = \sqrt{D[\Delta x]} = \frac{\delta}{2\sqrt{3}} . \qquad (4.37)$$

Статистические характеристики Δx зависят как от вида p(x), так и от величины δ_k. Оптимальное квантование по уровню может быть связано с неравномерным разбиением шкалы X сигнала.

Кодирование. Для удобства последующего преобразования дискретный сигнал подвергается кодированию. Большинство кодов основано на системах счисления, причем использующих позиционный принцип образования числа, при котором значение каждой цифры зависит от ее положения в числе.

Примером позиционной формы записи чисел является та, которой мы пользуемся (так называемая арабская форма чисел). Так, в числах 123 и 321 значение цифры 3, например, определяется ее положением в числе: в первом случае она обозначает три единицы (т. е. просто три), а во втором – три сотни (т. е. триста).

Тогда число N получается по формуле

$$N = \sum_{i=0}^{l} a_i m^i = a_l m^l + \ldots + a_1 m^1 + a_0 m^0 , \qquad (4.38)$$

где l – количество разрядов числа, уменьшенное на 1; i – порядок разряда; m – основание системы счисления; a_i – множитель, принимающий любые целочисленные значения от 0 до m–1, и

соответствующий цифре i-го порядка числа.

Например, для десятичного ($m=10$) числа 345 его полное значение рассчитывается по формуле

$$3\cdot 102 + 4\cdot 101 + 5\cdot 100 = 345 .$$

Римские числа являются примером полупозиционной системы образования числа: так, в числах IX и XI знак I обозначает в обоих случаях единицу (признак непозиционной системы), но, будучи расположенным слева от знака X (обозначающего десять), вычитается из десяти, а при расположении справа – прибавляется к десяти. В первом случае полное значение числа равно 9, во втором – 11.

В современной информатике используются в основном три системы счисления (все – позиционные): двоичная, шестнадцатеричная и десятичная.

Двоичная система счисления используется для кодирования дискретного сигнала, потребителем которого является вычислительная техника. Такое положение дел сложилось исторически, поскольку двоичный сигнал проще представлять на аппаратном уровне. В этой системе счисления для представления числа применяются два знака – 0 и 1.

Шестнадцатеричная система счисления используется для кодирования дискретного сигнала, потребителем которого является хорошо подготовленный пользователь – специалист в области информатики. В такой форме представляется содержимое любого файла, затребованное через интегрированные оболочки операционной системы, например, средствами Norton Commander в случае MS DOS. Используемые знаки для представления числа – десятичные цифры от 0 до 9 и буквы латинского алфавита – A, B, C, D, E, F.

Десятичная система счисления используется для кодирования дискретного сигнала, потребителем которого является так называемый конечный пользователь – неспециалист в области информатики (очевидно, что и любой человек может выступать в роли такого потребителя). Используемые знаки для представления числа – цифры от 0 до 9.

Соответствие между первыми несколькими натуральными числами всех трех систем счисления представлено в таблице перевода (табл. 4.1).

Для различения систем счисления, в которых представлены числа, в обозначение двоичных и шестнадцатеричных чисел вводят дополнительные реквизиты:

• для двоичных чисел – нижний индекс справа от числа в виде цифры 2 или букв B либо b (binary – двоичный), либо знак B или b справа от числа. Например, 1010002 = 101000b = =101000B = 101000B = 101000b;

• для шестнадцатеричных чисел - нижний индекс справа от числа в виде числа 16 или букв H либо h (hexadecimal – шестнадцатеричный), либо знак H или h справа от числа. Например, 3AB16 = 3ABH = 3ABh = 3ABH = 3ABh.

Для перевода чисел из одной системы счисления в другую существуют определенные правила. Они различаются в зависимости от формата числа – целое или правильная дробь. Для вещественных чисел используется комбинация правил перевода для целого числа и правильной дроби.

Правила перевода целых чисел. Результатом перевода целого числа всегда является целое число.

Перевод из десятичной системы счисления в двоичную и шестнадцатеричную:

а) исходное целое число делится на основание системы счисления, в которую переводится (на 2 – при переводе в двоичную систему счисления или на 16 – при переводе в шестнадцатеричную); получается частное и остаток;

б) если полученное частное меньше основания системы счисления, в которую выполняется перевод, процесс деления прекращается, переходят к шагу в. Иначе над частным выполняют действия, описанные в шаге а;

Таблица 4.1

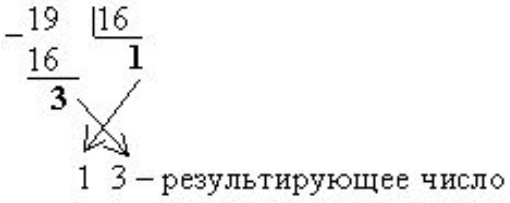
Таблица перевода

Десятичная система	Двоичная система	Шестнадцатеричная система
0	0	0
1	1	1
2	10	2
3	11	3
4	100	4
5	101	5
6	110	6
7	111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F
16	10000	10

в) все полученные остатки и последнее частное преобразуются в соответствии с таблицей перевода в цифры той системы счисления, в которую выполняется перевод;

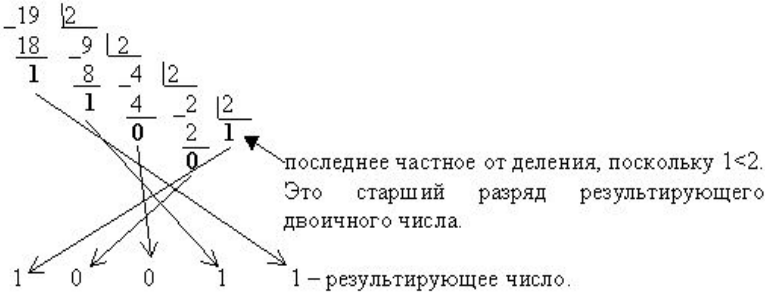
г) формируется результирующее число: его старший разряд – полученное последнее частное, каждый последующий младший разряд образуется из полученных остатков от деления, начиная с последнего и кончая первым. Таким образом, младший разряд полученного числа – первый остаток от деления, а старший – последнее частное.

Пример 4.4. Выполнить перевод числа 19 в шестнадцатеричную систему счисления:



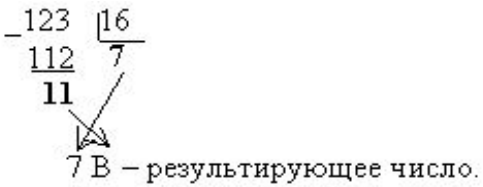
Таким образом, 19 = 1316.

Пример 4.5. Выполнить перевод числа 19 в двоичную систему счисления:



Таким образом, 19 = 100112.

Пример 4.6. Выполнить перевод числа 123 в шестнадцатеричную систему счисления:



Здесь остаток 11 преобразован в шестнадцатеричную цифру B

(табл. 4.1) и после этого данная цифра вошла в число. Таким образом, 123 = 7B16.

