

ВЫСШЕЕ ОБРАЗОВАНИЕ

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ЦИФРОВОЙ ОБРАБОТКИ И ПРЕДСТАВЛЕНИЯ СИГНАЛОВ



С.В. Умняшкин

С.В. Умняшкин

Теоретические основы цифровой обработки и представления сигналов

Допущено учебно-методическим объединением вузов
по университетскому политехническому образованию
в качестве учебного пособия
для студентов высших учебных заведений,
обучающихся по направлению 230100
«Информатика и вычислительная техника»

Москва
ИД «ФОРУМ» — ИНФРА-М
2009

УДК 519.6:621.391(075.8)

ББК 32.81я73

У54

Р е ц е н з е н т ы:

доктор технических наук, профессор *В.П. Дворкович*;
доктор физико-математических наук, профессор *В.А. Кастанов*

Умняшкин С. В.

У54 Теоретические основы цифровой обработки и представления сигналов: учеб. пособие. — М.: ИД «ФОРУМ»: ИНФРА-М, 2009. — 304 с. — (Высшее образование).

ISBN 978-5-8199-0357-5 (ИД «ФОРУМ»)
ISBN 978-5-16-003277-1 (ИНФРА-М)

Учебное пособие предназначено для студентов, обучающихся по направлению «Информатика и вычислительная техника» и специальности «Прикладная математика», включает в себя общие теоретические вопросы, связанные с цифровым представлением сигналов, основами анализа линейных дискретных систем. Значительное внимание удалено вопросам эффективного представления информации (скатия данных) и использования вейвлет-преобразований.

Предназначено для студентов вузов, может быть рекомендовано в качестве дополнительного материала для направлений радиотехнического и телекоммуникационного профиля.

УДК 519.6:621.391(075.8)

ББК 32.81я73

Рукопись подготовлена автором в рамках реализации приоритетного национального проекта «Образование» и при поддержке гранта Президента Российской Федерации МД-3452.2005.9.

ISBN 978-5-8199-0357-5 (ИД «ФОРУМ»)
ISBN 978-5-16-003277-1 (ИНФРА-М)

© Умняшкин С.В., 2008
© ИД «ФОРУМ», 2008

Оглавление

Предисловие.....	5
Глава 1. Элементы функционального анализа	7
1.1. Линейные нормированные пространства	7
1.2. Анализ в линейных нормированных пространствах	10
1.3. Банаховы пространства.....	13
1.4. Пространства со скалярным произведением	15
1.5. Аппроксимация в гильбертовом пространстве	18
1.6. Примеры ортогональных систем в пространстве L_2	25
Глава 2. Спектральное представление функций	33
2.1. Тригонометрические ряды Фурье. Интеграл Фурье.....	33
2.2. Обобщенное преобразование Фурье.....	40
2.3. Принцип неопределенности времени-частотного представления сигналов	44
2.4. Энергетический спектр. Спектр мощности.....	52
Глава 3. Дискретизация и квантование сигналов. Дискретные ортогональные преобразования.....	57
3.1. Преобразование непрерывных сигналов в дискретные.....	57
3.2. Дискретизация по критерию наибольшего отклонения	58
3.3. Частотный критерий выбора шага дискретизации	60
3.4. Спектр дискретного сигнала.....	66
3.5. Дискретное преобразование Фурье	71
3.6. Быстрое преобразование Фурье (БПФ). Алгоритм БПФ с прореживанием по времени	79
3.7. Алгоритм БПФ с прореживанием по частоте.....	93
3.8. Дискретное преобразование Уолша.....	98
3.9. Дискретное преобразование Хаара	101
3.10. Некоторые применения дискретных ортогональных преобразований	103
3.11. Квантование дискретных сигналов	106
Глава 4. Линейные дискретные системы	116
4.1. Z-преобразование	116
4.2. Линейные дискретные фильтры (ЛДФ).....	123
4.3. Соединения и структурные схемы фильтров	128
4.4. Устойчивость ЛДФ	133
4.5. Частотная характеристика ЛДФ	136
4.6. Синтез КИХ-фильтров по частотной характеристике	141
4.7. Нахождение отклика фильтра с использованием БПФ	146
4.8. Согласованный дискретный фильтр	149

Глава 5. Основы прикладной теории информации	155
5.1. Дискретный источник сообщений без памяти, количество информации. Энтропия.....	155
5.2. Основные теоремы о кодировании источника без памяти	162
5.3. Эффективное кодирование дискретного источника без памяти по методам Шэннона — Фано и Хаффмана.....	172
5.4. Кодирование длин серий	178
5.5. Арифметическое кодирование.....	180
5.6. Условная энтропия.....	189
5.7. Кодирование дискретного источника с памятью	193
5.8. Статистическое моделирование источника	201
5.9. Неопределенность непрерывного источника сообщений. Дифференциальная энтропия	202
5.10. Словарные методы кодирования	206
Глава 6. Теоретические основы применения ортогональных преобразований для представления дискретных сигналов	212
6.1. Корреляция как мера статистической зависимости данных. Преобразование Карунена — Лоэва	212
6.2. Эффективность использования дискретных ортогональных преобразований для кодирования коррелированных данных.....	216
6.3. ДПФ в вещественной форме. Дискретное преобразование Хартли	221
6.4. Дискретный марковский процесс первого порядка. Дискретное косинусное преобразование (ДКП).....	224
6.5. Компрессия изображений на основе двумерного ДКП.....	230
6.6. Аппроксимационный подход к выбору преобразований для кодирования дискретных сигналов. Частотная трактовка	236
6.7. Время-частотный анализ. Оконное преобразование Фурье	241
Глава 7. Вейвлет-преобразования и их приложения для обработки дискретных сигналов.....	246
7.1. Кратно-масштабный анализ	246
7.2. Проектирование функций на подпространства КМА	253
7.3. Вычисление вейвлет-преобразований	260
7.4. Квадратурно-зеркальные фильтры (КЗФ).....	265
7.5. Свойства КЗФ	270
7.6. Построение масштабирующих функций и вейвлетов по масштабирующим уравнениям.....	278
7.7. Примеры синтеза вейвлетов	283
7.8. Биортогональные вейвлет-преобразования	290
7.9. Применение вейвлет-преобразований для сжатия сигналов	292
7.10. Двумерные дискретные вейвлет-преобразования	297
Список литературы.....	301

Предисловие

В последние годы методы цифровой обработки сигналов (ЦОС) в радиотехнике, радиоэлектронике, системах связи, контроля и управления активно вытесняют методы аналоговой обработки. Этому способствует стремительно увеличивающаяся производительность вычислительной техники, которая проникает во все области человеческой деятельности. Важность изучения методов ЦОС трудно переоценить. На сегодняшний день вопросы, связанные с цифровой обработкой сигналов и кодированием информации, перестали быть узкоспециальными. Основы знаний в данной области требуются большинству инженеров, а для специалистов в области электронной техники и информационных технологий необходимо более глубокое понимание основных методов ЦОС и математической теории, лежащей в их основе. Предлагаемое вниманию читателя учебное пособие направлено на изучение этих вопросов.

Пособие предназначено в первую очередь для студентов, обучающихся по направлению «Информатика и вычислительная техника», а также по специальности «Прикладная математика». Однако при написании пособия автор старался опираться лишь на общий курс высшей математики, читаемый для всех инженерных специальностей. Поэтому в первую главу включен материал, большая часть которого должна быть хорошо знакома студентам, обучающимся по специальности «Прикладная математика» и прослушавшим курс функционального анализа. Во второй главе в качестве напоминания приведены краткие сведения об интеграле Фурье и его свойствах, на которые в дальнейшем будут производиться многочисленные ссылки. Кроме того, в первой главе вводятся популярные в ЦОС функциональные системы Уолша и Хаара; во второй главе рассматриваются обобщение интегрального преобразования Фурье, а также важный для спектрального анализа принцип неопределенности времени-частотного представления сигналов.

Третья глава посвящена вопросам дискретизации непрерывных сигналов и преобразований, прежде всего, преобразования

Фурье. Значительное внимание уделено частотным вопросам дискретизации и построению алгоритмов быстрого преобразования Фурье (БПФ); рассматривается квантование сигналов.

В четвертой главе вводятся основные понятия теории линейных дискретных систем (фильтров). Основное внимание уделено анализу фильтров. Также рассматриваются некоторые практические вопросы синтеза фильтров с конечной импульсной характеристикой, вычисления отклика фильтров с помощью БПФ, согласованной фильтрации.

Пятая глава представляет собой введение в теорию информации и помимо теоретических сведений содержит также описание методов эффективного кодирования, которые используются в современных алгоритмах сжатия данных.

В шестой и седьмой главах представлено более углубленное изложение разделов ЦОС, связанных с эффективным представлением сигналов. Фактически, здесь изучается тот теоретический базис, который составляет основу большинства современных цифровых методов сжатия аудио- и видеосигналов. В частности, седьмая глава целиком посвящена дискретным вейвлет-преобразованиям.

При изложении материала пособия в ссылках на формулы, теоремы используется двойная нумерация. Например, ссылка (3.1) означает формулу (1) главы 3. При нумерации (обозначении) аксиом и свойств используется значок « \diamond », например 1 \diamond . Начало и окончание доказательств теорем, решений примеров обозначаются соответственно символами \blacktriangleleft и \blacktriangleright .

Основу данного учебного пособия составляет курс лекций, читаемый автором на протяжении нескольких лет в МИЭТе — Московском государственном институте электронной техники (техническом университете), <http://www.miet.ru>. Большое значение в работе над книгой имело обсуждение содержания рукописи с коллегами — сотрудниками МИЭТА: доцентом В. В. Лесиным, профессорами А. С. Поступовым, Д. В. Незлиным и профессором МИЭМа В. М. Хаметовым. Автор выражает глубокую признательность рецензентам, внимательно ознакомившимся с текстом рукописи и высказавшим ряд ценных замечаний, которые были учтены при подготовке учебного пособия в печать.

Глава 1

Элементы функционального анализа

Функциональный анализ — раздел математики, который представляет собой абстрактное обобщение линейной алгебры и математического анализа. Рассмотрим понятия и методы функционального анализа, которые наиболее важны для теории обработки сигналов.

■ 1.1. Линейные нормированные пространства

Определение. Множество E элементов произвольной природы называется *линейным пространством*, если в нем определены операции сложения элементов $x + y$ и умножения элементов на скаляр λ (вещественное или комплексное число) λx , результатом которых является элемент из того же множества E , причем выполняются следующие аксиомы:

- 1°. $\forall x \in E, \forall y \in E: x + y = y + x$.
- 2°. $\forall x \in E, \forall y \in E, \forall z \in E: (x + y) + z = x + (y + z)$.
- 3°. $\exists \theta \in E, \forall x \in E: \theta + x = x$
(существование нулевого элемента θ).
- 4°. $\forall x \in E, \forall \lambda, \mu (\lambda, \mu — \text{скаляры}): \lambda(\mu x) = (\lambda\mu)x$.
- 5°. Умножение на скаляры $\lambda = 0$ и $\lambda = 1$: $0x = \theta, 1x = x$.
- 6°. $\forall x \in E, \forall \lambda, \mu: (\lambda + \mu)x = \lambda x + \mu x$.
- 7°. $\forall x \in E, \forall y \in E, \forall \lambda: \lambda(x + y) = \lambda x + \lambda y$.

Назовем *противоположным* элементом для $x \in E$ такой элемент $y \in E$, что $x + y = \theta$. Из аксиом 5° и 6° видим, что $y = (-1)x$ (элемент x , умноженный на число -1). Будем обозначать противоположный элемент: $-x$.

В курсе линейной алгебры изучаются линейные пространства \mathbb{R}^n арифметических векторов размерности n . Приведенное выше аксиоматическое определение обобщает понятие линейного пространства \mathbb{R}^n на множества произвольной природы. Аналогично

этому элементы любого линейного пространства будем называть *векторами*, а сами линейные пространства — *векторными пространствами*. В том же обычном смысле будем понимать термины «базис пространства», «линейная зависимость (независимость)» векторов и «размерность пространства». Напомним, что число n называется размерностью векторного пространства E (обозначается $n = \dim E$), если в E найдется n линейно независимых ненулевых элементов, а любые ненулевые ($n + 1$) элементов пространства E являются линейно зависимыми. Линейное пространство может иметь бесконечную размерность.

Пример 1.1. Пусть $C[a; b]$ — множество всех функций, непрерывных на отрезке $[a; b]$. Является ли это множество линейным пространством и если да, то какова его размерность?

◀ Выполнение аксиом 1°–7° очевидно, нулевым элементом является функция $f(x) = 0, \forall x \in [a; b]$. Покажем, что $C[a; b]$ — бесконечномерное пространство. Выберем из множества $C[a; b]$ n ненулевых элементов — функций $y_i(x) = x^{i-1}, i = 1, 2, \dots, n$. При любом числе n эти элементы являются линейно независимыми,

так как равенство нулю многочлена $\sum_{i=1}^n \lambda_i y_i = \sum_{i=1}^n \lambda_i x^{i-1} = 0$ для

всех точек отрезка $x \in [a; b]$ возможно лишь в случае $\lambda_i = 0, i = 1, \dots, n$. Поскольку число n можно выбрать как угодно большим, то $\dim(C[a; b]) = \infty$. ►

Лемма 1.1 (неравенство Минковского для интегралов). Пусть для $p \geq 1$ существуют интегралы $\int_a^b |u(x)|^p dx, \int_a^b |v(x)|^p dx$ (пределы интегрирования — не обязательно конечные). Тогда существует также интеграл $\int_a^b |u(x) + v(x)|^p dx$, причем верна оценка:

$$\left(\int_a^b |u(x) + v(x)|^p dx \right)^{1/p} \leq \left(\int_a^b |u(x)|^p dx \right)^{1/p} + \left(\int_a^b |v(x)|^p dx \right)^{1/p}. \quad (1.1)$$

Опустим доказательство леммы, которое носит технический характер¹.

¹ См.: Ефимов А.В., Золотарев Ю.Г., Терпигорева В.М. Математический анализ (специальные разделы). М., 1980.

Определение. Линейное пространство E называется *нормированным*, если каждому элементу $x \in E$ поставлено в соответствие вещественное число $\|x\|$, называемое *нормой*, для которой выполняются следующие аксиомы:

1°. Невырожденность нормы:

$$\forall x \in E \quad \|x\| \geq 0, \quad \|x\| = 0 \Leftrightarrow x = 0.$$

2°. Однородность нормы: $\forall \lambda \in \mathbb{R}, \forall x \in E \quad \|\lambda x\| = |\lambda| \cdot \|x\|$.

3°. Неравенство треугольника: $\forall x, y \in E \quad \|x + y\| \leq \|x\| + \|y\|$.

В одном и том же векторном пространстве E норму можно вводить различными способами.

Пример 1.2. Рассмотрим векторное пространство $C[a; b]$ из предыдущего примера. Докажите, что приводимые ниже способы вычисления нормы удовлетворяют аксиомам 1°–3°:

а) $\|x\| = \max_{t \in [a, b]} |x(t)|, \quad$ б) $\|x\| = \left(\int_a^b |x(t)|^p dt \right)^{1/p}$, где $p \geq 1$.

Указание. В пункте б) для доказательства аксиомы треугольника следует воспользоваться неравенством Минковского.

Определение. Расстоянием между элементами x, y нормированного векторного пространства E называется число $\rho(x, y) = \|x - y\|$.

На основании аксиом нормы легко показать, что введенное расстояние между элементами обладает следующими свойствами:

1°. $\forall x, y \in E: \rho(x, y) = \rho(y, x)$.

2°. $\forall x, y \in E: \rho(x, y) = 0 \Leftrightarrow x = y$.

3°. $\forall x, y, z \in E: \rho(x, y) \leq \rho(x, z) + \rho(y, z)$ (неравенство треугольника).

◀ Первое и второе свойства очевидны. Для неравенства треугольника

$$\begin{aligned} \rho(x, y) &= \|x - y\| = \|(x - z) + (z - y)\| \leq \\ &\leq \|x - z\| + \|z - y\| = \rho(x, z) + \rho(y, z). \end{aligned}$$



Расстояние между элементами называют *метрикой* пространства. Пространство (не обязательно нормированное), каждой паре x, y элементов которого поставлено в соответствие ве-

щественное число $\rho(x, y)$ (расстояние), обладающее свойствами 1° – 3° , называется *метрическим*¹.

Определения. В метрическом пространстве E *открытым шаром* радиуса $r > 0$ с центром $x_0 \in E$ называется множество $S_r(x_0) = \{x \in E | \rho(x, x_0) < r\}$; *замкнутым шаром* — множество $\bar{S}_r(x_0) = \{x \in E | \rho(x, x_0) \leq r\}$; *окрестностью* точки $x_0 \in E$ — открытый шар произвольного радиуса ε , т. е. $S_\varepsilon(x_0)$.

Понятия нормы, расстояния, окрестности являются исходными для построения анализа в линейных нормированных пространствах.

■ 1.2. Анализ в линейных нормированных пространствах

Определение. В линейном нормированном пространстве (ЛНП) E элемент $y \in E$ называется *пределом последовательности* $\{x_k\} \subset E$, если $\lim_{k \rightarrow \infty} \rho(y, x_k) = 0$. При этом говорят, что последовательность $\{x_k\}$ *сходится* к элементу y , и используют обозначение $\lim_{k \rightarrow \infty} x_k = y$.

Определение. Элемент a из ЛНП E называется *предельной точкой* множества $M \subset E$, если в любой окрестности a содержится хотя бы один элемент $x \in M$, $x \neq a$, т. е. $\forall r > 0 : (S_r(a) \setminus a) \cap M \neq \emptyset$.

Теорема 1.1. Для того чтобы элемент a из ЛНП E был предельной точкой множества $M \subset E$, необходимо и достаточно существование последовательности $\{x_k\} \subset M$, $x_k \neq a$, сходящейся к a : $\lim_{k \rightarrow \infty} x_k = a$.

◀ **Необходимость.** Возьмем сходящуюся к нулю числовую последовательность из ненулевых элементов, например $\varepsilon_k = 1/k$, $k = 1, 2, \dots$. Так как a — предельная точка M , то, по опре-

¹ Несмотря на то что метрическое пространство может и не быть нормированным, мы будем рассматривать только нормированные метрические пространства с метрикой $\rho(x, y) = \|x - y\|$.

делению, $\forall \varepsilon_k > 0 \quad \exists x_k \in M, x_k \neq a : x_k \in S_{\varepsilon_k}(a)$. Поскольку $\rho(x_k, a) = \|x_k - a\| < \varepsilon_k = 1/k$ и $\lim_{k \rightarrow \infty} \|x_k - a\| = 0$, то так построенная последовательность $\{x_k\}$ сходится к точке a : $\lim_{k \rightarrow \infty} x_k = a$.

Достаточность. Так как $\exists \{x_k\}, x_k \in M, x_k \neq a$, причем $\lim_{k \rightarrow \infty} x_k = a$, то, по определению предела, $\lim_{k \rightarrow \infty} \|x_k - a\| = 0$ и $\forall \varepsilon > 0 \exists N(\varepsilon) : \forall n > N(\varepsilon) \|x_n - a\| < \varepsilon$. То есть в любой ε -окрестности точки a содержатся элементы $x_n \in M, x_n \neq a$, поэтому точка a является предельной для множества M . ►

Определение. Пусть M — подмножество в ЛНП E , а M' — множество всех предельных точек M . Объединение множеств $\bar{M} = M \cup M'$ называется *замыканием* множества M . Если M содержит все свои предельные точки, т. е. $M' \subset M$, то множество M называется *замкнутым*.

Определение. Множество $M \subset E$ векторного пространства E называется *линейным многообразием*, если $\forall x, y \in M, \forall \lambda, \mu : (\lambda x + \mu y) \in M$.

Определение. Замкнутое линейное многообразие L в ЛНП E , $L \subset E$, называется *подпространством*.

Определение. *Расстоянием от точки x из ЛНП E до множества $L \subset E$* называется величина $\rho(x, L) = \inf_{u \in L} \|x - u\|$.

Для ограниченного снизу числового множества всегда найдется точная нижняя грань. Поскольку норма неотрицательна, то расстояние от точки до подмножества (подпространства) всегда существует. Расстояние $\rho(x, L)$ характеризует наилучшее приближение (т. е. *аппроксимацию*) элемента $x \in E$ элементами подмножества $L \subset E$.

Определение. Элемент $y \in L$, где L — подпространство из ЛНП E , называется *элементом наилучшего приближения* (ЭНП) для заданного элемента $x \in E$, если $\rho(x, L) = \|x - y\|$.

ЭНП существует не всегда и может быть не единственным.

Пример 1.3. Рассмотрим пространство \mathbb{R}^2 , т. е. множество упорядоченных пар вещественных чисел $x = (\xi_1, \xi_2)$, где $\xi_1 \in \mathbb{R}$, $\xi_2 \in \mathbb{R}$.

Введем норму: $\|x\| = |\xi_1| + |\xi_2|$ (убедитесь, что аксиомы нормы выполняются). Рассмотрим подмножество $L \subset \mathbb{R}^2$,

$$L = \{(\xi_1, \xi_2) \mid \xi_1 = \xi_2\} = \{(\alpha, \alpha) \mid \alpha \in \mathbb{R}\}.$$

Тогда:

1. L — подпространство в E ;

2. для $x = (-1, 1)$ имеем $\rho(x, L) = 2$, причем ЭНП — не единственный.

◀ 1. Множество L является линейным многообразием (убедитесь самостоятельно). Покажем, что L — замкнуто. Допустим противное: пусть существует элемент $y \notin L$, т. е. $y = (\beta_1, \beta_2)$, $\beta_1 \neq \beta_2$, который является предельной точкой множества L . Тогда для любой точки $u = (\alpha, \alpha)$ из множества L расстояние

$$\rho(y, u) = |\beta_1 - \alpha| + |\beta_2 - \alpha| \geq |\beta_1 - \beta_2| = r(y) > 0,$$

т.е. ограничено снизу положительной величиной $r = r(y)$. Следовательно, в окрестности $S_r(y)$ нет ни одного элемента из множества L , и произвольно выбранная точка $y \notin L$ не является предельной для L . Поэтому все предельные точки множества L могут содержаться только в самом этом множестве и L является замкнутым линейным многообразием (подпространством) в \mathbb{R}^2 .

$$2. \text{ Рассмотрим функцию } f(t) = |t+1| + |t-1| = \begin{cases} -2t, & t < -1; \\ 2, & -1 \leq t \leq 1; \\ 2t, & t > 1. \end{cases}$$

Очевидно, что $\inf_{t \in \mathbb{R}} f(t) = \min_{t \in \mathbb{R}} f(t) = 2$. Для расстояния от точки $x = (-1, 1)$ до подпространства L имеем:

$$\rho(x, L) = \inf_{u \in L} \|u - x\| = \inf_{\alpha \in \mathbb{R}} (|\alpha + 1| + |\alpha - 1|) = \inf_{\alpha \in \mathbb{R}} f(\alpha) = 2.$$

При этом элементами наилучшего приближения для x являются все точки отрезка $L^* = \{(\alpha, \alpha) \mid -1 \leq \alpha \leq 1\}$, $L^* \subset L$. ▶

■ 1.3. Банаховы пространства

Определение. Пусть X — ЛНП. Последовательность $\{x_n\} \subset X$ называется *фундаментальной*, если $\forall \varepsilon > 0 \exists N = N(\varepsilon): \forall n > N, \forall p \in \mathbb{N} \quad \|x_{n+p} - x_n\| < \varepsilon$. (\mathbb{N} — множество натуральных чисел.)

Напомним, что для случая $X = \mathbb{R}$ (множество действительных чисел) в курсе математического анализа был доказан *критерий Коши*: числовая последовательность $\{x_n\}$ сходится тогда и только тогда, когда она фундаментальна. Справедлив ли критерий Коши в произвольном ЛНП?

Лемма 1.2. Всякая сходящаяся в ЛНП X последовательность $\{x_n\}$ — фундаментальна.

◀ Так как последовательность $\{x_n\}$ сходится, то $\exists \lim_{n \rightarrow \infty} x_n = x$,

$x \in X$, и $\forall \varepsilon > 0 \exists N = N(\varepsilon), \forall n > N: \|x_n - x\| < \frac{\varepsilon}{2}$. Тогда также $\forall p \in \mathbb{N}:$

$$\|x_{n+p} - x\| < \frac{\varepsilon}{2}, \text{ поэтому}$$

$$\|x_{n+p} - x_n\| = \|x_{n+p} - x + x - x_n\| \leq \|x_{n+p} - x\| + \|x_n - x\| < \varepsilon,$$

где число $\varepsilon > 0$ может быть выбрано как угодно малым. Следовательно, последовательность $\{x_n\} \subset X$ является фундаментальной. ►

Задание. Покажите, что если последовательность $\{x_n\}$ — фундаментальна, то последовательность $\{\lambda x_n\}$ также фундаментальна.

Возникают вопросы: всякая ли фундаментальная последовательность $\{x_n\} \subset X$ сходится в произвольном ЛНП X ; для каждой ли фундаментальной последовательности существует предел $\lim_{n \rightarrow \infty} x_n = x \in X$?

Определение. ЛНП называется *полным*, если в нем сходится любая фундаментальная последовательность. Полное ЛНП называется *банаховым* (или пространством Банаха).

Пример 1.4. Простейший пример пространства Банаха — множество вещественных чисел \mathbb{R} с нормой $\|x\| = |x|$.

Пример 1.5. Пространство $\tilde{L}_2[0; T]$ непрерывных на отрезке $t \in [0; T]$ функций с нормой $\|x\| = \sqrt{\int_0^T |x(t)|^2 dt}$.

◀ Покажем, что это пространство неполно. Выберем на отрезке $t \in [0; T]$ кусочно-гладкую функцию $f(t)$, имеющую разрыв первого рода. Составим для этой функции тригонометрический ряд Фурье, тогда частичные суммы ряда

$$s_n(t) = \frac{a_0}{2} + \sum_{k=1}^n \left(a_k \cos \frac{2\pi k t}{T} + b_k \sin \frac{2\pi k t}{T} \right)$$

(непрерывные функции) будут сходиться в среднеквадратичном смысле к функции $f(t)$:

$$\lim_{n \rightarrow \infty} \sqrt{\int_0^T |f(t) - s_n(t)|^2 dt} = \lim_{n \rightarrow \infty} \|f - s_n\| = 0,$$

где $\{s_n\} \subset \tilde{L}_2[0; T]$, а $f \notin \tilde{L}_2[0; T]$. Это означает, что последовательность $\{s_n\}$ — фундаментальная в $\tilde{L}_2[0; T]$ (доказательство данного утверждения проводится аналогично схеме доказательства леммы 1.2). Однако в силу единственности предела последовательность $\{s_n\}$ не может сходиться к элементу пространства $\tilde{L}_2[0; T]$, так как выбранная нами функция $f(t)$ — разрывная. Отсюда следует, что пространство $\tilde{L}_2[0; T]$ не является полным. ►

Определения. Пусть X — ЛНП (не обязательно банахово), а $\{x_n\}$ — некоторая последовательность, $\{x_n\} \subset X$. Формально состав-

ленная сумма $\sum_{k=1}^{\infty} x_k$ называется *рядом* в X , а элемент $s_n = \sum_{k=1}^n x_k$ —

n-й частичной суммой ряда (заметим, что $\forall n: s_n \in X$, см. определение ЛНП). Ряд называется *сходящимся* по норме ЛНП X , если в X сходится последовательность элементов $\{s_n\}$, т. е. $\exists \lim_{n \rightarrow \infty} s_n = s \in X$.

Элемент s называется *суммой ряда*, а запись $s = \sum_{k=1}^{\infty} x_k$ означает, что ряд сходится по норме X и его сумма равна s .

■ 1.4. Пространства со скалярным произведением

Определение. Линейное пространство E называется *евклидовым*, если каждой паре его элементов $x, y \in E$ поставлено в соответствие вещественное число $\langle x, y \rangle$, называемое *скалярным произведением*, причем выполняются следующие аксиомы:

- 1°. $\forall x \in E: \langle x, x \rangle \geq 0$, причем $\langle x, x \rangle = 0 \Leftrightarrow x = \theta$.
- 2°. $\forall x, y \in E: \langle x, y \rangle = \langle y, x \rangle$.
- 3°. $\forall x, y \in E, \forall \lambda \in \mathbb{R}: \langle \lambda x, y \rangle = \lambda \langle x, y \rangle$.
- 4°. $\forall x, y, z \in E: \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

Заметим, что в данном определении ничего не говорится о нормированности пространства E . Однако евклидово пространство можно превратить в нормированное, если ввести норму следующим образом:

$$\|x\| = \sqrt{\langle x, x \rangle}. \quad (1.2)$$

Аксиомы нормы 1° и 2° при этом выполняются очевидным образом. Для доказательства выполнения аксиомы 3° (неравенства треугольника) предварительно рассмотрим следующую лемму.

Лемма 1.3. Норма, введенная в соответствии с определением (1.2), удовлетворяет неравенству Коши — Буняковского (или Шварца):

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

◀ Заметим, что $\forall \lambda \in \mathbb{R}: \langle x - \lambda y, x - \lambda y \rangle = \|x - \lambda y\|^2 \geq 0$. Поэтому

$$\begin{aligned} 0 &\leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle = \\ &= \|x\|^2 - 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2. \end{aligned}$$

Тогда дискриминант полученного квадратного трехчлена переменной λ : $4(\langle x, y \rangle)^2 - 4\|x\|^2 \cdot \|y\|^2 \leq 0$, что и доказывает неравенство Коши — Буняковского. ►

Докажем выполнение аксиомы треугольника. Так как

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2,$$

то, применяя к последнему выражению результат леммы 1.3, получаем:

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\|\cdot\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

или $\|x + y\| \leq \|x\| + \|y\|$.

Определения. Пусть E — линейное пространство с введенным скалярным произведением. *Ортогональными элементами* пространства E называются элементы $x, y \in E$, если $\langle x, y \rangle = 0$. Ортогональность элементов обозначается $x \perp y$. Очевидно, что нулевой элемент ортогонален всем элементам пространства. *Ортогональной системой* в E называется множество попарно ортогональных элементов $\{x_n\} \subset E$.

Теорема 1.2. Если $\{x_k\}_{k=1}^m$ — ортогональная система ненулевых элементов в евклидовом пространстве E , $\{x_k\}_{k=1}^m \subset E$, то элементы $\{x_k\}_{k=1}^m$ — линейно независимы.

◀ Допустим противное. Пусть элементы $\{x_k\}_{k=1}^m$ — линейно зависимы, т. е. существует такой набор чисел $\{\lambda_k\}_{k=1}^m$ (не все из них равны нулю), что $\sum_{k=1}^m \lambda_k x_k = \theta$. В силу ортогональности системы $\{x_k\}_{k=1}^m$ имеем $\forall j = 1, \dots, m$:

$$0 = \langle x_j, \theta \rangle = \left\langle x_j, \sum_{k=1}^m \lambda_k x_k \right\rangle = \sum_{k=1}^m \lambda_k \underbrace{\langle x_j, x_k \rangle}_{\neq 0} = \lambda_j \langle x_j, x_j \rangle.$$

Поэтому все коэффициенты $\{\lambda_k\}_{k=1}^m$ должны быть нулевыми, а это противоречит допущению о линейной зависимости элементов $\{x_k\}_{k=1}^m$. ►

Следствие. В n -мерном евклидовом пространстве ортогональная система из n ненулевых элементов образует базис.

В дальнейшем нам понадобятся два свойства скалярного произведения, которые устанавливаются в следующих леммах.

Лемма 1.4 (свойство непрерывности скалярного произведения). Пусть в евклидовом пространстве E заданы две сходящиеся последовательности: $\{x_n\} \subset E$, $\lim_{n \rightarrow \infty} x_n = x \in E$; $\{y_n\} \subset E$, $\lim_{n \rightarrow \infty} y_n = y \in E$.

Тогда числовая последовательность $\langle x_n, y_n \rangle$ также сходится и $\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle x, y \rangle$.

$$\begin{aligned} \blacktriangleleft | \langle x_n, y_n \rangle - \langle x, y \rangle | &= | \langle x_n - x, y_n \rangle + \langle x, y_n - y \rangle | \leq \\ &\leq | \langle x_n - x, y_n \rangle | + | \langle x, y_n - y \rangle | \leq \|x_n - x\| \cdot \|y_n\| + \|x\| \cdot \|y_n - y\| \end{aligned}$$

(последний переход осуществлен на основании леммы 1.3). Выражение в правой части неравенства стремится к нулю при $n \rightarrow \infty$. Действительно, сходящаяся числовая последовательность $\{\|y_n\|\}$ ограничена, кроме того, $\lim_{n \rightarrow \infty} \|x_n - x\| = \lim_{n \rightarrow \infty} \|y_n - y\| = 0$. Следовательно, $\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle x, y \rangle$. ►

Лемма 1.5 (равенство параллелограмма). Для любых элементов x, y евклидова пространства E и нормы (1.2) верно:

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= 2\|x\|^2 + 2\|y\|^2. \\ \blacktriangleleft \|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle = \dots \\ &= 2\|x\|^2 + 2\|y\|^2. \end{aligned}$$

Проделайте опущенные выкладки самостоятельно. ►

Определение. Пространством Гильберта (обычно обозначается H) называется евклидово пространство, которое полно в норме (1.2).

Пример 1.6. Пространство E^n арифметических векторов со скалярным произведением, определенным для векторов $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$ как $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^n x_k y_k$ — полное, т. е. гильбертово.

Пример 1.7. Гильбертово пространство $L_2[a; b]$.

◀ В примере 1.5 было рассмотрено пространство $\tilde{L}_2[0; T]$ непрерывных на отрезке $t \in [0; T]$ функций с нормой

$\|x\| = \sqrt{\int_0^T |x(t)|^2 dt}$ и было показано, что $\tilde{L}_2[0; T]$ не является полным. Можно также показать, что не является полным и пространство $\hat{L}_2[0; T]$ кусочно-непрерывных на отрезке $t \in [0; T]$ вещественных функций с нормой, определяемой тем же выражением.

Во многих теоретических вопросах рассматривают обобщение пространств $\tilde{L}_2[a; b]$ и $\hat{L}_2[a; b]$ — пространство $L_2[a; b]$ функций, для которых норма элемента определяется как $\|x\| = \left(\int_a^b |x(t)|^2 dt \right)^{1/2}$, но интеграл понимается в смысле Лебега.

Определенный интеграл Лебега представляет собой обобщение «традиционного» интеграла Римана и применим к более широкому классу функций. Теория интеграла Лебега выходит за рамки материала данного пособия¹, отметим лишь, что пространство $L_2[a; b]$ является полным, а значит, гильбертовым. Кроме того, любой элемент $x \in L_2[a; b]$ можно с какой угодно точностью $\varepsilon > 0$ приблизить по норме этого пространства элементом $\hat{x} \in \hat{L}_2[a; b]$, т. е. кусочно-непрерывной функцией: $\|\hat{x} - x\| < \varepsilon$.

В тех случаях, когда полнота является существенно важным свойством, необходимо рассматривать пространство $L_2[a; b]$. На практике для описания сигналов обычно ограничиваются множеством кусочно-непрерывных функций $\hat{L}_2[a; b] \subset L_2[a; b]$. Тогда при определении скалярного произведения $\langle x, y \rangle = \int_a^b x(t)y(t)dt$ и индуцируемой им нормы (1.2) определенный интеграл можно понимать в смысле Римана. ►

■ 1.5. Апроксимация в гильбертовом пространстве

Сформулируем задачу аппроксимации, которую будем рассматривать далее. Пусть H — гильбертово пространство, а L — подпространство в H , $L \subset H$. Для заданного элемента $x \in H$ необходимо найти элемент наилучшего приближения (ЭНП) $y \in L$, для которого $\rho(x, y) = \rho(x, L)$, т. е.

$$\|x - y\| = \inf_{u \in L} \|x - u\|. \quad (1.3)$$

Теорема 1.3. В гильбертовом пространстве существует, и притом единственный, ЭНП $y \in L$, который является решением задачи аппроксимации (1.3).

◀ Докажем сначала существование ЭНП. Обозначим $d = \inf_{u \in L} \|x - u\|$. Из определения точной нижней грани следует, что $\forall \varepsilon > 0 \exists u_\varepsilon \in L: d \leq \|x - u_\varepsilon\| < d + \varepsilon$. Тогда взяв числовую последовательность $\varepsilon_k = 1/k$, $k = 1, 2, \dots$, сможем построить последовательность элементов $\{u_k\} \subset L$ такую, что

$$d \leq \|x - u_k\| < d + \frac{1}{k}.$$

Покажем, что $\{u_k\}$ — фундаментальная последовательность. С использованием равенства параллелограмма (лемма 1.5) имеем

$$\begin{aligned} 2\|x - u_n\|^2 + 2\|x - u_m\|^2 &= \|u_m - u_n\|^2 + 4\left\|x - \frac{u_m + u_n}{2}\right\|^2 \geq \\ &\geq \|u_m - u_n\|^2 + 4d^2, \end{aligned}$$

поскольку элемент $v = \frac{u_m + u_n}{2} \in L$ и

$$\rho(x, v) = \|x - v\| \geq \inf_{u \in L} \|x - u\| = d.$$

Поэтому $\|u_m - u_n\|^2 \leq 2\|x - u_n\|^2 + 2\|x - u_m\|^2 - 4d^2$, и тогда

$$\begin{aligned} \|u_m - u_n\|^2 &\leq 2\left(d + \frac{1}{n}\right)^2 + 2\left(d + \frac{1}{m}\right)^2 - 4d^2 \leq 4\left(d + \frac{1}{N}\right)^2 - 4d^2 = \\ &= \frac{8d}{N} + \frac{4}{N^2} < \frac{8d + 4}{N}, \end{aligned}$$

где $N = \min(n, m)$.

¹ Подробнее см.: Треногин В.А. Функциональный анализ. М., 2003.

Величину $\|u_m - u_n\|$ можно ограничить как угодно малым значением за счет надлежащего выбора числа N , т. е. последовательность $\{u_k\}$ — фундаментальная, и вследствие полноты H $\exists \lim u_k = y \in H$. А поскольку сходящаяся последовательность $\{u_k\} \subset L$ и L — подпространство (т. е. замкнутое множество), то верно также: $y \in L$. Поэтому $\rho(x, y) = d$ и существование ЭНП доказано.