Перевод из двоичной и шестнадцатеричной систем счисления в десятичную осуществляется аналогично.

4.6. Определение энергии сигнала через отсчеты

При анализе и синтезе информационных систем часто используется выражение энергии сигнала, полученные через величины отсчетов.

Известно, что энергия равна (квадратичный эффект)

$$E = \int_{-\infty}^{\infty} f^2(t)dt . \qquad (4.39)$$

Воспользовавшись теоремой Котельникова, определим энергию сигнала через отсчеты.

Подставляя ряд Котельникова (4.15) в выражение энергии, получим

$$E = \int_{-\infty}^{\infty} f^2(t)dt = \int_{-\infty}^{\infty}\left[\sum_n f(\frac{n}{2F_c})\frac{\sin(2\pi F_c t - \pi n)}{2\pi F_c t - \pi n}\right]^2 dt . \qquad (4.40)$$

Квадрат суммы под интегралом можно представить в виде произведения двух сумм

$$E = \int_{-\infty}^{\infty}\left[\sum_к f(\frac{к}{2F_c})\frac{\sin(2\pi F_c t - \pi к)}{2\pi F_c t - \pi к}\right]\left[\sum_l f(\frac{l}{2F_c})\frac{\sin(2\pi F_c t - \pi l)}{2\pi F_c t - \pi l}\right]dt \qquad (4.41)$$

или в виде двойного ряда

$$E = \int_{-\infty}^{\infty}\left\{\sum_k\sum_l f(\frac{к}{2F_c}) f(\frac{l}{2F_c})\frac{\sin(2\pi F_c t - \pi к)}{2\pi F_c t - \pi к}\frac{\sin(2\pi F_c t - \pi l)}{2\pi F_c t - \pi l}\right\}dt . \qquad (4.42)$$

Изменяя порядок интегрирования и суммирования в (4.42) и, учитывая, что отсчеты $f(\frac{к}{2F_c})$ и $f(\frac{l}{2F_c})$ не зависят от переменной интегрирования t, получим

$$E = \sum_k\sum_l f(\frac{к}{2F_c}) f(\frac{l}{2F_c})\int_{-\infty}^{\infty}\frac{\sin(2\pi F_c t - \pi к)}{2\pi F_c t - \pi к}\frac{\sin(2\pi F_c t - \pi l)}{2\pi t - \pi l}dt . \qquad (4.43)$$

Вследствие ортогональности функции отсчетов при $k=1$ интеграл в выражении (4.43) равен $\frac{1}{2F_c}$. Окончательно, энергия дискретного сигнала

$$E=\frac{1}{2F_c}\sum_n\left[f(\frac{n}{2F_c})\right]^2. \qquad (4.44)$$

Если сигнал ограничен по времени интегралом [0,T], то количество отсчетов будет $2F_c\,T+1$, и справедлива теорема.

Теорема. Среднеквадратическое значение непрерывной функции f(t) равно среднему квадрату дискретных отсчетов (если значение $n=2F_cT$ достаточно велико),

$$\frac{1}{T}\int_0^T f^2(t)dt=\frac{1}{2F_cT}\sum_{n=0}^{2F_cT}f^2(\frac{n}{2F_c}). \qquad (4.45)$$

Если сигнал ограничен во времени $[t_1,t_2]$ ($T=t_2-t_1$), то количество отсчетов конечно и равно $n=2F_cT$. В этом случае, энергия сигнала равна

$$E=\frac{1}{2F_c}\sum_{n=0}^{2F_cT}\left[f^2(\frac{n}{2F_c})\right]^2. \qquad (4.46)$$

Контрольные вопросы:

1. В чем сущность процессов дискретизации и квантования?
2. Охарактеризуйте преимущества дискретной и цифровой передачи информации.
3. Сформулируйте общую постановку задачи дискретизации.
4. Сравните интерполяционные и экстраполяционные способы восстановления сигнала.
5. Что понимают под среднеквадратическим критерием восстановления сигнала?
6. Сформулируйте теорему Котельникова.
7. Поясните физическую возможность замены непрерывной

функции с ограниченным спектром совокупностью отсчетов.

8. Каковы теоретические неудобства использования функции с ограниченным спектром в качестве модели сигнала?
10. В чем трудности технической реализации способа передачи непрерывных сигналов, основанного на теореме В.А. Котельникова?
11. Какова процедура равномерной дискретизации по критерию наибольшего отклонения?
12. Укажите преимущества и недостатки адаптивной дискретизации.
13. Запишите выражение для среднеквадратической ошибки равномерного квантования.
14. Что такое шум квантования?
15. Как рассчитать шаг квантования сигнала при наличии помехи?
16. Определите энергию сигнала через дискретные отсчеты.

5. КАНАЛЫ И СИСТЕМЫ СВЯЗИ

Системой связи [3, 5] называют совокупность технических средств и математических методов, предназначенных для обмена сообщениями между пунктами (рис. 5.1). Она включает в себя передатчик (П), канал (К) и приемник (Пр).

Модуляцией называется изменение параметра переносчика в соответствии с изменением основного сигнала [11,18]. Основные виды сигналов-переносчиков приведены на рис. 5.2.

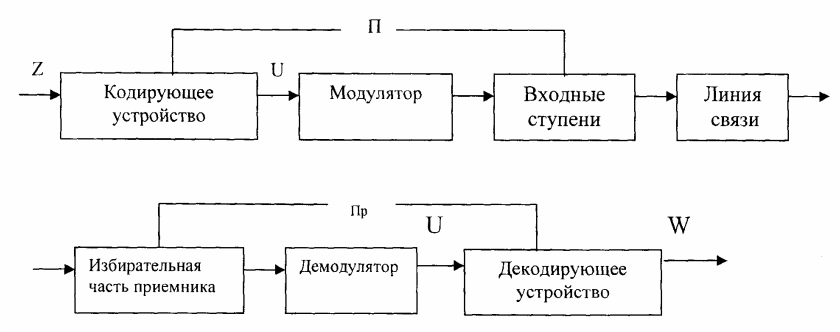


Рис. 5.1. Система связи

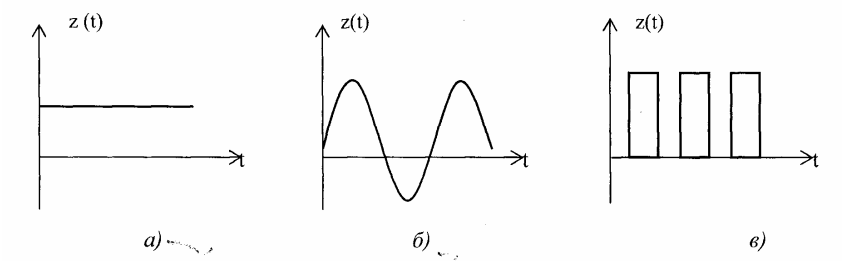


Рис. 5.2. Виды сигналов переносчиков информации:
а – постоянное напряжение; б – гармонические колебания (непрерывная модуляция); в – периодическая последовательность импульсов

5.1. Непрерывная модуляция

При непрерывной модуляции [7] переносчиком является гармоническое колебание непрерывной частоты

$$u = u_0 \cos(\omega_0 t + \varphi_0), \qquad (5.1)$$

где u_0 – амплитуда; ω_0 – частота и φ_0 – фаза колебания.