Покажем, что ЭНП y — единственный. Для этого допустим противное. Пусть наряду с y существует также другой ЭНП $\tilde{y} \in L$, т. е. $\rho(x, L) = \|x - y\| = \|x - \tilde{y}\| = d$, причем $\tilde{y} \neq y$. На основании равенства параллелограмма (лемма 1.5) получим:

$$\begin{aligned} 4d^2 &= 2\|x - y\|^2 + 2\|x - \tilde{y}\|^2 = \|y - \tilde{y}\|^2 + 4\left\|x - \frac{y + \tilde{y}}{2}\right\|^2 \geqslant \\ &\geqslant \|y - \tilde{y}\|^2 + 4d^2, \end{aligned}$$

отсюда $\|y - \tilde{y}\|^2 = 0$ и $y = \tilde{y}$, т. е. ЭНП — единственный. ►

Теорема 1.4. Пусть L — подпространство в гильбертовом пространстве H , $y \in L$ — ЭНП для заданного элемента $x \in H$. Тогда любой элемент $u \in L$ ортогонален элементу $v = x - y$: $v \perp u$, что обозначают также $v \perp L$.

◀ Допустим противное, т. е. $\exists u \in L: \langle x - y, u \rangle = \sigma \neq 0$. Тогда $u \neq 0$, и (см. аксиому 1° скалярного произведения) $\langle u, u \rangle > 0$. Рассмотрим элемент $\tilde{y} = y + \frac{\sigma}{\langle u, u \rangle} u$, который также лежит в подпространстве L : $\tilde{y} \in L$, т. к. $y \in L$, $u \in L$.

Имеем:

$$\begin{aligned} \|x - \tilde{y}\|^2 &= \left\langle \left(x - y \right) - \frac{\sigma}{\langle u, u \rangle} u, \left(x - y \right) - \frac{\sigma}{\langle u, u \rangle} u \right\rangle = \\ &= \langle x - y, x - y \rangle - 2 \left\langle x - y, \frac{\sigma}{\langle u, u \rangle} u \right\rangle + \left\langle \frac{\sigma}{\langle u, u \rangle} u, \frac{\sigma}{\langle u, u \rangle} u \right\rangle = \\ &= \|x - y\|^2 - \frac{2\sigma}{\langle u, u \rangle} \underbrace{\langle x - y, u \rangle}_{\sigma} + \frac{\sigma^2}{\langle u, u \rangle^2} \langle u, u \rangle = \|x - y\|^2 - \frac{\sigma^2}{\langle u, u \rangle}. \end{aligned}$$

Поскольку $\frac{\sigma^2}{\langle u, u \rangle} > 0$, то $\|x - \tilde{y}\|^2 < \|x - y\|^2$ и элемент y не является ЭНП. Получили противоречие, поэтому $\forall u \in L: \langle x - y, u \rangle = 0$. ►

Следствие из теорем 1.3, 1.4. Пусть L — подпространство в H . Тогда $\forall x \in H$ существует единственное разложение $x = y + z$, где $y \in L$, а $z \perp L$.

◀ Пусть $x = y + z$, ЭНП $y \in L$, $z \perp L$. Пусть существует также другое представление: $x = a + b$, где $a \in L$, $b \perp L$. Тогда $y - a + z - b = 0$, и

$$\begin{aligned} &\langle y - a + z - b, y - a \rangle = 0 = \\ &= \langle y - a, y - a \rangle + \langle z, y - a \rangle - \langle b, y - a \rangle = \|y - a\|^2, \end{aligned}$$

так как $(y - a) \in L$. Поэтому $y = a$ и $b = x - y$. ►

ЭНП $y \in L$ называют также *проекцией* элемента $x \in H$ на подпространство L . Для случая $H = E^3$, $L = E^2$ результат теоремы 1.4 хорошо известен и имеет несложную геометрическую интерпретацию (рис. 1.1).

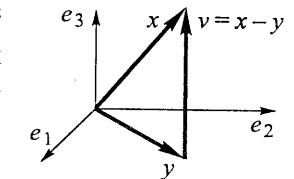


Рис. 1.1

Теорема 1.4 определяет способ нахождения ЭНП для $x \in H$ в случае конечной размерности подпространства L с заданным (не обязательно ортогональным) базисом $\{g_1, g_2, \dots, g_n\}$, $y = \sum_{j=1}^n \lambda_j g_j$. Поиск коэффициентов разложения $\{\lambda_j\}_{j=1}^n$ осуществляется следующим образом. Так как $\forall k: g_k \in L$, $\langle x - y, g_k \rangle = 0$, то

$$\left\langle x - \sum_{j=1}^n \lambda_j g_j, g_k \right\rangle = \langle x, g_k \rangle - \sum_{j=1}^n \lambda_j \langle g_j, g_k \rangle = 0$$

или

$$\sum_{j=1}^n \lambda_j \langle g_j, g_k \rangle = \langle x, g_k \rangle, \quad k = 1, 2, \dots, n. \quad (1.4)$$

Определитель системы линейных уравнений (1.4) есть определитель матрицы Грама $\mathbf{G} = \{\langle g_j, g_k \rangle\}_{k,j=1}^n$, причем $\det \mathbf{G} \neq 0$ в силу линейной независимости элементов $\{g_1, g_2, \dots, g_n\}$. (Напомним, что $\det \mathbf{G} = 0$ тогда и только тогда, когда элементы $\{g_1, g_2, \dots, g_n\}$ линейно зависимы.) Следовательно, система уравнений (1.4) имеет единственное решение — набор коэффициентов $\{\lambda_j\}_{j=1}^n$, который задает ЭНП $y = \sum_{j=1}^n \lambda_j g_j$.

Если же элементы базиса подпространства $\{g_1, g_2, \dots, g_n\} \subset L$ не только линейно независимы, но и ортогональны, то поиск коэффициентов $\{\lambda_j\}_{j=1}^n$ упрощается (убедитесь самостоятельно):

$$\lambda_j = \frac{\langle x, g_j \rangle}{\langle g_j, g_j \rangle}. \quad (1.5)$$

Определение. Пусть L — подпространство в H . Совокупность всех элементов из H , ортогональных к L , $L^\perp = \{x \in H \mid x \perp L\}$, называется *ортогональным дополнением* подпространства L .

Теорема 1.5. Пусть L — подпространство в гильбертовом пространстве H . Тогда L^\perp также является подпространством в H .

◀ Докажем, что L^\perp — замкнутое линейное многообразие.

Линейность. $\forall u \in L, \forall x, y \in L^\perp \forall \alpha, \beta$ — скаляров:

$$\langle u, \alpha x + \beta y \rangle = \alpha \langle u, x \rangle + \beta \langle u, y \rangle = 0,$$

т. е. для любой линейной комбинации $z = \alpha x + \beta y$ элементов из L^\perp имеем: $z \perp L$, следовательно, $z \in L^\perp$ и L^\perp — линейное многообразие.

Замкнутость. Пусть z — произвольная предельная точка множества L^\perp . Тогда по теореме 1.1 найдется последовательность $\{z_n\} \subset L^\perp, z_n \neq z: \lim_{n \rightarrow \infty} z_n = z$. Имеем $\forall u \in L: \langle u, z_n \rangle = 0$, но в силу непрерывности скалярного произведения (лемма 1.4)

$$\lim_{n \rightarrow \infty} \langle z_n, u \rangle = \langle z, u \rangle = 0.$$

Следовательно, $z \in L^\perp$ и множество L^\perp содержит все свои предельные точки, т. е. замкнуто. ►

Определение. Будем говорить, что гильбертово пространство H разлагается в ортогональную сумму подпространств

$$L_1, L_2, \dots, L_n,$$

и записывать это как $H = L_1 \oplus L_2 \oplus \dots \oplus L_n$, если:

1) все подпространства L_1, L_2, \dots, L_n попарно ортогональны, т. е.

$$\forall u \in L_i, \forall v \in L_j \quad \langle u, v \rangle = 0, \text{ если } i \neq j;$$

$$2) \forall x \in H \text{ существует разложение } x = \sum_{i=1}^n x_i, \text{ где } x_i \in L_i.$$

Заметим, что если L — подпространство в гильбертовом пространстве H , то $H = L \oplus L^\perp$, что вытекает непосредственно из определения L^\perp и следствия из теорем 1.3–1.5.

Теорема 1.6. Пусть в гильбертовом пространстве H задано конечномерное подпространство L с ортогональным базисом $\{g_1, g_2, \dots, g_n\}$, а $y = \sum_{j=1}^n \lambda_j g_j$ — ЭНП для заданного элемента $x \in H$. Тогда для ошибки приближения — вектора $x - y$ — справедливы равенства:

$$\|x - y\|^2 = \|x\|^2 - \|y\|^2 = \|x\|^2 - \sum_{j=1}^n \lambda_j^2 \|g_j\|^2.$$

◀ Поскольку (см. теорему 1.4) $\langle x - y, y \rangle = 0$, то

$$\langle x, y \rangle = \langle y, y \rangle = \|y\|^2,$$

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \langle x, x \rangle - 2 \langle x, y \rangle + \langle y, y \rangle = \|x\|^2 - \|y\|^2,$$

причем

$$\|y\|^2 = \left\langle \sum_{j=1}^n \lambda_j g_j, \sum_{m=1}^n \lambda_m g_m \right\rangle = \sum_{j=1}^n \lambda_j \sum_{m=1}^n \lambda_m \langle g_j, g_m \rangle = \sum_{j=1}^n \lambda_j^2 \|g_j\|^2. \blacktriangleright$$

Пусть теперь в H задана бесконечная последовательность ненулевых ортогональных векторов $\{\phi_k\}_{k=1}^\infty \subset H$. Это означает, что H — бесконечномерно, так как ортогональные элементы линейно независимы. Рассматривая первые элементы $\{\phi_k\}_{k=1}^n$ как базис, получаем некоторое линейное многообразие L_n , «натянутое» на $\{\phi_k\}_{k=1}^n$. Можно показать, что L_n — замкнуто, т. е. являет-

ся подпространством. Так как L_n — конечномерно, то с учетом теоремы 1.6 для ЭНП $y_n = \sum_{j=1}^n \lambda_j \varphi_j$, $y_n \in L_n$, имеем:

$$\|x - y_n\|^2 = \|x\|^2 - \|y_n\|^2 = \|x\|^2 - \sum_{j=1}^n \lambda_j^2 \|\varphi_j\|^2. \quad (1.6)$$

Числовая последовательность $s_n = \|y_n\|^2 = \sum_{j=1}^n \lambda_j^2 \|\varphi_j\|^2$ ограничена сверху, так как $\forall n s_n = \|x\|^2 - \|x - y_n\|^2 \leq \|x\|^2$, и является неубывающей ($s_{n+1} \geq s_n$). Поэтому $\{s_n\}$ — сходится. Сходимость последовательности частичных сумм $\{s_n\}$ означает сходимость ряда

$$s = \sum_{j=1}^{\infty} \lambda_j^2 \|\varphi_j\|^2, \text{ причем (см. равенство (1.6))}$$

$$\sum_{j=1}^{\infty} \lambda_j^2 \|\varphi_j\|^2 \leq \|x\|^2. \quad (1.7)$$

Выражение (1.7) называется *неравенством Бесселя*.

Определения. Ортогональная система $\{\varphi_k\}_{k=1}^{\infty} \subset H$ называется *полной* в гильбертовом пространстве H , если $\forall x \in H$ существует разложение

$$x = \sum_{k=1}^{\infty} \lambda_k \varphi_k, \quad (1.8)$$

т. е. $\lim_{n \rightarrow \infty} \left\| x - \sum_{k=1}^n \lambda_k \varphi_k \right\| = 0$. Ряд (1.8) называется *рядом Фурье* (по ортогональной системе $\{\varphi_k\}$), а числа $\{\lambda_k\}$ — *коэффициентами Фурье*.

Теорема 1.7. Пусть $\{g_k\}_{k=1}^{\infty} \subset H$ — полная ортогональная система в гильбертовом пространстве H . Тогда $\forall x \in H$ для коэффициентов Фурье $\{\lambda_k\}_{k=1}^{\infty}$ верна формула (1.5).

◀ Обозначим частичную сумму $x_n = \sum_{k=1}^n \lambda_k g_k$. В силу непре-

рывности скалярного произведения и ортогональности системы $\{g_k\}_{k=1}^{\infty}$ имеем:

$$\begin{aligned} < x, g_j > &= \lim_{n \rightarrow \infty} < x_n, g_j > = \lim_{n \rightarrow \infty} \left\langle \sum_{k=1}^n \lambda_k g_k, g_j \right\rangle = \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \lambda_k < g_k, g_j > = \sum_{k=1}^{\infty} \lambda_k < g_k, g_j > = \lambda_j < g_j, g_j >, \end{aligned}$$

откуда следует формула (1.5). ►

Теорема 1.8. Ортогональная система $\{\varphi_k\}_{k=1}^{\infty} \subset H$ является полной в гильбертовом пространстве H тогда и только тогда, когда $\forall x \in H$ неравенство (1.7) выполняется как равенство

$$\|x\|^2 = \sum_{j=1}^{\infty} \lambda_j^2 \|\varphi_j\|^2,$$

которое называется *равенством Парсеваля — Стеклова*.

◀ Действительно, понятие полной системы $\{\varphi_k\}_{k=1}^{\infty} \subset H$ означает, что

$$\forall x \in H: x = \sum_{k=1}^{\infty} \lambda_k \varphi_k \text{ и } \left\| x - \sum_{k=1}^{\infty} \lambda_k \varphi_k \right\|^2 = 0,$$

что эквивалентно равенству $\|x\|^2 - \sum_{j=1}^{\infty} \lambda_j^2 \|\varphi_j\|^2 = 0$, которое получается предельным переходом из соотношения (1.6). ►

Важнейшим примером гильбертова пространства является пространство функций $L_2[a; b]$ (см. пример 1.7). При этом под $L_2(\mathbb{R}) = L_2(-\infty; \infty)$ будем понимать пространство всех функций, интегрируемых с квадратом на всей числовой оси.

■ 1.6. Примеры ортогональных систем в пространстве L_2

Элементами в векторном пространстве L_2 являются функции. Приведем ряд примеров ортогональных функциональных базисов $\{\varphi_k\}$, которые нашли широкое применение для обработки сигналов.

Пример 1.8. Тригонометрическая система функций

$$\left\{1, \cos \frac{2\pi kt}{T}, \sin \frac{2\pi kt}{T}\right\}_{k=1}^{\infty}$$

является полной в пространстве $L_2[a; a+T]$ на любом отрезке $t \in [a; a+T]$ длины T .

Пример 1.9. Система ортогональных многочленов Лежандра $\{P_n(t)\}_{n=0}^{\infty}$, $P_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} (t^2 - 1)^n$, является полной в пространстве $L_2[-1; 1]$.

При цифровой обработке сигналов использование степенных многочленов для представления сигналов часто бывает более предпочтительным по сравнению с тригонометрическими функциями, так как реализация вычислений последних обычно более сложна. В этой связи еще более интересны базисы кусочно-постоянных функций.

Пример 1.10. Систему функций Радемахера $\{r_k(x)\}_{k=0}^{\infty}$ определим следующим образом. Для $x \in [0; 1)$ положим

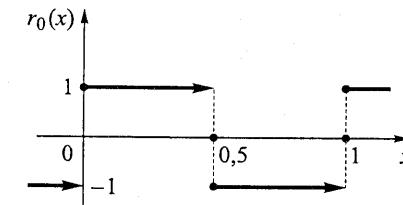
$$r_0(x) = \begin{cases} 1 & \text{при } x \in [0; 1/2]; \\ -1 & \text{при } x \in [1/2; 1) \end{cases}$$

и периодически продолжим $r_0(x)$ на всю числовую ось с периодом $T = 1$. Остальные функции системы определим так: $r_k(x) = r_0(2^k x)$, $k = 1, 2, \dots$ (рис. 1.2).

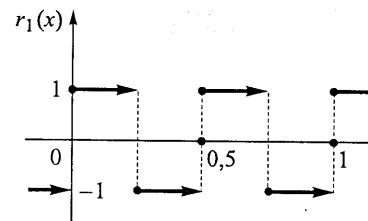
Для дальнейшего изложения удобно использовать следующее обозначение: $\Delta_m^n = \left[\frac{m}{2^n}; \frac{m+1}{2^n} \right)$, где $n \in \mathbb{Z}$, $m \in \mathbb{Z}$. Тогда из определения функций Радемахера и приведенных иллюстраций видны следующие свойства данной системы:

1°. Кусочное постоянство.

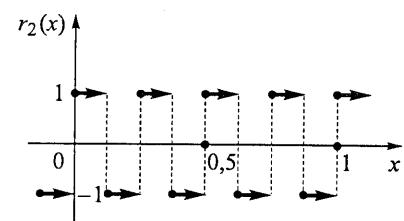
$$\forall x \in \Delta_m^{k+1}: r_k(x) = \text{const} = \begin{cases} 1, & \text{если } m \text{ — четное;} \\ -1, & \text{если } m \text{ — нечетное.} \end{cases}$$



а



б



в

Рис. 1.2

На более мелких подинтервалах, естественно, функции также постоянны:

$$\forall k \geq 0, \forall j \geq k+1, \forall x \in \Delta_m^j, m \in \mathbb{Z}: r_k(x) = \text{const}. \quad (1.9)$$

2°. Интеграл по периоду функции $r_k(x)$ равен нулю. Поэтому $\forall m \in \mathbb{Z} \int_{\Delta_m^k} r_k(x) dx = 0$ (как интеграл по одному периоду $T = |\Delta_m^k| = 2^{-k}$) и

$$\forall m \in \mathbb{Z}, \forall j = 0, \dots, k: \int_{\Delta_m^j} r_k(x) dx = 0, \quad (1.10)$$

как интеграл по N периодам, $N = |\Delta_m^j| / |\Delta_m^k| = \frac{2^{-j}}{2^{-k}} = 2^{k-j}$.

3°. Система функций $\{r_k(x)\}_{k=0}^{\infty}$ — ортонормированна на отрезке $x \in [0; 1]$.

◀ Очевидно, $\forall k: \langle r_k, r_k \rangle = \int_0^1 (r_k(x))^2 dx = 1$, т. е. функции нормированы. Покажем, что $\forall m \neq k: \langle r_k, r_m \rangle = \int_0^1 r_k(x) r_m(x) dx = 0$. Пусть для определенности $k > m$, тогда

$$\begin{aligned} \langle r_k, r_m \rangle &= \int_{\Delta_0^0} r_k(x) r_m(x) dx = \sum_{j=0}^{2^k-1} \int_{\Delta_j^k} r_k(x) \underbrace{r_m(x)}_{c(m, k, j)} dx = \\ &= \sum_{j=0}^{2^k-1} c(m, k, j) \underbrace{\int_{\Delta_j^k} r_k(x) dx}_{=0, \text{ (см. (1.10))}} = 0. \blacksquare \end{aligned}$$

Таким образом, система функций $\{r_k(x)\}_{k=0}^\infty$ является *ортонормированной*, но она не является базисом в пространстве $L_2[0; 1]$, поскольку не является полной.

Задание. Покажите по схеме, аналогичной доказательству свойства 3°, что ненулевой элемент $f(x) = r_0(x)r_1(x)$, $\|f(x)\|=1$, $f(x) \in L_2[0, 1]$ является ортогональным любой из функций Радемахера, т. е. $\forall k: \langle f, r_k \rangle = 0$. Следовательно, система $\{r_k(x)\}_{k=0}^\infty$ — неполная и не является базисом в $L_2[0; 1]$.

Пример 1.11. Систему функций Уолша $\{w_n(x)\}_{n=0}^\infty$ определим следующим образом. Представим целое число $n \geq 0$ в виде двоичного разложения: $n = \sum_{k=0}^{l(n)} n_k 2^k$, $n_k \in \{0, 1\}$. Тогда функции системы Уолша выражаются с помощью функций Радемахера следующим образом:

$$w_n(x) = \prod_{k=0}^l (r_k(x))^{n_k} = \prod_{k: n_k=1} r_k(x), \quad (1.11)$$

где конечное число $l = l(n)$ определяется номером n функции Уолша, $n < 2^l$. Таким образом, функция Уолша $w_n(x)$ определяется как произведение функций Радемахера с номерами, которые со-

ответствуют единичным коэффициентам в двоичном разложении числа n . При этом если все коэффициенты $\{n_k\}$ двоичного разложения равны нулю, то считаем последнее произведение в (1.11) равным единице, т. е. $w_0(x) = 1$. Поясним определение системы $\{w_n(x)\}_{n=0}^\infty$ построением ее первых функций (см. табл.). График функции $w_3(x)$ приведен на рис. 1.3.

n	n_2	n_1	n_0	$w_n(x)$
0	0	0	0	$w_0(x) = 1$
1	0	0	1	$w_1(x) = r_0(x)$
2	0	1	0	$w_2(x) = r_1(x)$
3	0	1	1	$w_3(x) = r_0(x)r_1(x)$
4	1	0	0	$w_4(x) = r_2(x)$

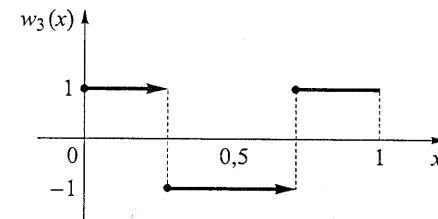


Рис. 1.3

Замечание. Очевидно, что функции системы Уолша имеют период $T = 1$.

Задание. Постройте самостоятельно графики функций

$$w_3(x), \dots, w_7(x).$$

Теорема 1.9. На полуинтервале $x \in [0; 1)$ система Уолша ортонормированна.

◀ Очевидно, что $\forall n: \langle w_n, w_n \rangle = 1$. Пусть теперь $k \neq n$:

$$\begin{aligned} \langle w_k, w_n \rangle &= \int_0^1 \left(\prod_{j: k_j=1} r_j(x) \right) \left(\prod_{j: n_j=1} r_j(x) \right) dx = \\ &= \int_0^1 \left(\prod_{j: k_j=n_j=1} (r_j(x))^2 \right) \left(\prod_{j: n_j \neq k_j} r_j(x) \right) dx = \int_0^1 \prod_{j: n_j \neq k_j} r_j(x) dx. \end{aligned}$$

Поскольку $n \neq k$ и не все коэффициенты n_j, k_j одинаковы, то в по-
дынтегральном произведении имеется, по крайней мере, один
сомножитель. Положим, что $\tilde{j} = \max_{n_j \neq k_j} j$, и продолжим преобразо-

вания:

$$\langle w_k, w_n \rangle = \int_0^1 r_{\tilde{j}}(x) \prod_{\substack{j: n_j \neq k_j \\ j < \tilde{j}}} r_j(x) dx = \sum_{m=0}^{2^{\tilde{j}}-1} \int_{\Delta_m^{\tilde{j}}} r_{\tilde{j}}(x) \underbrace{\prod_{\substack{j: n_j \neq k_j \\ j < \tilde{j}}} r_j(x)}_{\text{константа } c(k, n, m) \text{ (см. (1.9))}} dx =$$

$$\sum_{m=0}^{2^{\tilde{j}}-1} \underbrace{c(k, n, m)}_{\pm 1} \underbrace{\int_{\Delta_m^{\tilde{j}}} r_{\tilde{j}}(x) dx}_{= 0 \text{ (см. (1.10))}} = 0.$$

Теорема 1.10. Система Уолша (1.11) — полная в пространстве $L_2[0; 1]$. Примем утверждение теоремы без доказательства.

Так как функции системы Уолша принимают лишь два значения ± 1 , они очень удобны для программных вычислений и для аппаратной реализации в цифровой аппаратуре.

Пример 1.12. Систему функций Хаара $\{h_n(x)\}_{n=0}^\infty$ определим на полуинтервале $x \in [0; 1]$ следующим образом. Положим $h_0(x) = 1$. Для $n > 0$ номер базисной функции $h_n(x)$ представим следующим образом: $n = 2^k + m$, где целые числа $k \geq 0$, $0 \leq m \leq 2^k - 1$ однозначно определяются по номеру $n > 0$. Тогда

$$h_n(x) = \begin{cases} 2^{k/2} & \text{при } x \in \Delta_{2m}^{k+1}; \\ -2^{k/2} & \text{при } x \in \Delta_{2m+1}^{k+1}; \\ 0 & \text{при } x \notin \Delta_m^k = \Delta_{2m}^{k+1} \cup \Delta_{2m+1}^{k+1}. \end{cases} \quad (1.12)$$

Приведем графики первых функций системы $\{h_n(x)\}_{n=0}^\infty$, см. рис. 1.4.

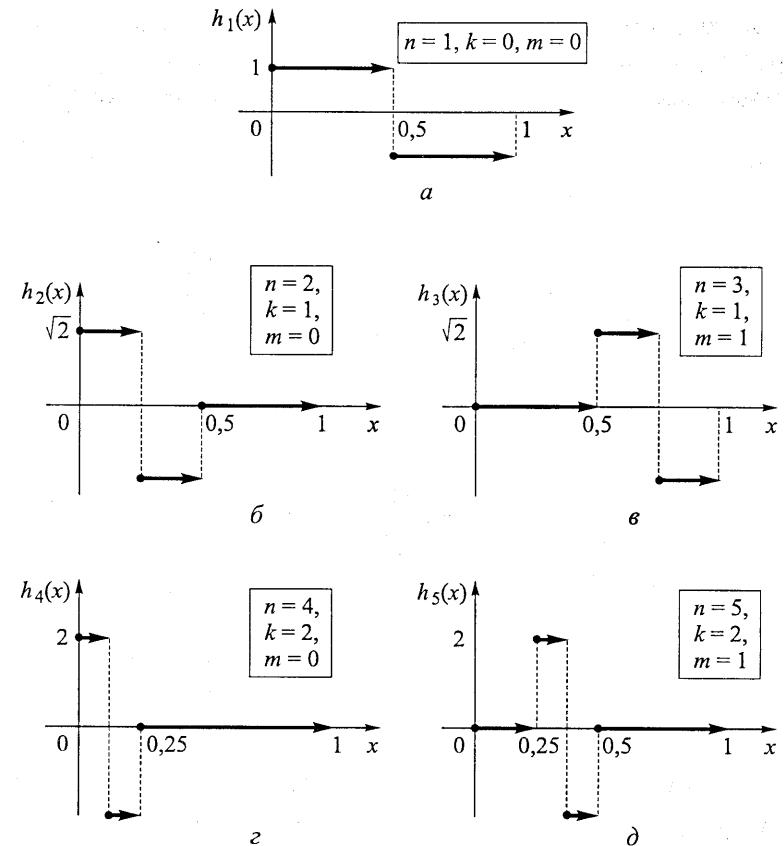


Рис. 1.4

Рассмотренным ранее свойствам системы Радемахера во многом аналогичны следующие очевидные свойства системы Хаара. Для $n > 0$, $n = 2^k + m$, $k \geq 0$, $0 \leq m \leq 2^k - 1$:

1°. $\forall j \geq k+1$, $\forall l \in \{0, 1, \dots, 2^j - 1\}$, $\forall x \in \Delta_l^j$

$$h_n(x) = \text{const} \in \{0, -2^{k/2}, 2^{k/2}\}.$$

2°. $\forall j \in \{0, \dots, k\}$, $\forall l \in \{0, 1, \dots, 2^j - 1\}$ $\int_{\Delta_l^j} h_n(x) dx = 0$.

Теорема 1.11. Система функций Хаара $\{h_n(x)\}_{n=0}^{\infty}$ — ортонормированна на интервале $x \in [0; 1]$.

◀ В соответствии с определением (1.12)

$$\langle h_n(x), h_n(x) \rangle = \int_{\Delta_m^k} 2^{k/2} 2^{k/2} dx = 1, \text{ где } n = 2^k + m.$$

Рассмотрим скалярное произведение $\langle h_n(x), h_v(x) \rangle$, где $n = 2^k + m$, $v = 2^\gamma + \mu$, причем $n \neq v$. Возможны два случая.

Случай 1. Пусть $k \neq \gamma$, для определенности положим $k \geq \gamma + 1$.

Тогда

$$\langle h_n(x), h_v(x) \rangle = \sum_{l=0}^{2^k-1} \int_{\Delta_l^k} h_n(x) \underbrace{h_v(x)}_{c(v,k,l)=\text{const}} dx = \sum_{l=0}^{2^k-1} c(v, k, l) \underbrace{\int_{\Delta_l^k} h_n(x) dx}_0 = 0,$$

как следует из приведенных выше свойств системы Хаара.

Случай 2. Пусть $\gamma = k$, но $m \neq \mu$. Так как $h_n(x) = 0$ при $x \notin \Delta_m^k$, $h_v(x) = 0$ при $x \notin \Delta_\mu^k$ (см. определение (1.12)), то $\forall x \in [0; 1]$ $h_n(x)h_v(x) = 0$, поскольку для $m \neq \mu$ имеем $\Delta_m^k \cap \Delta_\mu^k = \emptyset$. Поэтому вновь $\langle h_n(x), h_v(x) \rangle = 0$. Таким образом, система Хаара является ортонормированной. ►

Теорема 1.12. Система Хаара (1.12) является полной в пространстве $L_2[0; 1]$. Примем утверждение теоремы без доказательства.

Глава 2 Спектральное представление функций

Широкий класс устройств в электротехнике и радиоэлектронике описывается математической моделью линейной инвариантной во времени (ЛИВ) системы, преобразующей входной сигнал — функцию времени $x(t)$ — в выходной сигнал $y(t) = L\{x(t)\}$ по правилу, определяемому формулой интегральной свертки:

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau, \text{ где } h(\tau) — \text{импульсная характеристика сист-$$

емы. Несложно доказать, что для таких ЛИВ-систем функции вида $e^{i\omega t}$ являются *собственными*, т. е. $L\{e^{i\omega t}\} = Ae^{i\omega t}$, где $A = A(\omega)$ — некоторое комплексное число. Это означает, что гармонические колебания не изменяют своей формы при прохождении через ЛИВ-систему, измениться может лишь их амплитуда и фаза. По этой причине базис гармонических колебаний наиболее удобен для анализа ЛИВ-систем, а *спектральное* (или *частотное*) представление функций с помощью преобразования Фурье является исключительно важным математическим аппаратом, используемым для анализа и синтеза систем обработки сигналов.

■ 2.1. Тригонометрические ряды Фурье. Интеграл Фурье

Напомним следующую теорему.

Теорема 2.1. Если функция $f(t)$ имеет период T и является кусочно-гладкой на периоде, то ее ряд Фурье¹ сходится к функции $f(t)$ в каждой точке ее непрерывности и к значению $\frac{1}{2}(f(t+0) + f(t-0))$ в точках разрыва, т. е.

¹ Если не оговаривается, какая система функций рассматривается в качестве базиса для построения ряда Фурье (см. раздел 1.5), то обычно подразумевается тригонометрическая система.

$$\frac{f(t+0) + f(t-0)}{2} = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos \frac{2\pi k}{T} t + b_k \sin \frac{2\pi k}{T} t \right), \quad (2.1)$$

где коэффициенты Фурье находятся по формулам:

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos \left(\frac{2\pi k}{T} t \right) dt, \quad k = 0, 1, \dots; \quad (2.2)$$

$$b_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin \left(\frac{2\pi k}{T} t \right) dt, \quad k = 1, 2, \dots$$

Задание. Убедитесь, что формулы (2.2) являются частным случаем (1.5) для $f \in L_2[-T/2; T/2]$ (см. пример 1.8).

Часто более удобной является комплексная форма записи ряда Фурье (2.1):

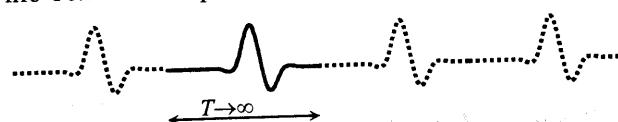
$$\frac{f(t+0) + f(t-0)}{2} = \sum_{k=-\infty}^{\infty} c_k e^{i \frac{2\pi k}{T} t}; \quad (2.3)$$

$$c_k = \frac{1}{T} \int_{-T/2}^{T/2} f(t) \exp \left(-i \frac{2\pi k}{T} t \right) dt, \quad k = 0, \pm 1, \pm 2, \dots \quad (2.4)$$

Реальные сигналы, чаще всего, представляют собой апериодические функции, искусственная периодизация которых, необходимая для корректного использования разложений (2.1) или (2.3), не представляет собой неоднозначную процедуру, приводящую к исчезновению сигнала. Поступим следующим образом. Обозначим $v_k = k/T$, $\Delta v_k = v_{k+1} - v_k = 1/T = \Delta v$, тогда с использованием данных обозначений из (2.3) и (2.4) получим

$$\sum_{k=-\infty}^{\infty} c_k e^{i \frac{2\pi k}{T} t} = \sum_{k=-\infty}^{\infty} \left(\int_{-T/2}^{T/2} f(u) \exp(-i 2\pi v_k u) du \right) e^{i 2\pi v_k t} \Delta v.$$

Далее непериодический сигнал представим как периодический с бесконечно большим периодом.



Предположим, что существует интеграл

$$S(v_k) = \lim_{T \rightarrow \infty} c_k T = \int_{-\infty}^{+\infty} f(u) \exp(-i 2\pi v_k u) du.$$

При формальном переходе к пределу при $T \rightarrow \infty$ в ряде из (2.3) получим

$$\lim_{T \rightarrow \infty} \sum_{k=-\infty}^{\infty} c_k e^{i \frac{2\pi k}{T} t} = \lim_{\Delta v \rightarrow 0} \sum_{k=-\infty}^{\infty} S(v_k) e^{i 2\pi v_k t} \Delta v = \int_{-\infty}^{+\infty} S(v) \exp(2\pi i v t) dv,$$

в случае существования последнего интеграла, который понимается в смысле главного значения по Коши,

$$\int_{-\infty}^{+\infty} S(v) \exp(2\pi i v t) dv = \lim_{N \rightarrow \infty} \int_{-N}^{+N} S(v) \exp(2\pi i v t) dv.$$

Данный интеграл носит название интеграла Фурье. Условия, при которых функция допускает представление в виде интеграла Фурье, определяет следующая теорема.

Теорема 2.2. Если функция $f(t)$ абсолютно интегрируема на всей числовой оси, т. е. $\int |f(t)| dt < \infty$, и является кусочно-

гладкой на любом конечном отрезке $t \in [a; b] \subset (-\infty; \infty)$, то ее можно представить в виде интеграла Фурье:

$$f(t) = \lim_{N \rightarrow \infty} \int_{-N}^{+N} S(v) \exp(2\pi i v t) dv = \int_{-\infty}^{+\infty} S(v) e^{2\pi i v t} dv \quad (2.5)$$

(считаем, что в точках разрыва $f(t) = \frac{f(t+0) + f(t-0)}{2}$), где

$$S(v) = \int_{-\infty}^{+\infty} f(t) e^{-2\pi i v t} dt. \quad (2.6)$$

Функция (2.6) называется *частотным спектром*, или *спектральной плотностью*, или *спектральной характеристикой* функции (сигнала) $f(t)$. Выражения (2.6) и (2.5) называют соот-

всегда прямым и обратным преобразованиями Фурье сигнала $f(t)$. Их записывают также с использованием в качестве аргумента спектральной плотности циклической частоты $\omega = 2\pi\nu$:

$$S(\nu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt, \quad f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} S(\nu) e^{i\omega t} d\nu.$$

Таким образом, при выполнении условий теоремы 2.2 сигнал можно описать как во *временной области*, т. е. через функцию времени $f(t)$, так и в *частотной области*, через функцию частоты $S(\nu)$. Оба представления взаимно однозначно соответствуют друг другу: $f(t) \leftrightarrow S(\nu)$.

Отметим ряд важных свойств интегрального преобразования Фурье:

1°. Сопряженная симметрия. Для любой вещественной функции $f(t)$: $S(\nu) = \overline{S(-\nu)}$ (докажите самостоятельно).

2°. Линейность. $\forall x(t) \leftrightarrow S_x(\nu)$, $\forall y(t) \leftrightarrow S_y(\nu)$, $\forall \alpha, \beta$:

$$f(t) = \alpha x(t) + \beta y(t) \leftrightarrow S(\nu) = \alpha S_x(\nu) + \beta S_y(\nu).$$

3°. Изменение масштаба. $\forall f(t) \leftrightarrow S(\nu)$, $\forall \alpha \neq 0$:

$$f(\alpha t) \leftrightarrow S_\alpha(\nu) = \frac{1}{|\alpha|} S(\nu/\alpha).$$

$$\begin{aligned} \blacktriangleleft S_\alpha(\nu) &= \int_{-\infty}^{+\infty} f(\alpha t) e^{-2\pi i \nu t} dt = \frac{1}{\alpha} \int_{-\infty}^{+\infty} f(u) e^{-2\pi i \frac{\nu}{\alpha} u} du = \\ &= \begin{cases} \frac{1}{\alpha} \int_{-\infty}^{+\infty} f(u) e^{-2\pi i (\nu/\alpha) u} du, & \text{при } \alpha > 0 \\ \frac{1}{\alpha} \int_{+\infty}^{-\infty} f(u) e^{-2\pi i (\nu/\alpha) u} du, & \text{при } \alpha < 0 \end{cases} = \frac{1}{|\alpha|} S(\nu/\alpha). \blacksquare \end{aligned}$$

4°. Задержка сигнала. $\forall f(t) \leftrightarrow S(\nu)$, $\forall t_0$:

$$f(t - t_0) \leftrightarrow e^{-2\pi i \nu t_0} S(\nu) \quad (\text{докажите самостоятельно}).$$

5°. Сдвиг спектра. $\forall f(t) \leftrightarrow S(\nu)$, $\forall \alpha$:

$$f(t) e^{2\pi i \nu t \alpha} \leftrightarrow S(\nu - \alpha) \quad (\text{докажите самостоятельно}).$$

6°. Свертка сигналов. $\forall u(t) \leftrightarrow S_u(\nu)$, $\forall w(t) \leftrightarrow S_w(\nu)$:

$$f(x) = \int_{-\infty}^{\infty} u(t) w(x-t) dt \leftrightarrow S(\nu) = S_u(\nu) S_w(\nu).$$

$$\begin{aligned} \blacktriangleleft S(\nu) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} u(t) w(x-t) dt \right) e^{-2\pi i \nu x} dx = \int_{-\infty}^{\infty} u(t) \underbrace{\int_{-\infty}^{\infty} w(x-t) e^{-2\pi i \nu x} dx}_{\substack{(Cm. свойство 4^o)}} dt = \\ &= \int_{-\infty}^{\infty} u(t) S_w(\nu) e^{-2\pi i \nu t} dt = S_u(\nu) S_w(\nu). \blacksquare \end{aligned}$$

7°. Произведение сигналов. $\forall u(t) \leftrightarrow S_u(\nu)$, $\forall w(t) \leftrightarrow S_w(\nu)$:

$$f(t) = u(t) w(t) \leftrightarrow S(\nu) = \int_{-\infty}^{\infty} S_u(x) S_w(\nu - x) dx$$

(докажите самостоятельно аналогично доказательству свойства 6°).