При амплитудной модуляции (АМ) амплитуда колебания изменяется относительно некоторого значения u_0 в зависимости от сигнала $x(t)$

$$u = u_0 + \Delta u \cdot x(t), \qquad (5.2)$$

где постоянный коэффициент Δu характеризует влияние сигнала $x(t)$ на амплитуду.

При частотной модуляции (ЧМ) частота колебания изменяется относительно некоторого значения ω_0 в зависимости от сигнала $x(t)$:

$$\omega = \omega_0 + \Delta\omega \cdot x(t), \qquad (5.3)$$

где постоянный коэффициент $\Delta\omega$ характеризует влияние сигнала $x(t)$ на частоту.

При фазовой модуляции (ФМ) фаза колебания изменяется относительно некоторого значения φ_0 в зависимости от сигнала $x(t)$:

$$\varphi = \varphi_0 + \Delta\varphi \cdot x(t), \qquad (5.4)$$

где постоянный коэффициент $\Delta\varphi$ характеризует влияние сигнала $x(t)$ на частоту.

Подставляя (5.2) в (5.1) получаем для амплитудной модуляции

$$u_{\text{АМ}} = u_0[1 + \frac{\Delta u}{u_0} x(t)]\cos(\omega_0 t + \varphi_0). \qquad (5.5)$$

Подставляя (5.4) в (5.1) получаем для фазовой модуляции

$$u_{\text{ФМ}} = u_0 \cos(\omega_0 t + \varphi_0 + \Delta\varphi \cdot x(t)). \qquad (5.6)$$

Если частотной модуляции подвергается входной синусоидальный сигнал $x(t) = x \cdot \sin\Omega t$, то будем иметь

$$u_{чм} = u_0 \cos(\omega_0 t - \frac{\Delta\omega \cdot x}{\Omega}\cos\Omega t), \qquad (5.7)$$

где $\frac{\Delta\omega}{\Omega}$ – индекс модуляции, $\Delta\omega \cdot x$ – девиация частоты.

При малом индексе модуляции полоса частот модулированного колебания приблизительно равна удвоенному значению частоты модулирующего колебания. При большом индексе модуляции полоса частот модулированного колебания близка к удвоенному значению девиации частоты.

Пример 5.1. Рассмотрим сигнал, изменяющийся по закону $x(t) = bt$ (рис. 5.3, *а*)

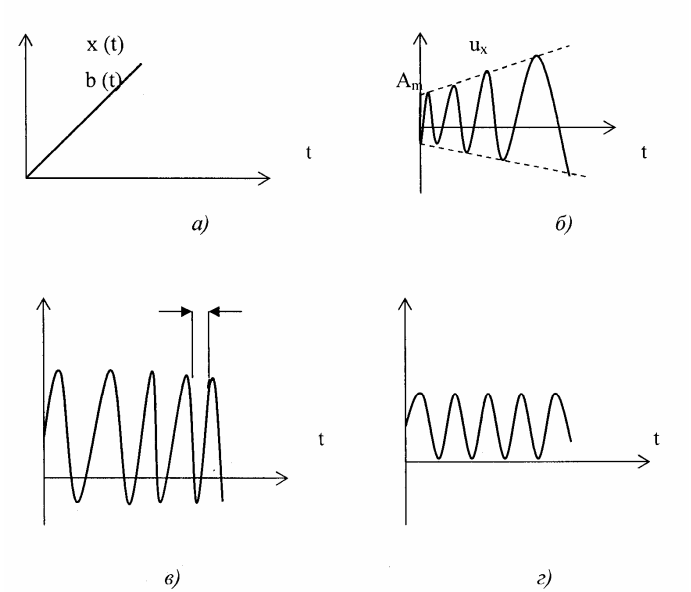


Рис.5.3. Первоначальный и модулированный сигналы
а – первоначальный сигнал; б – амплитудная модуляция; в – фазовая модуляция; г – частотная модуляция

Амплитудно-модулированный сигнал

$$u_{АМ} = u_0(1 - at)\cos\omega_0 t,$$

где $a = b\dfrac{\Delta u}{u_0}$ (рис. 5.3 *б*).

Фазомодулированный сигнал

$$u_{\text{ФМ}} = u_0 \cos(\omega_0 + a)t,$$

где $a = \Delta\varphi_b$ (рис. 5.3 *в*).

Частотно-модулированный сигнал

$$u_{\text{ЧМ}} = u_0 \cos[(\omega_0 + at)t],$$

где $a = \dfrac{\Delta\omega b}{2}$ (рис. 5.3 *г*).

5.2. Импульсная модуляция

Переносчиком при импульсной модуляции [11] является последовательность импульсов (в идеальном случае прямоугольной формы), имеющих параметры (рис. 5.4): h – уровень; τ_u – длительность; ω – частота; ν – фаза.

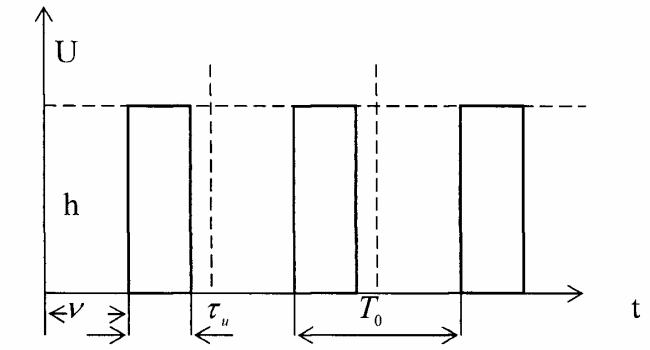


Рис. 5.4. Прямоугольная последовательность импульсов

Последовательность импульсов может быть представлена следующим рядом:

$$u = \sum_{k=0}^{\infty} h\delta_{\lambda}\left(t - k\frac{2\pi}{w} + v\right), \qquad (5.8)$$

где $\delta_{\lambda}(t)$ – «прямоугольная» функция, обладающая свойством

$$\delta_{\lambda}(t) = \begin{cases} 1, \text{при} & \frac{2k\pi}{\omega} \le t \le \frac{2k\pi}{\omega} + \tau_u \\ 0, \text{при} & \frac{2k\pi}{\omega} + \tau_u \le t \le \frac{2(k+1)\pi}{\omega} \end{cases}. \qquad (5.9)$$