8°. Равенство Парсеваля. $\forall f(t) \leftrightarrow S(\nu)$:

$$E = \int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |S(\nu)|^2 d\nu$$

(величину E называют *энергией сигнала*).

$$\begin{aligned} E &= \int_{-\infty}^{\infty} \overline{f(t)} f(t) dt = \int_{-\infty}^{\infty} \overline{f(t)} \int_{-\infty}^{\infty} S(\nu) e^{2\pi i \nu t} d\nu dt = \int_{-\infty}^{\infty} S(\nu) \int_{-\infty}^{\infty} \overline{f(t)} e^{2\pi i \nu t} dt d\nu = \\ &= \int_{-\infty}^{\infty} S(\nu) \int_{-\infty}^{\infty} f(t) e^{-2\pi i \nu t} dt d\nu = \int_{-\infty}^{\infty} S(\nu) \overline{S(\nu)} d\nu = \int_{-\infty}^{\infty} |S(\nu)|^2 d\nu. \blacksquare \end{aligned}$$

9°. Дифференцирование во временной области. Если $f(t) \leftrightarrow S(\nu)$ и функция $f(t)$ дифференцируема, причем существует интеграл

$$\int_{-\infty}^{+\infty} v S(v) e^{2\pi i \nu t} dv, \quad \text{то} \quad f'(t) \leftrightarrow 2\pi i \nu S(\nu).$$

$$\int f'(t) e^{-2\pi i \nu t} dt = f(t) e^{-2\pi i \nu t} \Big|_{t=-\infty}^{t=\infty} - \int_{-\infty}^{\infty} (-2\pi i \nu) f(t) e^{-2\pi i \nu t} dt = 2\pi i \nu S(\nu),$$

так как несобственный интеграл $S(v) = \int_{-\infty}^{+\infty} f(t) e^{-2\pi i vt} dt$ существует, а необходимым условием его сходимости является $\lim_{t \rightarrow \pm\infty} f(t) e^{-2\pi i vt} = 0$. ►

10°. Дифференцирование в частотной области. Если $f(t) \leftrightarrow S(v)$ и функция $S(v)$ дифференцируема, причем существует интеграл

$$\int_{-\infty}^{+\infty} t f(t) e^{-2\pi i vt} dt, \text{ то } -2\pi i t f(t) \leftrightarrow S'(v)$$

(докажите самостоятельно аналогично доказательству свойства 9°).

11°. Свойство непрерывности. Если для нормы $\|x(t)\| = \int_{-\infty}^{+\infty} |x(t)| dt$ некоторая последовательность $\{f_n(t)\}$ сходится,

$$\lim_{n \rightarrow \infty} f_n(t) = f(t), \text{ т. е. } \lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} |f_n(t) - f(t)| dt = 0,$$

то в этой норме

$$\lim_{n \rightarrow \infty} S_n(v) = S(v), \text{ где } f_n(t) \leftrightarrow S_n(v), \quad f(t) \leftrightarrow S(v),$$

причем последовательность $\{S_n(v)\}$ сходится равномерно. Прием данное утверждение без доказательства.

Определения. Амплитудным спектром сигнала $f(t)$ называется модуль $|S(v)|$ спектральной плотности $S(v)$; фазовым спектром — ее аргумент, взятый с противоположным знаком¹ $\phi(v) = -\arg S(v)$.

Амплитудный и фазовый спектры позволяют записать спектральную плотность в показательной форме: $S(v) = |S(v)|e^{-i\phi(v)}$. Будем считать по определению, что $\phi(v) \in (-\pi; \pi]$.

Замечание. Из свойства 4° следует, что сдвиги сигнала в временной области в частотной области влияют лишь на фазовый спектр, но не изменяют амплитудный спектр сигнала.

¹ Нередко фазовым спектром называют аргумент (без смены знака) $\phi(v) = \arg S(v)$.

Пример 2.1. Найти амплитудный и фазовый спектры сигнала, представляющего собой прямоугольный импульс длительности T :

$$f(t) = \begin{cases} 1/\sqrt{T} & \text{при } t \in [0; T]; \\ 0 & \text{при } t \notin [0; T]. \end{cases}$$

◀ Найдем сначала спектральную плотность функции $g(t) = f(t + 0,5T)$:

$$S_g(v) = \int_{-\infty}^{+\infty} g(t) e^{-2\pi i vt} dt = \frac{1}{\sqrt{T}} \int_{-T/2}^{+T/2} e^{-2\pi i vt} dt = \frac{e^{i\pi v T} - e^{-i\pi v T}}{2\pi i v \sqrt{T}} = \frac{\sin \pi v T}{\pi v \sqrt{T}}.$$

Так как $f(t) = g(t - 0,5T)$, то на основании свойства 4° получим $S_f(v) = e^{-\pi i v T} \frac{\sin \pi v T}{\pi v \sqrt{T}}$, откуда находим амплитудный спектр:

$$|S_f(v)| = \frac{1}{\sqrt{T}} \left| \frac{\sin \pi v T}{\pi v} \right|.$$

Для фазового спектра $\phi(v)$ рассмотрим сначала частоты $v \geq 0$.

При $v = k/T$, $k = 1, 2, \dots$ $\arg S_f(v)$ не определен, так как $S_f(v) = 0$.

При $v \in (2k/T; (2k+1)/T)$, $k = 0, 1, \dots$, имеем

$$|S_f(v)| = \sin \pi v T, \\ S_f(v) = \frac{\sin \pi v T}{\pi v \sqrt{T}} e^{-\pi i v T} = \underbrace{\left| \frac{\sin \pi v T}{\pi v \sqrt{T}} \right|}_{|S_f(v)|} e^{-\pi i v T} = |S_f(v)| e^{-i\phi(v)},$$

т. е. $e^{-i\phi(v)} = e^{-i\pi v T}$ и $\phi(v) = \arg(e^{i\pi v T})$, причем $\phi(v) = \phi(v + \frac{2}{T})$.

При $v \in ((2k+1)/T; (2k+2)/T)$, $k = 0, 1, \dots$, имеем

$$|S_f(v)| = -\sin \pi v T, \\ S_f(v) = \frac{\sin \pi v T}{\pi v \sqrt{T}} e^{-\pi i v T} = -\underbrace{\left| \frac{\sin \pi v T}{\pi v \sqrt{T}} \right|}_{|S_f(v)|} e^{-\pi i v T} = |S_f(v)| e^{-i\phi(v)},$$

т. е. $e^{-i\phi(v)} = -e^{-i\pi v T} = e^{-i\pi v T + i\pi}$ и $\phi(v) = \arg(e^{i\pi(vT-1)})$, и вновь $\phi(v) = v + \frac{2}{T}$.

Таким образом, для неотрицательных частот достаточно привести один период фазового спектра $\phi(v)$. Для $v \in [0; 2/T]$ получаем

$$\phi(v) = \begin{cases} \pi v T & \text{при } v \in (0; 1/T); \\ \pi v T - \pi & \text{при } v \in (1/T; 2/T). \end{cases}$$

Вид функции $\phi(v)$ для $v < 0$ находим на основании свойства 1° интеграла Фурье, которое означает, что $\phi(-v) = -\phi(v)$. Графики амплитудного и фазового спектров приведены на рис. 2.1. ►

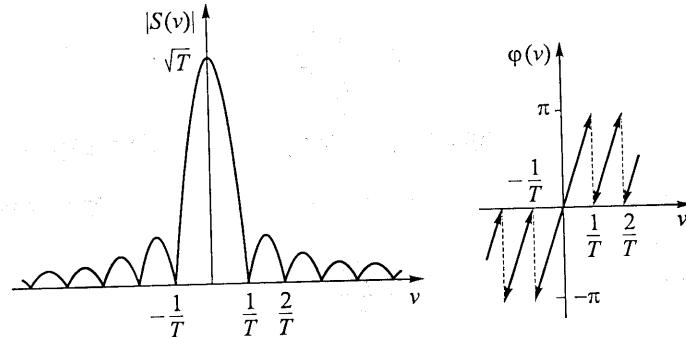


Рис. 2.1

■ 2.2. Обобщенное преобразование Фурье

При выполнении условий, определяемых теоремой 2.2, между сигналом и его спектром существует взаимно однозначное соответствие:

$$S(v) = \int_{-\infty}^{+\infty} f(t) e^{-2\pi i v t} dt \leftrightarrow f(t) = \int_{-\infty}^{+\infty} S(v) e^{2\pi i v t} dv.$$

Рассмотрим функцию $g(t) = 1$. Условия теоремы 2.2, очевидно, для нее не выполнены, и спектр $S_g(v)$, т. е. понимаемый в традиционном смысле интеграл $\int_{-\infty}^{+\infty} e^{-2\pi i v t} dt$, не существует. Однако если положить, что $S_g(v) = \delta(v)$ (δ -функция Дирака), то запись обратного преобразования Фурье не вызывает никаких затруднений и дает точное восстановление функции $g(t)$:

$$g(t) = \int_{-\infty}^{+\infty} S_g(v) e^{2\pi i v t} dv = \int_{-\infty}^{+\infty} \delta(v) e^{2\pi i v t} dv = 1.$$

Для того чтобы расширить класс функций, для которых применимо интегральное преобразование Фурье, положим, по определению, что

$$\int_{-\infty}^{+\infty} e^{-2\pi i v t} dt = \delta(v). \quad (2.7)$$

Замечание. Эквивалентными выражению (2.7) являются следующие определения. В силу вещественности δ -функции $\int_{-\infty}^{+\infty} e^{2\pi i v t} dt = \overline{\delta(v)} = \delta(v)$. В силу симметрии выражения (2.7) относительно переменных v и t : $\int_{-\infty}^{+\infty} e^{2\pi i v t} dv = \int_{-\infty}^{+\infty} e^{-2\pi i v t} dv = \delta(t)$.

Определение (2.7) не противоречит естественным представлениям о δ -функции. Покажем, что интеграл $\int_{-\infty}^{+\infty} e^{-2\pi i v t} dt$, действительно, проявляет свойства δ -функции. Пусть $\phi(t)$ — некоторая непрерывная в окрестности точки $t = 0$ функция, отвечающая условиям теоремы 2.2. Тогда

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{-2\pi i v t} dv \right) \phi(t) dt = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \phi(t) e^{-2\pi i v t} dt \right) dv = \int_{-\infty}^{\infty} S_{\phi}(v) dv = \phi(0),$$

что соответствует поведению δ -функции: $\int_{-\infty}^{+\infty} \delta(t) \phi(t) dt = \phi(0)$.

Определение (2.7) расширяет область применимости интегральных преобразований (2.5) и (2.6), превращая ряды Фурье для периодических функций (см. формулы (2.3), (2.4)) в частный случай интегральных преобразований. Действительно, исходя из выражения (2.7), функции $\phi_k(t) = e^{\frac{i2\pi kt}{T}}$ соответствует обобщенный спектр

$$S_{\phi_k}(v) = \int_{-\infty}^{\infty} e^{-2\pi i t \left(v - \frac{k}{T}\right)} dt = \delta\left(v - \frac{k}{T}\right).$$

Поэтому для произвольной функции периода T , представленной в виде ряда (2.3), $f(t) = \sum_{k=-\infty}^{\infty} c_k \phi_k(t)$, обобщенное интегральное преобразование дает спектр

$$\begin{aligned} S(v) &= \int_{-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} c_k \phi_k(t) \right) e^{-2\pi i vt} dt = \\ &= \sum_{k=-\infty}^{\infty} c_k \int_{-\infty}^{\infty} \phi_k(t) e^{-2\pi i vt} dt = \sum_{k=-\infty}^{\infty} c_k \delta\left(v - \frac{k}{T}\right), \end{aligned}$$

по которому функция может быть восстановлена в результате обратного преобразования Фурье:

$$\begin{aligned} f(t) &= \int_{-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} c_k \delta\left(v - \frac{k}{T}\right) \right) e^{2\pi i vt} dv = \sum_{k=-\infty}^{\infty} c_k \left(\int_{-\infty}^{\infty} \delta\left(v - \frac{k}{T}\right) e^{2\pi i vt} dv \right) = \\ &= \sum_{k=-\infty}^{\infty} c_k e^{\frac{i2\pi kt}{T}} = \sum_{k=-\infty}^{\infty} c_k \phi_k(t). \end{aligned}$$

Таким образом, интегралы в преобразовании Фурье (2.5) и (2.6) будем понимать обобщенно. Для обозначения обобщенного преобразования Фурье далее будем использовать следующие сокращенные записи:

$$\begin{aligned} S(v) &= \Phi\{f(t)\} \text{ — прямое преобразование,} \\ f(t) &= \Phi^{-1}\{S(v)\} \text{ — обратное.} \end{aligned}$$

Из рассмотренного выше следует важное наблюдение: если функция (сигнал) $f(t)$ является T -периодической, то ее спектр $S(v) = \sum_{k=-\infty}^{\infty} c_k \delta\left(v - \frac{k}{T}\right)$ является *решетчатой* функцией, т. е. принимает ненулевые (вообще говоря, бесконечные!) значения лишь для определенных значений частоты, а именно, $v_k = k/T$, $k \in \mathbb{Z}$. Таким образом, спектр периодических функций полностью характеризуется набором коэффициентов (2.4) ряда (2.3):

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{i2\pi kt}{T}} \Leftrightarrow S(v) = \sum_{k=-\infty}^{\infty} c_k \delta\left(v - \frac{k}{T}\right). \quad (2.8)$$

Поэтому под амплитудным спектром для периодических функций понимают набор модулей коэффициентов Фурье $\{|c_k|\}$, а под фазовым спектром — набор их аргументов, взятых с противоположным знаком, $\{-\arg(c_k)\}$.

Преобразования (2.5) и (2.6) имеют сходную природу и отличаются только знаком при мнимом показателе подынтегральной экспоненты. Вследствие этого преобразования (2.5), (2.6) обладают и сходными дуальными свойствами, которые мы уже наблюдали. Например, произведению функций соответствует свертка в области преобразований (см. свойства 6°, 7° интеграла Фурье). Установив, что периодическому сигналу соответствует «линейчатый» спектр, мы можем утверждать, что для решетчатых функций вида $f(t) = \sum_{k=-\infty}^{\infty} f_k \delta(t - k\Delta t)$, принимающих ненулевые значения лишь для равноотстоящих значений дискретного аргумента $t_k = k\Delta t$, в частотной области спектр имеет период: $S(v) = S\left(v + \frac{1}{\Delta t}\right)$, т. е. периодическая функция в одной области соответствует решетчатой функции в другой, и наоборот.

Пример 2.2. Показать, что функции $G(t) = \Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t)$ в частотной области соответствует спектр $S_G(v) = \sum_{k=-\infty}^{\infty} \delta\left(v - \frac{k}{\Delta t}\right)$.

◀ Заметим, что функция $G(t)$ имеет период $T = \Delta t$, и на основании соотношения (2.8)

$$S_G(v) = \sum_{k=-\infty}^{\infty} c_k \delta\left(v - \frac{k}{\Delta t}\right), \quad \text{где (см. формулу (2.4))}$$

$$c_k = \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} G(t) \exp\left(-i \frac{2\pi k}{\Delta t} t\right) dt = \int_{-\Delta t/2}^{\Delta t/2} \delta(t) \exp\left(-i \frac{2\pi k}{\Delta t} t\right) dt = 1. \blacksquare$$

Заметим, что из соотношения (2.8) следует также, что разложение функции $G(t)$ из приведенного примера в ряд Фурье (2.3) имеет вид:

$$\Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t) = \sum_{k=-\infty}^{\infty} \exp\left(i \frac{2\pi k}{\Delta t} t\right). \quad (2.9)$$

■ 2.3. Принцип неопределенности времени-частотного представления сигналов

Определение. Носителем Ω функции $f(x)$ называется замыкание множества аргументов x , при которых $f(x)$ принимает ненулевые значения, т. е. $\Omega = \overline{\{x | f(x) \neq 0\}}$. Обозначается: $\Omega = \text{supp } f(x)$. Говорят, что функция имеет ограниченный носитель, если существует конечный отрезок $[a; b]$, содержащий носитель $\text{supp } f(x) \subset [a; b]$.

Например, носителем функции $h_5(x)$ системы Хаара (см. пример 1.12) является отрезок $x \in [1/4; 1/2]$.

Теорема 2.3. Никакая пара ненулевых функций $f(t) \leftrightarrow S(v)$ не может одновременно иметь ограниченные носители во временной и частотной областях.

◀ Докажем, что из ограниченности носителя спектральной плотности $S(v)$ в частотной области следует неограниченность носителя функции $f(t) \neq 0$ во временной области.

Допустим противное: пусть $\text{supp } S(v) \subset [-b; b]$ (носитель ограничен), и функция $f(t) \neq 0$ также имеет ограниченный носи-

тель. Тогда $f(t) = \int_{-b}^b S(v) e^{2\pi i v t} dv$, причем найдется некоторый

отрезок $t \in [c; d]$ для которого $f(t) = 0$. Применяя n -кратно свойство 9° интеграла Фурье, для точки $t_0 = (c + d)/2$ имеем $\forall n \in \mathbb{N}$

$$f^{(n)}(t_0) = (2\pi i)^n \int_{-b}^b v^n S(v) e^{2\pi i v t_0} dv = 0.$$

Так как $\forall t \in \mathbb{R}$,

$$f(t) = \int_{-b}^b S(v) e^{2\pi i v(t-t_0)} e^{2\pi i v t_0} dv \quad \text{и} \quad e^{2\pi i v(t-t_0)} = \sum_{n=0}^{\infty} \frac{(2\pi i v(t-t_0))^n}{n!},$$

то, подставляя разложение экспоненты под знак интеграла и проводя почлененное интегрирование ряда, получаем $\forall t \in \mathbb{R}$

$$f(t) = \sum_{n=0}^{\infty} \frac{(2\pi i(t-t_0))^n}{n!} \int_{-b}^b v^n S(v) e^{2\pi i v t_0} dv = \sum_{n=0}^{\infty} \frac{(t-t_0)^n}{n!} f^{(n)}(t_0) = 0.$$

Получили противоречие условию $f(t) \neq 0$, следовательно, функция $f(t)$ имеет неограниченный носитель.

Аналогично показывается, что из ограниченности носителя функции $f(t)$ следует неограниченность носителя спектральной плотности $S(v) \neq 0$. \blacksquare

Из теоремы 2.3 следует, что сигналы конечной длительности имеют бесконечную частотную полосу (носитель спектральной плотности). Так, в примере 2.1 рассматривалась функция, имеющая конечный носитель во временной области, $\text{supp } f(t) = [0; 1]$, а

в частотной области носитель спектра $S(v) = e^{-\pi i v T} \frac{\sin \pi v T}{\sqrt{T} \pi v}$ совпадал со всей числовой осью, $\text{supp } S(v) = (-\infty; \infty)$. Однако на практике почти всегда необходимо задаваться требованиями конечной частотной полосы. То есть для произвольного сигнала (функции) $f(t)$ требуется каким-то образом определить его частотную полосу $v \in [v_1; v_2]$, вполне характеризующую сигнал в

частотной области. Тогда под *шириной полосы* спектра понимается величина $\Delta v = v_2 - v_1$.

Единого строгого подхода для определения частотной полосы сигнала, реально имеющего бесконечную ширину спектра, нет. На практике обычно выбирают на оси частот такой отрезок $v \in [v_1; v_2]$, который содержит основную часть энергии сигнала, т. е.

$$E = \int_{-\infty}^{\infty} |S(v)|^2 dv \approx E_{\Delta v} = \int_{v_1}^{v_2} |S(v)|^2 dv.$$

Разность $E - E_{\Delta v}$ характеризует величину тех искажений, которые связаны с искусственным «усечением» полосы. Действительно, если обозначить усеченный спектр

$$\tilde{S}(v) = \begin{cases} S(v) & \text{при } v \in [v_1; v_2]; \\ 0 & \text{при } v \notin [v_1; v_2] \end{cases}$$

и соответствующий ему искаженный сигнал $\tilde{f}(t)$, то, очевидно, в силу свойства 2° преобразования Фурье имеем

$$S(v) - \tilde{S}(v) = \Phi(f(t) - \tilde{f}(t)),$$

а на основании свойства 8° энергия ошибки $\varepsilon(t) = f(t) - \tilde{f}(t)$:

$$\begin{aligned} \int_{-\infty}^{\infty} |\varepsilon(t)|^2 dt &= \int_{-\infty}^{\infty} |f(t) - \tilde{f}(t)|^2 dt = \int_{-\infty}^{\infty} |S(v) - \tilde{S}(v)|^2 dv = \\ &= \int_{-\infty}^{v_1} |S(v)|^2 dv + \int_{v_2}^{\infty} |S(v)|^2 dv = \int_{-\infty}^{\infty} |S(v)|^2 dv - \int_{v_1}^{v_2} |S(v)|^2 dv = E - E_{\Delta v}. \end{aligned}$$

Так, для примера 2.1 в качестве полосы сигнала можно было бы взять $v \in \left[-\frac{1}{T}; \frac{1}{T}\right]$ (тогда $E_{\Delta v}/E = 0,90$) или $v \in \left[-\frac{2}{T}; \frac{2}{T}\right]$ (тогда $E_{\Delta v}/E = 0,95$).

Для пояснения принципа неопределенности времени-частотного представления вещественных сигналов более удобен иной подход к определению ширины полосы в частотной области и длительности сигнала во временной области.

Положим, что энергия сигнала единичная, т. е.

$$\int_{-\infty}^{\infty} |S(v)|^2 dv = \int_{-\infty}^{\infty} |f(t)|^2 dt = 1.$$

Тогда по своему физическому смыслу функция $\rho(v) = |S(v)|^2$ представляет собой плотность распределения энергии в частотной области, причем для вещественных сигналов $\rho(v) = \rho(-v)$, в силу свойства 1° спектральной плотности. По аналогии с понятием дисперсии характеризовать ширину спектра Δv будем по величине

$$(\Delta v)^2 = \int_{-\infty}^{\infty} v^2 |S(v)|^2 dv, \quad (2.10)$$

которая представляет собой меру «разброса» энергии в частотной области относительно начала координат, т. е. постоянной составляющей $v = 0$, которая для вещественных сигналов всегда является средним значением распределения энергии в спектре:

$$\int_{-\infty}^{\infty} v |S(v)|^2 dv = 0. \quad \text{Естественно называть определенную в формуле (2.10) частотную полосу } [-\Delta v/2; \Delta v/2] \text{ среднеквадратичной частотной полосой.}$$

Во временной области функцией плотности распределения энергии сигнала является $\gamma(t) = |f(t)|^2$. Аналогично, в качестве меры длительности сигнала Δt возьмем величину среднеквадратичной длительности, т. е.

$$(\Delta t)^2 = \int_{-\infty}^{\infty} (t - m_t)^2 |f(t)|^2 dt = \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt - m_t^2, \quad (2.11)$$

где $m_t = \int_{-\infty}^{\infty} t |f(t)|^2 dt$ — среднее значение для распределения

энергии сигнала во временной области. Среднеквадратичные величины Δv и Δt , определенные с помощью формул (2.10) и (2.11), называют также *эффективными* значениями (ширины полосы и длительности сигнала).

Формулы (2.10) и (2.11) характеризуют локализацию энергии сигнала: чем меньше среднеквадратичная полоса (длительность), тем выше локализация энергии в частотной (временной) области. *Принцип неопределенности* гласит, что добиться высокой локализации энергии одновременно и во временной, и в частотной областях нельзя. Так, верна следующая теорема.

Теорема 2.4. Для любых дифференцируемых вещественных сигналов единичной энергии произведение полосы (2.10) и длительности (2.11) ограничено снизу:

$$\Delta t \Delta v \geq \frac{1}{4\pi}. \quad (2.12)$$

◀ Пусть, для упрощения изложения, во временной области имеем $m_t = \int_{-\infty}^{\infty} t |f(t)|^2 dt = 0$ (в необходимых случаях выполняется сдвиг сигнала по оси времени, не изменяющий его амплитудный спектр (см. свойство 4° преобразования Фурье)). Заметим, что существование интеграла $m_t = \int_{-\infty}^{\infty} t |f(t)|^2 dt$ означает также выполнение необходимого условия его сходимости: $\lim_{t \rightarrow \pm\infty} t |f(t)|^2 = 0$.

Поскольку $f'(t) \leftrightarrow 2\pi i v S(v)$ (свойство 9°), то на основании равенства Парсеваля (свойство 8°)

$$\int_{-\infty}^{\infty} |f'(t)|^2 dt = (2\pi)^2 \int_{-\infty}^{\infty} v^2 |S(v)|^2 dv.$$

Тогда

$$(\Delta v)^2 (\Delta t)^2 = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} |f'(t)|^2 dt \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt = \frac{1}{(2\pi)^2} \|f'(t)\|^2 \cdot \|t f(t)\|^2,$$

где в пространстве $L_2(\mathbb{R})$ норма вещественной функции индуцирована скалярным произведением (см. раздел 1.4)

$$\langle f(t), g(t) \rangle = \int_{-\infty}^{\infty} f(t) g(t) dt.$$

Так как на основании неравенства Коши—Буняковского (лемма 1.3)

$$\|f'(t)\| \cdot \|t f(t)\| \geq |\langle f'(t), t f(t) \rangle| = \left| \int_{-\infty}^{\infty} f'(t) t f(t) dt \right|,$$

то

$$\Delta v \Delta t \geq \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} f'(t) t f(t) dt \right| = \frac{1}{4\pi} \left| \underbrace{t |f(t)|^2}_{0-0} \Big|_{t=-\infty}^{t=\infty} - \int_{-\infty}^{\infty} |f(t)|^2 dt \right| = \frac{1}{4\pi}. \blacksquare$$

Задание. Убедитесь, что для сигнала с энергией

$$E = \int_{-\infty}^{+\infty} |f(t)|^2 dt$$

оценка (2.12) принимает следующий вид: $\Delta t \Delta v \geq \frac{E}{4\pi}$.

Пример 2.3. Показать, что точное значение оценки (2.12) достигается для *гауссова импульса*, т. е. для функции вида $f(t) = C e^{-kt^2}$, где C, k — некоторые константы ($k > 0$).

◀ Для построения рассуждений достаточно рассмотреть случай единичной энергии сигнала, положив $C = \sqrt[4]{2k/\pi}$ (убедитесь, что в этом случае энергия сигнала единичная, $\int_{-\infty}^{\infty} |f(t)|^2 dt = 1$, воспользовавшись равенством $\int_{-\infty}^{+\infty} \exp(-t^2/2) dt = \sqrt{2\pi}$).

Для упрощения выкладок проведем доказательство только для $k = 1/4$, т. е. для несложно обобщаемого (с использованием свойства 3° преобразования Фурье) на другие значения $k > 0$ случая $f(t) = \frac{1}{\sqrt[4]{2\pi}} e^{-t^2/4}$.

Поскольку $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 \exp(-t^2/2) dt = 1$ (как выражение для дисперсии стандартного нормального закона), то и для эфек-

тивной длительности сигнала во временной области также получим $(\Delta t)^2 = \int_{-\infty}^{+\infty} t^2 |f(t)|^2 dt = 1$, т. е. $\Delta t = 1$.

Для производной спектральной плотности имеем:

$$\begin{aligned} S'(v) &= \frac{1}{\sqrt[4]{2\pi}} \int_{-\infty}^{+\infty} e^{-t^2/4} e^{-2\pi v t} (-2\pi i t) dt = \frac{4\pi i}{\sqrt[4]{2\pi}} \int_{-\infty}^{+\infty} e^{-2\pi v t} d\left(e^{-t^2/4}\right) = \\ &= \frac{4\pi i}{\sqrt[4]{2\pi}} \left(e^{-2\pi v t - (t^2/4)} \Big|_{t=-\infty}^{t=\infty} - \int_{-\infty}^{+\infty} e^{-t^2/4} (-2\pi i v) e^{-2\pi v t} dt \right) = -8\pi^2 v S(v). \end{aligned}$$

Отсюда получаем следующее дифференциальное уравнение: $(\ln S(v))' = -8\pi^2 v$, интегрируя которое, находим: $S(v) = N e^{-4\pi^2 v^2}$, где нормировочная константа N определяется из равенства Парсеваля:

$$1 = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\infty}^{+\infty} |S(v)|^2 dv = \frac{N^2}{\sqrt{4\pi}} \underbrace{\int_{-\infty}^{+\infty} e^{-\frac{(4\pi v)^2}{2}} d(4\pi v)}_{\sqrt{2\pi}} = N^2 / \sqrt{8\pi}.$$

Окончательно получаем спектральную плотность

$$S(v) = \sqrt{8\pi} e^{-4\pi^2 v^2}.$$

Тогда для эффективной полосы имеем:

$$\begin{aligned} (\Delta v)^2 &= \sqrt{8\pi} \int_{-\infty}^{+\infty} v^2 e^{-8\pi^2 v^2} dv = \\ &= \frac{\sqrt{8\pi}}{4\pi(4\pi)^2} \underbrace{\int_{-\infty}^{+\infty} (4\pi v)^2 e^{-(4\pi v)^2/2} d(4\pi v)}_{\sqrt{2\pi}} = \frac{1}{(4\pi)^2}. \end{aligned}$$

Отсюда $\Delta v = \frac{1}{4\pi}$ и $\Delta v \Delta t = \frac{1}{4\pi}$. ►

Задание. Используя свойства 2° и 3° интеграла Фурье и результаты примера 2.3, убедитесь, что для произвольного числа $k > 0$ гауссову импульсу $f(t) = \sqrt[4]{\frac{2k}{\pi}} e^{-kt^2}$ соответствует спектральная плотность $S(v) = \sqrt[4]{\frac{2\pi}{k}} e^{-\frac{\pi^2 v^2}{k}}$.

Заметим, что для сигнала любой формы его «сжатие» по аргументу во временной области приводит к такому же масштабному «растяжению» по аргументу в частотной области (см. свойство 3° преобразования Фурье). Таким образом, чем меньше ширина полосы сигнала, тем выше его длительность, и наоборот. Построив прямоугольную систему координат, осями которой являются время и частота, каждому сигналу на полученной плоскости «Время × Частота» можно будет поставить в соответствие некоторую прямоугольную область Ω локализации сигнала, со сторонами Δt и Δv . При этом, если понимать Δt и Δv в смысле эффективных значений, то для сигнала единичной энергии площадь данной области в соответствии с принципом неопределенности не может быть меньше величины $\frac{1}{4\pi}$.

Пример 2.4. Изобразить на плоскости «Время × Частота» область локализации $\Omega = \Delta t \times \Delta v$ сигнала из примера 2.1 для эффективных значений полосы и длительности (2.10), (2.11).

◀ Во временной области получаем среднее значение распределения энергии $m_t = \int_{-\infty}^{\infty} t |f(t)|^2 dt = \frac{1}{T} \int_0^T t dt = \frac{T}{2}$, а для эффективной длительности сигнала Δt имеем

$$(\Delta t)^2 = \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt - m_t^2 = \int_0^T t^2 dt - \frac{T^2}{4} = \frac{T^2}{12}, \text{ т. е. } \Delta t = \frac{T}{2\sqrt{3}}.$$

В частотной области

$$(\Delta v)^2 = \lim_{F \rightarrow \infty} \int_{-F}^F v^2 |S(v)|^2 dv = \lim_{F \rightarrow \infty} \int_{-F}^F v^2 \frac{\sin^2 \pi v T}{T(\pi v)^2} dv = \infty,$$

т. е. эффективная полоса сигнала в данном случае неограничена, $v \in (-\infty; \infty)$. Область $\Omega = \Delta t \times \Delta v$ показана на рис. 2.2. ►

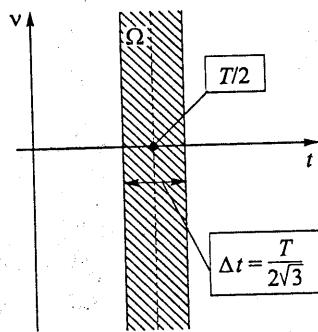


Рис. 2.2

■ 2.4. Энергетический спектр. Спектр мощности

Пусть $f(t)$ — функция из пространства $L_2(\mathbb{R})$ (см. пример 1.7), т. е. $E = \int_{-\infty}^{+\infty} f^2(t) dt < \infty$. Для детерминированного сигнала $f(t) \in L_2(\mathbb{R})$ вторым начальным моментом назовем интеграл

$$B(\tau) = \int_{-\infty}^{+\infty} f(t)f(t-\tau) dt. \quad (2.13)$$

$B(\tau)$ характеризует «степень сходства» сигнала и его сдвига по аргументу: чем больше значение интеграла (2.13), тем это сходство выше. Так, максимальное «сходство» наблюдается при $\tau = 0$, когда значение функции (2.13) максимально и равно энергии сигнала E , вообще же $\forall \tau : |B(\tau)| \leq B(0) = E$. Действительно, понимая под скалярным произведением функций $f(t)$, $g(t)$ (элементов пространства $L_2(\mathbb{R})$) (см. пример 1.7) величину $\langle f, g \rangle = \int_{-\infty}^{+\infty} f(t)g(t) dt$, на основании неравенства Коши—Буняковского (лемма 1.3) получаем для функции (2.13) $\forall \tau$

$$|B(\tau)| = |\langle f(t), f(t-\tau) \rangle| \leq \|f(t)\| \cdot \|f(t-\tau)\| = \|f(t)\|^2 = B(0),$$

так как

$$\forall \tau : \|f(t-\tau)\| = \sqrt{\int_{-\infty}^{+\infty} f(t-\tau)^2 dt} = \sqrt{\int_{-\infty}^{+\infty} f(t)^2 dt} = \|f(t)\|.$$

Убедитесь самостоятельно, что интеграл (2.13) — четная функция относительно параметра τ .

Найдем преобразование Фурье для второго начального момента.

$$\begin{aligned} \Phi(B(\tau)) &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(t)f(t-\tau) dt \right) e^{-2\pi i v \tau} d\tau = \\ &= \int_{-\infty}^{+\infty} f(t) \left(\int_{-\infty}^{+\infty} f(t+\tau) e^{2\pi i v \tau} d\tau \right) dt = \\ &= \int_{-\infty}^{+\infty} f(t) \left(\int_{-\infty}^{+\infty} f(t+\tau) e^{2\pi i v(\tau+t)} d\tau \right) e^{-2\pi i vt} dt = \\ &= \int_{-\infty}^{+\infty} f(t) \overline{S_f(v)} e^{-2\pi i vt} dt = \overline{S_f(v)} S_f(v) = |S_f(v)|^2, \end{aligned} \quad (2.14)$$

где $S_f(v) = \Phi(f(t))$. Таким образом, для $f(t) \in L_2(\mathbb{R})$ преобразование Фурье функции (13) представляет собой квадрат амплитудного спектра сигнала. Функция $\Phi(B(\tau)) = |S_f(v)|^2$ называется *энергетическим спектром* сигнала.

Для многих важных моделей детерминированных сигналов энергия $E = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \infty$, т.е. интеграл (2.13) для $\tau = 0$ не существует (например, для периодических сигналов). В этом случае вместо $B(\tau)$ (2.13) будем рассматривать функцию

$$R(\tau) = \lim_{T \rightarrow \infty} R_T(\tau), \quad (2.15)$$

где $R_T(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} f(t)f(t-\tau) dt$, если предел в формуле (2.15)

существует по норме $\|x(t)\| = \int_{-\infty}^{+\infty} |x(t)| dt$, т. е.

$$\lim_{T \rightarrow \infty} \int_{-\infty}^{+\infty} |R_T(\tau) - R(\tau)| d\tau = 0.$$

Тогда $R(0)$ представляет собой среднюю энергию сигнала в единицу времени, т. е. *мощность*. (Убедитесь самостоятельно, что (2.15) — также четная функция и $|R(\tau)| \leq R(0)$.) В силу свойства 11° преобразования Фурье

$$s_f(v) = \Phi(R(\tau)) = \lim_{T \rightarrow \infty} \Phi(R_T(\tau)). \quad (2.16)$$

Поэтому (см. свойство 4° преобразования Фурье)

$$\begin{aligned} s_f(v) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{+\infty} \left(\int_{-T/2}^{T/2} f(t) f(t-\tau) dt \right) e^{-2i\pi v \tau} d\tau = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) \left(\int_{-\infty}^{\infty} f(t-\tau) e^{-2i\pi v \tau} d\tau \right) dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) \left(\overline{\int_{-\infty}^{\infty} f(\tau+t) e^{-2i\pi v \tau} d\tau} \right) dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) \overline{S_f(v)} e^{-2i\pi v t} dt = \lim_{T \rightarrow \infty} \frac{S_{f,T}(v) \overline{S_f(v)}}{T}, \end{aligned} \quad (2.17)$$

где $S_{f,T}(v) = \int_{-T/2}^{T/2} f(t) e^{-2\pi i v t} dt$. Назовем функцию $s_f(v)$ в формулках (2.16), (2.17) *спектром мощности* сигнала $f(t)$.

Для стационарного в широком смысле случайного процесса (сигнала) $X(t)$ вторым начальным моментом называется функция

$$R_X(\tau) = M(X(t), X(t-\tau)) = \text{cov}(X(t), X(t-\tau)) + m_X^2, \quad (2.18)$$

где $m_X = M(X(t))$, а ковариационной функцией (КФ) называется

$$K_X(\tau) = \text{cov}(X(t), X(t-\tau)). \quad (2.19)$$

Для эргодического процесса функцию (2.18) можно найти по формуле (2.15), полагая $f(t) = X(t)$.