Изменяя любой из параметров переносчика можно получить четыре вида импульсной модуляции:

1. Амплитудно-импульсная модуляция (АИМ)

$$h = h_0 + \Delta h \cdot x(t), \qquad (5.10)$$

Δh выбирается таким образом, чтобы

$$h_0 > \Delta h\left|X_{max}\right|. \qquad (5.11)$$

2. Широтно-импульсная модуляция (ШИМ)

$$\tau_u = \tau_{u0} + \Delta\tau_u x(t), \qquad (5.12)$$

$$\tau_{u0} > \Delta\tau_u\left|X_{max}\right|. \qquad (5.13)$$

3. Частотно-импульсная модуляция (ЧИМ)

$$\omega = \omega_0 + \Delta\omega \cdot x(t). \qquad (5.14)$$

4. Фазово-импульсная модуляция (ФИМ)

$$v = v_0 + \Delta v \cdot x(t). \qquad (5.15)$$

5.3. Цифровые методы модуляции

Цифровые методы модуляции [3,18] основаны на трех преобразованиях полезных непрерывных сигналов:

1) дискретизации;
2) квантовании;
3) кодировании.

Четвертое преобразование – модуляцию используют при передаче сигналов в многоканальных системах.

Наиболее распространенными цифровыми видами модуляции являются импульсно-кодовая модуляция, дельта- модуляция и комбинированные виды.

Достоинства цифровых способов передачи информации:

1. Слабое влияние нестабильности характеристик аппаратуры на качество передачи информации.
2. Высокая помехоустойчивость.
3. Возможность восстановления сигналов в узлах связи сетей.
4. Универсальная форма представления сигналов для различных сообщений (речь, телевизионное изображение, дискретные данные и т. д.).
5. Простое согласование этих систем с ЦВМ и электронными автоматическими телефонными станциями.
6. Возможность автоматизации передачи и обработки сигналов с помощью ЦВМ.

Области применения цифровых видов модуляции:

Уплотнение действующих соединительных линий между АТС; построение интегральных сетей связи с электронными АТС; радиорелейные линии связи прямой видимости и дальние линии связи; спутниковые системы связи; телеметрия. Временные диаграммы сигналов в системе с импульсно-кодовой модуляцией ИКМ приведены на рис. 5.5.

5.4. Спектральный анализ модулированных колебаний

Ширина спектра модулированного сигнала зависит от типа модуляции. В соответствии с этим полезно оценить спектры сигналов, передаваемых с помощью различных типов модуляции, т. к. шириной спектра определяется потенциальная помехоустойчивость передачи сигнала на физическом уровне [3,4].

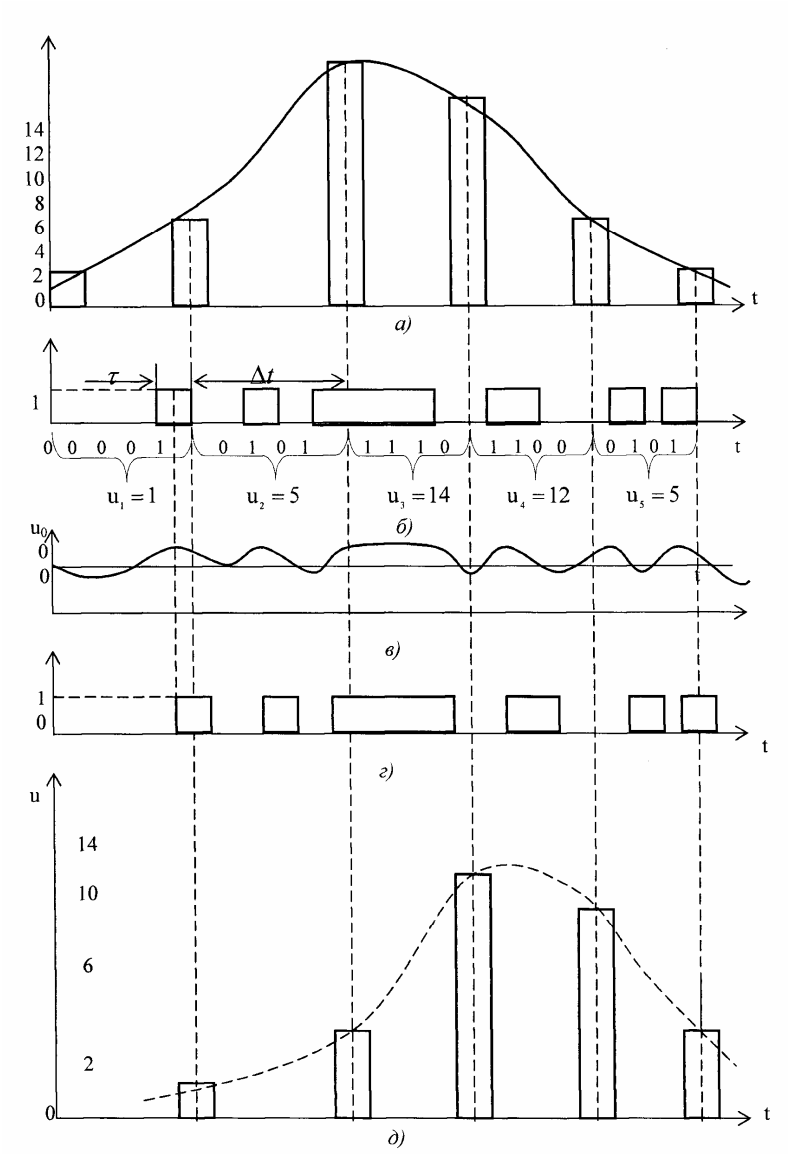


Рис. 5.5. Временные диаграммы сигналов в системе с ИКМ

Рассмотрим спектры периодических сигналов при

различных видах непрерывной модуляции.

а) Амплитудно-модулированный гармонический сигнал

$$x(t) = A(t)\sin(\omega_0 t + \varphi_0). \qquad (5.16)$$

При амплитудной модуляции амплитуда изменяется по определенному закону

$$A(t) = A_0 + \delta A\, f(t), \qquad (5.17)$$

где A_0 – постоянная составляющая амплитуд; δA – наибольшее изменение амплитуды при модуляции; $f(t)$ – нормированная функция (изменяющаяся в пределах от – 1 до +1). Так как модулируемый параметр сигнала (в данном случае амплитуда) является непосредственным переносчиком информации, то функция $f(t)$ выражает закон изменения во времени передаваемого сообщения.

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

$$x(t) = A_0[1 + m_A f(t)]\sin(\omega_0 t + \varphi_0), \qquad (5.18)$$

где $m_A = \dfrac{\delta_A}{A_0}$ – глубина амплитудной модуляции.

Рассмотрим частный случай, когда функция $f(t)$ изменяется по гармоническому закону [6]

$$f(t) = \cos\Omega t, \qquad (5.19)$$

причем $\Omega \prec\prec \omega_0$.