Функции (2.18) и (2.19) являются четными и для стационарного сигнала зависят только от расстояния τ между сечениями процесса по оси времени, но не от самого момента наблюдения t . Заметим, что $R_X(0)$ также имеет смысл мощности сигнала, причем

$$R_X(0) = M(X^2(t)) = \sigma_X^2 + m_X^2,$$

где σ_X^2 — дисперсия процесса. Таким образом, для центрированного процесса ($m_X = 0$) дисперсия характеризует мощность сигнала.

Для случайного процесса $X(t)$, заданного при $t \in \mathbb{R}$, можно рассмотреть аналог спектра мощности (2.16), который следует определить (см. выражение (2.17)) через среднеквадратический предел¹:

$$s_X(v) = \text{l.i.m.}_{T \rightarrow \infty} \frac{S_{X,T}(v) \overline{S_X(v)}}{T}, \quad (2.20)$$

где случайные функции:

$$S_X(v) = \Phi(X(t)), \quad S_{X,T}(v) = \int_{-T/2}^{T/2} X(t) e^{-2\pi i v t} dt.$$

Найдем математическое ожидание случайной функции $s_X(v)$ из выражения (2.20). Рассматривая вещественный случайный процесс $X(t)$, имеем:

$$\begin{aligned} M(s_X(v)) &= M\left(\text{l.i.m.}_{T \rightarrow \infty} \frac{S_{X,T}(v) \overline{S_X(v)}}{T}\right) = \lim_{T \rightarrow \infty} M\left(\frac{S_{X,T}(v) \overline{S_X(v)}}{T}\right) = \\ &= \lim_{T \rightarrow \infty} M\left(\frac{1}{T} \int_{-T/2}^{T/2} X(t) e^{-2\pi i v t} dt \int_{-\infty}^{\infty} X(u) e^{2\pi i v u} du\right) = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-\infty}^{\infty} M(X(t)X(u)) e^{-2\pi v i(t-u)} du dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left(\int_{-\infty}^{\infty} R_X(u-t) e^{2\pi i v(u-t)} du \right) dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left(\int_{-\infty}^{\infty} R_X(u) e^{2\pi i v u} du \right) dt = \end{aligned}$$

¹ Напомним, запись $\text{l.i.m.}_{t \rightarrow t_0} X(t) = Y$ означает, что $\lim_{t \rightarrow t_0} M((X(t) - Y)^2) = 0$, при этом $M\left(\text{l.i.m.}_{t \rightarrow t_0} X(t)\right) = \lim_{t \rightarrow t_0} M(X(t))$.

$$= \int_{-\infty}^{\infty} R_X(u) e^{2\pi i v u} du = \int_{-\infty}^{\infty} R_X(u) e^{-2\pi i v u} du, \quad (2.21)$$

где функция $R_X(t)$ определена по формуле (2.18), а последний переход осуществлен на основании четности функции $R_X(t)$. Таким образом, $M(s_X(v)) = \Phi(R_X(t))$.

По определению, спектром мощности $\tilde{s}_X(v)$ для стационарного случайного процесса $X(t)$ называют преобразование Фурье КФ (2.19), т. е.

$$\tilde{s}_X(v) = \Phi(K_X(\tau)), \quad K_X(\tau) = \Phi^{-1}(\tilde{s}_X(v)).$$

Пара приведенных соотношений, связывающих КФ сигнала и его спектр мощности посредством преобразования Фурье, носит название *теоремы Винера — Хинчина*. Для центрированного процесса $K_X(\tau) = R_X(\tau)$, и, как было показано выше (см. (2.20)),

$$\tilde{s}_X(v) = \Phi(R_X(\tau)) = M\left(\lim_{T \rightarrow \infty} \frac{S_{X,T}(v) \overline{S_{X,T}(v)}}{T} \right).$$

Если спектр мощности стационарного случайного сигнала постоянен, $\forall v \in \mathbb{R} : \tilde{s}_X(v) = C > 0$, то такой сигнал называют *белым шумом*. КФ (2.19) белого шума представляет собой δ -функцию: $K_X(\tau) = \Phi^{-1}\{\tilde{s}_X(v) = C\} = C\delta(\tau)$ (см. замечание к формуле (2.7)). Таким образом, даже как угодно близкие сечения белого шума некоррелированы, так как

$$\text{cov}(X(t), X(t - \tau)) = C\delta(\tau). \quad (2.22)$$

Белый шум представляет собой физически нереализуемый процесс, поскольку его дисперсия бесконечна: $\sigma_X^2 = C\delta(0)$, однако является полезной математической моделью, которая часто используется на практике.

Глава 3

Дискретизация и квантование сигналов. Дискретные ортогональные преобразования

■ 3.1. Преобразование непрерывных сигналов в дискретные

Под термином *сигнал* традиционно понимается физический процесс (например, изменяющееся во времени напряжение), отображающий некоторую информацию (сообщение). Математически сигнал описывается функцией $f(t)$ определенного вида. Характеризовать виды сигналов будем с использованием следующих понятий.

Аналоговый сигнал описывается непрерывной (или кусочно-непрерывной) функцией $f(t)$.

Дискретный сигнал — функция дискретного аргумента $y = y(nT)$ с областью определения $D = \{nT | n \in \mathbb{Z}\}$. Значения $y(nT)$ называются *отсчетами* дискретного сигнала, а величина T , представляющая собой расстояние по оси аргумента между соседними отсчетами, называется *периодом (шагом) дискретизации*. Будем использовать также обозначения $y(nT) = y(n) = y_n$. Таким образом, дискретный сигнал описывается последовательностью отсчетов $\{y_n\}$. Последовательность может быть и конечной, тогда будем использовать запись вида $y(n) = \{y_0, y_1, \dots, y_N\}$, например $y(n) = \{0; 0,5; -0,7; 1\}$.

Под *цифровым* сигналом будем понимать дискретный сигнал, который может принимать значения из конечного множества чисел, $\forall n \ y(n) \in \{h_1, \dots, h_M\}$. Возможные значения сигнала h_j ($j = 1, \dots, M$) называем *уровнями квантования* сигнала.

Реальные физические процессы почти всегда описываются аналоговыми сигналами $f(t)$, причем интервалы наблюдения конечны, $t \in [a, b]$. Для последующей цифровой обработки аналоговый сигнал, прежде всего, необходимо преобразовать в дискретный. Для этого в простейшем случае из аналогового сигнала $f(t)$

производится равномерная выборка с шагом дискретизации $T = \Delta t$: $f_n = f(a + n\Delta t)$, $n = 0, 1, \dots, N$. Для восстановления непрерывного (аналогового) сигнала по отсчетам $\{f_0, f_1, \dots, f_N\}$ чаще всего используются интерполяционные методы, например, восстановленный сигнал можно представить в виде интерполяционного многочлена $\tilde{f}(t) = \sum_{k=0}^N a_k t^k$, или в более общем виде как

$\tilde{f}(t) = \sum_{k=0}^N c_k \varphi_k(t)$, где $\{\varphi_k(t)\}$ — некоторая система функций на отрезке $t \in [a; b]$.

При дискретизации сигнала в первую очередь возникает вопрос: каким выбирать шаг аргумента Δt , т. е. как часто надо брать отсчеты для формирования выборки $\{f_i\}$. Очевидно, что для сокращения объема данных и, соответственно, упрощения обработки не следует выбирать отсчеты слишком часто, т. е. шаг дискретизации Δt следует сделать по возможности большим. С другой стороны, увеличивая Δt , мы можем неадекватно описать исходный аналоговый сигнал через дискретное представление, т. е. получим неприемлемую ошибку $\|f(t) - \tilde{f}(t)\|$ для восстановленного сигнала $\tilde{f}(t)$.

Для того чтобы сформулировать критерий выбора величины шага дискретизации Δt , необходимо задаться какими-либо ограничениями, накладываемыми на функцию $f(t)$.

■ 3.2. Дискретизация по критерию наибольшего отклонения

Для заданного способа R восстановления непрерывного сигнала по дискретному, $\tilde{f}(t) = R\{f_k\}$, сформулируем задачу выбора периода дискретизации Δt для формирования отсчетов $f_k = f(k\Delta t)$ в следующем виде: выбрать по заданному уровню ошибки ε_0 шаг Δt так, чтобы

$$\|f(t) - \tilde{f}(t)\| \leq \varepsilon_0. \quad (3.1)$$

В качестве нормы в выражении (3.1) обычно берется наибольшее отклонение $\|\varepsilon(t)\| = \max_{t \in [a; b]} |\varepsilon(t)|$, а для восстановления не-

прерывного сигнала $\tilde{f}(t)$ по дискретному $\{f_k\}$ используются степенные интерполяционные многочлены. Известно, что по $n+1$ точкам $\{(t_k, f_k)\}_{k=0}^{k=n}$ можно построить единственный интерполяционный многочлен n -й степени, $\tilde{f}(t) = P_n(t) = \sum_{k=0}^n a_k t^k$, а для ошибки интерполяции верно:

$$|\varepsilon(t)| = |f(t) - \tilde{f}(t)| \leq \frac{M_{n+1}}{(n+1)!} \left| \prod_{k=0}^n (t - t_k) \right|, \quad (3.2)$$

$$\|\varepsilon(t)\| = \max_{t \in [a; b]} |\varepsilon(t)| \leq \frac{M_{n+1}}{(n+1)!} \max_{t \in [a; b]} \left| \prod_{k=0}^n (t - t_k) \right|, \quad (3.3)$$

где $M_{n+1} = \max_{t \in [a; b]} |f^{(n+1)}(t)|$ — максимум модуля $(n+1)$ -й производной, $[a; b]$ — отрезок, на котором производится интерполяция.

Задание большого числа узлов интерполяции (т. е. высокой степени интерполяционного полинома) нередко влечет недопустимые для практических приложений задержки восстановления аналогового сигнала: поскольку для построения интерполяционного полинома требуется знание всех $n+1$ отсчетов сигнала, то задержка времени восстановления не может быть меньше $T = n\Delta t$. Кроме того, повышение порядка интерполяции (т. е. степени используемого интерполяционного полинома) влечет за собой и увеличение объема необходимых вычислений, ужесточая аппаратурные требования. Поэтому на практике обычно используются интерполяторы не более, чем третьего порядка.

Итак, положим, что восстановление аналогового сигнала происходит с использованием степенного интерполяционного полинома. Тогда на основании формул (3.1) и (3.3) получим следующий критерий выбора шага дискретизации:

$$\max_{t \in [a; b]} \left| \prod_{k=0}^n (t - t_k) \right| \leq \frac{(n+1)!}{M_{n+1}} \varepsilon_0. \quad (3.4)$$

Для использования данного критерия необходимо знать оценку максимума модуля $(n+1)$ -й производной интерполируемого сигнала.

Пример 3.1. Пусть для восстановления сигнала используется линейная интерполяция, т. е. порядок полинома $n = 1$. Известно значение $M_2 = \max_{t \in (-\infty; \infty)} |f''(t)|$, и задано ограничение на ошибку восстановления сигнала: $\max_{t \in (-\infty; \infty)} |f(t) - \tilde{f}(t)| \leq \varepsilon_0$. Каким следует выбрать шаг дискретизации Δt ?

◀ Все участки линейной интерполяции $[t_k; t_{k+1}]$ имеют одинаковую длину, равную $\Delta t = t_{k+1} - t_k$, и для каждого такого участка должно выполняться неравенство (3.4), т.е.

$$\max_{t \in [t_k; t_{k+1}]} |(t - t_k)(t - t_{k+1})| \leq \frac{2}{M_2} \varepsilon_0.$$

Так как

$$\max_{t \in [t_k; t_{k+1}]} |(t - t_k)(t - t_{k+1})| = \left(\frac{t_{k+1} - t_k}{2} \right)^2 = \frac{(\Delta t)^2}{4}, \text{ то } \Delta t \leq \sqrt{\frac{8\varepsilon_0}{M_2}}.$$

■ 3.3. Частотный критерий выбора шага дискретизации

Прежде чем формулировать соответствующий критерий для выбора шага дискретизации Δt , рассмотрим следующий пример.

Пример 3.2 (эффект наложения частот). Рассмотрим два аналоговых сигнала, представляющих собой гармонические колебания: $\varphi_1(t) = \cos(\pi(\beta + \alpha)t + \gamma)$, $\varphi_2(t) = \cos(\pi(\beta - \alpha)t - \gamma)$, где α, β, γ — любые вещественные числа. Тогда выбрав шаг дискретизации $\Delta t = 1/\beta$, мы не сможем по отсчетам в точках $t_k = k\Delta t$, $k \in \mathbb{Z}$, отличить сигналы $\varphi_1(t)$ и $\varphi_2(t)$, хотя гармонические колебания имеют различные частоты: $v_1 = \frac{\beta + \alpha}{2}$, $v_2 = \frac{\beta - \alpha}{2}$ ($\omega_1 = \pi(\beta + \alpha)$, $\omega_2 = \pi(\beta - \alpha)$).

◀ Так как $\cos x - \cos y = 2 \sin \frac{x+y}{2} \sin \frac{y-x}{2}$, то для отсчетов в точках $t_k = k\Delta t = k/\beta$ имеем:

$$\varphi_2(t_k) - \varphi_1(t_k) = 2 \underbrace{\sin(\pi\beta t_k)}_{\sin \pi k = 0} \sin(\pi\alpha t_k + \gamma) = 0.$$

Иллюстрация эффекта наложения частот для случая $\alpha = 2$, $\beta = 8$, $\gamma = 0$ приведена на рис. 3.1. ►

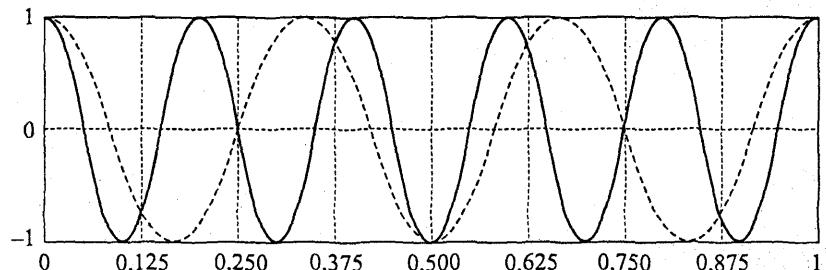


Рис. 3.1. Эффект наложения частот: функции $\varphi_1(t) = \cos(10\pi t)$ (сплошная линия) и $\varphi_2(t) = \cos(6\pi t)$ (пунктирная линия) совпадают в точках $t_k = k/8$, $k \in \mathbb{Z}$

Если спектр $S(v)$ функции $f(t)$ имеет ограниченный носитель, то ответ на вопрос о выборе величины шага дискретизации Δt дает следующая *теорема отсчетов*, известная в отечественной литературе как теорема Котельникова, а в зарубежной — как теорема Найквиста или теорема Шеннона.

Теорема 3.1. Если функция $f(t)$ имеет спектр с ограниченной частотной полосой, т. е. $S(v) = 0$ при $|v| \geq F$, то $f(t)$ может быть точно восстановлена по своим значениям в точках $t_k = k\Delta t$ ($k \in \mathbb{Z}$), где $\Delta t = \frac{1}{2F}$, с помощью ряда Котельникова:

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\Delta t) \frac{\sin(2\pi F(t - k\Delta t))}{2\pi F(t - k\Delta t)}. \quad (3.5)$$

◀ Выразим функцию через ее спектр, $f(t) = \int_{-F}^F S(v) e^{2\pi i vt} dv$,

тогда с использованием обозначения $\Delta t = \frac{1}{2F}$ получим $\forall k \in \mathbb{Z}$:

$$f(k\Delta t) = \int_{-F}^F S(v) e^{\frac{2\pi i v k}{2F}} dv.$$

Спектральную плотность $S(v)$, принимающую ненулевые значения только на интервале $v \in (-F; F)$, периодически продолжим (с периодом $2F$) на всю ось частот, и полученную в результате функцию $\tilde{S}(v)$ представим в виде ряда Фурье (2.3):

$$\tilde{S}(v) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{2\pi i v k}{2F}}, \text{ где } c_k = \frac{1}{2F} \int_{-F}^F S(v) e^{\frac{-2\pi i v k}{2F}} dv.$$

Сравнение полученной записи для c_k с выражением для $f(k\Delta t)$

показывает, что $c_k = \frac{1}{2F} f(-k\Delta t)$. Тогда

$$\tilde{S}(v) = \frac{1}{2F} \sum_{k=-\infty}^{\infty} f(-k\Delta t) e^{\frac{2\pi i v k}{2F}} = \frac{1}{2F} \sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-\frac{2\pi i v k \Delta t}{2F}},$$

$$f(t) = \int_{-F}^F S(v) e^{2\pi i vt} dv = \frac{1}{2F} \int_{-F}^F \left(\sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-\frac{2\pi i v k \Delta t}{2F}} \right) e^{2\pi i vt} dv.$$

Поменяв в последнем выражении порядок интегрирования и суммирования, получим:

$$f(t) = \frac{1}{2F} \sum_{k=-\infty}^{\infty} f(k\Delta t) \left(\int_{-F}^F e^{2\pi i v(t-k\Delta t)} dv \right).$$

Отсюда, учитывая, что $\int_{-F}^F e^{2\pi i v(t-k\Delta t)} dv = \frac{\sin(2\pi F(t-k\Delta t))}{\pi(t-k\Delta t)}$, полу-

чаем утверждение теоремы. ▶

Таким образом, с помощью формулы (3.5) для функции с ограниченным носителем спектра можно найти ее точное значение в любой точке числовой оси $t \in (-\infty; \infty)$ по отсчетам $f_k = f(t_k)$ в точках $t_k = k\Delta t$, поэтому отсчеты полностью определяют функцию. Формула (3.5) представляет собой функциональный ряд $f(t) = \sum_{k=-\infty}^{\infty} \lambda_k \phi_k(t)$ по системе функций

$$\left\{ \phi_k(t) = \frac{\sin(2\pi F(t - k\Delta t))}{2\pi F(t - k\Delta t)} \right\}_{k \in \mathbb{Z}}, \quad (3.6)$$

представляющих собой сдвиги по оси абсцисс на величины $k\Delta t = \frac{k}{2F}$ функции $\phi_0(t) = \phi(t) = \frac{\sin 2\pi F t}{2\pi F t}$. Базисные функции $\{\phi_k\}$ обращаются в ноль во всех точках $t_m = m\Delta t$ ($m \in \mathbb{Z}$), за исключением точки $t_k = k\Delta t$, для которой $\phi_k(k\Delta t) = 1$ (см. рис. 3.2), т. е. в представлении ряда (3.5) для $f(t)$ получаем, что в точках $t = t_m = m\Delta t$ в сумме (3.5) имеется единственное ненулевое слагаемое для $k = m$.

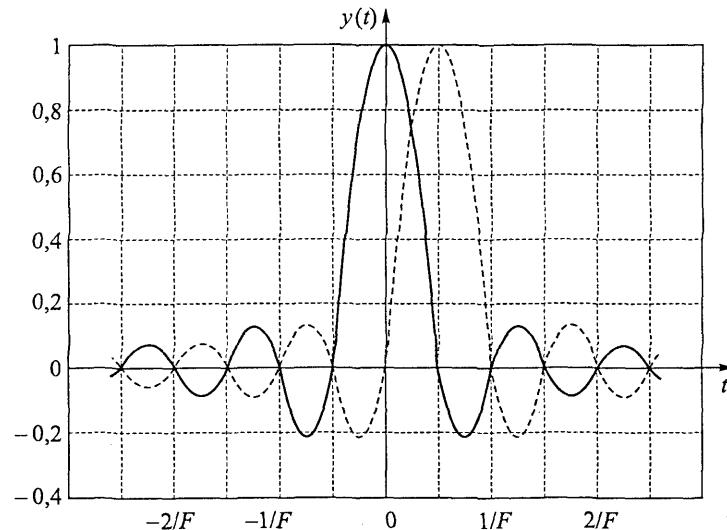


Рис. 3.2. Графики базисных функций $\phi_0(t)$ (сплошная линия) и $\phi_1(t)$ (пунктирная линия) ряда Котельникова (3.5)

Удобным свойством системы функций (3.6) является ее ортогональность на числовой оси $t \in (-\infty; \infty)$.

Теорема 3.2. Для системы функций (3.6) при $\Delta t = \frac{1}{2F}$ верно:

$$\langle \phi_k(t), \phi_m(t) \rangle = \int_{-\infty}^{\infty} \phi_k(t) \phi_m(t) dx = \begin{cases} \frac{1}{2F} & \text{при } k = m, \\ 0 & \text{при } k \neq m. \end{cases}$$

◀ Заметим, что спектр функции $\phi_0(t) = \phi(t) = \frac{\sin 2\pi F t}{2\pi F t}$ имеет вид:

$$S_\phi(v) = \begin{cases} \frac{1}{2F} & \text{при } v \in [-F; F]; \\ 0 & \text{при } v \notin [-F; F]. \end{cases}$$

Действительно, выполняя обратное преобразование Фурье, имеем:

$$\int_{-\infty}^{\infty} S_\phi(v) e^{2\pi i vt} dv = \frac{1}{2F} \int_{-F}^F e^{2\pi i vt} dv = \frac{\sin 2\pi F t}{2\pi F t}.$$

Теперь рассмотрим скалярное произведение

$$\langle \phi_k(t), \phi_m(t) \rangle = \int_{-\infty}^{\infty} \underbrace{\phi(t - k\Delta t)}_{\text{Обозн. } x} \phi(t - m\Delta t) dt = \int_{-\infty}^{\infty} \phi(x) \phi(x - (m - k)\Delta t) dx$$

и обозначим $u = (m - k)\Delta t$. В силу четности функции $\phi(x - u) = \phi(u - x)$, поэтому $g(u) = \langle \phi_k(t), \phi_m(t) \rangle = \int_{-\infty}^{+\infty} \phi(x) \phi(u - x) dx$ — свертка функций $\phi(t)$ и $\phi(t)$. Тогда в соответствии со свойством 6° интеграла Фурье для спектра функции $g(u)$ получим:

$$S_g(v) = S_\phi(v) S_\phi(v) = \frac{1}{2F} S_\phi(v)$$

и, в силу линейности преобразования Фурье (свойство 2°), $g(u) = \frac{1}{2F} \phi(u)$. Поэтому (см. рис. 3.2)

$$\langle \phi_k(t), \phi_m(t) \rangle = g((m - k)\Delta t) = \frac{1}{2F} \phi((m - k)\Delta t) = \begin{cases} \frac{1}{2F} & \text{при } k = m; \\ 0 & \text{при } k \neq m. \end{cases}$$

Таким образом, сдвиги $\phi(t)$ образуют функциональную систему, которая является ортогональным базисом для представления функций, спектр которых ограничен частотной полосой $v \in \left(-\frac{1}{2\Delta t}, \frac{1}{2\Delta t}\right)$. Коэффициенты Фурье ряда Котельникова (3.5),

т. е. коэффициенты разложения по базису $\{\phi_k\}$, находятся очень просто, равномерной выборкой отсчетов анализируемой функции: $\lambda_k = f(k\Delta t)$, $k \in \mathbb{Z}$. По этому набору отсчетов можно восстановить значение функции в любой точке числовой оси $t \in (-\infty; \infty)$ — соответствующее правило восстановления определяет ряд (3.5), который можно рассматривать как интерполяционную формулу, имеющую при правильном выборе периода дискретизации нулевую ошибку интерполяции для функций с ограниченным носителем спектра.

Итак, при шаге дискретизации $\Delta t = \frac{1}{2F}$ можно точно восстановить аналоговый сигнал по его отсчетам, если спектр сигнала ограничен: $\max_{v: S(v) \neq 0} |v| = F_{\max} < F$. Отсюда окончательно критерий

для выбора шага дискретизации аналогового сигнала при известной максимальной («верхней») частоте F_{\max} спектрального представления принимает вид:

$$\Delta t < \frac{1}{2F_{\max}}. \quad (3.7)$$

Величину $f_d = 1/\Delta t$ называют *частотой дискретизации*. В этих обозначениях условие (3.7) можно переписать в эквивалентном виде:

$$f_d > 2F_{\max}.$$

Неравенство (3.7) требует ряда комментариев. Прежде всего, следует заметить, что ограничение на полосу спектров реальных сигналов всегда является определенным допущением, так как для сигналов конечной длительности (с ограниченным носителем во временной области) в частотной области спектр имеет неограниченный носитель $\Omega_v = (-\infty; \infty)$, как это следует из теоремы 2.3.

Чаще всего для определения ширины полосы спектра верхнюю частоту F_{\max} выбирают так, чтобы в энергетическом смысле (см. раздел 2.3) ограничение на ширину полосы не вносило

больших искажений: $\int_{-F_{\max}}^{F_{\max}} |S(v)|^2 dv = \beta \int_{-\infty}^{\infty} |S(v)|^2 dv$, принимая

значение β близким к единице (например, $\beta = 0,95$). Вне зависимости от того из каких соображений была определена верхняя частота F_{\max} , перед дискретизацией сигнала целесообразно проводить низкочастотную фильтрацию с заданной частотой среза (полного подавления) $f_{\text{реза}} = F_{\max}$. В противном случае при восстановлении сигнала могут возникнуть недопустимые искажения.

Пример 3.3. Для представления речи достаточно полосы звуковых колебаний с верхней частотой $F_{\max} = 3500$ Гц. В этом случае используется стандартная частота дискретизации $f_d = 1/\Delta t = 8000$ Гц, что удовлетворяет условию (3.7). Однако перед дискретизацией звуковой сигнал обязательно надо подвергнуть НЧ-фильтрации с частотой полного подавления

$$f_{\text{реза}} \leq \frac{1}{2\Delta t} = \frac{f_d}{2} = 4000 \text{ Гц.}$$

В противном случае может возникнуть эффект наложения (смешения) частот. Например, если в звуковом сигнале имелась гармоническая составляющая (тон) частотой $f = 6000$ Гц и предварительная НЧ-фильтрация не выполнялась, то при восстановлении сигнала (обратном преобразовании из дискретного вида в аналоговый по формуле (3.5)) может появиться тон с частотой $f = 2000$ Гц, которого, возможно, в исходном звуковом сигнале вообще не было!

◀ См. пример 3.2, положив $\beta = 8000$, $\alpha = 4000$. ▶

Описанный эффект наложения частот (появления посторонних тонов) можно самостоятельно смоделировать с использованием мультимедийных средств персональных компьютеров.

■ 3.4. Спектр дискретного сигнала

Итак, выбрав определенным образом шаг дискретизации Δt , можно получить из аналогового сигнала дискретный: $f(t) \rightarrow \{f_k = f(k\Delta t)\}$. Каким образом распространить понятие «частотный спектр» на дискретный сигнал — набор отсчетов $\{f_k\}$? Естественно попытаться связать данное понятие с непрерывным сигналом $\tilde{f}(t)$, который можно восстановить по дискретному с использованием формулы (3.5): $\{f_k\} \rightarrow \tilde{f}(t)$.

Если исходный непрерывный сигнал $f(t)$ имел ограниченную частотную полосу и интервал дискретизации был выбран в соответствии с условием (3.7), то использование формулы (3.5) дает точное восстановление аналогового сигнала, т. е. $f(t) = \tilde{f}(t)$, поэтому в силу свойств 2° и 4° преобразования Фурье (см. раздел 2.1) спектр $S(v) = \Phi\{f(t)\} = \Phi\{\tilde{f}(t)\}$ можно записать так:

$$\begin{aligned} S(v) &= \sum_{k=-\infty}^{\infty} f(k\Delta t) \Phi\left\{\frac{\sin 2\pi F(t - k\Delta t)}{2\pi F(t - k\Delta t)}\right\} = \\ &= \Phi\left\{\frac{\sin 2\pi F t}{2\pi F t}\right\} \sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-2\pi i v k \Delta t}. \end{aligned}$$

Так как (см. доказательство теоремы 3.2)

$$\Phi\left\{\frac{\sin 2\pi F t}{2\pi F t}\right\} = \begin{cases} \frac{1}{2F}, & v \in [-F; F]; \\ 0, & v \notin [-F; F]; \end{cases} \quad \frac{1}{2F} = \Delta t,$$

то с использованием обозначения для функции «окна»

$$\Pi_F(v) = \begin{cases} 1, & v \in [-F; F]; \\ 0, & v \notin [-F; F] \end{cases}$$

имеем:

$$S(v) = \Pi_F(v) \sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-2\pi i v k \Delta t} \Delta t = \Pi_F(v) \tilde{S}(v), \quad (3.8)$$

$$\tilde{S}(v) = \sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-2\pi i v k \Delta t} \Delta t. \quad (3.9)$$

Выражение (3.9) уже встречалось при доказательстве теоремы 3.1 и представляет собой периодическое продолжение ограниченного спектра $S(v)$ на всю частотную ось. Получаем, что для функции со спектром, ограниченным частотной полосой $|v| < F$, выражение (3.9) для $v \in [-F; F]$ совпадает (см. (3.8)) с интегралом Фурье $S(v) = \int_{-\infty}^{+\infty} f(t) e^{-2\pi i v t} dt$. Заметим также, что выражение (3.9) представляет собой квадратурную формулу прямого

угольников для интеграла Фурье, в которой узлы взяты в точках $t_k = k\Delta t$ — отсчетах дискретного сигнала.

Выражение (3.9) будем называть *спектром дискретного сигнала* $\{f_k\}$. Действительно, если преобразовать дискретный сигнал $\{f(k\Delta t)\}$ в решетчатую функцию

$$\tilde{f}(t) = f(t)\Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t) = \Delta t \sum_{k=-\infty}^{\infty} f(k\Delta t) \delta(t - k\Delta t),$$

то для обобщенного преобразования Фурье получим: $\Phi\{\tilde{f}\} = \tilde{S}(v)$. Отличие выражения (3.8) от (3.9) состоит в том, что спектр дискретного сигнала (3.9) является периодическим продолжением спектра непрерывного сигнала $S(v)$ из области $v \in [-F; F]$ на всю частотную ось, что мы уже наблюдали при доказательстве теоремы 3.1.

Действительно, решетчатую функцию $\tilde{f}(t) = f(t)\Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t)$

на основании соотношения (2.9) можно представить в виде

$$\tilde{f}(t) = f(t) \sum_{k=-\infty}^{\infty} \exp\left(i \frac{2\pi k}{\Delta t} t\right),$$

а ее спектр (2.6), с учетом свойств 2° и 5° преобразования Фурье, записать так:

$$\tilde{S}(v) = \Phi\{\tilde{f}(t)\} = \sum_{k=-\infty}^{\infty} \Phi\left\{f(t) \exp\left(i \frac{2\pi k}{\Delta t} t\right)\right\} = \sum_{k=-\infty}^{\infty} S\left(v - \frac{k}{\Delta t}\right). \quad (3.10)$$

То есть спектр дискретного сигнала представляет собой сумму сдвинутых копий спектра непрерывного сигнала. Как видно из выражений (3.9) и (3.10), $\tilde{S}(v) = \tilde{S}\left(v + \frac{1}{\Delta t}\right)$, $\Delta t = \frac{1}{2F}$ (см. рис. 3.3).

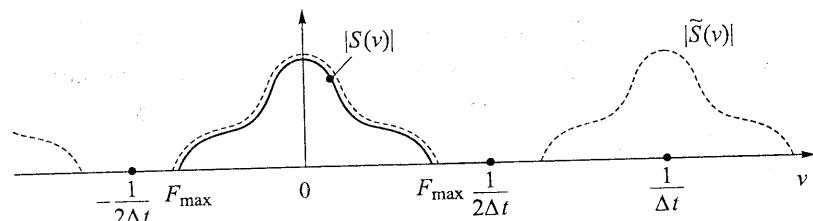


Рис. 3.3. Правильный выбор частоты дискретизации: $f_d = \frac{1}{\Delta t} > 2F_{\max}$

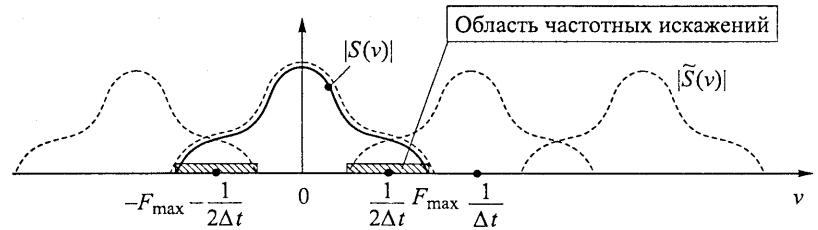


Рис. 3.4. Неправильный выбор частоты дискретизации: $f_d = \frac{1}{\Delta t} < 2F_{\max}$

Если частота дискретизации была выбрана с нарушением условия (3.7), то период функции (3.9), (3.10) окажется меньше полосы $[-F; F]$ исходного сигнала и $S(v) \neq \Pi_F(v)\tilde{S}(v)$, при этом копии спектра непрерывного сигнала перекрываются и возникают необратимые искажения сигнала, связанные с эффектом наложения частот (см. рис. 3.4 и примеры 3.2, 3.3). Вследствие этого при восстановлении непрерывного сигнала по отсчетам $\{f(k\Delta t)\}$ с помощью формулы (3.5) произойдет «подмена» гармоник спектра с частотами $v = \frac{f_d + \alpha}{2}$ на гармоники с частотой $v' = \frac{f_d - \alpha}{2}$.

Спектр дискретного сигнала, определяемый по формулам (3.9) и (3.10), соответствует обобщенному спектру решетчатой функции $\tilde{f}(t) = f(t)\Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t)$. На практике при дискретизации сигнала приходится иметь дело со стробирующими (дискретизирующими) импульсами конечной длительности, отличными от δ -функции. Рассмотрим более общий случай — функцию $\hat{f}(t) = f(t)\Delta t \sum_{k=-\infty}^{\infty} \sigma(t - k\Delta t)$, где $\sigma(t)$ — стробирующий импульс, длительность которого τ — конечна, но достаточно мала для того, чтобы считать дискретизируемую функцию $f(t)$ постоянной на любом интервале длительности τ : $t \in [(k - \frac{1}{2})\tau; (k + \frac{1}{2})\tau]$,

$k \in \mathbb{Z}$. Тогда снова можно записать: $\hat{f}(t) = \Delta t \sum_{k=-\infty}^{\infty} f(k\Delta t) \sigma(t - k\Delta t)$.

С использованием свойств 2° и 4° преобразования Фурье получаем:

$$\Phi\{\tilde{f}(t)\} = \Delta t \sum_{k=-\infty}^{\infty} f(k\Delta t) S_{\sigma}(\nu) e^{-2\pi i \nu k \Delta t} = S_{\sigma}(\nu) \tilde{S}(\nu),$$

где $\tilde{S}(\nu)$ — исходный спектр дискретного сигнала (3.9), а $S_{\sigma}(\nu) = \Phi\{\sigma(t)\}$ — спектр стробирующего импульса. Видим, что отличие формы дискретизирующего импульса от δ -функции (когда $S_{\sigma}(\nu) \neq 1$) влечет мультиплексивные искажения в спектре дискретного сигнала. Данными искажениями можно пренебречь, если для полосы частот $\nu \in \left[0; \frac{1}{2\Delta t}\right]$ (которая характеризует непрерывный сигнал, см. рис. 3.3) имеем $S_{\sigma}(\nu) \approx 1$. Это достигается, если длительность стробирующего импульса τ выбрана значительно меньше периода дискретизации Δt .

Пример 3.4. Оценить искажения спектра дискретного сигнала, которые обусловлены прямоугольным стробирующим импульсом

$$\sigma(t) = \begin{cases} 1/\tau, & t \in [-\tau/2; \tau/2]; \\ 0, & t \notin [-\tau/2; \tau/2] \end{cases} \quad \text{для } \tau = \Delta t/3, \tau = \Delta t/10.$$

◀ $S_{\sigma}(\nu) = \frac{\sin \pi \nu \tau}{\pi \nu \tau}$ (убедитесь). В полосе частот $\nu \in [0; 1/\tau]$ функция $S_{\sigma}(\nu)$ монотонно убывает, причем $S_{\sigma}(0) = 1$, $S_{\sigma}(1/\tau) = 0$. Максимальное отклонение $S_{\sigma}(\nu)$ от единицы, определяемое минимальным значением $S_{\sigma}(\nu)$ в интересующей нас области $\nu \in \left[0; \frac{1}{2\Delta t}\right]$, будет наблюдаться для верхней частоты $\nu = \frac{1}{2\Delta t}$, поэтому максимальные искажения оцениваем по величине

$$s_{\min} = S_{\sigma}\left(\frac{1}{2\Delta t}\right) = \frac{\sin \frac{\pi \tau}{2\Delta t}}{\frac{\pi \tau}{2\Delta t}}.$$

Для $\tau = \Delta t/3$ получим $s_{\min} \approx 0,955$; для $\tau = \Delta t/10$ — $s_{\min} \approx 0,996$. ►

■ 3.5. Дискретное преобразование Фурье

Положим, что сигнал $f(t)$ наблюдается на конечном интервале, т. е. после дискретизации (с должным выбором шага Δt) получим и конечный набор отсчетов $\{f_k = f(k\Delta t)\}_{k=0}^{N-1}$. Тогда, согласно (3.9), получим спектр

$$\tilde{S}(\nu) = \Delta t \sum_{k=0}^{N-1} f_k e^{-2\pi i \nu k \Delta t}, \quad (3.11)$$

который соответствует сигналу (решетчатой функции) $\Phi^{-1}\{\tilde{S}(\nu)\} = \tilde{f}(t) = \Delta t \sum_{k=0}^{N-1} f_k \delta(t - k\Delta t)$. Так как каждое слагаемое в формуле (3.11) непрерывно по переменной ν , то $\tilde{S}(\nu)$ — также непрерывная функция. На ее периоде $\nu \in [0; 1/\Delta t]$ по оси частоты возьмем отсчеты $\tilde{s}_j = \tilde{S}(\nu_j)$ в точках $\nu_j = j/(N\Delta t)$, $j = 0, \dots, N-1$. Тогда $\nu_j = j\Delta\nu$, где $\Delta\nu = \nu_{j+1} - \nu_j = 1/(N\Delta t)$ — шаг частотной дискретизации. Рассмотрим полученный набор частотных отсчетов $\{\tilde{s}_j = \tilde{S}(\nu_j)\}_{j=0}^{N-1}$ (см. (3.11)):

$$\tilde{s}_j = \Delta t \sum_{k=0}^{N-1} f_k e^{-\frac{2\pi i}{N} kj}, \quad j = 0, 1, \dots, N-1. \quad (3.12)$$

Вектор (вообще говоря, с комплексными компонентами) $(\tilde{s}_0, \dots, \tilde{s}_{N-1})$ назовем *дискретным спектром*, или *дискретным преобразованием Фурье* (ДПФ) числового вектора (f_0, \dots, f_{N-1}) . Заметим, что можно рассматривать значения частотных отсчетов (3.12) и для любого целого j , при этом $\tilde{s}_j = \tilde{s}_{j+N}$.

Выполнив дискретизацию непрерывного по аргументу спектра (3.11) с шагом $\Delta\nu = \frac{1}{N\Delta t}$, представим дискретный спектр

$(\tilde{s}_0, \dots, \tilde{s}_{N-1})$ в виде решетчатой функции, аналогично тому, как это было проделано в разделе 3.4 с сигналом $f(t)$ во временной

области. С учетом периодичности дискретного спектра ($\tilde{s}_k = \tilde{s}_{k+N}$) имеем:

$$\begin{aligned}\tilde{S}(\nu) &= \Delta\nu \sum_{k=-\infty}^{\infty} \tilde{s}_k \delta(\nu - k\Delta\nu) = \tilde{S}(\nu) \Delta\nu \sum_{k=-\infty}^{\infty} \delta(\nu - k\Delta\nu) = \\ &= \tilde{S}(\nu) \sum_{k=-\infty}^{\infty} \exp\left(\frac{2\pi i}{\Delta\nu} kv\right),\end{aligned}$$

где последний переход осуществлен на основании формулы (2.9). Полученному дискретному спектру (решетчатой функции) во временной области соответствует сигнал (см. раздел 2.1, свойство 4° преобразования Фурье)

$$\begin{aligned}\tilde{f}(t) &= \Phi^{-1}\{\tilde{S}(\nu)\} = \Phi^{-1}\left\{\sum_{k=-\infty}^{\infty} \tilde{s}_k \exp\left(\frac{2\pi i}{\Delta\nu} kv\right)\right\} = \\ &= \sum_{k=-\infty}^{\infty} \Phi^{-1}\left\{\tilde{s}_k \exp\left(\frac{2\pi i}{\Delta\nu} kv\right)\right\} = \sum_{k=-\infty}^{\infty} \tilde{f}\left(t + \frac{k}{\Delta\nu}\right) = \sum_{k=-\infty}^{\infty} \tilde{f}(t + kN\Delta t),\end{aligned}$$

где $\tilde{f}(t) = \Delta t \sum_{k=0}^{N-1} f(k\Delta t) \delta(t - k\Delta t)$. Таким образом, дискретизация спектра влечет периодическое продолжение с периодом $T = N\Delta t$ конечного во временной области дискретного сигнала $\tilde{f}(t)$ на всю числовую ось, т. е. исходно конечная последовательность $\{f_k = f(k\Delta t)\}_{k=0}^{N-1}$ становится бесконечной: $\{f_k = f(k\Delta t)\}_{k=-\infty}^{\infty}$, $f_k = f_{k+N}$.

Замечание. Величина $T = N\Delta t$ представляет собой длину интервала наблюдения дискретного сигнала во временной области. С использованием данного обозначения формула для ДПФ (3.12) принимает вид:

$$\tilde{s}_j = \sum_{k=0}^{N-1} f(k\Delta t) e^{-\frac{2\pi i}{T} j(k\Delta t)} \Delta t.$$

Таким образом, с точностью до множителя $1/T$ ДПФ (3.12) представляется собой квадратурную формулу прямоугольников для ин-

теграла (2.4) — коэффициента c_j ряда Фурье (2.3) функции периода T :

$$c_j = \frac{1}{T} \int_0^T f(t) \exp\left(-i \frac{2\pi j}{T} t\right) dt \approx \frac{1}{T} \sum_{k=0}^{N-1} f(k\Delta t) e^{-\frac{2\pi i}{T} j(k\Delta t)} \Delta t = \frac{1}{T} \tilde{s}_j.$$

Приближенное равенство здесь можно записать, когда период $T_j = T/j$ гармонического колебания $\omega_j(t) = \exp\left(-i \frac{2\pi j}{T} t\right)$ много больше интервала дискретизации Δt : $T_j = T/j = N\Delta t/j \gg \Delta t$, т. е. когда $j \ll N$.

Теорема 3.3 (обратное ДПФ). Вектор (f_0, \dots, f_{N-1}) можно восстановить по дискретному спектру (3.12) с помощью обратного ДПФ (ОДПФ), которое определяется формулой

$$\begin{aligned}f_k &= \Delta\nu \sum_{j=0}^{N-1} \tilde{s}_j e^{\frac{2\pi i}{N} kj}, \\ k &= 0, 1, \dots, N-1,\end{aligned}\tag{3.13}$$

$$\text{где } \Delta\nu = \frac{1}{N\Delta t}.$$

◀ Рассмотрим правую часть равенства (3.13):

$$a(k) = \Delta\nu \sum_{j=0}^{N-1} \tilde{s}_j e^{\frac{2\pi i}{N} kj}.$$

Имеем (см. (3.12)):

$$a(k) = \Delta\nu \sum_{j=0}^{N-1} \left(\Delta t \sum_{m=0}^{N-1} f_m e^{-\frac{2\pi i}{N} mj} \right) e^{\frac{2\pi i}{N} kj} = \underbrace{\Delta\nu \Delta t}_{1/N} \sum_{m=0}^{N-1} f_m \sum_{j=0}^{N-1} e^{\frac{2\pi i}{N} j(k-m)}.$$

обозначим $v_{k,m}$

Очевидно, что при $k = m$ имеем $v_{k,m} = N$. Если же $k \neq m$, то для всех возможных значений индексов $k, m = 0, \dots, N-1$ получаем, что величина $q = e^{\frac{2\pi i}{N}(k-m)} \neq 1$, тогда по формуле для суммы геометрической прогрессии при $k \neq m$ имеем:

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

Поэтому окончательно $\forall k, m = 0, \dots, N-1$: $v_{k,m} = \begin{cases} N & \text{при } k = m; \\ 0 & \text{при } k \neq m \end{cases}$
и $a(k) = \frac{1}{N} \sum_{m=0}^{N-1} f_m v_{k,m} = f_k$. ►

Из теоремы 3.3 следует, что если дискретный сигнал описывается конечной последовательностью $\{f_0, \dots, f_{N-1}\}$, то восстановить его по спектру $\tilde{S}(v)$, который представляет собой непрерывную периодическую функцию (см. (3.11)), можно, взяв лишь конечное число отсчетов спектральной плотности: $\tilde{s}_j = \tilde{S}\left(\frac{j}{N\Delta t}\right)$, $j = 0, \dots, N-1$.

Итак, формулы (3.12) и (3.13) — взаимно однозначные преобразования векторов (вообще говоря, комплекснозначных): $(f_0, \dots, f_{N-1}) \leftrightarrow (\tilde{s}_0, \dots, \tilde{s}_{N-1})$. Если теперь отвлечься от природы этих векторов (временных и частотных отсчетов), то величины Δt в (3.12) и Δv в (3.13) есть просто некоторые масштабные множители, такие, что $\Delta t \Delta v = 1/N$. Для симметрии положим $\Delta t = \Delta v = 1/\sqrt{N}$ и парой ДПФ-ОДПФ будем называть преобразования векторов $(x_0, \dots, x_{N-1}) \leftrightarrow (y_0, \dots, y_{N-1})$, где ДПФ и ОДПФ определяются соответственно:

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N} kj}, \quad (3.14)$$

$$k = 0, \dots, N-1;$$

$$x_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} y_k e^{\frac{2\pi i}{N} kj}, \quad (3.15)$$

$$j = 0, \dots, N-1.$$

Как уже отмечалось, в радиотехнике и электронике частотное описание аналоговых сигналов во многих случаях более предпочтительно. Мы показали, что спектр дискретного сигнала опреде-

ляется выражением (3.9), причем для конечных интервалов наблюдения сигнала последний можно выразить через ОДПФ (3.15) частотных отсчетов спектральной функции (3.9), которые находятся с помощью ДПФ (3.14). Таким образом, определяемый через ДПФ (3.14) дискретный спектр является аналогом интегрального преобразования Фурье (2.6). Точно также, термин «амплитудный спектр» употребляется для модулей $\{|y_k|\}_{k=0}^{N-1}$ спектральных коэффициентов (3.14), «фазовый спектр» — для аргументов, взятых с противоположным знаком, $\{-\arg y_k\}_{k=0}^{N-1}$.

Пусть $\mathbf{Y} = (y_0, \dots, y_{N-1})$ — дискретный спектр, полученный из вектора $\mathbf{X} = (x_0, \dots, x_{N-1})$ в результате ДПФ (3.14), $\mathbf{X} \leftrightarrow \mathbf{Y}$. Отметим ряд важных свойств ДПФ.

1°. Сопряженная симметрия. Для любого вещественного вектора $\mathbf{X} = \bar{\mathbf{X}}$ в векторе-спектре $y_0 = \bar{y}_0$ (компоненты вещественная), а $\forall k = 1, \dots, N-1$: $y_k = \bar{y}_{N-k}$.

◀ Докажите самостоятельно. ►

2°. Линейность. $\forall \mathbf{X}_1 \leftrightarrow \mathbf{Y}_1, \forall \mathbf{X}_2 \leftrightarrow \mathbf{Y}_2, \forall \alpha, \beta$:

$$\mathbf{X} = \alpha \mathbf{X}_1 + \beta \mathbf{X}_2 \leftrightarrow \mathbf{Y} = \alpha \mathbf{Y}_1 + \beta \mathbf{Y}_2.$$

3°. Циклический сдвиг вправо. Сдвигу компонент вектора-сигнала

$$\mathbf{X} = (x_0, x_1, \dots, x_{N-1}) \rightarrow \mathbf{X}' = (x_{N-1}, x_0, x_1, \dots, x_{N-2})$$

соответствует такое изменение в области дискретного спектра

$$\mathbf{Y} = (y_0, y_1, \dots, y_{N-1}) \rightarrow \mathbf{Y}' = (y'_0, y'_1, \dots, y'_{N-1}),$$

$$\text{что } \forall k: y'_k = e^{-\frac{i2\pi k}{N}} y_k.$$

$$\blacktriangleleft y'_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x'_j e^{-\frac{2\pi i}{N} kj} = \frac{1}{\sqrt{N}} \left(x_{N-1} + \sum_{j=0}^{N-2} x_j e^{-\frac{2\pi i}{N} k(j+1)} \right) =$$

$$\frac{e^{-\frac{2\pi i}{N} k}}{\sqrt{N}} \left(x_{N-1} e^{-\frac{2\pi i}{N} (N-1)k} + \sum_{j=0}^{N-2} x_j e^{-\frac{2\pi i}{N} kj} \right) = \frac{e^{-\frac{2\pi i}{N} k}}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N} kj} = e^{-\frac{2\pi i}{N} k} y_k. \blacktriangleright$$

Следствие. Циклический сдвиг вправо на m позиций. При изменении вектора-сигнала

$$\mathbf{X} = (x_0, x_1, \dots, x_{N-1}) \rightarrow \mathbf{X}' = (x_{N-m}, \dots, x_{N-2}, x_{N-1}, x_0, \dots, x_{N-m-1})$$

в векторе-спектре \mathbf{Y}' имеем

$$\forall k: y'_k = e^{-\frac{2\pi k}{N}m} y_k.$$

4°. Циклический сдвиг влево на m позиций. Сдвигу компонент вектора:

$$\mathbf{X} = (x_0, x_1, \dots, x_{N-1}) \rightarrow \mathbf{X}' = (x_m, x_{m+1}, \dots, x_{N-1}, x_0, x_1, \dots, x_{m-1})$$

соответствует такое изменение в области дискретного спектра

$$\mathbf{Y} = (y_0, y_1, \dots, y_{N-1}) \rightarrow \mathbf{Y}' = (y'_0, y'_1, \dots, y'_{N-1}),$$

$$\text{что } \forall k: y'_k = e^{\frac{i2\pi k}{N}m} y_k.$$

◀ Докажите самостоятельно. ►

Следствие из свойств 3° и 4°. Инвариантность амплитудного спектра относительно циклических сдвигов, $\forall k: |y_k| = |y'_k|$.

Определение. Под дискретной сверткой векторов $\mathbf{A} = (a_0, a_1, \dots, a_{N-1})$ и $\mathbf{B} = (b_0, b_1, \dots, b_{N-1})$ будем понимать вектор $\mathbf{C} = (c_0, c_1, \dots, c_{2N-1})$ вдвое большей размерности, компоненты которого определяются по формуле

$$c_k = \sum_{j=0}^{N-1} a_j b_{k-j} = \sum_{j=0}^{N-1} a_{k-j} b_j \quad (3.16)$$

($k = 0, 1, \dots, 2N-1$), где предполагается, что $a_m = b_m = 0$, если значение индекса m выходит за допустимые границы (при $m < 0$ или $m > N-1$).

5°. Свойство ДПФ свертки векторов.

Обозначим $\check{\mathbf{A}} = (a_0, \dots, a_{N-1}, \underbrace{0, \dots, 0}_{N \text{ нулей}})$, $\check{\mathbf{B}} = (b_0, \dots, b_{N-1}, \underbrace{0, \dots, 0}_{N \text{ нулей}})$,

$\hat{\mathbf{A}} = (\hat{a}_0, \hat{a}_1, \dots, \hat{a}_{2N-1})$ — ДПФ вектора $\check{\mathbf{A}}$,

$$\hat{a}_k = \frac{1}{\sqrt{2N}} \sum_{j=0}^{2N-1} \tilde{a}_j e^{-\frac{2\pi i}{2N}kj},$$

$\check{\mathbf{B}} = (\hat{b}_0, \hat{b}_1, \dots, \hat{b}_{2N-1})$ — ДПФ вектора $\check{\mathbf{B}}$,

$$\hat{b}_k = \frac{1}{\sqrt{2N}} \sum_{j=0}^{2N-1} \tilde{b}_j e^{-\frac{2\pi i}{2N}kj}.$$

Тогда для вектора \mathbf{C} — свертки (3.16) векторов \mathbf{A} и \mathbf{B} — ДПФ $\hat{\mathbf{C}} = (\hat{c}_0, \hat{c}_1, \dots, \hat{c}_{2N-1})$ можно найти перемножением компонент векторов $\hat{\mathbf{A}}$ и $\hat{\mathbf{B}}$, $\forall k: \hat{c}_k = \sqrt{2N} \hat{a}_k \hat{b}_k$.

◀ Обозначим $\omega = e^{-\frac{2\pi i}{2N}}$. Для ДПФ вектора-свертки имеем:

$$\begin{aligned} \hat{c}_k &= \frac{1}{\sqrt{2N}} \sum_{j=0}^{2N-1} c_j \omega^{kj} = \frac{1}{\sqrt{2N}} \sum_{j=0}^{2N-1} \left(\sum_{m=0}^{N-1} a_m b_{j-m} \right) \omega^{kj} = \frac{1}{\sqrt{2N}} \sum_{m=0}^{N-1} a_m \sum_{j=0}^{2N-1} b_{j-m} \omega^{kj} = \\ &= \frac{1}{\sqrt{2N}} \sum_{m=0}^{N-1} a_m \sum_{l=-m}^{2N-1-m} b_l \omega^{k(l+m)} = \frac{1}{\sqrt{2N}} \sum_{m=0}^{N-1} a_m \omega^{km} \sum_{l=0}^{N-1} b_l \omega^{kl}, \end{aligned}$$

поскольку $b_l = 0$ при $l < 0$ или $l > N-1$. Далее замечаем, что $\sum_{l=0}^{N-1} b_l \omega^{kl} = \sum_{l=0}^{2N-1} b_l \omega^{kl} = \sqrt{2N} \hat{b}_k$ и, аналогично, $\sum_{m=0}^{N-1} a_m \omega^{km} = \sqrt{2N} \hat{a}_k$. ►

6°. Совмещенное ДПФ вещественных векторов.

Пусть вектор $\mathbf{B} = (b_0, \dots, b_{N-1})$ представляет собой ДПФ (3.14) вектора $\mathbf{A} = (a_0, \dots, a_{N-1})$, а вектор $\mathbf{V} = (v_0, \dots, v_{N-1})$ — ДПФ (3.14) вектора $\mathbf{U} = (u_0, \dots, u_{N-1})$, причем компоненты векторов \mathbf{A} и \mathbf{U} — вещественные. Тогда, вычислив ДПФ (3.14) вектора $\mathbf{X} = \mathbf{A} + i\mathbf{U}$, получим спектральный вектор \mathbf{Y} такой, что

$$b_0 = \operatorname{Re} y_0, \quad v_0 = \operatorname{Im} y_0; \quad \forall k = 1, \dots, N-1:$$

$$b_k = \frac{\operatorname{Re} y_k + \operatorname{Re} y_{N-k}}{2} + i \frac{\operatorname{Im} y_k - \operatorname{Im} y_{N-k}}{2};$$

$$v_k = \frac{\operatorname{Im} y_k + \operatorname{Im} y_{N-k}}{2} + i \frac{\operatorname{Re} y_{N-k} - \operatorname{Re} y_k}{2}.$$

◀ В силу линейности ДПФ (свойство 2°) имеем:

$$\mathbf{Y} = \mathbf{B} + i\mathbf{V} = (\operatorname{Re} \mathbf{B} - \operatorname{Im} \mathbf{V}) + i(\operatorname{Re} \mathbf{V} + \operatorname{Im} \mathbf{B}).$$

Отсюда на основании свойства 1° ДПФ получаем:

$$b_0 = \operatorname{Re} y_0, \quad v_0 = \operatorname{Im} v_0;$$

$$\forall k = 1, \dots, N-1:$$

$$y_k = \underbrace{(\operatorname{Re} b_k - \operatorname{Im} v_k)}_{\operatorname{Re} y_k} + i \underbrace{(\operatorname{Re} v_k + \operatorname{Im} b_k)}_{\operatorname{Im} y_k},$$

$$y_{N-k} = \underbrace{(\operatorname{Re} b_k + \operatorname{Im} v_k)}_{\operatorname{Re} y_{N-k}} + i \underbrace{(\operatorname{Re} v_k - \operatorname{Im} b_k)}_{\operatorname{Im} y_{N-k}},$$

откуда $\forall k = 1, \dots, N-1$:

$$\operatorname{Re} b_k = \frac{1}{2}(\operatorname{Re} y_k + \operatorname{Re} y_{N-k}); \quad \operatorname{Im} b_k = \frac{1}{2}(\operatorname{Im} y_k - \operatorname{Im} y_{N-k});$$

$$\operatorname{Re} v_k = \frac{1}{2}(\operatorname{Im} y_k + \operatorname{Im} y_{N-k}); \quad \operatorname{Im} v_k = \frac{1}{2}(\operatorname{Re} y_{N-k} - \operatorname{Re} y_k). \blacktriangleright$$

ДПФ (3.14) и ОДПФ (3.15) удобно записывать в матричном виде, представив \mathbf{X} и \mathbf{Y} как векторы-столбцы: $\mathbf{X} = (x_0, \dots, x_{N-1})^T$, $\mathbf{Y} = (y_0, \dots, y_{N-1})^T$. Тогда ДПФ и ОДПФ, соответственно, принимают вид: $\mathbf{Y} = \mathbf{W}\mathbf{X}$, $\mathbf{X} = \bar{\mathbf{W}}\mathbf{Y}$, где матрица ДПФ

$$\mathbf{W} = \left\{ w_{k,j} = \frac{1}{\sqrt{N}} e^{-\frac{2\pi i}{N} kj} \right\}_{k,j=0}^{N-1},$$

а матрица ОДПФ (убедитесь!)

$$\mathbf{W}^{-1} = \bar{\mathbf{W}} = \left\{ \bar{w}_{k,j} = \frac{1}{\sqrt{N}} e^{\frac{2\pi i}{N} kj} \right\}_{k,j=0}^{N-1}.$$

Определение. Комплексная матрица \mathbf{W} называется *унитарной*, если $\mathbf{W}^{-1} = \mathbf{W}^*$ (здесь $\mathbf{W}^* = \bar{\mathbf{W}}^T$, матрица \mathbf{W}^* называется *сопряженной* к матрице \mathbf{W}). Соответствующее дискретное преобразование $\mathbf{B} = \mathbf{W}\mathbf{A}$ также называется *унитарным*.

Очевидно, ДПФ — унитарное преобразование. В частном случае, для вещественных унитарных преобразований, $\mathbf{W}^{-1} = \mathbf{W}^T$, и вместо понятия «унитарность» используется термин «ортогональность». Важнейшее свойство унитарных (ортогональных) преобразований — сохранение евклидовой нормы векторов — раскрывается в следующей лемме.

Лемма 3.1. Матрица \mathbf{W} унитарна тогда и только тогда, когда для любого вектора \mathbf{A} его евклидова норма равна норме вектора $\mathbf{B} = \mathbf{WA}$: $\|\mathbf{A}\|_E = \|\mathbf{B}\|_E$.

◀ Действительно, если матрица \mathbf{W} унитарна, то для любого вектора \mathbf{A}

$$\begin{aligned} \|\mathbf{B}\|_E^2 &= \sum_{k=0}^{N-1} |b_k|^2 = \sum_{k=0}^{N-1} \bar{b}_k b_k = \bar{\mathbf{B}}^T \mathbf{B} = (\bar{\mathbf{W}} \mathbf{A})^T (\mathbf{W} \mathbf{A}) = \\ &= \bar{\mathbf{A}}^T \underbrace{\bar{\mathbf{W}}^T \mathbf{W} \mathbf{A}}_{\mathbf{E}} = \bar{\mathbf{A}}^T \mathbf{A} = \|\mathbf{A}\|_E^2, \end{aligned}$$

и наоборот, выполнение данного равенства для произвольного вектора \mathbf{A} означает, что $\bar{\mathbf{W}}^T \mathbf{W} = \mathbf{E}$, т. е. матрица \mathbf{W} — унитарна. ►

Свойство сохранения евклидовой нормы преобразования $\mathbf{B} = \mathbf{W}\mathbf{A}$, $\sum_{k=0}^{N-1} |a_k|^2 = \sum_{k=0}^{N-1} |b_k|^2$ представляет собой дискретный аналог равенства Парсеваля.

■ 3.6. Быстрое преобразование Фурье (БПФ).

Алгоритм БПФ с прореживанием по времени

Обозначим $x(k) = x_k$, $y(k) = y_k$ и запишем ДПФ (3.14):

$$y(k) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x(j) e^{-\frac{2\pi i}{N} kj}, \quad k = 0, \dots, N-1. \quad (3.17)$$

Хотя преобразование (3.17) определено для $k = 0, 1, \dots, N-1$, его можно рассматривать и для любого целого значения индекса k , при этом $y(k) = y(k+N)$, т. е. период последовательности $\{y_k\}_{k \in \mathbb{Z}}$ равен размерности ДПФ.

Рассмотрим ДПФ размерности $N = 2^n$. Обозначив

$$\omega_n = \exp(-2\pi i / 2^n),$$

запишем «ненормированное» ДПФ, $\tilde{y}(k) = \sqrt{N} y(k)$:

$$\begin{aligned} \tilde{y}(k) &= \sum_{j=0}^{2^n-1} x(j) \omega_n^{kj}, \\ k &= 0, 1, \dots, 2^n - 1. \end{aligned} \quad (3.18)$$

Вычисление каждого коэффициента $\tilde{y}(k)$, как легко видеть, требует около 2^n комплексных сложений с умножениями. Поэтому для реализации всего преобразования (3.18) необходимо около $2^n \times 2^n = 2^{2n}$ комплексных сложений с умножениями.

Пусть

$$\mathbf{X}_0 = (x(0), x(2), \dots, x(2^n - 2)) = (x_0(0), x_0(1), \dots, x_0(2^{n-1} - 1))$$

— вектор четных, а

$$\mathbf{X}_1 = (x(1), x(3), \dots, x(2^n - 1)) = (x_1(0), x_1(1), \dots, x_1(2^{n-1} - 1))$$

— нечетных компонент исходного вектора \mathbf{X} . Тогда, замечая, что $\omega_n^2 = \omega_{n-1}$, из (3.18) получаем:

$$\begin{aligned} \tilde{y}(k) &= \sum_{j=0}^{2^{n-1}-1} (x(2j) \omega_n^{k \cdot 2j} + x(2j+1) \omega_n^{k \cdot (2j+1)}) = \\ &= \underbrace{\sum_{j=0}^{2^{n-1}-1} x_0(j) \omega_{n-1}^{kj}}_{\text{Обозначим } \tilde{y}_0(k)} + \underbrace{\omega_n^k \sum_{j=0}^{2^{n-1}-1} x_1(j) \omega_{n-1}^{kj}}_{\text{Обозначим } \tilde{y}_1(k)}, \end{aligned}$$

где $\tilde{y}_0(k)$ — k -й элемент ДПФ (3.18) (размерности $N/2 = 2^{n-1}$) вектора \mathbf{X}_0 , а $\tilde{y}_1(k)$ — k -й элемент для ДПФ (3.18) вектора \mathbf{X}_1 .

Учитывая, что $\tilde{y}_0(k) = \tilde{y}_0(k + 2^{n-1})$, $\tilde{y}_1(k) = \tilde{y}_1(k + 2^{n-1})$, и $\omega_n^{2^{n-1}+k} = e^{-\frac{2\pi i}{2^n}(2^{n-1}+k)} = -\omega_n^k$, последнее выражение для $\tilde{y}(k)$ можем записать следующим образом:

$$\begin{cases} \tilde{y}(k) = \tilde{y}_0(k) + \omega_n^k \tilde{y}_1(k); \\ \tilde{y}(k + 2^{n-1}) = \tilde{y}_0(k) - \omega_n^k \tilde{y}_1(k); \\ k = 0, 1, \dots, 2^{n-1} - 1. \end{cases} \quad (3.19)$$

С помощью формул (3.19) мы выразили коэффициенты $\tilde{\mathbf{Y}} = (\tilde{y}(0), \tilde{y}(1), \dots, \tilde{y}(2^{n-1} - 1))$ ДПФ (3.18) размерности 2^n через коэффициенты ДПФ размерности 2^{n-1} ,

$$\begin{aligned} \tilde{\mathbf{Y}}_0 &= (\tilde{y}_0(0), \tilde{y}_0(1), \dots, \tilde{y}_0(2^{n-1} - 1)) \text{ и} \\ \tilde{\mathbf{Y}}_1 &= (\tilde{y}_1(0), \tilde{y}_1(1), \dots, \tilde{y}_1(2^{n-1} - 1)), \end{aligned}$$

которые получены из векторов \mathbf{X}_0 и \mathbf{X}_1 соответственно.

Таким образом, вычисление 2^n -точечного (т. е. размерности $N = 2^n$) ДПФ можно осуществить, выполнив предварительно два 2^{n-1} -точечных ДПФ. Вычисление спектров $\tilde{\mathbf{Y}}_0$ и $\tilde{\mathbf{Y}}_1$ размерности 2^{n-1} согласно (3.18) требует около $2^{2(n-1)}$ комплексных умножений в каждом случае, т. е. всего 2^{2n-1} комплексных умножений. Еще 2^{n-1} умножений требуется выполнить при дальнейших вычислениях по (3.19). Итого, вычисление ДПФ (3.18) размерности 2^n с использованием (3.19) потребует порядка $2^{2n-1} + 2^{n-1} \sim 2^{2n-1}$ операций комплексного умножения, т. е. примерно в два раза меньше, чем при непосредственной реализации ДПФ по формуле (3.18). Граф вычислений 8-точечного ДПФ по описанной схеме приведен на рис. 3.5.

Формулу (3.19) можно рекуррентно применять и далее, для вычисления ДПФ меньших размерностей. Например, дальнейшее использование формулы (3.19) для ДПФ размерности $N = 8$ дает полный граф вычислений, который изображен на рис. 3.6.

В общем случае при размерности преобразования $N = 2^n$ процедура сведения ДПФ к двум ДПФ меньшей размерности выполняется n раз. В соответствующем алгоритме *быстрого преобразования Фурье* (БПФ) имеем n шагов (см. рис. 3.6). На j -м шаге 2^{n-j}

раз выполняются вычисления, аналогичные (3.19), результатом которых являются 2^{n-j} «частичных» ДПФ размерности 2^j , выполненных над некоторыми подмножествами отсчетов исходного вектора \mathbf{X} . Каждый шаг требует (см. (3.19)) около $N/2 = 2^{n-1}$ комплексных умножений, итого в алгоритме БПФ имеем около $n2^{n-1}$ комплексных умножений, вместо 2^{2n} при непосредственной реализации ДПФ по формуле (3.18). С точки зрения реализации на вычислительных устройствах комплексное умножение является значительно более сложной операцией по сравнению с комплексным сложением или вычитанием, поэтому сокращение объема вычислений при использовании алгоритма БПФ можно оценить именно по количеству необходимых комплексных умножений. Например, для размерности $N = 2^{10} = 1024$ получим выигрыши примерно в 200 раз.

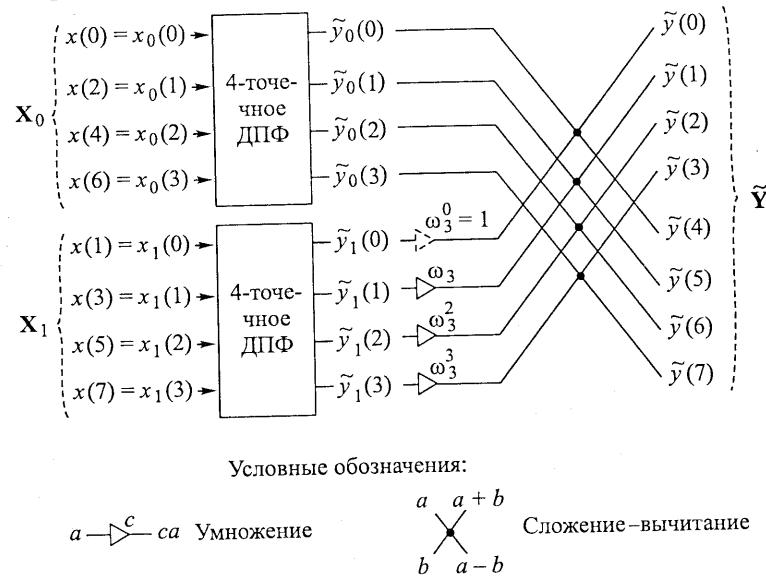


Рис. 3.5. Граф вычислений ДПФ, определяемый формулой (3.19), на примере преобразования размерности $N = 8$, $\omega_n = \exp(-2\pi i/2^n)$

Как видно из формулы (3.19) и примеров, приведенных на рис. 3.5 и 3.6, базовой операцией БПФ на j -м шаге является так называемая «бабочка» (см. рис. 3.7).

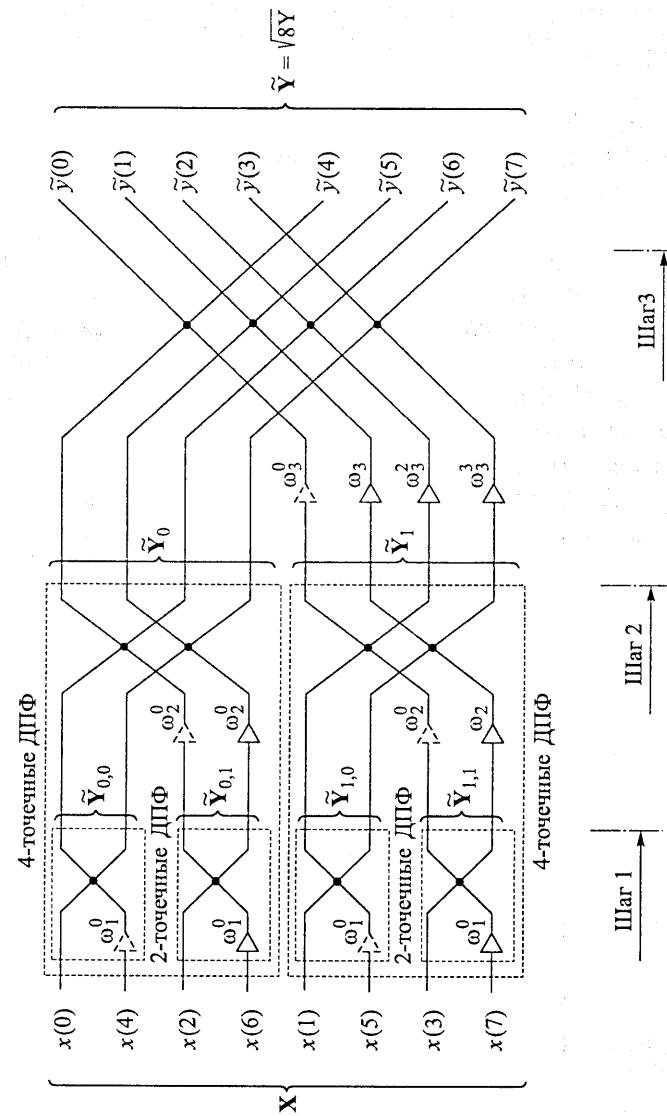


Рис. 3.6. Граф быстрого вычисления 8-точечного ДПФ (обозначения см. рис. 3.5)

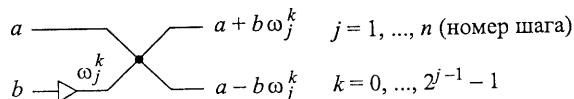


Рис. 3.7. Элементарные операции алгоритма БПФ на j -м шаге

Алгоритм БПФ, основанный на рекуррентном использовании формулы (3.19), называется алгоритмом БПФ с прореживанием по времени, так как (см. также рис. 3.5) вычисление ДПФ размерности N сводится к обработке векторов $\mathbf{X}_0, \mathbf{X}_1$, полученных «прореживанием» (выбором каждого второго отсчета, только четных или только нечетных номеров) вектора \mathbf{X} .

Отметим, что при обработке вещественных сигналов получаемые дискретные спектры обладают двойной избыточностью данных (см. свойство 1° ДПФ), что можно использовать для дополнительного ускорения вычислений путем использования схемы совмещенного ДПФ (см. свойство 6°).

Запишем ДПФ и ОДПФ в матричном виде, исключив нормировочный множитель $1/\sqrt{N}$ из структуры матриц преобразований:

$$\begin{aligned}\mathbf{Y} &= \frac{1}{\sqrt{N}} \mathbf{W} \mathbf{X}, \\ \mathbf{X} &= \frac{1}{\sqrt{N}} \bar{\mathbf{W}} \mathbf{Y},\end{aligned}\quad (3.20)$$

где

$$\mathbf{W} = \left\{ w_{l,m} = \exp\left(-\frac{2\pi i}{N} lm\right) \right\}_{l,m=0}^{N-1}.$$

Тогда алгоритм БПФ можно трактовать как представление матрицы ДПФ из (3.20) в виде произведения слабозаполненных (т. е. состоящих в основном из нулевых элементов) матриц:

$\mathbf{W} = \mathbf{W}_n \mathbf{W}_{n-1} \cdots \mathbf{W}_1$, где каждая матрица \mathbf{W}_j соответствует j -му шагу алгоритма БПФ

$$\sqrt{N} \mathbf{Y} = \tilde{\mathbf{Y}} = \underbrace{\mathbf{W}_n \mathbf{W}_{n-1} \cdots \mathbf{W}_1}_{\text{Шаг 1}} \underbrace{\mathbf{X}}_{\text{Шаг } n}$$

и содержит в каждой строке лишь два ненулевых элемента: 1 и ω_j^k , $k = 0, \dots, 2^{j-1} - 1$. Представление матрицы ДПФ в виде произведения слабозаполненных матриц называют *факторизацией*.

Пример 3.5. Представить матрицу ДПФ размерности $N = 8$ в факторизованном виде, соответствующем алгоритму БПФ с прореживанием по времени.

◀ Для $N = 2^3$ получаем три матрицы-сомножителя в разложении матрицы ДПФ, $\mathbf{Y} = \frac{1}{\sqrt{8}} \mathbf{W}_3 \mathbf{W}_2 \mathbf{W}_1 \mathbf{X}$ и, соответственно, три шага алгоритма БПФ (см. рис. 3.6). Структура матрицы \mathbf{W}_3 соответствует последнему шагу алгоритма БПФ; в использованных ранее обозначениях: $\tilde{\mathbf{Y}} = \mathbf{W}_3 \begin{pmatrix} \tilde{\mathbf{Y}}_0 \\ \tilde{\mathbf{Y}}_1 \end{pmatrix}$, где $\tilde{\mathbf{Y}} = \sqrt{8} \mathbf{Y}$, а векторы $\tilde{\mathbf{Y}}_0$ и

$\tilde{\mathbf{Y}}_1$, есть соответственно результаты 4-точечных ненормированных ДПФ (3.18), выполненных над четными \mathbf{X}_0 и нечетными \mathbf{X}_1 компонентами вектора \mathbf{X} . Несложно убедиться, что граф рисунка 3.5 (см. также (3.19)) соответствует следующей матричной операции умножения:

$$\tilde{\mathbf{Y}} = \mathbf{W}_3 \begin{pmatrix} \tilde{\mathbf{Y}}_0 \\ \tilde{\mathbf{Y}}_1 \end{pmatrix} = \mathbf{W}_3 \begin{pmatrix} \tilde{y}_0(0) \\ \tilde{y}_0(1) \\ \tilde{y}_0(2) \\ \tilde{y}_0(3) \\ \tilde{y}_1(0) \\ \tilde{y}_1(1) \\ \tilde{y}_1(2) \\ \tilde{y}_1(3) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \omega_3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \omega_3^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \omega_3^3 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -\omega_3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -\omega_3^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -\omega_3^3 \end{pmatrix} \begin{pmatrix} \tilde{y}_0(0) \\ \tilde{y}_0(1) \\ \tilde{y}_0(2) \\ \tilde{y}_0(3) \\ \tilde{y}_1(0) \\ \tilde{y}_1(1) \\ \tilde{y}_1(2) \\ \tilde{y}_1(3) \end{pmatrix},$$

где $\omega_3 = \exp(-i\pi/4)$.