Тогда выражение (5.18) примет вид

$$x(t) = A_0[1 + m_A \cos\Omega t]\sin(\omega_0 t + \varphi_0) = A_0 \sin(\omega_0 t + \varphi_0) +$$
$$+ \frac{A_0 m_A}{2}\sin[(\omega_0 - \Omega)t + \varphi_0] + \frac{A_0 m_A}{2}\sin[(\omega_0 + \Omega)t + \varphi_0]. \qquad (5.20)$$

Как видно из выражения (5.20), спектр сигнала x(t) состоит из трех гармонических составляющих: несущей с частотой ω_0 и двух боковых – нижней с частотой $\omega_0 - \Omega$ и верхней с частотой $\omega_0 + \Omega$ (рис. 5.6, *а*). Ширина спектра сигнала $\Delta\omega = 2\Omega$.

Если модуляция гармонического сигнала производится по более сложному закону, причем спектр огибающей амплитуды находится в диапазоне частот от $\Omega_{\text{мин}}$ до $\Omega_{\text{макс}}$, то можно показать, что в спектре амплитудно-модулированного сигнала вместо двух боковых частот будут две боковые полосы частот (рис. 5.6, *б*). Причем, нижняя боковая полоса частот будет находиться в пределах $\omega_0 - \Omega_{\text{макс}}$, $\omega_0 - \Omega_{\text{мин}}$, а верхняя боковая полоса частот – в пределах $\omega_0 + \Omega_{\text{макс}}$, $\omega_0 + \Omega_{\text{макс}}$.

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

Для уменьшения полосы частот модулированного сигнала, а следовательно, и требуемой полосы пропускания канала связи иногда используется так называемая однополосная передача. При такой передаче с помощью фильтров производится подавление несущей частоты и частот одной из боковых полос, т. е. передача производится в полосе частот $\Omega_{\text{макс}} - \Omega_{\text{мин}}$. Это позволяет более чем в 2 раза сократить требуемую полосу канала связи.

б) Частотно-модулированный гармонический сигнал

При частотной модуляции частота сигнала изменяется в общем виде по закону

$$\omega(t) = \omega_0[1 + m_\omega f(t)], \qquad (5.21)$$

где ω_0 – постоянная составляющая частоты; $m_\omega = \dfrac{\delta\omega}{\omega_0}$ – глубина частотной модуляции; $\delta\omega$ – наибольшее изменение частоты при модуляции (девиация частот); f(t*)* – закон модуляции.

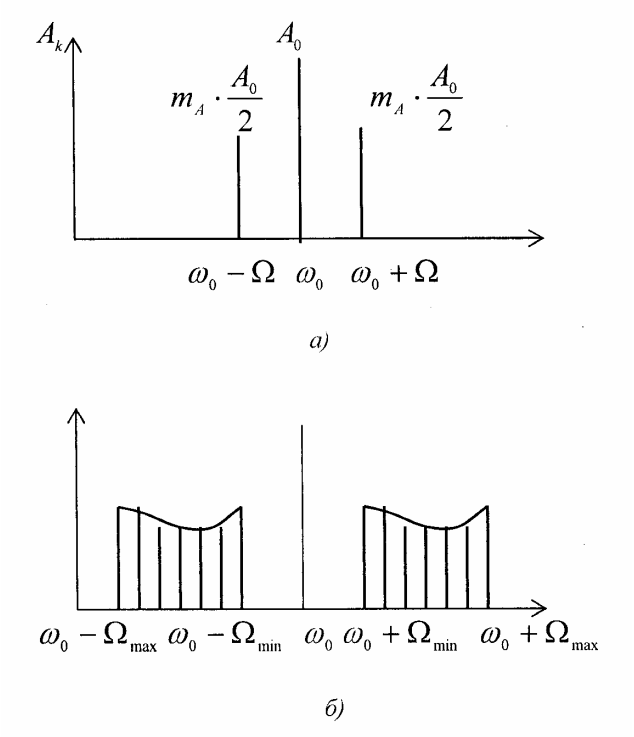


Рис. 5.6. Спектры гармонического сигнала

Частотно-модулированный сигнал можно представить в виде

$$x(t) = A_0 \sin\varphi(t) = A_0 \sin[\int_0^t \omega(t)dt + \varphi_0] =$$

$$= A_0 \sin[\omega_0 t + \omega_0 m_\omega \int_0^t f(t)dt + \varphi_0], \qquad (5.22)$$

где $\varphi(t)$ – текущая фаза гармонического сигнала; φ_0 – начальная фаза сигнала.

В частном случае, когда модуляция частоты осуществляется по гармоническому закону, частотно-модулированный сигнал может быть записан в виде

$$x(t) = A_0 \sin[\omega_0 t + \omega_0 m_\omega \int_0^t \cos\Omega t dt + \varphi_0] = A_0 \sin(\omega_0 t + \beta \sin\Omega t) =$$

$$= A_0 \sin\omega_0 t \cos[\beta \sin\Omega t] + A_0 \cos\omega_0 t \sin[\beta \sin\Omega t], \quad (5.23)$$

где $\beta = \dfrac{\delta\omega}{\Omega}$ – индекс частотной модуляции.

Сигнал такого вида представляется следующим образом:

$$x(t) = A_0 \{j_0(\beta)\sin\omega_0 t + \sum_{k=1}^{\infty} j_k(\beta)[\sin(\omega_0 + k\Omega)t + (-1)^k \sin(\omega_0 - k\Omega)t]\}, \quad (5.24)$$

где $J_o(\beta)$ – функция Бесселя нулевого порядка; $J_k(\beta)$ – функция Бесселя k -го порядка.

Спектр амплитуд модулированного по частоте сигнала приведен на рис. 5.7.

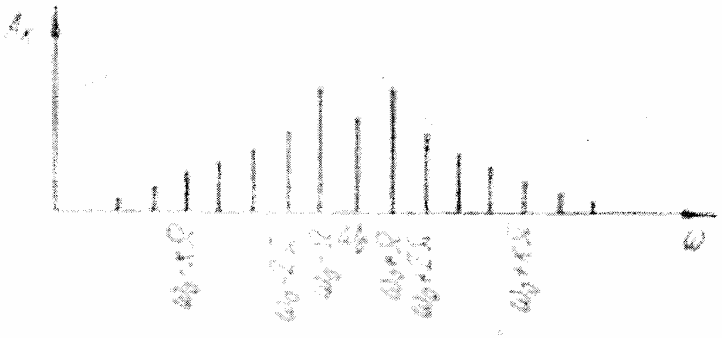


Рис.5.7. Спектр частотно-модулированного сигнала

Спектр дискретный состоит из колебаний с несущей частотой ω_0 и бесконечного количества верхних и нижних боковых составляющих, с частотами $\omega_0 + k\Omega$ и $\omega_0 - k\Omega$. Распределение амплитуд гармонических составляющих зависит от индекса частотной модуляции. При очень малом индексе частотной модуляции ($\beta \rightarrow 0$) спектр частотно-модулированного сигнала практически не отличается от спектра амплитудно-модулированного сигнала, т. е. он состоит всего из трех гармонических составляющих: колебаний с несущей частотой ω_0 и двух боковых с частотами $\omega_0 - \Omega$ и $\omega_0 + \Omega$. В этом случае полоса пропускания канала связи должна

быть равна 2Ω.