В свою очередь (см. рис. 3.6 и (3.19))

$$\tilde{\mathbf{Y}}_0 = \begin{pmatrix} \tilde{y}_0(0) \\ \tilde{y}_0(1) \\ \tilde{y}_0(2) \\ \tilde{y}_0(3) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \omega_2 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\omega_2 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{Y}}_{0,0} \\ \tilde{\mathbf{Y}}_{0,1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \omega_2 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\omega_2 \end{pmatrix} \begin{pmatrix} \tilde{y}_{0,0}(0) \\ \tilde{y}_{0,0}(1) \\ \tilde{y}_{0,1}(0) \\ \tilde{y}_{0,1}(1) \end{pmatrix},$$

$$\tilde{\mathbf{Y}}_1 = \begin{pmatrix} \tilde{y}_1(0) \\ \tilde{y}_1(1) \\ \tilde{y}_1(2) \\ \tilde{y}_1(3) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \omega_2 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\omega_2 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{Y}}_{1,0} \\ \tilde{\mathbf{Y}}_{1,1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \omega_2 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\omega_2 \end{pmatrix} \begin{pmatrix} \tilde{y}_{1,0}(0) \\ \tilde{y}_{1,0}(1) \\ \tilde{y}_{1,1}(0) \\ \tilde{y}_{1,1}(1) \end{pmatrix},$$

где $\omega_2 = \exp(-i\pi/2) = -i$. Отсюда для матрицы \mathbf{W}_2 получаем структуру:

$$\begin{pmatrix} \tilde{\mathbf{Y}}_0 \\ \tilde{\mathbf{Y}}_1 \end{pmatrix} = \begin{pmatrix} \tilde{y}_0(0) \\ \tilde{y}_0(1) \\ \tilde{y}_0(2) \\ \tilde{y}_0(3) \\ \tilde{y}_1(0) \\ \tilde{y}_1(1) \\ \tilde{y}_1(2) \\ \tilde{y}_1(3) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -i & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -i \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & i \end{pmatrix} \underbrace{\mathbf{w}_2}_{\mathbf{w}_2} \begin{pmatrix} \tilde{y}_{0,0}(0) \\ \tilde{y}_{0,0}(1) \\ \tilde{y}_{0,1}(0) \\ \tilde{y}_{0,1}(1) \\ \tilde{y}_{1,0}(0) \\ \tilde{y}_{1,0}(1) \\ \tilde{y}_{1,1}(0) \\ \tilde{y}_{1,1}(1) \end{pmatrix}$$

Аналогично, по графу рисунка 3.6 получаем структуру матрицы \mathbf{W}_1 :

$$\begin{pmatrix} \tilde{\mathbf{Y}}_{0,0} \\ \tilde{\mathbf{Y}}_{0,1} \\ \tilde{\mathbf{Y}}_{1,0} \\ \tilde{\mathbf{Y}}_{1,1} \end{pmatrix} = \begin{pmatrix} \tilde{y}_{0,0,0} \\ \tilde{y}_{0,0,1} \\ \tilde{y}_{0,1,0} \\ \tilde{y}_{0,1,1} \\ \tilde{y}_{1,0,0} \\ \tilde{y}_{1,0,1} \\ \tilde{y}_{1,1,0} \\ \tilde{y}_{1,1,1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \underbrace{\mathbf{w}_1}_{\mathbf{w}_1} \begin{pmatrix} \tilde{y}_{0,0,0} = x(0) \\ \tilde{y}_{0,0,1} = x(4) \\ \tilde{y}_{0,1,0} = x(2) \\ \tilde{y}_{0,1,1} = x(6) \\ \tilde{y}_{1,0,0} = x(1) \\ \tilde{y}_{1,0,1} = x(5) \\ \tilde{y}_{1,1,0} = x(3) \\ \tilde{y}_{1,1,1} = x(7) \end{pmatrix}$$

где вектор-столбец

$$\begin{aligned} \tilde{\mathbf{X}} &= (\tilde{x}(0), \tilde{x}(1), \tilde{x}(2), \tilde{x}(3), \tilde{x}(4), \tilde{x}(5), \tilde{x}(6), \tilde{x}(7))^T = \\ &= (\tilde{y}_{0,0,0}, \tilde{y}_{0,0,1}, \tilde{y}_{0,1,0}, \tilde{y}_{0,1,1}, \tilde{y}_{1,0,0}, \tilde{y}_{1,0,1}, \tilde{y}_{1,1,0}, \tilde{y}_{1,1,1})^T \end{aligned}$$

получен из вектора $\mathbf{X} = (x(0), x(1), x(2), x(3), x(4), x(5), x(6), x(7))^T$ перестановкой пар компонент, индексы которых образуют двоич-

но-инверсную пару (k, k^*) : $\mathbf{X} \xrightarrow{x(k) \leftrightarrow x(k^*)} \tilde{\mathbf{X}}$. Двоично-инверсным для целого числа $k \in \{0, 1, \dots, 2^n - 1\}$ мы называем число k^* , которое получается в результате обратной записи k в двоичном представлении, т. е. если

$$k = \sum_{j=0}^{n-1} k_j 2^j = k_{n-1} k_{n-2} \dots k_0, \quad k_j \in \{0, 1\},$$

$$\text{то } k^* = \sum_{j=0}^{n-1} k_{n-1-j} 2^j = k_0 k_1 \dots k_{n-1}.$$

Перестановку компонент с двоично-инверсными номерами также можно представить в виде матричного умножения, $\tilde{\mathbf{X}} = \mathbf{C}\mathbf{X}$. Для рассматриваемого нами примера 8-точечного преобразования

$$\tilde{\mathbf{X}} = \begin{pmatrix} x(0) \\ x(4) \\ x(2) \\ x(6) \\ x(1) \\ x(5) \\ x(3) \\ x(7) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \underbrace{\mathbf{C}}_{\mathbf{C}} \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \\ x(4) \\ x(5) \\ x(6) \\ x(7) \end{pmatrix} \blacktriangleright$$

Обобщая результаты примера 3.5 на случай размерности ДПФ $N = 2^n$ для произвольного числа $n \in \mathbb{N}$, структуру матрицы \mathbf{W} не-нормированного преобразования $\tilde{\mathbf{Y}} = \mathbf{W}\mathbf{X}$, которая определяет алгоритм БПФ с прореживанием по времени, можем представить в следующем виде:

$$\mathbf{W} = \mathbf{W}_n \cdot \mathbf{W}_{n-1} \cdot \dots \cdot \mathbf{W}_1 \cdot \mathbf{C}, \quad (3.21)$$

где $\mathbf{C} = \{c_{k,j}\}_{k,j=0}^{2^n-1}$ — симметрическая матрица перестановок элем-

ментов с двоично-инверсными номерами: $c_{k,j} = \begin{cases} 1, & \text{если } k = j^*, \\ 0, & \text{если } k \neq j^*. \end{cases}$

Матрица \mathbf{W}_k имеет блочный вид:

$$\mathbf{W}_k = \begin{pmatrix} \mathbf{B}_k^{(0)} \\ \mathbf{B}_k^{(1)} \\ \vdots \\ \mathbf{B}_k^{(2^{n-k}-1)} \end{pmatrix}, \quad k=1, \dots, n,$$

а каждый блок $\mathbf{B}_k^{(j)}$ ($j = 0, 1, \dots, 2^{n-k}-1$) состоит из 2^k строк и может быть записан как

$$\mathbf{B}_k^{(j)} = \left[\begin{array}{c|cc|cc|cc|cc|c} 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & \omega_k & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & \omega_k^2 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & \dots & \omega_k^{2^{k-1}-1} & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 1 & 0 & 0 & \dots & 0 & -1 & 0 & 0 & \dots & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & 1 & 0 & \dots & 0 & 0 & -\omega_k & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & 0 & 1 & \dots & 0 & 0 & 0 & -\omega_k^2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & \dots & -\omega_k^{2^{k-1}} & \vdots & \vdots & \vdots \end{array} \right] \begin{matrix} 2^{k-1} \\ \text{строк} \\ 2^{k-1} \\ \text{строк} \\ 2^{k-1} \\ \text{строк} \\ 2^{k-1} \\ \text{строк} \\ 2^{k-1} \\ \text{строк} \end{matrix}$$

$j \cdot 2^k$ нулевых столбцов 2^{k-1} столбцов 2^{k-1} столбцов $2^n - (j+1) \cdot 2^k$ нулевых столбцов

При использовании данного способа факторизации матрицы ДПФ реализация БПФ осуществляется (после перестановки элементов входного вектора \mathbf{X} с двоично-инверсными индексами) посредством трех вложенных циклов.

1. Цикл по матрицам \mathbf{W}_k , $k = 1, \dots, n$;
 2. Цикл по блокам $\mathbf{B}_k^{(j)}$ матрицы \mathbf{W}_k , $j = 0, \dots, 2^{n-k}-1$;
 3. Цикл по строкам блока $\mathbf{B}_k^{(j)}$, $l = 0, \dots, 2^{k-1}-1$:
- $$\begin{cases} y(j \cdot 2^k + l) = x(j \cdot 2^k + l) + \omega_k^l \cdot x(j \cdot 2^k + 2^{k-1} + l); \\ y(j \cdot 2^k + 2^{k-1} + l) = x(j \cdot 2^k + l) - \omega_k^l \cdot x(j \cdot 2^k + 2^{k-1} + l). \end{cases}$$

В приведенной структуре вычислений под вектором \mathbf{X} на k -м шаге для $k = 2, \dots, n$ понимается вектор \mathbf{Y} , полученный на $(k-1)$ -м шаге.

Описанный способ вычисления БПФ позволяет экономно использовать память ЭВМ, так как получаемые после выполнения «бабочки» значения $y(j \cdot 2^k + l)$ и $y(j \cdot 2^k + 2^{k-1} + l)$ можно сохранять на месте элементов $x(j \cdot 2^k + l)$ и $x(j \cdot 2^k + 2^{k-1} + l)$, которые больше не будут участвовать в вычислениях. Поэтому достаточно выделить память для хранения только одного вектора размерности N , при этом результирующий вектор $\tilde{\mathbf{Y}}$ будет получен в результате последовательного замещения компонент «входного» вектора \mathbf{X} .

Алгоритм БПФ с прореживанием по времени можно реализовать, исключив операции перестановок элементов исходного вектора \mathbf{X} с двоично-инверсными номерами. Рассмотрим это на примере 8-точечного ДПФ. Если схему сведения одного ДПФ к двум ДПФ меньшей размерности (см. рис. 3.5) представить в несколько ином виде (рис. 3.8) и рекуррентно применить ее далее, то получим полный граф вычислений, приведенный на рис. 3.9. Отметим, что исключение операций перестановок элементов входного вектора достигается ценой двойного увеличения объема памяти, необходимой при реализации БПФ, так как для векторов \mathbf{X} и $\tilde{\mathbf{Y}}$ приходится выделять уже различные участки памяти.

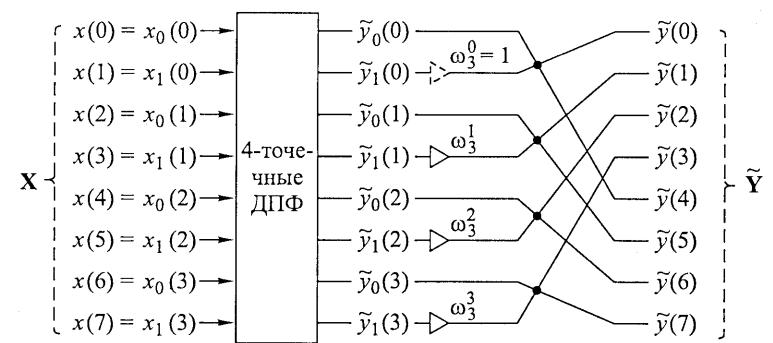


Рис. 3.8. Вариант графа вычислений ДПФ, определяемый формулой (3.19), на примере преобразования размерности $N = 8$ (обозначения см. рис. 3.5)

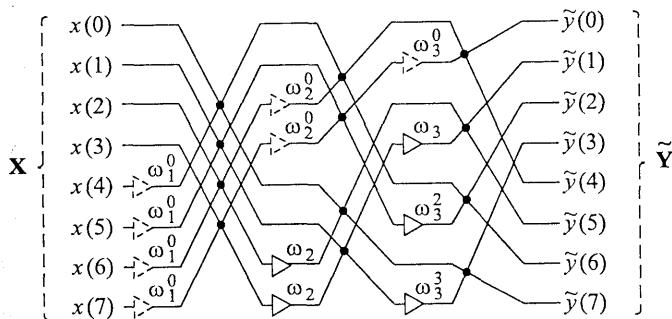


Рис. 3.9. Граф вычислений БПФ (алгоритм с прореживанием по времени) без перестановок элементов входного вектора с двоично-инверсными индексами, на примере преобразования размерности $N = 8$ (обозначения см. рис. 3.5)

Пример 3.6. Представить матрицу ДПФ размерности $N = 8$ в факторизованном виде, соответствующем алгоритму БПФ с прореживанием по времени без перестановок элементов входного вектора.

◀ Для $N = 2^3$ вновь получаем три сомножителя в разложении матрицы ненормированного ДПФ: $\mathbf{W} = \mathbf{W}_3 \mathbf{W}_2 \mathbf{W}_1$ и, соответственно, три шага алгоритма БПФ. По аналогии с тем, как были получены матрицы-сомножители в примере 3.5 по графу, представленному на рис. 3.6, получаем по графу рис. 3.9 следующие структуры матриц факторизованного представления:

$$\mathbf{W}_3 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \omega_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \omega_3^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \omega_3^3 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\omega_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\omega_3^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\omega_3^3 \end{pmatrix},$$

$$\mathbf{W}_2 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \omega_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \omega_2 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -\omega_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -\omega_2 \end{pmatrix},$$

$$\mathbf{W}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Вновь можно заметить, что матрицы-сомножители факторизованного представления матрицы ДПФ имеют блочную структуру. Обобщая результаты примера 3.6 на случай размерности $N = 2^n$, структуру матрицы ненормированного ДПФ для алгоритма БПФ с прореживанием по времени (без перестановок элементов входного вектора) можем представить в виде

$$\mathbf{W} = \mathbf{W}_n \cdot \mathbf{W}_{n-1} \cdot \dots \cdot \mathbf{W}_1, \quad (3.22)$$

$$\text{где } \mathbf{W}_k = \begin{pmatrix} \mathbf{B}_k^{(0)}(+) \\ \mathbf{B}_k^{(1)}(+) \\ \vdots \\ \mathbf{B}_k^{(2^{k-1}-1)}(+) \\ \hline \mathbf{B}_k^{(0)}(-) \\ \mathbf{B}_k^{(1)}(-) \\ \vdots \\ \mathbf{B}_k^{(2^{k-1}-1)}(-) \end{pmatrix}, \quad k = 1, \dots, n,$$

а каждый блок $B_k^{(j)}(\pm)$ ($j = 0, 1, \dots, 2^{k-1} - 1$) может быть представлен следующим образом:

$$B_k^{(j)}(\pm) = \left[\begin{array}{c|ccccc|ccccc|c} 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & \pm \omega_k^j & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & \pm \omega_k^j & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \pm \omega_k^j & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & \pm \omega_k^j & 0 & 0 & 0 \end{array} \right]_{\text{строк}}^{2^{n-k}}$$

$j \cdot 2^{n-k+1}$
нулевых столбцов 2^{n-k} столбцов 2^{n-k} столбцов $2^n - (j+1) \cdot 2^{n-k+1}$
нулевых столбцов

Вычисление БПФ вновь реализуется посредством трех вложенных циклов (по матрицам факторизованного представления, по блокам матриц, по строкам блоков). Фактически, вторая форма алгоритма БПФ с прореживанием по времени, определяемая разложением (3.22), получается из первой формы (3.21) путем определенных перестановок строк в матрицах факторизации.

Кратко остановимся на построении быстрого алгоритма обратного ДПФ. Все проведенные рассуждения для прямого ДПФ (3.18) можно в точности повторить, вместо $\omega_j = \exp\left(-\frac{2\pi i}{2^j}\right)$ по-

ложив $\omega_j = \exp\left(\frac{2\pi i}{2^j}\right)$, верными останутся и аналитическое выражение (3.19), и примеры графов вычислений, приведенные на рис. 3.5–3.8. Однако, имея вычислительную процедуру нахождения ДПФ $\tilde{\mathbf{Y}} = \mathbf{W}\mathbf{X}$, можно осуществить и умножение на матрицу $\overline{\mathbf{W}}$, необходимое при ОДПФ (см. (3.20)), поскольку $\sqrt{N}\overline{\mathbf{X}} = \mathbf{W}\overline{\mathbf{Y}}$. Реализация процедуры $\mathbf{B} = \overline{\mathbf{W}}\mathbf{A}$ производится путем вычисления матричного произведения $\overline{\mathbf{B}} = \mathbf{W}\mathbf{A}$ и сводится к следующей последовательности действий:

1. $\mathbf{U} \leftarrow \overline{\mathbf{A}}$ (сформировать вектор \mathbf{U} из комплексно сопряженных компонент вектора \mathbf{A}).
2. $\mathbf{V} = \mathbf{W}\mathbf{U}$ (используя имеющуюся процедуру реализации ДПФ, выполнить умножение матрицы на вектор).
3. $\mathbf{B} \leftarrow \overline{\mathbf{V}}$ (сформировать вектор \mathbf{B} из комплексно сопряженных компонент вектора \mathbf{V}).

Таким образом, реализовав быстрый алгоритм умножения на матрицу \mathbf{W} (прямое БПФ), сразу получаем и процедуру быстрого алгоритма умножения на матрицу $\overline{\mathbf{W}}$ (обратное БПФ).

Отметим, что при обработке вещественных сигналов для дополнительного двукратного ускорения вычислений как прямого, так и обратного БПФ может использоваться свойство 6° ДПФ.

■ 3.7. Алгоритм БПФ с прореживанием по частоте

Ненормированное ДПФ (3.18) можно записать также в следующем виде:

$$\tilde{y}(k) = \sum_{j=0}^{2^{n-1}-1} x(j)\omega_n^{kj} + \sum_{j=0}^{2^{n-1}-1} x(j+2^{n-1})\omega_n^{k(j+2^{n-1})},$$

где $\omega_n = \exp(-2\pi i/2^n)$. Выделив множитель

$$\omega_n^{2^{n-1}k} = e^{-\frac{2\pi i}{2^n}2^{n-1}k} = e^{-\pi ik} = (-1)^k$$

из второй суммы, можем записать:

$$\tilde{y}(k) = \sum_{j=0}^{2^{n-1}-1} (x(j) + (-1)^k x(j+2^{n-1}))\omega_n^{kj}.$$

Отсюда, учитывая что $\omega_n^2 = \omega_{n-1}$, получаем:

$$\begin{cases} \tilde{y}(2k) = \sum_{j=0}^{2^{n-1}-1} (x(j) + x(j+2^{n-1}))\omega_{n-1}^{kj}; \\ \tilde{y}(2k+1) = \sum_{j=0}^{2^{n-1}-1} (x(j) - x(j+2^{n-1}))\omega_n^j\omega_{n-1}^{kj}; \end{cases} \quad k = 0, \dots, 2^{n-1} - 1. \quad (3.23)$$

Четные компоненты вектора $\tilde{\mathbf{Y}}$ (первая строка системы (3.23)) представляют собой ДПФ размерности $N/2 = 2^{n-1}$, выполненное над вектором

$$(x(0) + x(2^{n-1}), x(1) + x(2^{n-1} + 1), \dots, x(2^{n-1} - 1) + x(2^n - 1)),$$

а компоненты вектора $\tilde{\mathbf{Y}}$ с нечетными номерами есть ДПФ (размерности $N/2 = 2^{n-1}$) вектора

$$\begin{aligned} ((x(0) - x(2^{n-1}))\omega_n^0, (x(1) - x(2^{n-1} + 1))\omega_n^1, \dots, \\ (x(2^{n-1} - 1) - x(2^n - 1))\omega_n^{2^{n-1}-1}). \end{aligned}$$

Таким образом, мы вновь свели вычисление ДПФ размерности $N = 2^n$ к реализации двух ДПФ размерности $N/2 = 2^{n-1}$ каждое. Для случая $N = 8$ соответствующую схему вычислений можно представить в виде графа, приведенного на рис. 3.10.

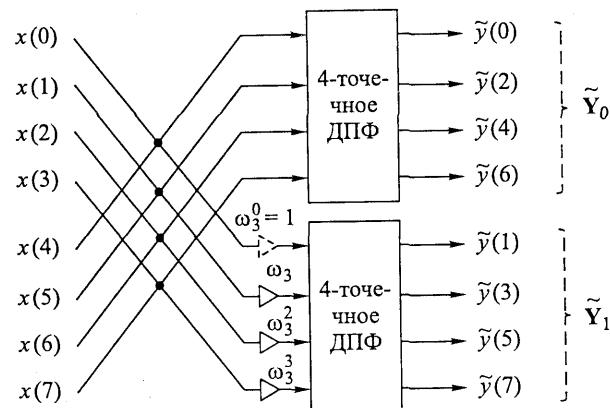


Рис. 3.10. Граф вычислений ДПФ, определяемый формулой (3.23), на примере преобразования размерности $N = 8$ (обозначения см. рис. 3.5)

Формулы (3.23) определяют схему вычислений БПФ с прореживанием по частоте, которую можно рекурсивно применять далее для ДПФ векторов

$$\begin{aligned} \tilde{\mathbf{Y}}_0 &= (\tilde{y}(0), \tilde{y}(2), \dots, \tilde{y}(2^n - 2)) = (\tilde{y}_0(0), \tilde{y}_0(1), \dots, \tilde{y}_0(2^{n-1} - 1)) \\ \text{и } \tilde{\mathbf{Y}}_1 &= (\tilde{y}(1), \tilde{y}(3), \dots, \tilde{y}(2^n - 1)) = (\tilde{y}_1(0), \tilde{y}_1(1), \dots, \tilde{y}_1(2^{n-1} - 1)). \end{aligned}$$

Так, для случая $N = 8$ полный график вычислений БПФ будет иметь вид, приведенный на рис. 3.11. Как и в алгоритме с прореживани-

ем по времени, при размерности преобразования $N = 2^n$ процедура сведения ДПФ к двум ДПФ меньшей размерности выполняется n раз, т. е. в алгоритме БПФ с прореживанием по частоте также имеем n шагов (рис. 3.11).

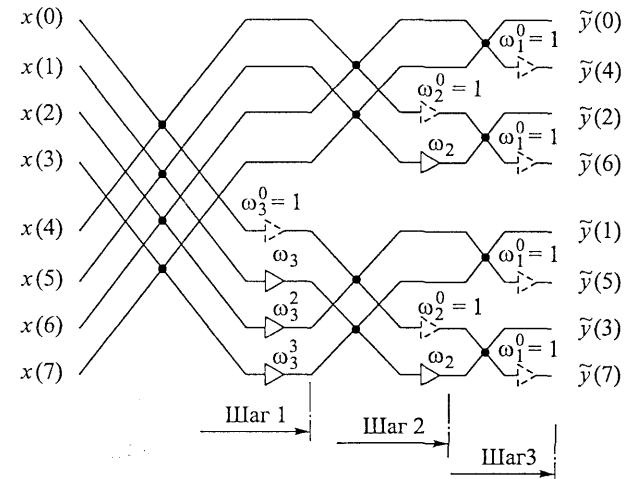


Рис. 3.11. Граф 8-точечного БПФ, алгоритм с прореживанием по частоте (обозначения см. рис. 3.5)

Как видно из формулы (3.23) и примеров, приведенных на рис. 3.10 и 3.11, базовой операцией БПФ на j -м шаге для алгоритма с прореживанием по частоте является «бабочка», которая имеет вид, приведенный на рис. 3.12 (сравните с рис. 3.7).

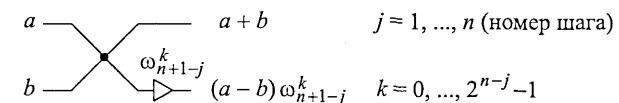


Рис. 3.12. Элементарные операции алгоритма БПФ с прореживанием по частоте на j -м шаге

Фактически (см. рис. 3.6 и 3.11) алгоритм с прореживанием по частоте имеет график, который получается при проходе в обратном направлении графа алгоритма БПФ с прореживанием по времени с заменой узловых операций (см. рис. 3.7) на «бабочки» (рис. 3.12). В матричном виде факторизация матрицы ДПФ, опре-

деляемая алгоритмом с прореживанием по частоте, может быть получена непосредственно из представления (3.21) (алгоритма с прореживанием по времени), если учесть, что матрица ДПФ — симметрическая. А именно,

$$\mathbf{W} = \mathbf{W}^T = (\mathbf{W}_n \cdot \mathbf{W}_{n-1} \cdot \dots \cdot \mathbf{W}_1 \cdot \mathbf{C})^T = \mathbf{C} \cdot \mathbf{W}_1^T \cdot \dots \cdot \mathbf{W}_{n-1}^T \cdot \mathbf{W}_n^T,$$

где симметрическая матрица \mathbf{C} представляет собой матрицу перестановок элементов с двоично-инверсными индексами, а матрица \mathbf{W}_j определяется по формуле (3.21), $j = 1, \dots, n$.

Задание. Воспользовавшись результатами примера 3.5, выпишите факторизованное представление для матрицы ДПФ размерности $N = 8$, которое определяет алгоритм БПФ с прореживанием по частоте.

Для алгоритма БПФ с прореживанием по частоте также можно выделить две модификации.

Первая, рассмотренная выше, позволяет производить вычисления с использованием $\sim N$ ячеек памяти ЭВМ, но после окончания арифметических вычислений требует выполнения перестановок пар элементов вектора-спектра $\tilde{\mathbf{Y}}$, имеющих двоично-инверсные номера.

Вторая модификация алгоритма БПФ с прореживанием по частоте позволяет избежать перестановок элементов вектора-спектра, но требует при вычислениях порядка $\sim 2N$ ячеек памяти ЭВМ — соответствующая факторизация матрицы ДПФ задается соотношением:

$$\mathbf{W} = \mathbf{W}^T = \mathbf{W}_1^T \mathbf{W}_2^T \cdot \dots \cdot \mathbf{W}_n^T,$$

где матрицы \mathbf{W}_j ($j = 1, \dots, n$) определены в представлении (3.22).

Пример 3.7. Представить матрицу ДПФ размерности $N = 8$ в факторизованном виде, соответствующем алгоритму БПФ с прореживанием по частоте без перестановок элементов входного вектора.

◀ Воспользуемся результатами и обозначениями примера 3.6, получим:

$$\mathbf{W} = \mathbf{W}_1^T \mathbf{W}_2^T \mathbf{W}_3^T =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_2 & 0 & 0 & 0 & -\omega_2 \\ 0 & 0 & 0 & \omega_2 & 0 & 0 & 0 & -\omega_2 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \omega_3 & 0 & 0 & 0 & -\omega_3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \omega_3^2 & 0 & 0 & 0 & -\omega_3^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \omega_3^3 & 0 & 0 & 0 & -\omega_3^3 \end{pmatrix}.$$

Граф вычислений алгоритма БПФ с прореживанием по частоте, который не требует перестановок элементов вектора-спектра, получается при обратном проходе графа соответствующего (без

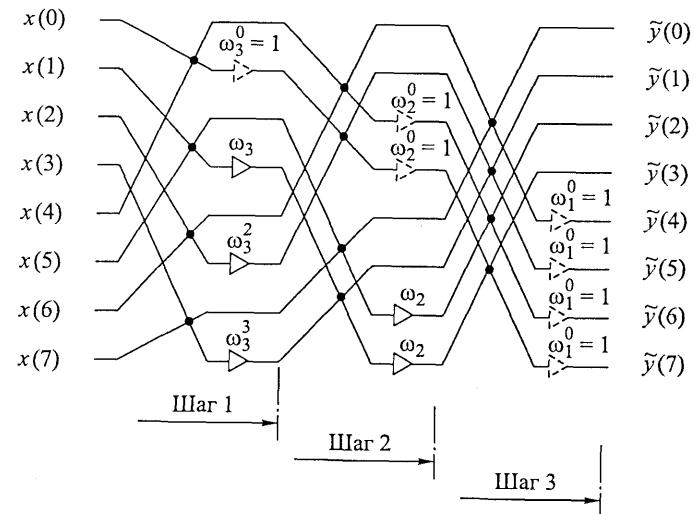


Рис. 3.13. Граф 8-точечного БПФ, алгоритм с прореживанием по частоте, не требующий перестановок элементов вектора-спектра (обозначения см. рис. 3.5)

перестановок) алгоритма с прореживанием по времени, с заменой узловых операций рис. 3.7 на операции рис. 3.12. Например, для случая $N = 8$ получим полный граф вычислений, приведенный на рис. 3.13 (сравните с рис. 3.9).

■ 3.8. Дискретное преобразование Уолша

Определение. Дискретным преобразованием Уолша (ДПУ) вектора $\mathbf{X} = (x_0, x_1, \dots, x_{2^n-1})$ назовем вектор $\mathbf{Y} = (y_0, y_1, \dots, y_{2^n-1})$, компоненты которого находятся по формуле

$$y_k = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x_j w_k(j/2^n), \quad (3.24)$$

$$k = 0, 1, \dots, 2^n - 1,$$

где $w_k(t)$ — функции системы Уолша (см. раздел 1.6).

Замечание. Если x_j — отсчеты некоторой функции $f(t)$, определенной на отрезке $t \in [0; 1]$, т. е. $x_j = f(j/2^n)$, то (3.24) можно рассматривать как формулу для приближенного вычисления коэффициентов Фурье c_k по системе Уолша функции $f(t)$ (с точностью до множителя $1/\sqrt{2^n}$). Действительно, используя формулу прямоугольников для приближенного вычисления интеграла, получим:

$$c_k = \langle f(t), w_k(t) \rangle = \int_0^1 f(t) w_k(t) dt \approx \sum_{j=0}^{2^n-1} x_j w_k(j/2^n) \frac{1}{2^n} = \frac{1}{\sqrt{2^n}} y_k.$$

ДПУ часто удобнее представлять в матричном виде, $\mathbf{Y} = \mathbf{W}\mathbf{X}$, где \mathbf{X} и \mathbf{Y} — векторы-столбцы, а матрица ДПУ

$$\mathbf{W} = \left\{ \underbrace{w_{k,j}}_{\text{Эл-т матрицы}} = \frac{1}{\sqrt{2^n}} \underbrace{w_k(j/2^n)}_{\text{Отсчет функции Уолша}} \right\}_{k,j=0}^{2^n-1}.$$

Лемма 3.2. Матрица ДПУ — симметрическая, т. е. $\mathbf{W} = \mathbf{W}^T$.

◀ Любое вещественное неотрицательное число $t \in \mathbb{R}$ можно записать в виде двоичного разложения:

$$t = \sum_{m=-\infty}^{\infty} t_m 2^m = \underbrace{\sum_{m=0}^{M(t)} t_m 2^m}_{[x] \text{ — целая часть}} + \underbrace{\sum_{m=-\infty}^{-1} t_m 2^m}_{\{x\} \text{ — дробная часть}},$$

где $\forall m t_m \in \{0; 1\}$, $t_m \equiv [t/2^m] \pmod{2}$. Тогда для $t \geq 0$ функцию $r_0(t)$ системы Радемахера (определение системы см. в разделе 1.6) можно определить следующим образом (см. рис. 3.14):

$$r_0(t) = r_0 \left(\sum_{m=-\infty}^{\infty} t_m 2^m \right) = \begin{cases} 1 & \text{при } t_{-1} = 0; \\ -1 & \text{при } t_{-1} = 1. \end{cases}$$

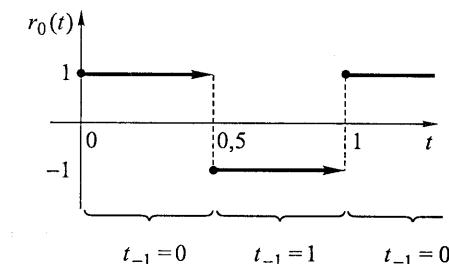


Рис. 3.14

Так как $r_m(t) = r_0(2^m t)$, то для точек $t_j = j/2^n$, $j = 0, \dots, 2^n - 1$ можем записать:

$$r_m(j/2^n) = r_0 \left(2^{m-n} \sum_{\mu=0}^{n-1} j_\mu 2^\mu \right) = r_0 \left(\sum_{\mu=0}^{n-1} j_\mu 2^{\mu+m-n} \right) = r_0 \left(\sum_{l=m-n}^{m-1} \frac{j_{l-m+n}}{t_l} 2^l \right).$$

Отсюда $r_m(j/2^n) = \begin{cases} 1, & \text{при } t_{-1} = j_{-1-m+n} = 0 \\ -1, & \text{при } t_{-1} = j_{-1-m+n} = 1 \end{cases} = (-1)^{j_{n-1-m}}$. Поэтому для функций системы Уолша (см. формулу (1.11)) при $k = 0, \dots, 2^n - 1$ имеем:

$$w_k\left(\frac{j}{2^n}\right) = \prod_{m=0}^{n-1} \left(r_m\left(\frac{j}{2^n}\right)\right)^{k_m} = \prod_{m=0}^{n-1} (-1)^{j \frac{n-1-m}{\mu} k_m} = \prod_{\mu=0}^{n-1} (-1)^{k_{n-1-\mu} j_\mu} = w_j\left(\frac{k}{2^n}\right),$$

т. е. $w_{k,j} = \frac{1}{\sqrt{2^n}} w_k(j/2^n) = \frac{1}{\sqrt{2^n}} w_j(k/2^n) = w_{j,k}$. ►

Теорема 3.4. Матрица ДПУ — ортогональна.

◀ Нужно доказать, что $\mathbf{W}^{-1} = \mathbf{W}^T$. Для этого обозначим $\mathbf{U} = \mathbf{W}^T \mathbf{W} = \mathbf{W} \mathbf{W}^T$ и покажем, что матрица \mathbf{U} — единичная, т. е. $u_{k,j} = \begin{cases} 1 & \text{при } k=j; \\ 0 & \text{при } k \neq j. \end{cases}$ Рассмотрим элемент k -й строки j -го столбца матрицы \mathbf{U} , который представляет собой скалярное произведение k -й строки матрицы \mathbf{W} и j -го столбца матрицы \mathbf{W} :

$$u_{k,j} = \sum_{m=0}^{2^n-1} w_{k,m} w_{m,j} = \sum_{m=0}^{2^n-1} w_{k,m} w_{j,m} = \sum_{m=0}^{2^n-1} \frac{1}{2^n} w_k\left(\frac{m}{2^n}\right) w_j\left(\frac{m}{2^n}\right).$$

Поскольку $\forall k = 0, \dots, n-1: r_k(t) = \text{const}$ при $t \in \Delta_m^n = \left[\frac{m}{2^n}, \frac{m+1}{2^n}\right)$ (см. формулу (1.9)), то на этих же полуинтервалах Δ_m^n также

$\forall v < 2^n: w_v(t) = \prod_{k=0}^{n-1} (r_k(t))^{v_k} = \text{const}$. Поэтому последнее выражение для $u_{k,j}$ можем переписать как

$$\begin{aligned} u_{k,j} &= \sum_{m=0}^{2^n-1} \int_{\substack{\Delta_m^n \\ \text{const}=w_k\left(\frac{m}{2^n}\right) w_j\left(\frac{m}{2^n}\right)}} w_k(t) w_j(t) dt = \\ &= \langle w_k(t), w_j(t) \rangle = \begin{cases} 1 & \text{при } k=j; \\ 0 & \text{при } k \neq j \end{cases} \end{aligned}$$

в силу ортонормированности системы Уолша (см. теорему 1.9). ►

Следствие. Обратная матрица ДПУ $\mathbf{W}^{-1} = \mathbf{W}$, т. е. формула обратного ДПУ совпадает с формулой прямого ДПУ (3.24): $\mathbf{X} = \mathbf{WY}$, или

$$x_k = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} y_j w_k(j/2^n), \quad k = 0, 1, \dots, 2^n - 1.$$

Пример 3.8. Найдем матрицу ДПУ для преобразования размерности $N = 2^2 = 4$.

◀ При доказательстве леммы 3.2 мы получили общее выражение:

$$w_{k,j} = w_{j,k} = \frac{1}{\sqrt{2^n}} w_k\left(\frac{j}{2^n}\right) = \frac{1}{\sqrt{2^n}} \prod_{m=0}^{n-1} (-1)^{j_{n-1-m} k_m}.$$

В нашем случае $n = 2$, поэтому

$$\begin{aligned} \mathbf{W} &= \left(w_{k,j} = \frac{1}{2} (-1)^{j_0 k_1 + j_1 k_0} \right)_{k,j=0}^3 \\ \text{Получаем: } \mathbf{W} &= \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}. \end{aligned} \quad \blacktriangleright$$

■ 3.9. Дискретное преобразование Хаара

Определение. Дискретным преобразованием Хаара (ДПХ) вектора $\mathbf{X} = (x_0, x_1, \dots, x_{2^n-1})$ назовем вектор $\mathbf{Y} = (y_0, y_1, \dots, y_{2^n-1})$, компоненты которого находятся по формуле

$$\begin{aligned} y_k &= \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x_j h_k(j/2^n), \\ k &= 0, 1, \dots, 2^n - 1, \end{aligned} \tag{3.25}$$

где $h_k(t)$ — функции системы Хаара (см. раздел 1.6).

Замечание. Выражение (3.25) также (см. замечание к (3.24)) можно рассматривать как формулу для приближенного вычисле-

ния коэффициентов Фурье — Хаара функции $f(t)$ (с точностью до множителя $1/\sqrt{2^n}$):

$$c_k = \langle f(t), h_k(t) \rangle = \int_0^1 f(t) h_k(t) dt \approx \sum_{j=0}^{2^n-1} x_j h_k(j/2^n) \frac{1}{2^n} = \frac{1}{\sqrt{2^n}} y_k,$$

где c_k представляет собой k -й коэффициент Фурье разложения функции $f(t)$ по системе Хаара.

Теорема 3.5. Матрица $\mathbf{H} = \left\{ \frac{1}{\sqrt{2^n}} h_k \left(\frac{j}{2^n} \right) \right\}_{k,j=0}^{2^n-1}$, определяющая ДПХ $\mathbf{Y} = \mathbf{H}\mathbf{X}$ — ортогональна, т. е. $\mathbf{H}^T = \mathbf{H}^{-1}$.

◀ Обозначим $\mathbf{U} = \mathbf{H}\mathbf{H}^T$ и покажем, что матрица \mathbf{U} — единичная. Рассмотрим элемент k -й строки j -го столбца матрицы \mathbf{U} , который представляет собой скалярное произведение k -й строки матрицы \mathbf{H} и j -го столбца матрицы \mathbf{H}^T (т. е. j -й строки матрицы \mathbf{H}):

$$u_{k,j} = \sum_{m=0}^{2^n-1} h_k \left(\frac{m}{2^n} \right) h_j \left(\frac{m}{2^n} \right) \frac{1}{2^n}.$$

Поскольку $\forall k = 0, \dots, 2^n-1: h_k(t) = \text{const}$ при $t \in \Delta_m^n = \left[\frac{m}{2^n}; \frac{m+1}{2^n} \right]$

(см. свойство 1° системы Хаара, пример 1.12), то выражение для $u_{k,j}$ можем переписать так:

$$u_{k,j} = \sum_{m=0}^{2^n-1} \underbrace{\int_{\Delta_m^n} h_k(t) h_j(t) dt}_{\text{const}} = \int_0^1 h_k(t) h_j(t) dt = \langle h_k(t), h_j(t) \rangle = \begin{cases} 1 & \text{при } k = j; \\ 0 & \text{при } k \neq j \end{cases}$$

в силу ортонормированности системы Хаара (см. теорему 1.11). ►

Следствие. Обратное к ДПХ (3.25) преобразование записывается в виде:

$$x_j = \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} y_k h_k(j/2^n), \quad j = 0, 1, \dots, 2^n - 1.$$

Пример 3.9. Найдем матрицу ДПХ для преобразования размерности $N = 2^2 = 4$.

◀ Построив по определению (см. пример 1.12) первые функции $h_0(t), h_1(t), h_2(t), h_3(t)$, находим значения $h_k(j/4)$, $k, j = 0, 1, 2, 3$. Получаем следующую структуру матрицы ДПХ:

$$\mathbf{H} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}.$$

ДПХ и ДПУ также обладают быстрыми алгоритмами вычислений. Причем данные преобразования вычисляются намного быстрее ДПФ, так как реализуются только за счет операций сложения и вычитания чисел, не требуя операций умножения или деления (за исключением окончательной нормировки коэффициентов, $\sim N$ умножений, что на практике требуется не всегда).

■ 3.10. Некоторые применения дискретных ортогональных преобразований

Кратко рассмотрим некоторые применения дискретных ортогональных преобразований.

Пример 3.10. Вычисление свертки векторов с использованием БПФ.

Для вектора-свертки $\mathbf{C} = (c_0, c_1, \dots, c_{2N-1})$ векторов

$$\mathbf{A} = (a_0, a_1, \dots, a_{N-1}) \text{ и } \mathbf{B} = (b_0, b_1, \dots, b_{N-1})$$

была определена формула (3.16), выписав которую по отдельности для каждой компоненты, можем подсчитать число умножений, необходимых для вычисления всего вектора \mathbf{C} . Получаем (см. табл.) $2(1 + 2 + \dots + (N-1)) + N = N^2$ умножений (вещественных или комплексных, в зависимости от характера векторов \mathbf{A} и \mathbf{B}).

Компонента свертки, формула вычисления	Число умножений
$c_0 = a_0 b_0$	1
$c_1 = a_0 b_1 + a_1 b_0$	2
...	...
$c_k = \sum_{j=0}^k a_j b_{k-j}$	$k+1$
...	...
$c_{N-1} = \sum_{j=0}^{N-1} a_j b_{N-1-j}$	N
$c_N = \sum_{j=1}^{N-1} a_j b_{N-j}$	$N-1$
$c_{N+1} = \sum_{j=2}^{N-1} a_j b_{N+1-j}$	$N-2$
...	...
$c_{2N-3} = a_{N-2} b_{N-1} + a_{N-1} b_{N-2}$	2
$c_{2N-2} = a_{N-1} b_{N-1}$	1

Положим, что $N = 2^n$. Используя свойство 5° ДПФ, можно вычислить вектор-свертку следующим образом:

1. Сформируем векторы

$$\check{\mathbf{A}} = (a_0, \dots, a_{N-1}, \underbrace{0, \dots, 0}_{N \text{ нулей}}), \quad \check{\mathbf{B}} = (b_0, \dots, b_{N-1}, \underbrace{0, \dots, 0}_{N \text{ нулей}})$$

и с использованием алгоритма БПФ получим векторы $\check{\mathbf{A}} = (\hat{a}_0, \hat{a}_1, \dots, \hat{a}_{2N-1})$ — ДПФ вектора $\check{\mathbf{A}}$ и $\check{\mathbf{B}} = (\hat{b}_0, \hat{b}_1, \dots, \hat{b}_{2N-1})$ — ДПФ вектора $\check{\mathbf{B}}$. Если $N = 2^n$, то каждое вычисление двух БПФ размерности $2N = 2^{n+1}$ потребует (см. раздел 3.6) примерно $(n+1)2^n$ комплексных умножений, т. е. всего около $(n+1)2^{n+1} = 2N \log_2(2N)$ умножений.

2. Вычислим вектор $\hat{\mathbf{C}} = (\hat{c}_0, \hat{c}_1, \dots, \hat{c}_{2N-1})$ (ДПФ вектора \mathbf{C}) перемножением компонент векторов $\check{\mathbf{A}}$ и $\check{\mathbf{B}}$, $\forall k : \hat{c}_k = \sqrt{2N} \hat{a}_k \hat{b}_k$. Без учета нормировки (умножения на $\sqrt{2N}$) для этого необходимо $2N$ комплексных умножений.

3. Вектор-свертку $\mathbf{C} = (c_0, c_1, \dots, c_{2N-1})$ получим далее в результате ОДПФ вектора $\hat{\mathbf{C}} = (\hat{c}_0, \hat{c}_1, \dots, \hat{c}_{2N-1})$ по быстрому алгоритму, для чего необходимо примерно $(n+1)2^n = N \log_2(2N)$ комплексных умножений.

Таким образом, общее число комплексных умножений, необходимых для вычисления свертки по описанной схеме, составит величину $\approx 3N \log_2(2N) + 2N = N(2 + 3 \log_2(2N)) \sim N \log N$, что при больших значениях N существенно меньше, чем N^2 .

Пример 3.11. Сжатие данных с потерями информации.

Пусть необходимо закодировать дискретный сигнал, представленный в виде вектора данных \mathbf{X} . Для некоторых ортогональных преобразований с матрицей \mathbf{F} и исходных векторов данных заранее известно, что часть компонент $\{y_k\}$ вектора $\mathbf{Y} = \mathbf{F}\mathbf{X}$ с большой вероятностью принимает малые по модулю значения. Если вместо вектора \mathbf{X} кодировать вектор $\check{\mathbf{Y}}$ меньшей размерности, образованный из \mathbf{Y} отбрасыванием указанных компонент, то происходит сжатие данных. При восстановлении данных отброшенные компоненты вектора \mathbf{Y} заменяются нулями, в результате чего получается вектор $\check{\mathbf{Y}}$, который отличается от исходного \mathbf{Y} наличием принудительно обнуленных компонент, «зон» нулей. Поэтому восстановленный вектор $\check{\mathbf{X}} = \mathbf{F}^{-1}\check{\mathbf{Y}}$ также отличается от исходного вектора \mathbf{X} . Если характеризовать потерю данных евклидовой нормой ошибки $\|\mathbf{X} - \check{\mathbf{X}}\|$, то в силу ортогональности используемого преобразования, так как $\mathbf{Y} - \check{\mathbf{Y}} = \mathbf{F}(\mathbf{X} - \check{\mathbf{X}})$, имеем: $\|\mathbf{X} - \check{\mathbf{X}}\|^2 = \|\mathbf{Y} - \check{\mathbf{Y}}\|^2 = \sum_{k \in \Theta} |y_k|^2$, где Θ — множество индексов обнуленных компонент вектора \mathbf{Y} . Поскольку для ортогональных преобразований величину ошибки кодирования можно анализировать в области преобразованных векторов \mathbf{Y} и $\check{\mathbf{Y}}$, такие пре-

образования используются для сжатия данных чаще всего. Подробнее это будет рассмотрено в главе 6.

■ 3.11. Квантование дискретных сигналов

До сих пор предполагалось, что полученные из непрерывного сигнала $f(t)$ дискретные отсчеты $\{f_k = f(k\Delta t)\}$ могут принимать любые значения из некоторого диапазона $[x_{\min}; x_{\max}]$. Для представления сигнала в цифровом виде дискретный сигнал дополнительно необходимо *проквантовать*, т. е. заменить каждый отсчет f_k значением из некоторого конечного множества $\Omega = \{d_j\}_{j=0}^{N-1}$ возможных уровней сигнала в соответствии с определенным правилом Q : $f_k \xrightarrow{Q} \hat{f}_k \in \Omega$. Возникает вопрос: как выбрать правило квантования Q ? Отсчеты дискретного сигнала часто описывают в терминах случайных величин, поэтому процесс квантования удобно представить как преобразование случайной величины непрерывного типа X в случайную величину дискретного типа: $\hat{X} = Q(X)$. В зависимости от того, насколько полна исходная информация о законе распределения случайной величины X и каковы технические возможности реализации квантователя (как устройства или программной процедуры), возможны и различные подходы к выбору правила (функции) Q при заданном числе уровней квантования N .

Равномерное квантование

Обычно равномерное квантование используется тогда, когда о случайной величине X известно лишь то, что она попадает в некоторый диапазон $X \in [x_{\min}; x_{\max}]$ и (или) необходимо обеспечить простейшую реализацию квантователя (т. е. устройства, реализующего правило Q). Для технического удобства изложения в качестве диапазона возможных значений случайной величины непрерывного типа выберем полуинтервал $X \in [x_{\min}; x_{\max}]$.

При равномерном квантовании диапазон $[x_{\min}; x_{\max}]$ разбивается на N равных интервалов длины $q = (x_{\max} - x_{\min})/N$:

$$\Delta_j = [x_{\min} + jq; x_{\min} + (j+1)q], \quad j = 0, \dots, N-1.$$

В качестве уровней квантования $\{d_j\}_{j=0}^{N-1}$ выбираются середины интервалов, $d_j = x_{\min} + (j + 0,5)q$. Правило квантования имеет простой вид: если $X \in \Delta_j$, то $Q(X) = d_j$. Очевидно, что ошибка квантования при этом: $|X - \hat{X}| \leq q/2$. Часто эту ошибку характеризуют некоторыми средними значениями, например используется величина

$$\overline{\varepsilon^2} = M[(X - \hat{X})^2], \quad (3.26)$$

где M означает математическое ожидание, а горизонтальная черта — операцию усреднения. Известно, что при достаточно малой величине дискрета квантования q ошибку квантования $\varepsilon = X - \hat{X}$ можно считать случайной величиной, подчиняющейся равномерному закону распределения на отрезке $[-q/2; q/2]$ вне зависимости от того, какой закон распределения имеет случайная величина X . Для равномерного закона распределения $\varepsilon \sim R[-q/2; q/2]$ математическое ожидание $M(\varepsilon) = 0$, а дисперсия ошибки квантования $D(\varepsilon) = \overline{\varepsilon^2} = q^2/12$.

Оптимальное квантование Ллойда — Макса

Вновь положим, что случайная величина непрерывного типа X принимает значения из какого-то диапазона $[x_{\min}; x_{\max}]$, причем известна ее функция плотности распределения вероятностей $f_X(x)$; $f_X(x) > 0$ при $X \in [x_{\min}; x_{\max}]$ и $f_X(x) = 0$ при $X \notin [x_{\min}; x_{\max}]$. Разобьем интервал $[x_{\min}; x_{\max}]$ на N , вообще говоря, неравных частей:

$$\Delta_j = [t_j; t_{j+1}), \quad j = 0, \dots, N-1; \quad t_0 = x_{\min}, \quad t_N = x_{\max}.$$

Внутри каждого полуинтервала Δ_j расположим точку d_j — уровень квантования. Правило квантования имеет тот же вид, что и ранее: если $X \in \Delta_j$, то $Q(X) = d_j$. Однако теперь параметры квантования — пороги $\{t_j\}_{j=1}^{N-1}$ и уровни $\{d_j\}_{j=0}^{N-1}$ — будем искать с точки зрения минимизации величины (3.26) как функции данных параметров:

$$\begin{aligned}\overline{\varepsilon^2} &= g(t_1, \dots, t_{N-1}; d_0, \dots, d_{N-1}) = M[(X - Q(X))^2] = \\ &= \int_0^{t_N} (x - Q(x))^2 f_X(x) dx = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} (x - d_j)^2 f_X(x) dx.\end{aligned}\quad (3.27)$$

Необходимым условием минимума функции $g(\dots)$ из выражения (3.27) является равенство нулю ее частных производных:

$$\begin{aligned}\frac{\partial g(\dots)}{\partial t_j} &= \frac{\partial}{\partial t_j} \left(\dots + \int_{t_{j-1}}^{t_j} (x - d_{j-1})^2 f_X(x) dx + \int_{t_j}^{t_{j+1}} (x - d_j)^2 f_X(x) dx + \dots \right) = \\ &= (t_j - d_{j-1})^2 f_X(t_j) - (t_j - d_j)^2 f_X(t_j) = 0.\end{aligned}$$

Отсюда, так как $f_X(t_j) \neq 0$, получаем: $|t_j - d_{j-1}| = |t_j - d_j|$. Поскольку $d_{j-1} \leq t_j \leq d_j$, то раскрывая модули как $t_j - d_{j-1} = d_j - t_j$, получим:

$$t_j = \frac{d_{j-1} + d_j}{2}, \quad j = 1, \dots, N-1. \quad (3.28)$$

Расположение порогов $\{t_j\}_{j=1}^{N-1}$ посередине между соседними уровнями $\{d_j\}_{j=0}^{N-1}$ означает, что округление случайной величины $\hat{X} = Q(X)$ происходит до ближайшего уровня квантования d_j : $|d_j - X| = \min_{d \in \{d_0, \dots, d_{N-1}\}} |d - X|$.

Для частных производных функции $g(\dots)$ из выражения (3.27) по переменным $\{d_j\}_{j=0}^{N-1}$ получим:

$$\frac{\partial g(\dots)}{\partial d_j} = -2 \int_{t_j}^{t_{j+1}} (x - d_j) f_X(x) dx = 0.$$

Отсюда

$$d_j = \int_{t_j}^{t_{j+1}} x f_X(x) dx / \int_{t_j}^{t_{j+1}} f_X(x) dx, \quad j = 0, \dots, N-1. \quad (3.29)$$

Заметим, что величина d_j из формулы (3.29) представляет собой математическое ожидание случайной величины X при условии, что последняя попала в интервал Δ_j : $d_j = M(X | X \in \Delta_j)$.

Решение системы уравнений (3.28), (3.29) определяет *квантователь Ллойда — Макса* и, как можно показать, дает минимальное значение ошибки (3.27). В общем случае решение системы уравнений (3.28), (3.29) необходимо искать численными методами. Причем квантователь Ллойда — Макса полностью определяется значениями уровней $\{d_j\}_{j=0}^{N-1}$, поскольку для любой функции плотности распределения $f_X(x)$ пороги квантования $\{t_j\}_{j=1}^{N-1}$ всегда находятся в точности посередине между оптимальными уровнями квантования (см. формулу (3.28)).

Отметим также, что все рассуждения, которые привели к построению системы уравнений (3.28), (3.29), сохраняются в силе и для более общего случая, когда функция плотности распределения принимает ненулевые значения на всей числовой оси и $t_0 = x_{\min} = -\infty$, $t_N = x_{\max} = +\infty$.

Пример 3.12. Найти параметры оптимального квантователя Ллойда — Макса для квантования с N уровнями случайной величины X , имеющей равномерное распределение на интервале $[a; b]$, т. е. $X \sim R[a; b]$.

◀ Функция плотности распределения имеет в данном случае вид:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{при } x \in [a; b]; \\ 0 & \text{при } x \notin [a; b]. \end{cases}$$

Поэтому для уравнений (3.29) получаем:

$$d_j = \int_{t_j}^{t_{j+1}} \frac{x}{b-a} dx \sqrt{\int_{t_j}^{t_{j+1}} \frac{1}{b-a} dx} = \frac{t_{j+1}^2 - t_j^2}{2(b-a)} = \frac{t_{j+1} + t_j}{2}, \quad j=0, \dots, N-1.$$

С учетом формулы (3.28) и условий $t_0 = a$, $t_N = b$, для параметров оптимального квантователя Ллойда — Макса получим окончательно следующую систему уравнений, которая является линейной и легко разрешима:

$$\begin{cases} t_0 = a; \quad t_N = b; \\ t_j = \frac{d_{j-1} + d_j}{2}, \quad j=1, \dots, N-1; \\ d_j = \frac{t_j + t_{j+1}}{2}, \quad j=0, \dots, N-1. \end{cases}$$

Решение системы — очевидно, оно определяет рассмотренный выше равномерный квантователь, который для равномерно распределенной случайной величины является оптимальным в смысле минимума ошибки (3.26). ►

Итак, процесс квантования отсчетов дискретного сигнала можно рассматривать как преобразование случайной величины непрерывного типа X в случайную величину дискретного типа $\hat{X} = Q(X)$. Трактуя уровни квантования $\{d_j\}_{j=0}^{N-1}$ как символы некоторого алфавита, цифровой сигнал можно считать «словом», составленным из этого алфавита, «буквы» которого представляют собой проквантованные отсчеты дискретного сигнала. Такое слово, или дискретное сообщение, можно представить в двоичном виде. Вопросы эффективного двоичного кодирования дискретных сообщений мы рассмотрим позднее в главе 5.

Выше мы рассмотрели способы квантования, при которых элементы исходного набора данных (например, отсчеты дискретного сигнала) обрабатываются независимо друг от друга. Такое покомпонентное квантование называют *скалярным*.

Векторное квантование

При *векторном* квантовании из исходного набора данных для последующей обработки выбираются уже не отдельные эл-

ементы, а некоторые M -компонентные блоки, которые можно рассматривать как векторы из пространства \mathbb{R}^M . Векторное квантование является эффективным способом сжатия и символьного описания данных, представленных в упорядоченных наборах (векторах). При векторном квантовании задается некоторое правило Q , по которому каждому вектору $\mathbf{X} \in \mathbb{R}^M$ ставится в соответствие некоторый вектор \mathbf{V} из того же пространства \mathbb{R}^M . Множество всех N возможных значений кодовых векторов $C = \{\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}\}$ называется *кодовой книгой*, а число векторов N — *объемом* кодовой книги. После векторного квантования передаче или хранению вместо исходного вектора \mathbf{X} подвергается индекс j вектора $\mathbf{V}_j = \hat{\mathbf{X}} = Q(\mathbf{X})$. При декодировании используется та же кодовая книга, что и при кодировании. (Положив размерность обрабатываемых векторов $M = 1$, получим из векторного квантования скалярное, а кодовая книга будет представлять собой ни что иное, как множество уровней квантования $C = \{d_0, \dots, d_{N-1}\}$.)

При векторном квантовании, в отличие от квантования скалярного, могут быть учтены межкомпонентные зависимости обрабатываемого набора данных, поэтому по сравнению со скалярным квантованием векторное квантование обладает более высоким потенциалом с точки зрения эффективного представления данных. Рассмотрим это на простом примере.

Пример 3.13. Пусть вектор $(X, Y) \in \mathbb{R}^2$ имеет равномерное распределение в области

$$G = \{0 \leq x \leq 1; 0 \leq y \leq 1\} \cup \{-1 \leq x < 0; -1 \leq y < 0\},$$

т. е. $f_{X,Y}(x, y) = \begin{cases} 1/2, & (x, y) \in G; \\ 0, & (x, y) \notin G. \end{cases}$ Необходимо найти параметры скалярного (независимого) квантования Ллойда — Макса с двумя уровнями для каждой компоненты и определить объем кодовой книги, которая позволяет осуществлять векторное квантование с той же или меньшей ошибкой.

► Для нахождения оптимальных параметров квантования Ллойда — Макса необходимо знать законы распределения компо-

ент. Найдем плотность распределения вероятностей для компоненты X :

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \begin{cases} 1/2, & x \in [-1;1]; \\ 0, & x \notin [-1;1]. \end{cases}$$

Воспользовавшись результатами решения примера 3.12, находим два оптимальных уровня квантования для компоненты X : $d_0 = -1/2$, $d_1 = 1/2$. В силу симметрии функции $f_{X,Y}(x,y)$ относительно аргументов, функция плотности распределения для компоненты y имеет тот же вид, что и для компоненты x , поэтому теми же будут и уровни оптимального квантования Ллойда — Макса. При независимой обработке компонент результат квантования представляется в виде пары двоичных «символов» $\hat{X} \in \{-1/2; 1/2\}$ и $\hat{Y} \in \{-1/2; 1/2\}$. Вместе с тем очевидно (см. рис. 3.15), что при совместном рассмотрении компонент вектора (X, Y) результат квантования (\hat{X}, \hat{Y}) можно представить одним двоичным символом, который является номером кодового вектора в кодовой книге $C = \{\mathbf{V}_0 = (-1/2; -1/2), \mathbf{V}_1 = (1/2; 1/2)\}$. Объем кодовой книги $N = 2$, при этом ошибка векторного квантования совпадает с ошибкой скалярного квантования. ►

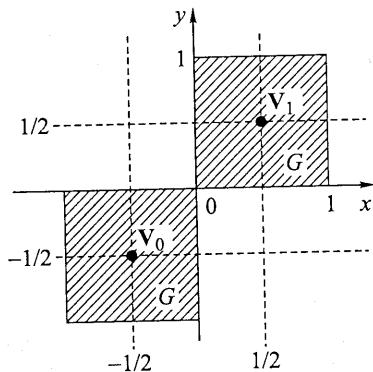


Рис. 3.15

Ошибку, возникающую в результате векторного квантования, характеризуют некоторой нормой $\|\mathbf{X} - \mathbf{V}_j\|$. Для каждого входно-

го вектора \mathbf{X} векторный квантователь \mathcal{Q} ставит в соответствие выходной вектор $\mathbf{V}_j^* = \mathcal{Q}(\mathbf{X})$, который выбирается из кодовой книги по критерию максимальной близости:

$$\|\mathbf{X} - \mathbf{V}_j^*\| = \min_{C=\{\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}\}} \|\mathbf{X} - \mathbf{V}_j\|$$

(в случае скалярного квантования это определялось условием (3.28)). Таким образом, квантователь (функция \mathcal{Q}) полностью определяется кодовой книгой C . Поэтому построение оптимального векторного квантователя для заданного объема кодовой книги N можно понимать как отыскание таких векторов $C = \{\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}\}$, которые дают минимальное значение математического ожидания ошибки квантования:

$$\bar{e}(\mathcal{Q}) = \bar{e}(\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}) = M(\|\mathbf{X} - \mathcal{Q}(\mathbf{X})\|) \rightarrow \min_{\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}}.$$

В качестве векторной нормы чаще всего используется евклидова норма, а кодовая книга ищется по критерию

$$\bar{e}^2(\mathcal{Q}) = \bar{e}^2(\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}) = M(\|\mathbf{X} - \mathcal{Q}(\mathbf{X})\|_E^2) \rightarrow \min_{\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}}. \quad (3.30)$$

Соответствующий квантователь называют *обобщенным квантователем Ллойда — Макса*. При известной функции плотности совместного распределения $f(x_1, \dots, x_M)$ компонент обрабатываемого вектора \mathbf{X} поиск оптимального векторного квантователя по критерию (3.30) означает минимизацию функции

$$\bar{e}^2(\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}) = \int_{\mathbb{R}^M} \|\mathbf{X} - \mathcal{Q}(\mathbf{X})\|_E^2 f(x_1, \dots, x_M) dx_1 \dots dx_M. \quad (3.31)$$

Обратите внимание на то, что для $M = 1$ выражение (3.31) переходит в частный случай — формулу (3.27), где надо положить $t_0 = -\infty$, $t_N = +\infty$.

Аналитическое выражение для функции плотности совместного распределения может быть неизвестно, но даже при заданной функции $f(x_1, \dots, x_M)$ нахождение векторов $C = \{\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}\}$, которые обращают в минимум ошибку (3.31), для большинства задач, возникающих на практике, представляет значительные

сложности. Поэтому для построения кодовой книги используются квазиоптимальные методы.

Рассмотрим кратко классический, наиболее известный, алгоритм построения кодовой книги *Linde—Buzo—Gray* (LBG). Кодовая книга $C = \{\mathbf{V}_k\}_{k=0}^{N-1}$ строится итерационно по некоторой заданной обучающей последовательности векторов $A = \{\mathbf{a}_0, \dots, \mathbf{a}_{L-1}\}$ — выборке объема $L \gg N$ из генеральной совокупности с неизвестным законом распределения.

Шаг 1. Инициализация.

Задать некоторую начальную кодовую книгу $C_0 = \{\mathbf{V}_k^{(0)}\}_{k=0}^{N-1}$.

Положить среднюю ошибку квантования $E_0 = +\infty$.

Установить счетчик итераций $i = 0$.

Шаг 2. Построить разбиение $\Omega_i = \{\omega_k\}_{k=0}^{N-1}$ множества обучающих векторов A , где ячейки разбиения определяются по текущей кодовой книге $C_i = \{\mathbf{V}_k^{(i)}\}_{k=0}^{N-1}$ следующим образом:

$$\omega_k = \left\{ \mathbf{a}_m \in A \mid \left\| \mathbf{a}_m - \mathbf{V}_k^{(i)} \right\| = \min_{j=0, \dots, N-1} \left\| \mathbf{a}_m - \mathbf{V}_j^{(i)} \right\| \right\}.$$

Шаг 3. Вычислить средний квадрат ошибки квантования:

$$E_i = \frac{1}{L} \sum_{k=0}^{N-1} \sum_{\mathbf{a}_m \in \omega_k} \left\| \mathbf{a}_m - \mathbf{V}_k^{(i)} \right\|^2.$$

Шаг 4. Проверка возможности останова. Если $(E_{i-1} - E_i)/E_i < \varepsilon$, то прекратить вычисления, положив $C = C_i$, иначе перейти на шаг 5.

Шаг 5. Найти новые кодовые векторы усреднением по ячейкам:

$$\mathbf{V}_k^{(i)} = \frac{1}{|\{\mathbf{a}_m \in \omega_k\}|} \sum_{\mathbf{a}_m \in \omega_k} \mathbf{a}_m, \quad k = 0, \dots, N-1,$$

где $|\{\mathbf{a}_m \in \omega_k\}|$ — количество векторов обучающей последовательности A , попавших в ячейку ω_k .

Положить $i = i + 1$ и перейти на шаг 2.

Конец.

Как видно из приведенного описания алгоритма LBG, его идея состоит в следующем. По заданной каким-то образом начальной кодовой книге (шаг 1) множество обучающих векторов A разбивается на классы (подмножества) $\{\omega_k\}_{k=0}^{N-1}$ по принципу минимального отклонения от соответствующих векторов кодовой книги $\{\mathbf{V}_k^{(0)}\}_{k=0}^{N-1}$ (шаг 2). Затем кодовые векторы заменяются усредненными векторами, найденными для каждого класса $\{\omega_k\}_{k=0}^{N-1}$ (шаг 5). Процедура повторяется до тех пор, пока изменение ошибки квантования обучающей последовательности не станет малым (шаги 3, 4). К сожалению, результаты работы алгоритма LBG во многом зависят от того, насколько удачно была выбрана начальная кодовая книга $C^0 = \{\mathbf{V}_k^0\}_{k=0}^{N-1}$. Алгоритм не дает гарантии, что для заданной последовательности обучающих векторов итоговая ошибка (см. шаг 3)

$$E(\mathbf{V}_0, \dots, \mathbf{V}_{N-1}) = \frac{1}{L} \sum_{k=0}^{N-1} \sum_{\mathbf{a}_m \in \omega_k} \left\| \mathbf{a}_m - \mathbf{V}_k \right\|^2$$

достигнет своего глобального (а не локального) минимума.

Несмотря на то что векторное квантование потенциально позволяет более эффективно представить информацию, которая заключена в обрабатываемых данных, применение процедуры векторного квантования на практике существенно ограничено, прежде всего, вычислительной сложностью поиска векторов в кодовой книге. Сложным, в общем случае, вопросом является также построение оптимальной кодовой книги. Поэтому скалярное квантование используется значительно чаще, векторное квантование применяется в основном для обработки векторов небольшой размерности и с малым объемом кодовых книг. Для снижения отрицательного эффекта (наподобие рассмотренного в примере 3.13), который проявляется при скалярном квантовании и связан с зависимостью компонент обрабатываемых данных, чаще используются другие приемы, которые будут подробнее рассмотрены далее в главе 6.

Глава 4

Линейные дискретные системы

В предыдущей главе мы рассмотрели вопросы преобразования аналогового сигнала в дискретный: $f(t) \rightarrow \{f_n = f(n\Delta t)\}_{n=-\infty}^{\infty}$. Величина Δt , определенная, например, в соответствии с критерием (3.7), характеризует масштаб по оси абсцисс, выбранный для представления дискретного сигнала во временной области. Однако, не меняя общности рассуждений, мы можем для упрощения изложения считать $\Delta t = 1$ (проведя своего рода «нормировку» масштаба времени дискретного сигнала). Тогда для спектра дискретного сигнала (3.9) верхняя частота сигнала $F_{\max} = \frac{1}{2\Delta t}$ (см. (3.7)) будет соответствовать нормированной частоте $v = 1/2$, а период спектральной функции (3.9) (см. рис. 3.3) будет единичным. Обозначив через $\tilde{S}_0(v)$ спектр дискретного сигнала для $\Delta t = 1$, спектр $\tilde{S}(v)$ для $\Delta t \neq 1$ всегда можно выразить с помощью соотношения $\tilde{S}(v) = \Delta t \tilde{S}_0(v\Delta t)$, которое непосредственно следует из формулы (3.9). В данной главе будем полагать, что $\Delta t = 1$ и $f_n = f(n)$.

■ 4.1. Z-преобразование

Определение. Пусть $\{x(n)\}_{n=-\infty}^{\infty}$ — последовательность комплексных чисел, причем $x(n) = 0$ при $n < 0$, и существует верхний предел $\limsup_{n \rightarrow \infty} \sqrt[n]{|x(n)|} = r < \infty$. Z-преобразованием этой последовательности назовем следующую функцию $X(z)$ комплексной переменной z :

$$X(z) = \begin{cases} \phi(z) = \sum_{n=0}^{\infty} x(n)z^{-n} \text{ для } |z| > r; \\ \text{аналитическое продолжение функции } \phi(z) \text{ для } |z| \leq r. \end{cases} \quad (4.1)$$

Будем использовать обозначение $X(z) = Z\{x(n)\}$.

Поясним определение. Функция $\phi(z)$, фигурирующая в выражении (4.1), представляет собой сумму ряда Лорана специального вида: $\sum_{n=0}^{\infty} x(n)z^{-n}$. В общем случае ряд Лорана имеет следующий вид:

$$\sum_{n=-\infty}^{\infty} c_n(z - z_0)^n, \quad (4.2)$$

где z_0 — точка комплексной плоскости, c_n — комплексные числа.

Напомним следующую теорему.

Теорема 4.1. Если коэффициенты ряда Лорана (4.2) таковы, что $\limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} = r < R = \frac{1}{\liminf_{n \rightarrow \infty} \sqrt[n]{|c_n|}}$, то ряд (4.2) сходится абсолютно

в кольце $r < |z - z_0| < R$, в котором сумма $f(z) = \sum_{n=-\infty}^{\infty} c_n(z - z_0)^n$ является аналитической функцией, а коэффициенты ряда выражаются с помощью контурных интегралов:

$$c_n = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz, \quad (4.3)$$

где γ — произвольный замкнутый контур (обход контура — в положительном направлении), целиком лежащий в кольце $r < |z - z_0| < R$ и охватывающий точку z_0 .

В соответствии с теоремой 4.1 ряд $\sum_{n=0}^{\infty} x(n)z^{-n}$ абсолютно сходится в области $r < |z| (R = \infty)$, где его сумма $\phi(z)$ представляет собой аналитическую функцию. Особые точки функции $X(z)$

могут быть только в круге $|z| \leq r$, в который аналитически продолжена функция $\varphi(z)$.

Пример 4.1. Найти Z-преобразование последовательности $x(n) = n$, $n = 0, 1, \dots$

◀ Для $|z| > \lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$ имеем:

$$\begin{aligned} X(z) &= \sum_{n=0}^{\infty} nz^{-n} = -z \sum_{n=0}^{\infty} \frac{d}{dz} z^{-n} = -z \frac{d}{dz} \left(\sum_{n=0}^{\infty} z^{-n} \right) = \\ &= -z \frac{d}{dz} \left(\frac{z}{z-1} \right) = \frac{z}{(z-1)^2}. \end{aligned}$$

Заметим, что использование формулы для суммы геометрической прогрессии было корректно только для $|z| > 1$. Однако мы считаем полученное выражение верным и для $|z| \leq 1$, аналитически продолжая функцию $X(z)$, т. е. $\forall z \in \mathbb{C}$ (множество комплексных чисел): $Z\{x(n)\} = z/(z-1)^2$. ►

Из теоремы 4.1 и формулы (4.3) непосредственно следует, что элементы последовательности $x(n)$ могут быть восстановлены по ее Z-образу $X(z)$ с помощью формулы

$$x(n) = \frac{1}{2\pi i} \oint_{\gamma} X(z) z^{n-1} dz, \quad (4.4)$$

где γ — произвольный замкнутый контур, лежащий в области аналитичности функции $X(z)$ и охватывающий все ее особые точки. Соответствие между последовательностью $\{x(n)\}_{n=0}^{\infty}$ и ее Z-образом $X(z) = Z\{x(n)\}$ обозначим так: $x(n) \leftrightarrow X(z)$.

Рассмотрим некоторые свойства Z-преобразования.

1°. Линейность.

$$\forall \alpha, \beta: \alpha x(n) + \beta y(n) \leftrightarrow \alpha Z\{x(n)\} + \beta Z\{y(n)\} = \alpha X(z) + \beta Y(z).$$

2°. Задержка последовательности на $N \geq 0$ отсчетов:

$$x(n-N) \leftrightarrow z^{-N} X(z).$$

◀ Так как $x(n) = 0$ при $n < 0$, имеем:

$$\begin{aligned} Z\{x(n-N)\} &= \sum_{n=0}^{\infty} x(n-N) z^{-n} = \sum_{k=-N}^{\infty} x(k) z^{-k-N} = \\ &= z^{-N} \sum_{k=0}^{\infty} x(k) z^{-k} = z^{-N} X(z). \end{aligned}$$

3°. Опережающий сдвиг последовательности на $M \geq 0$ отсчетов:

$$x(n+M) \leftrightarrow z^M (X(z) - x(0) - x(1)z^{-1} - \dots - x(M-1)z^{-(M-1)}).$$

$$\begin{aligned} \blacktriangleleft \sum_{n=0}^{\infty} x(n+M) z^{-n} &= \sum_{k=M}^{\infty} x(k) z^{-k+M} = \\ &= z^M \left(\sum_{k=0}^{\infty} x(k) z^{-k} - \sum_{k=0}^{M-1} x(k) z^{-k} \right). \end{aligned}$$

4°. Умножение последовательности на a^n : $a^n x(n) \leftrightarrow X(z/a)$.

$$\blacktriangleleft \sum_{n=0}^{\infty} a^n x(n) z^{-n} = \sum_{n=0}^{\infty} x(n) (z/a)^{-n} = X(z/a). \quad \blacktriangleright$$

5°. Умножение последовательности на n :

$$nx(n) \leftrightarrow -z \frac{d}{dz} X(z).$$

◀ Фактически мы уже воспользовались этим свойством в примере 4.1.

$$\begin{aligned} \sum_{n=0}^{\infty} nx(n) z^{-n} &= -z \sum_{n=0}^{\infty} x(n) \frac{d}{dz} z^{-n} = \\ &= -z \frac{d}{dz} \left(\sum_{n=0}^{\infty} x(n) z^{-n} \right) = -z \frac{d}{dz} X(z). \end{aligned}$$

Определение. Сверткой бесконечных последовательностей $x(n)$, $y(n)$, назовем последовательность $u(n)$, элементы которой находятся по формуле $u(n) = \sum_{k=0}^{\infty} x(k)y(n-k)$. Для свертки используем обозначение $u(n) = x(n)^*y(n)$.

Замечание. Так как вновь считается, что $x(n) = y(n) = 0$ при $n < 0$, то $u(n) = 0$ при $n < 0$ и, как несложно показать, следующие определения свертки являются эквивалентными:

$$\begin{aligned} u(n) &= \sum_{k=-\infty}^{\infty} x(k)y(n-k) = \sum_{k=0}^n x(k)y(n-k) = \\ &= \sum_{k=0}^n y(k)x(n-k) = \sum_{k=-\infty}^{\infty} y(k)x(n-k). \end{aligned}$$

6°. Z-преобразование свертки: $x(n)*y(n) \leftrightarrow X(z)Y(z)$.

◀ Используя доказанное выше свойство 2°, получаем:

$$\begin{aligned} Z\{x(n)*y(n)\} &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\infty} x(k)y(n-k) \right) z^{-n} = \sum_{k=0}^{\infty} x(k) \sum_{n=0}^{\infty} y(n-k) z^{-n} = \\ &= \sum_{k=0}^{\infty} x(k) z^{-k} Y(z) = X(z)Y(z). \blacksquare \end{aligned}$$

Для вычисления интегралов (4.4) при обращении Z-преобразования удобно использовать следующую теорему о вычетах.