С ростом индекса модуляции β увеличивается вес боковых составляющих и соответственно требуемая полоса пропускания канала связи. Если ограничиться в спектре только составляющими, амплитуда которых составляет не менее 5 – 10 % от амплитуды несущей A_o до модуляции, то ширина спектра этих составляющих, а следовательно, и требуемая полоса канала связи будут составлять $2(\delta\omega + \Omega)$.

При больших индексах модуляции (β >>1) ширина спектра сигнала практически равна удвоенной величине девиации частоты $2\delta\omega$ и не зависит от частоты модулирующего сигнала Ω.

в) Фазомодулированный гармонический сигнал

$$x(t) = A_0 \sin[\omega_0 t + \varphi(t)]. \qquad (5.25)$$

При фазовой модуляции фаза гармонического сигнала изменяется по закону модулирующего сигнала *f (t)*

$$\varphi(t) = \varphi_0[1 + m_\varphi f(t)], \qquad (5.26)$$

где $m_\varphi = \dfrac{\delta\varphi}{\varphi_0}$ – глубина фазовой модуляции; $\delta\varphi$ – наибольшее изменение фазы при модуляции (индекс фазовой модуляции).

Сигнал с фазовой модуляцией можно представить в следующем виде:

$$x(t) = A_0 \sin[\omega_0 t + \varphi_0 m_\varphi f(t) + \varphi_0]. \qquad (5.27)$$

Так как мгновенное значение частоты является производной от фазы по времени $\omega = \dfrac{d\varphi}{dt}$, то при фазовой модуляции оно может быть представлено в виде

$$\omega(t) = \omega_0 + \varphi_0 m_\varphi \frac{df(t)}{dt}. \qquad (5.28)$$

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

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

$$\varphi(t) = \omega_0 t + \varphi_0 m_\varphi \cos \Omega t + \varphi_0 , \qquad (5.29)$$

а фазомодулированное колебание опишется выражением

$$x(t) = A_0 \sin[\omega_0 t + \varphi_0 m_\varphi \cos \Omega t + \varphi_0] . \qquad (5.30)$$

Мгновенное значение частоты фазомодулированного колебания

$$\omega(t) = \frac{d\varphi(t)}{dt} = \omega_0 - \delta\varphi\Omega \sin \Omega t = \omega_0 - \omega_в \sin \Omega t , \qquad (5.31)$$

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

Сравнение выражений (5.27) и (5.23) показывает, что при гармоническом модулирующем сигнале выражение, описывающее частотно-модулированное колебание, отличается от такого же для фазомодулированного колебания только фазой гармонической функции, определяющей изменение полной фазы высокочастотного колебания. В связи с этим по внешнему виду сигнала нельзя заключить, как модулирован сигнал – по частоте или фазе.

Однако при изменении частоты модуляции Ω проявляется различие между частотной и фазовой модуляциями.

При частотной модуляции величина девиации частоты $\omega_в$ зависит только от амплитуды модулирующего сигнала и не зависит от частоты модуляции Ω. Величина же индекса β убывает с ростом частоты Ω.

При фазовой модуляции величина индекса модуляции определяется выражением $\beta = \frac{\omega_в}{\Omega} = \delta\varphi$, т. е. зависит только от амплитуды модулирующего сигнала и не зависит от частоты модуляции.

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

Спектр фазомодулированного сигнала состоит из колебаний с несущей частотой ω_0 и бесконечного количества верхних и нижних

боковых составляющих с частотами $\omega_0 \pm k\Omega$. Амплитуды боковых составляющих определяются только индексом модуляции β. При очень малом значении индекса фазовой модуляции β спектр фазомодулированного сигнала состоит из несущей частоты ω_0 и двух боковых частот $\omega_0 - \Omega$ и $\omega_0 + \Omega$. Ширина спектра равна 2Ω. При $\beta >> 1$ практическая ширина спектра сигнала равна удвоенной девиации частоты $\Delta\omega = 2\omega_в$, но зависит как от амплитуды, так и от частоты модулирующего сигнала, в то время как при частотной модуляции ширина спектра в данном случае зависит только от амплитуды модулирующего сигнала.

Практическая ширина спектра частотно-модулированного и фазомодулированного сигнала примерно в β раз шире спектра амплитудно-модулированного сигнала при такой же частоте модуляции.

5.5. Согласование физических характеристик сигнала и канала

Независимо от назначения непрерывного канала его можно охарактеризовать тремя основными параметрами [9]: временем, в течение которого он предоставляется для передачи сигнала T_K; шириной полосы пропускания сигнала $F_к$ и допустимым превышением сигнала над помехой в канале H_K. Объем (емкость канала) $V_к = T_к \cdot F_к \cdot H_к$.

Передаваемый сигнал также характеризуется тремя параметрами:

а) длительностью T_c;

б) шириной спектра F_c;

в) превышением над помехой H_c.

Характеристика H_c определяется формулой

$$H_C = \log(P_u / P_\xi), \qquad (5.32)$$

где P_u – средняя мощность передаваемого сигнала, P_ξ – средняя мощность помехи в канале.

Емкость сигнала определяется

$$V_c = T_c \cdot F_c \cdot H_c . \qquad (5.33)$$

Тогда необходимое условие неискаженной передачи сигнала по данному каналу определяется условием [2, 3]

$$V_c \le V_к . \qquad (5.34)$$

Достаточные условия неискаженной передачи:

$$T_c \le T_к , \quad F_c \le F_к , \quad H_c \le H_к . \qquad (5.35)$$

Если $F_c > F_к$, то ширину спектра можно уменьшить за счет увеличения длительности, а в случае, если $T_c > T_к$, то согласование осуществляется за счет расширения спектра сигнала.

Пример 5.2. Записанный на магнитную ленту телефонный сигнал, имеющий ширину спектра 3 кГц, необходимо передать по каналу связи с шириной полосы пропускания 300 Гц. Это можно сделать, воспроизводя при передаче сигнал со скоростью в 10 раз меньшей той скорости, на которой он был записан. При этом все частоты спектра исходного сигнала уменьшаются в 10 раз, но во столько же раз увеличивается время передачи. При приеме сигнал записывается на магнитную ленту и затем воспроизводится со скоростью в 10 раз большей. Можно передавать сигналы быстрее, чем они создавались, если полоса пропускания канала шире ширины спектра сигнала. Большой интерес представляет собой возможность "обмена" динамического диапазона на полосу пропускания. Например, используя широкополосные, помехоустойчивые виды модуляции, можно передавать сообщение с динамическим диапазоном 60 дБ по каналу, в котором сигнал превышает помеху всего лишь на 20 дБ. При этом используется полоса пропускания канала в несколько раз более широкая, чем спектр передаваемого сообщения.

5.6. Многоканальные системы связи

Систему связи называют *многоканальной,* если она обеспечивает передачу нескольких сообщений по одной общей линии связи.

В многоканальной системе связи несколько сообщений передаются по одной линии связи [6, 16]. Структурная схема такой системы представлена на рис. 5.8.

Сообщения от нескольких источников $\lambda_1(t)$, $\lambda_2(t)$, … , $\lambda_n(t)$ преобразуются модуляторами M_1 , M_2 , … , M_n в электрические сигналы $S_1(t)$, $S_2(t)$, … , $S_n(t)$, которые в устройстве уплотнения (объединения) объединяются в групповой сигнал S(t), передаваемый по линии связи. В результате воздействия помехи n(t) искажается, и на вход аппаратуры разъединения подается искаженный групповой сигнал $\hat{S}(t)$, из которого выделяются копии индивидуальных сигналов $\hat{S}_1(t)$, $\hat{S}_2(t)$, … , $\hat{S}_n(t)$, преобразуемые канальными демодуляторами $Д_1, Д_2, \ldots, Д_n$ копии соответствующих сообщений $\hat{\lambda}_1(t), \hat{\lambda}_2(t), \ldots, \hat{\lambda}_n(t)$. В основном используются принципы частотного и временного разделения каналов. Первый, как правило, используется в аналоговых системах связи, второй – в дискретных и цифровых системах, но возможны их комбинации.

В реальном канале сигнал в процессе его передачи претерпевает искажение, и сообщение воспроизводится с некоторой ошибкой [4].

Существуют две причины ошибок: искажения, вносимые каналом, и помехи, воздействующие на сигнал. Частотные и временные характеристики определяют так называемые линейные искажения, но канал может вносить и нелинейные искажения, обусловленные нелинейностью (преднамеренной или непреднамеренной) характеристик отдельных звеньев канала. Линейные и нелинейные искажения обусловлены известными характеристиками канала и могут быть (по крайней мере в принципе) устранены в результате применения тех или иных методов коррекции. В отличие от искажений, помехи имеют случайный характер, заранее не известны и не могут быть полностью исключены.

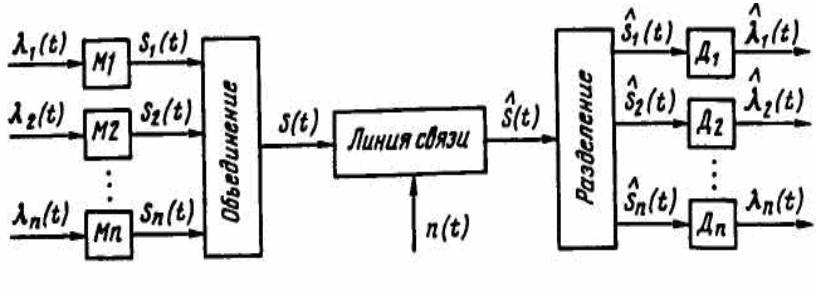


Рис. 5.8. Многоканальные системы связи

Помеха – любое случайное воздействие на сигнал, которое ухудшает верность воспроизведения передаваемых сообщений. Помехи разнообразны по происхождению и физическим свойствам:

– атмосферные (в радиоканалах, обусловлены электрическими процессами в атмосфере, энергия их в основном сосредоточена в диапазонах длинных и средних волн);

– индустриальные (возникают в результате резких изменений тока в электрических цепях различных устройств, например, электрического транспорта, системах зажигания двигателей и т. п.);

– помехи от посторонних радиостанций и радиоканалов (возникают из-за нарушения регламента рабочих частот, недостаточной стабильности элементов схем, плохой фильтрации гармоник сигнала, нелинейных процессов в каналах "перекрестные помехи");

– импульсные шумы и прерывания связи (в проводной связи импульсные шумы определяются, например, работой устройства автоматической коммутации каналов);

– внутренние шумы аппаратуры (присутствуют во всех диапазонах, обусловлены хаотическим движением носителей зарядов в усилительных элементах, резисторах; особенно проявляются на сверхвысоких частотах, где остальные помехи невелики).

В общем случае влияние помехи n(t), на передаваемый сигнал S (λ, t), может быть описано оператором y (t) = V [S (λ, t), n(t)]. В частном случае, когда этот оператор выражает сумму – y (t) = =S(λ,t)+ n(t), то помеху называют аддитивной. Если оператор выражает произведение y(t) = S (λ, t) n(t), то помеху называют мультипликативной.

Мультипликативная помеха возникает из-за изменения во времени амплитудных и фазовых характеристик канала, например при прохождении сигнала через среду с изменяющимися во времени параметрами, при многолучевом распространении радиоволны и т. д.

Среди аддитивных помех различного происхождения особое место занимают флуктуационные (флюктуационные) помехи, которые, как правило, являются непрерывными, случайными, нормальными (гауссовскими) процессами. Источником таких помех являются входные элементы или цепи приемных устройств.

Импульсные помехи представляет собой хаотическую последовательность коротких импульсов (как правило, это помехи индустриального происхождения, например, на железных дорогах от устройств контактной сети).

Между сигналом и помехой отсутствует принципиальное различие, они существуют в единстве. Например, излучение радиопередатчика является полезным сигналом для "своего" приемника и помехой для других приемников.

Контрольные вопросы:

1. В чем состоит назначение и содержание процедур модуляции и демодуляции?
2. Запишите выражение для амплитудно-модулированного колебания.
3. Приведите выражение для частотно-модулированного колебания.
4. Каковы особенности импульсных видов модуляции?
5. Запишите выражение для широтно-импульсного модулированного колебания.
6. Поясните смысл операций: дискретизация по времени, квантование по уровню.
7. В чем причина высокой помехоустойчивости при цифровых методах модуляции?
8. Назовите области применения цифровых методов модуляции.
9. Как изменяется спектр модулирующего сигнала при амплитудной модуляции?

10. Как изменяется спектр модулирующего сигнала при частотной модуляции?
11. Сравните методы частотной и амплитудной модуляции по помехоустойчивости и
сложности реализации.
12. Определите условия неискаженной передачи сигнала по каналу.
13. Какие в основном используются принципы разделения каналов?
14. Какие помехи называют аддитивными?
15. Какие помехи называют мультипликативными?

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Акулиничев Ю.П., Дроздова В.И.* Сборник задач по теории информации. Томск: Изд-во Томского Университета, 1976.

2. *Борисов Ю.П., Пеннин П.И.* Основы многоканальной передачи информации. М.: Связь, 1967.

3. *Дмитриев В.И.* Прикладная теория информации. М.: Высш. шк., 1989.

4. *Зюко А.Г.* Помехоустойчивость и эффективность систем связи. М.: Связь, 1972.

5. *Зюко А.Г.* Элементы теории передачи информации. М.: Техника, 1969.

6. *Коган И.М.* Прикладная теория информации. М.: Радио и связь, 1981.

7. *Липкин И.А.* Основы статистической радиотехники, теории информации и кодирования. М.: Радио и связь, 1978.

8. *Игнатов В.А.* Теория информации и передачи сигналов: Учебник для высших учебных заведений. М.: Радио и связь, 1991.

9. *Советов Б.Я.* Теория информации. Л.: Изд-во ЛГУ, 1985.

10. *Советов Б.Я.* Информационные технологии: Учебник для вузов по специальности «Автоматизированные системы обработки информации и управления». М.: Высш. шк.,1994.

11. *Кузьмин И.В., Кедрус В.А.* Основы теории информации и кодирования. Киев: Вищ. шк., 1977.

12. *Лебедев Д.С.* Учеб. пособие по курсу «Основы теории информации». М.: Изд-во МГУ, 1966.

13. *Темников Ф.Е., Афонин В.А., Дмитриев В.И.* Теоретические основы информационной техники. М.: Энергия, 1979.

14. *Горелов Г.В., Фомин А.Ф.* и *др.* Теория передачи сигналов на железнодорожном транспорте. М.: Транспорт, 1999.

15. *Пеннин П.И.* Системы передачи цифровой информации. М.: Радио и связь, 1976.

16. *Пугачев В.С.* Теория случайных процессов и ее применение к задачам автоматического управления. М.: Изд-во технико-теоретической литературы, 1957.

17. Основы кибернетики. Теория кибернетических систем/ Под ред. *К.А. Пупкова*. М.: Высш. шк., 1976.

18. *Солодов А.В.* Теория информации ее применение к задачам автоматического управления. М.: Наука, 1967.

19. *Тарасенко Ф.П.* Введение в курс теории информации. Томск: Изд-во ТГУ, 1963.

20. *Финк Л.М.* Теория передачи дискретных сообщений. М.: Советское радио, 1970.

21. *Шеннон К.* Работы по теории информации и кибернетике. М.: Изд-во иностранной литературы,1963.

22. *Харкевич А.А.* Борьба с помехами. М.: Физматгиз, 1965.

23. *Хартли Р.* Передача информации. Теория информации и ее приложения/ Под ред. А.А. Харкевича. М.: Физматгиз, 1959.

24. *Хэмминг Р.В.* Теория кодирования и теория информации. М.: Радио и связь, 1983.

25. *Четвериков В.Н.* Преобразование и передача информации в АСУ. М.: Высш. шк., 1974.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ ……. 3
1. ОСНОВНЫЕ ПОЛОЖЕНИЯ ТЕОРИИ ИНФОРМАЦИИ ……... 5
1.1. Теория информации, и ее место среди дисциплин специальности ……. 5
1.2. Основные понятия и определения теории информации ……..... 8
1.3. Виды информации. Этапы обращения информации …….. 10
Контрольные вопросы ……..... 14
2. ИЗМЕРЕНИЕ ИНФОРМАЦИИ ……... …15
2.1. Структурные меры измерения информации ..……..... 15
2.1.1. Геометрическая мера …….. 15
2.1.2. Комбинаторная мера ……..... 16
2.1.3. Аддитивная мера (мера хартли)……. 18
2.2. Семантические меры измерения информации……..... 19
2.2.1. Содержательность информации ……... 20
2.2.2. Целесообразность информации …… 21
2.2.3. Динамическая энтропия……24
2.3. Статистические меры измерения информации ……... 25
2.3.1. Вероятность и информация …… 25
2.3.2. Свойства энтропии ансамбля……..... 28
2.3.3. Условная энтропия и энтропия объединения ……..... 29
2.3.4. Количество информации и избыточность …… 34
2.3.5. Дифференциальная энтропия …… 37
2.3.6. Эпсилон-энтропия ……..... 41
2.3.7. Информационные потери при передаче по дискретным каналам связи с шумами …….. 42
2.3.8. Скорость передачи информации по дискретному каналу связи ……. 44
2.3.9. Пропускная способность дискретного канала ……..... 46
Контрольные вопросы …….. ..48
3. МАТЕМАТИЧЕСКИЕ МОДЕЛИ СИГНАЛОВ ……. 50
3.1. Сигнал и его модели …….. ..50
3.2. Временная форма представления сигнала ……..... 53
3.3. Частотная форма представления сигнала…… 55
3.4. Распределение энергии сигнала в спектре …….. 60

3.5. Случайный процесс как модель сигнала ……..... 63
3.5.1. Вероятностные характеристики случайных процессов…… 66
3.5.2. Стационарные случайные процессы …… 70
3.5.3. Спектральная плотность стационарного
случайного процесса …… 78
3.5.4. Эффективная ширина спектра случайного процесса …….. 83
Контрольные вопросы ……..... 84
4. КВАНТОВАНИЕ ИНФОРМАЦИИ…… 86
4.1. Общая постановка задачи дискретизации информации…… 86
4.2. Предельная дискретизация непрерывных сообщений…..... 90
4.3. Методы восстановления непрерывного сигнала……..... 94
4.4. Дискретизация по критерию наибольшего
отклонения …….....100
4.5. Квантование сигналов по уровню ……..104
4.6. Определение энергии сигнала через отсчеты ……..... 113
Контрольные вопросы …….. 114
5. КАНАЛЫ И СИСТЕМЫ СВЯЗИ ……..116
5.1. Непрерывная модуляция ……117
5.2. Импульсная модуляция …….....119
5.3. Цифровые методы модуляции …….. 120
5.4. Спектральный анализ модулированных колебаний…….. 121
5.5. Согласование физических характеристик сигнала
и канала …… 129
5.6. Многоканальные системы связи ……131
Контрольные вопросы …… 133
БИБЛИОГРАФИЧЕСКИЙ СПИСОК ……...135

Учебное издание

Бекмурза Хусаинович Айтчанов,
Татьяна Владимировна Яскевич

ТЕОРИЯ ИНФОРМАЦИИ

Учебное пособие

РО НТИЦ *З.А. Губайдулина*
Редактор *Ш.Т. Касымбекова*
Компьтерный набор
и верстка *А.Н. Оразалиева*

Подписано в печать 18 .11. 2009 г.
Тираж 500 экз. Формат 60x84 1/16. Бумага типографская №1.
Объем 8,7 уч.– изд. л. Заказ № 623 . Цена договорная

Издание Казахского национального технического университета
имени К.И. Сатпаева
Издательский центр КазНТУ,
Алматы, Ладыгина, 32