Теорема 4.2. Пусть $f(z)$ — функция, аналитическая в некоторой области D , за исключением точек z_1, \dots, z_N ($z_k \neq \infty$, $k = 1, \dots, N$), лежащих в этой области. Тогда если γ — замкнутый контур, целиком лежащий в области D и охватывающий точки $\{z_k\}_{k=1}^N$, то

$$\frac{1}{2\pi i} \oint_{\gamma} f(z) dz = \sum_{k=1}^N \text{res}[f(z), z_k], \quad (4.5)$$

где $\text{res}[f(z), z_k]$ — вычет функции $f(z)$ в точке z_k .

Таким образом (см. (4.4)), для обращения Z-преобразования в формулу (4.5) нужно подставить функцию $f(z) = X(z)z^{n-1}$.

Напомним, что если в окрестности некоторой точки z_0 известно разложение функции в ряд Лорана (4.2), $f(z) = \sum_{k=-\infty}^{\infty} c_k (z - z_0)^k$, то (см. (4.3)) известен и вычет этой функции в данной точке:

$\text{res}[f(z), z_0] = c_{-1}$. На практике часто используются следующие способы вычисления вычетов.

1. Если z_0 — полюс порядка m (это означает, что в ряде (4.2) $c_{-m} \neq 0$, а для $j > m$ $c_{-j} = 0$), то

$$c_{-1} = \text{res}[f(z), z_0] = \frac{1}{(m-1)!} \lim_{z \rightarrow z_0} \frac{d^{m-1}[(z - z_0)^m f(z)]}{dz^{m-1}}.$$

В частном случае для полюса первого порядка

$$c_{-1} = \text{res}[f(z), z_0] = \lim_{z \rightarrow z_0} ((z - z_0)f(z)).$$

2. Если в некоторой окрестности точки z_0 $f(z) = u(z)/v(z)$ и $u(z_0) \neq 0$, $v(z_0) = 0$, $\frac{d}{dz} v(z_0) \neq 0$, то $c_{-1} = \text{res}[f(z), z_0] = u(z_0)/v'(z_0)$.

Пример 4.2. Найти последовательность $x(n)$ по ее Z-образу

$$X(z) = \frac{z}{z^2 + 1}.$$

◀ Функция $X(z)$ имеет две особые точки — полюсы первого порядка: $z = \pm i$. Используя формулы (4.4) и (4.5), получаем:

$$\begin{aligned} x(n) &= \frac{1}{2\pi i} \oint_{\gamma} \frac{z^n}{z^2 + 1} dz = \text{res}\left[\frac{z^n}{z^2 + 1}, z = i\right] + \text{res}\left[\frac{z^n}{z^2 + 1}, z = -i\right] = \\ &= \lim_{z \rightarrow i} \frac{z^n(z - i)}{(z + i)(z - i)} + \lim_{z \rightarrow -i} \frac{z^n(z + i)}{(z + i)(z - i)} = \frac{i^n - (-i)^n}{2i} = \sin \frac{\pi n}{2}. \blacksquare \end{aligned}$$

С использованием математического аппарата Z-преобразования удобно решать разностные уравнения.

Определение. Линейным разностным уравнением N -го порядка с постоянными коэффициентами называется уравнение

$$\begin{aligned} \sum_{k=0}^N a_k x(n+k) &= f(n), \\ n &= 0, 1, \dots, \end{aligned} \quad (4.6)$$

где $a_k, f(n)$ — заданные числа (вообще говоря, комплексные), $\{x(n)\}$ — искомая последовательность (решение уравнения). Не умаляя общности, можно положить $a_N = 1$.

Разностные уравнения являются дискретными аналогами дифференциальных уравнений. Для того чтобы уравнение (4.6) имело единственное решение, необходимо задать, например, « начальные условия», т. е. значения N первых элементов искомой последовательности: $x(0) = x_0, \dots, x(N - 1) = x_{N-1}$. Тогда можно выразить оставшиеся члены последовательности через найденные ранее, записав (4.6) в рекуррентном виде:

$$x(N + n) = f(n) - \sum_{k=0}^{N-1} a_k x(n+k), \quad n = 0, 1, \dots$$

Такая запись определяет некоторую расчетную процедуру. Получить аналитическое выражение для решения уравнения (4.6) можно с использованием свойства 3° Z-преобразования. Рассмотрим это на примере.

Пример 4.3. Решить уравнение $x(j+2) - 2x(j+1) + x(j) = -4$ с условиями: $x(0) = 0, x(1) = 2$.

◀ Так как

$$\begin{aligned} Z\{x(j+2)\} &= z^2 \left(X(z) - x(0) - \frac{x(1)}{z} \right) = z^2 \left(X(z) - \frac{2}{z} \right), \\ Z\{x(j+1)\} &= z(X(z) - x(0)) = zX(z), \quad Z\{-4\} = \sum_{n=0}^{\infty} (-4)z^{-n} = \frac{-4z}{z-1}, \end{aligned}$$

то для исходного разностного уравнения получаем следующий Z-образ:

$$z^2 \left(X(z) - \frac{2}{z} \right) - 2zX(z) + X(z) = -\frac{4z}{z-1},$$

откуда

$$X(z) = \frac{2z^2 - 6z}{(z-1)^3}.$$

Обращая Z-преобразование при помощи формул (4.4), (4.5), находим:

$$x(n) = \text{res} \left[\frac{2z^2 - 6z}{(z-1)^3} z^{n-1}, z=1 \right] = \frac{1}{2} \lim_{z \rightarrow 1} \frac{d^2((2z-6)z^n)}{dz^2} = 2n(2-n). \blacktriangleright$$

■ 4.2. Линейные дискретные фильтры (ЛДФ)

Определение. Линейным дискретным фильтром (линейной дискретной системой) назовем устройство, которое преобразует входную последовательность $x(n)$ в выходную $y(n)$ по правилу, определяемому для любого целого значения n следующим разностным уравнением

$$\sum_{m=0}^M a_m y(n-m) = \sum_{k=0}^N b_k x(n-k), \quad (4.7)$$

где a_m, b_k — некоторые константы (вообще говоря, комплексные). Вновь полагаем, что $x(n) = y(n) = 0$ при $n < 0$.

Таким образом, параметрами системы являются наборы коэффициентов $\{a_m\}_{m=0}^M, \{b_k\}_{k=0}^N$. Постоянство параметров характеризует свойство *стационарности* системы. Линейность системы, алгоритм работы которой определяется уравнением (4.7), очевидна. Действительно, если пары последовательностей $\{x_1(n), y_1(n)\}$ и $\{x_2(n), y_2(n)\}$ удовлетворяют уравнению (4.7), то и пара последовательностей

$$\{x(n) = \alpha x_1(n) + \beta x_2(n), y(n) = \alpha y_1(n) + \beta y_2(n)\},$$

где α, β — произвольные числа, также ему удовлетворяет.

Положив $a_0 = 1$ (при необходимости все коэффициенты уравнения (4.7) делим на a_0), уравнение (4.7) можно переписать в другом виде:

$$y(n) = -\sum_{m=1}^M a_m y(n-m) + \sum_{k=0}^N b_k x(n-k), \quad (4.8)$$

что задает в виде рекуррентной формулы некоторую расчетную процедуру нахождения очередного элемента выходной последо-

вательности по уже известным элементам входной и выходной последовательностей.

Преобразование уравнения (4.7) к виду (4.8) корректно, поскольку если в уравнении (4.7) коэффициент $a_0 = 0$, то обязательно $b_0 = 0$. Тогда перенумеруем коэффициенты в (4.7) так, чтобы получить $a_0 \neq 0$. Действительно, при $a_0 = 0$ и $b_0 \neq 0$ уравнение (4.8) превращается в уравнение физически нереализуемой системы

$$(полагая уже $a_1 = 1$): y(n-1) = -\sum_{m=2}^M a_m y(n-m) + \sum_{k=0}^N b_k x(n-k), \text{ где}$$

отсчет выходной последовательности $y(n-1)$ зависит от еще не поступившего на вход системы отсчета $x(n)$.

Если в уравнении (4.8) хотя бы один коэффициент $a_m \neq 0$, то фильтр называется *рекурсивным*, в противном случае — *нерекурсивным*.

Пример 4.4. Найти реакцию (отклик) $y(n)$ рекурсивного фильтра, описываемого разностным уравнением $y(n) = ay(n-1) + x(n)$, на входное воздействие $x(n) = \tilde{\delta}(n) = \begin{cases} 1, & n=0 \\ 0, & n \neq 0 \end{cases}$ при условии, что

$y(n) = 0$ для $n < 0$.

$$\blacktriangleleft y(0) = 1, y(1) = ay(0) + x(1) = a, \dots, y(n) = ay(n-1) = a^n, \dots \blacktriangleright$$

Пример 4.5. Найти реакцию нерекурсивного фильтра, описанного разностным уравнением $y(n) = x(n) + bx(n-1)$, на входное воздействие $x(n) = \tilde{\delta}(n)$.

$$\blacktriangleleft y(0) = 1, y(1) = b, \forall n > 1: y(n) = 0. \blacktriangleright$$

Определение. Реакция (отклик) ЛДФ на единичное воздействие $x(n) = \tilde{\delta}(n) = \begin{cases} 1, & n=0 \\ 0, & n \neq 0 \end{cases}$ (при условии, что $y(n) = 0$ для $n < 0$) называется *импульсной характеристикой* (ИХ) этого ЛДФ. Для ИХ используется обозначение $h(n)$, т. е. $h(n) = y(n)|_{x(n)=\tilde{\delta}(n)}$.

В примерах 4.4 и 4.5 как раз и были найдены импульсные характеристики некоторых фильтров. В примере 4.4 получена *бесконечная ИХ* ($\forall M \exists m > M: h(m) \neq 0$), а в примере 4.5 рассмотрен

фильтр с *конечной ИХ* ($\exists M \forall m > M: h(m) = 0$). Фильтры с конечной ИХ называют КИХ-фильтрами, с бесконечной — БИХ-фильтрами.

Зная ИХ фильтра $h(n)$, можно найти реакцию фильтра $y(n)$ на произвольное входное воздействие $x(n)$ следующим образом. Представим входной сигнал как

$$x(n) = \sum_{m=0}^{\infty} x(m) \underbrace{\tilde{\delta}(n-m)}_{0 \text{ при } n \neq m}.$$

Вследствие стационарности фильтра его реакция на сигнал $\tilde{\delta}(n-m)$ ($n = 0, 1, \dots$) представляет собой последовательность $h(n-m)$ ($n = 0, 1, \dots$). Поэтому в силу линейности фильтра реакцию на сигнал $x(n) = \sum_{m=0}^{\infty} x(m) \tilde{\delta}(n-m)$ получим $y(n) = \sum_{m=0}^{\infty} x(m) h(n-m)$.

Таким образом, по известной ИХ отклик фильтра находится с помощью дискретной свертки:

$$y(n) = \sum_{m=0}^{\infty} x(m) h(n-m) = \sum_{m=0}^n x(m) h(n-m) = \sum_{m=0}^n h(m) x(n-m). \quad (4.9)$$

Пример 4.6. По заданной ИХ $h(n)$ найти реакцию фильтра на входное воздействие $x(n) = 1$.

\blacktriangleleft По формуле (4.9) получим $y(n) = \sum_{m=0}^n h(m)$ — данный выходной сигнал (отклик на воздействие $x(n) = 1$) называют также *переходной характеристикой*. \blacktriangleright

Другой важнейшей характеристикой фильтра является передаточная функция.

Определение. *Передаточной функцией* ЛДФ называется отношение Z -образов выходной и входной последовательностей: $H(z) = Y(z)/X(z)$.

Пример 4.7. По заданным входной $x(n) = \{1, 0, 1, 2, 0, 0, 0, \dots\}$ и выходной $y(n) = \{0, 1, 2, 3, 0, 0, 0, \dots\}$ последовательностям найти передаточную функцию фильтра.

◀ По определению:

$$H(z) = \frac{Z\{y(n)\}}{Z\{x(n)\}} = \frac{z^{-1} + 2z^{-2} + 3z^{-3}}{1 + z^{-2} + 2z^{-3}} = \frac{z^2 + 2z + 3}{z^3 + z + 2}. ▶$$

Воспользуемся свойствами 1°, 2° Z-преобразования и найдем

Z-образ уравнения (4.8): $Y(z) = -\sum_{m=1}^M a_m Y(z)z^{-m} + \sum_{k=0}^N b_k X(z)z^{-k}$.

Отсюда

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^N b_k z^{-k}}{1 + \sum_{m=1}^M a_m z^{-m}}. \quad (4.10)$$

Таким образом, передаточная функция не зависит от входного воздействия и характеризует свойства самого ЛДФ. Зная алгоритм работы ЛДФ, т. е. наборы коэффициентов в разностном уравнении (4.8), можно сразу записать передаточную функцию в виде (4.10). По известной передаточной функции отклика фильтра на произвольное входное воздействие можно определить, найдя на начала Z-образ входной последовательности $X(z) = Z\{x(n)\}$. Тогда, в силу свойства 6° Z-преобразования, из уравнения (4.9) получаем $Y(z) = H(z)X(z)$ и $y(n) = Z^{-1}\{Y(z)\}$.

Пример 4.8. Найти передаточную функцию для фильтра из примера 4.4.

$$◀ H(z) = \frac{1}{1 + (-az^{-1})} = \frac{z}{z - a}. ▶$$

Между передаточной функцией и импульсной характеристики существует очевидная связь. Так как передаточная функция не зависит от входного воздействия, возьмем $x(n) = \tilde{\delta}(n)$. Тогда $H(z) = Z\{h(n)\}/Z\{\tilde{\delta}(n)\} = Z\{h(n)\}$. Таким образом, передаточная функция ЛДФ представляет собой Z-преобразование ИХ.

Пример 4.9. Для ЛДФ, описываемого уравнением

$$y(n) = -2y(n-1) - y(n-2) + x(n) + x(n-1),$$

найти:

а) передаточную функцию фильтра;

б) импульсную характеристику фильтра;

в) отклик фильтра на входное воздействие $x(n) = 3^n$ с использованием уравнения свертки (4.9), а также посредством обращения Z-образа отклика фильтра.

◀ а) Сравнивая заданное разностное уравнение с уравнением в общем виде (4.8), находим ненулевые коэффициенты ЛДФ: $a_1 = 2$, $a_2 = 1$, $b_0 = 1$, $b_1 = 1$. Поэтому с учетом выражения (4.10)

$$H(z) = \frac{1 + z^{-1}}{1 + 2z^{-1} + z^{-2}} = \frac{z}{z + 1};$$

б) $h(n) = Z^{-1}\{H(z)\} = \text{res}\left[\frac{z^n}{z + 1}, z = -1\right] = (-1)^n$;

в) Используя формулу свертки (4.9), находим

$$\begin{aligned} y(n) &= \sum_{k=0}^n x(k)h(n-k) = \sum_{k=0}^n 3^k (-1)^{n-k} = (-1)^n \sum_{k=0}^n (-3)^k = \\ &= (-1)^n \frac{1 - (-3)^{n+1}}{4} = \frac{(-1)^n + 3^{n+1}}{4}. \end{aligned}$$

Выполнив Z-преобразование входного сигнала, получаем:

$$X(z) = Z\{3^n\} = \frac{z}{z - 3} \quad \text{и} \quad Y(z) = H(z)X(z) = \frac{z^2}{(z + 1)(z - 3)}.$$

Отсюда

$$\begin{aligned} y(n) &= Z^{-1}\left\{\frac{z^2}{(z + 1)(z - 3)}\right\} = \text{res}\left[\frac{z^{n+1}}{(z + 1)(z - 3)}, z = -1\right] + \\ &+ \text{res}\left[\frac{z^{n+1}}{(z + 1)(z - 3)}, z = 3\right] = \frac{(-1)^{n+1}}{-4} + \frac{3^{n+1}}{4} = \frac{(-1)^n + 3^{n+1}}{4}. ▶ \end{aligned}$$

Из формулы (4.10) следует важный для последующего рассмотрения факт: передаточная функция ЛДФ может быть представлена в виде отношения некоторых степенных многочленов неотрицательной степени: $H(z) = P(z)/Q(z)$ (для этого необходимо умножить числитель и знаменатель передаточной функции (4.10) на величину $z^{\max(M, N)}$). Поэтому передаточные функции ЛДФ, оп-

ределяемых по (4.7) или (4.8), имеют конечное число нулей — точек комплексной плоскости, в которых $P(z) = 0$, и полюсов — точек, для которых $Q(z) = 0$.

■ 4.3. Соединения и структурные схемы фильтров

1. Последовательное соединение фильтров (рис. 4.1, а). Очевидно, что $Y(z) = H_2(z)V(z) = H_2(z)H_1(z)X(z)$, т. е. передаточная функция

$$H(z) = H_2(z)H_1(z).$$

2. Параллельное соединение фильтров (рис. 4.1, б). Очевидно, что

$$Y(z) = H_1(z)X(z) + H_2(z)X(z).$$

Отсюда передаточная функция:

$$H(z) = H_1(z) + H_2(z).$$

3. Соединение с обратной связью (рис. 4.1, в). Выполняя несложные преобразования, получим:

$$\begin{aligned} Y(z) &= H_1(z)(X(z) + U(z)) = \\ &= H_1(z)(X(z) + H_2(z)Y(z)). \end{aligned}$$

Отсюда

$$H(z) = \frac{H_1(z)}{1 - H_1(z)H_2(z)}.$$

Рассмотрим теперь структурные схемы фильтров, которые соответствуют заданному набору коэффициентов $\{a_m\}_{m=1}^M$, $\{b_k\}_{k=0}^N$, определяющих ЛДФ в соответствии с формулами (4.8) и (4.10). Схемы ЛДФ будем строить из элементов, выполняющих задержку последовательности на один

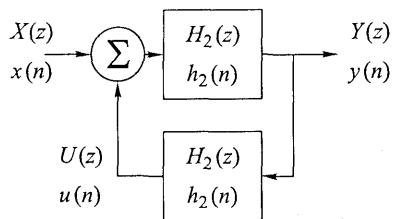
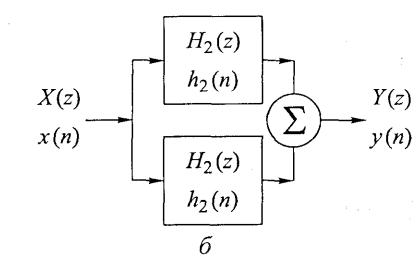
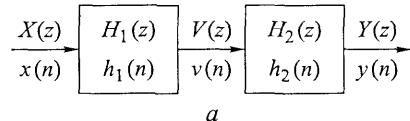


Рис. 4.1

с формулами (4.8) и (4.10). Схемы ЛДФ будем строить из элементов, выполняющих задержку последовательности на один

отсчет, т. е. работающих по правилу $y(n) = x(n-1)$ и имеющих передаточную функцию $H(z) = z^{-1}$, а также из элементов-умножителей: $y(n) = cx(n)$, для которых $H(z) = c$.

Прямая форма структурной схемы ЛДФ реализуется непосредственно по формуле (4.8), когда выходная последовательность формируется в виде линейной комбинации с коэффициентами $\{b_k\}_{k=0}^N$ и $\{a_m\}_{m=1}^M$ из задержанных элементов входной и выходной последовательностей (рис. 4.2).

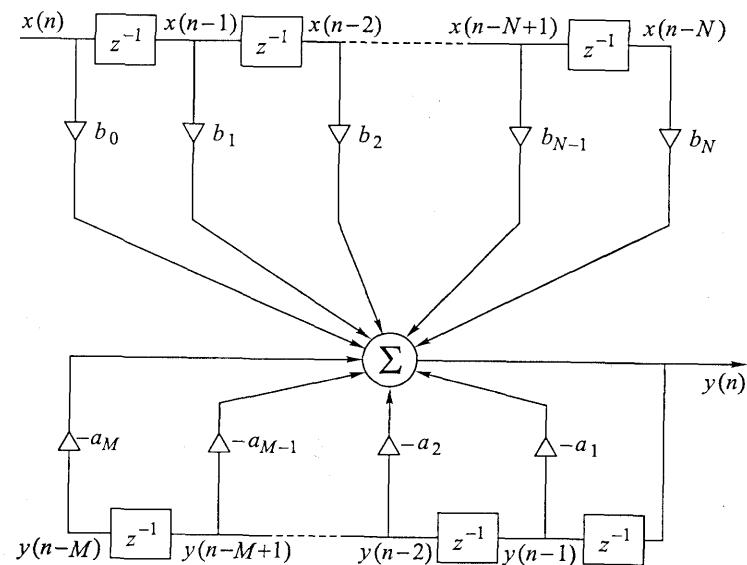


Рис. 4.2

Пример 4.10. Найти передаточную функцию для фильтра, имеющего структурную схему, представленную на рис. 4.3.

$$\blacktriangleleft H(z) = \frac{1 + 3z^{-1} + 2z^{-2}}{1 + 0,4z^{-1}} = \frac{z^2 + 3z + 2}{z^2 + 0,4z}. \triangleright$$

Прямая каноническая форма — форма структурной схемы, содержащей минимальное число элемен-

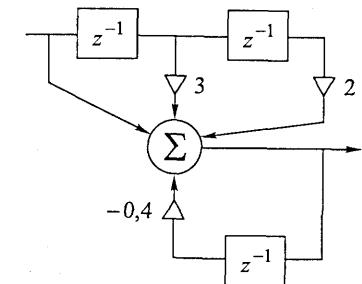


Рис. 4.3

тов задержки. Для ее получения представим передаточную функцию (4.10) как результат последовательного соединения фильтров (рис. 4.4):

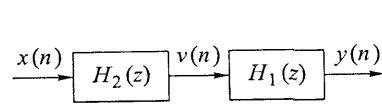


Рис. 4.4

$$H(z) = \frac{\sum_{k=0}^N b_k z^{-k}}{H_1(z)} \cdot \frac{1}{1 + \sum_{m=1}^M a_m z^{-m}} = \frac{1}{H_2(z)}$$

Обозначим $V(z) = H_2(z)X(z)$, $Y(z) = H_1(z)V(z)$ (см. рис. 4.4). Передаточным функциям $H_2(z)$, $H_1(z)$ соответствуют разностные уравнения:

$$v(n) = x(n) - \sum_{m=1}^M a_m v(n-m),$$

$$y(n) = \sum_{k=0}^N b_k v(n-k).$$

В фильтрах, реализующих данные разностные уравнения, имеет место только задержка сигнала $v(n)$, поэтому можно использовать один набор элементов задержки. Этому соответствует структурная схема фильтра (рис. 4.5), где $L = \max(N, M)$.

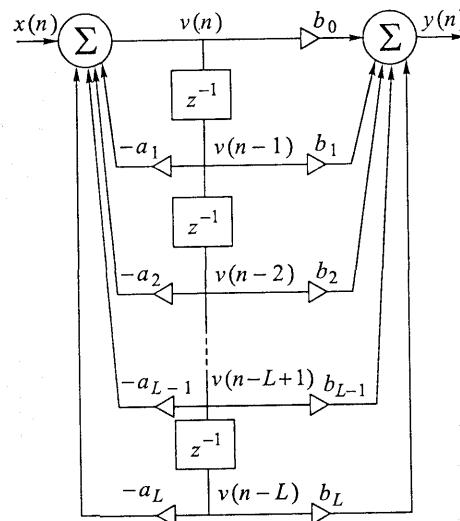


Рис. 4.5

Пример 4.11. Построить прямую каноническую форму структурной схемы для фильтра из примера 4.10.

◀ Проделайте самостоятельно, результат приведен на рис. 4.6. ►

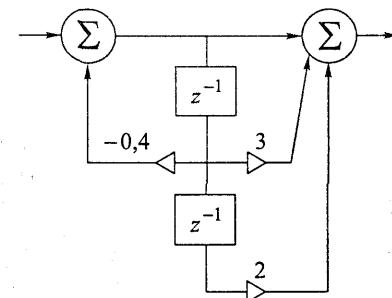


Рис. 4.6

Если в разностном уравнении (4.8) отсутствует зависимость отсчета выходной последовательности $y(n)$ от ее предыдущих от-

$$\text{счетов } y(n-k) (k \geq 1), \text{ т. е. } y(n) = \sum_{k=0}^N b_k x(n-k), \quad H(z) = \sum_{k=0}^N b_k z^{-k},$$

то, очевидно, соответствующий фильтр будет КИХ-фильтром, импульсная характеристика которого $h(n) = \{b_0, b_1, \dots, b_N, 0, 0, \dots\}$. Такие фильтры не имеют в структурной схеме обратной связи и являются нерекурсивными. БИХ-фильтры могут быть только рекурсивными, т. е. фильтрами с обратной связью, и в разностном уравнении (4.8) по крайней мере один коэффициент $a_m \neq 0$. Однако наличие обратной связи еще не говорит о том, что фильтр имеет бесконечную импульсную характеристику.

Пример 4.12. Найти передаточную функцию, импульсную характеристику и упростить структурную схему фильтра, изображенную на рис. 4.7.

◀ Убедитесь самостоятельно в том, что $H(z) = 1$, $h(n) = \delta(n)$. ►

Синтезировать КИХ-фильтр по заданной ИХ

$$h(n) = \{h(0), h(1), \dots, h(N), 0, 0, \dots\}$$

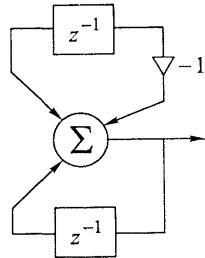


Рис. 4.7

очень просто, для этого надо выбрать коэффициенты $\{b_k = h(k)\}_{k=0}^N$. Синтез рекурсивных фильтров методологически более сложен, однако, при аппроксимации ИХ (в том числе, конечных) использование рекурсивных схем фильтров часто позволяет приближенно реализовать заданную ИХ меньшим числом элементов задержки и умножителей.

При аппаратной реализации в интегральных схемах используется также *транспонированная форма* структурной схемы фильтра. Для получения транспонированной формы представим фильтр в виде схемы, приведенной на рис. 4.8, где $L = \max(M, N)$. Эта структурная схема, очевидно, непосредственно реализует уравнение (4.8) и может быть представлена как преобразование прямой формы, при котором меняются местами операции умножения и задержки, а общий сумматор заменяется цепочкой сумматоров.

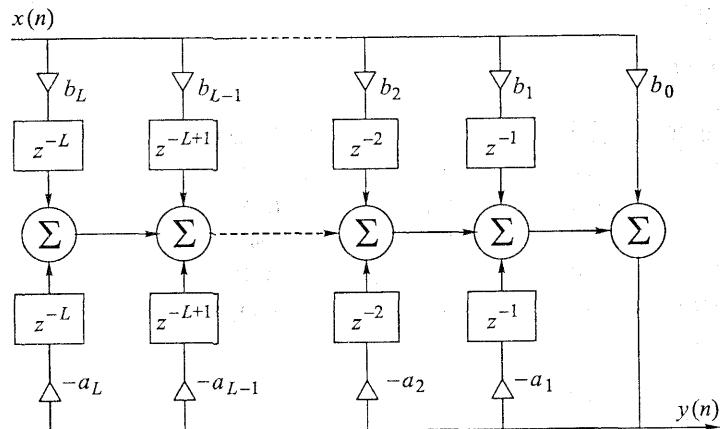


Рис. 4.8

В полученной схеме можно поменять местами операции суммирования и задержки, в итоге получится транспонированная форма фильтра (рис. 4.9).

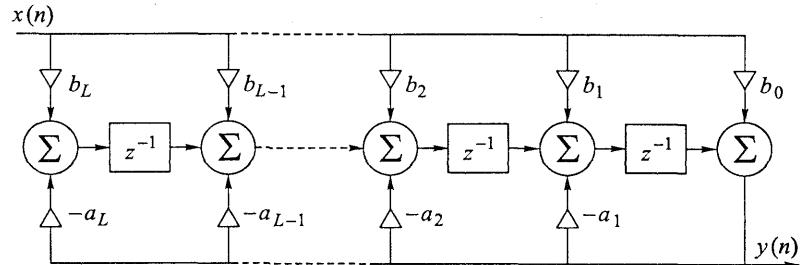


Рис. 4.9. Транспонированная форма ЛДФ

Задание. Постройте транспонированную форму для фильтра из примера 4.10.

■ 4.4. Устойчивость ЛДФ

Определение. ЛДФ называется *устойчивым*, если при любой ограниченной входной последовательности $x(n)$ ($|x(n)| \leq C_1 < \infty$) выходная последовательность также является ограниченной, т. е. $|y(n)| \leq C_2 < \infty$.

В *неустойчивой* системе при некоторых ограниченных входных сигналах выходной сигнал неограниченно возрастает.

Теорема 4.3. Необходимым и достаточным условием устойчивости ЛДФ является ограниченность суммы ряда

$$\sum_{n=0}^{\infty} |h(n)| \leq C < \infty. \quad (4.11)$$

◀ Докажем сначала достаточность условия (4.11) для устойчивости ЛДФ. Если выполняется неравенство (4.11), то для любого входного воздействия, такого что $|x(n)| \leq C_1 < \infty$, получаем отклик ЛДФ $\forall n$

$$|v(n)| = \left| \sum_{k=0}^n x(n-k)h(k) \right| \leq \sum_{k=0}^n |x(n-k)||h(k)| \leq C_1 \sum_{k=0}^n |h(k)| \leq C_1 C < \infty,$$

т. е. выходная последовательность ограничена, фильтр устойчив.

Докажем необходимость условия (4.11) для устойчивости ЛДФ. Допустим противное: фильтр устойчив, но условие (4.11) не выполняется, последовательность частичных сумм $S_j = \sum_{n=0}^j |h(n)|$ неограниченно возрастает с увеличением номера j , т. е. $\forall C > 0 \exists J = J(C): \sum_{n=0}^J |h(n)| > C$. Выбрав такой номер J по произвольному как угодно большому значению $C > 0$, рассмотрим реакцию ЛДФ на следующее (ограниченное!) входное воздействие, которое строится по ИХ данного ЛДФ:

$$x(n) = \begin{cases} 1 & \text{при } h(J-n) \geq 0; \\ -1 & \text{при } h(J-n) < 0. \end{cases}$$

Тогда для выходной последовательности получаем:

$$|y(J)| = \left| \sum_{n=0}^J x(n)h(J-n) \right| = \sum_{n=0}^J |h(J-n)| = \sum_{n=0}^J |h(n)| > C.$$

Таким образом, для как угодно большого значения C можно подобрать ограниченное входное воздействие, при котором существует элемент $y(J)$ выходной последовательности, для которого $|y(J)| > C$, т. е. фильтр не является устойчивым. Следовательно, предположение о том, что для устойчивого ЛДФ условие (4.11) может не выполняться, является ложным. ►

Следствие 1. Необходимое условие устойчивости ЛДФ: $\lim_{n \rightarrow \infty} h(n) = 0$.

Следствие 2. КИХ-фильтры всегда устойчивы.

Пример 4.13. При каких значениях параметра a фильтр из примера 4.4 является устойчивым?

◀ При решении примера 4.4 было установлено, что для данного фильтра ИХ $h(n) = a^n$. Условие (4.11) примет вид: $\sum_{n=0}^{\infty} |a|^n < \infty$.

Отсюда $|a| < 1$. ►

Вместо проверки условия (4.11) часто оказывается удобнее использовать другой критерий устойчивости.

Теорема 4.4. Для того чтобы ЛДФ был устойчив, необходимо и достаточно, чтобы все полюсы его передаточной функции $H(z)$ лежали на комплексной плоскости внутри единичного круга $|z| < 1$.

◀ Необходимость. Так как фильтр устойчив, то для $|z| \geq 1$ передаточная функция ограничена:

$$|H(z)| = \left| \sum_{n=0}^{\infty} h(n)z^{-n} \right| \leq \sum_{n=0}^{\infty} |h(n)| |z^{-n}| \leq \sum_{n=0}^{\infty} |h(n)| < C.$$

Следовательно, в области $|z| \geq 1$ у передаточной функции нет полюсов.

Достаточность. В соответствии с выражением (4.4) ИХ

$$h(n) = \frac{1}{2\pi i} \oint_{\gamma} H(z) z^{n-1} dz,$$

где γ — замкнутый контур, охватывающий все особые точки подынтегральной функции. Как отмечалось ранее, передаточную функцию ЛДФ можно представить в виде отношения степенных многочленов, $H(z) = P(z)/Q(z)$. Особыми точками подынтегральной функции могут быть лишь нули многочлена $Q(z)$, а при $n = 0$, возможно, и точка $z = 0$. Так как число нулей $\{z_k | Q(z_k) = 0\}_{k=1}^L$ конечно, то $\exists r_1 = \max_{k=1, \dots, L} |z_k|$, причем $r_1 < 1$. В качестве контура интегрирования γ , охватывающего все особые точки функции $H(z)z^{n-1}$, возьмем окружность $\gamma = \{z | |z| = r\}$, где $r = \frac{1+r_1}{2} < 1$.

На этой окружности $Q(z) \neq 0$, передаточная функция не имеет полюсов и потому ограничена, т. е. $\exists M < \infty$:

$$\max_{|z|=r} |H(z)| \leq \max_{|z|=r} |P(z)| / \min_{|z|=r} |Q(z)| \leq M.$$

Выполняя замену переменной интегрирования $z = r e^{i\phi}$, для ИХ получаем следующую оценку:

$$|h(n)| = \left| \frac{1}{2\pi i} \oint_{|z|=r} H(z) z^{n-1} dz \right| = \left| \frac{1}{2\pi i} \int_0^{2\pi} H(re^{i\phi}) ir^n e^{i\phi n} d\phi \right| \leqslant \frac{r^n}{2\pi} \int_0^{2\pi} M |e^{i\phi n}| d\phi = Mr^n.$$

Поскольку $r < 1$, получаем для условия (4.11):

$$\sum_{n=0}^{\infty} |h(n)| \leq M \sum_{n=0}^{\infty} r^n = \frac{M}{1-r} < \infty,$$

т. е. ЛДФ устойчив. ►

Пример 4.14. Определить по передаточной функции, при каких значениях параметра a фильтр из примера 4.4 является устойчивым.

◀ Ранее было найдено, что $H(z) = \frac{z}{z-a}$ (см. пример 4.8).

Единственная особая точка передаточной функции — полюс первого порядка $z = a$. Требование теоремы 4.4 приводит к уже установленному в примере 4.13 условию устойчивости фильтра: $|a| < 1$. ►

■ 4.5. Частотная характеристика ЛДФ

Определение. Частотной характеристикой (ЧХ) ЛДФ назовем комплекснозначную функцию вещественной переменной ω : $K(\omega) = H(e^{i\omega})$, которая получается при подстановке в передаточную функцию аргумента $z = e^{i\omega}$.

Иными словами, частотная характеристика определяется выражением $K(\omega) = \sum_{k=0}^{\infty} h(k) e^{-i\omega k}$ — для устойчивых ЛДФ данный ряд сходится абсолютно при любых значениях ω , так как $\sum_{k=0}^{\infty} |h(k) e^{-i\omega k}| = \sum_{k=0}^{\infty} |h(k)| \leq C < \infty$.

Рассмотрим в качестве входной последовательности отсчеты комплексного гармонического сигнала, а именно, $x(n) = e^{i\omega n}$ ($n \geq 0$). Тогда найденный по формуле свертки отклик ЛДФ равен

$$y(n) = \sum_{k=0}^n h(k) \underbrace{e^{i\omega(n-k)}}_{x(n-k)} = \underbrace{e^{i\omega n}}_{x(n)} \sum_{k=0}^n h(k) e^{-i\omega k} \xrightarrow{n \rightarrow \infty} x(n) K(\omega). \quad (4.12)$$

То есть в установившемся режиме (при больших значениях номера n) на выходе ЛДФ получаем гармоническое колебание той же частоты ω , а его амплитуда и фаза определяются ЧХ, которую можно также записать в показательной форме: $K(\omega) = |K(\omega)| e^{-i\phi(\omega)}$. Тогда изменение амплитуды гармонического колебания на выходе ЛДФ в установившемся режиме характеризуется модулем частотной характеристики, или амплитудно-частотной характеристикой (АЧХ) $|K(\omega)|$, а фазы — фазочастотной характеристикой (ФЧХ) $\phi(\omega) = -\arg K(\omega)$. Рассмотрим это на примере.

Пример 4.15. Определить отклик фильтра из примера 4.4 для значения параметра $a = \frac{1}{\sqrt{3}}$ в установившемся режиме при подаче на вход последовательности $x(n) = \{1, 0, -1, 0, 1, 0, -1, 0, \dots\}$.

◀ Для данного фильтра $H(z) = \frac{z}{z-a}$, откуда $K(\omega) = \frac{1}{1 - \frac{e^{-i\omega}}{\sqrt{3}}}$.

Заданное входное воздействие $x(n) = \cos \frac{\pi n}{2} = \operatorname{Re}(e^{i\pi n/2})$. В силу линейности ЛДФ и вещественности его коэффициентов выходной сигнал $y(n)$ будет представлять собой действительную часть отклика фильтра на комплексное гармоническое колебание частоты $\omega = \pi/2$: $\tilde{x}(n) = e^{i\pi n/2}$, $y(n) = \operatorname{Re} \tilde{y}(n)$. В установившемся режиме $\tilde{y}(n) = K(\pi/2) \tilde{x}(n)$. Поэтому для $n \gg 1$:

$$y(n) = \operatorname{Re} \left(\frac{e^{\frac{i\pi n}{2}}}{1 - \frac{e^{-i\pi/2}}{\sqrt{3}}} \right) = \operatorname{Re} \left(\frac{e^{\frac{i\pi n}{2}}}{1 + \frac{i}{\sqrt{3}}} \right) = \operatorname{Re} \left(\frac{e^{\frac{i\pi n}{2}}}{\frac{2}{\sqrt{3}} e^{\frac{i\pi}{6}}} \right) =$$

$$= \operatorname{Re} \left(\underbrace{e^{\frac{i\pi n}{2}}}_{\tilde{x}(n)} \underbrace{\frac{\sqrt{3}}{2} e^{-\frac{i\pi}{6}}}_{K(\pi/2)} \right) = \frac{\sqrt{3}}{2} \cos \left(\frac{\pi n}{2} - \frac{\pi}{6} \right).$$

Заметим, что по АЧХ и ФЧХ полученный ответ можно было записать сразу. Так, изменение амплитуды характеризуется АЧХ и для $\omega = \pi/2$ имеем $|K(\pi/2)| = \sqrt{3}/2$, а задержка фазы выходного сигнала определяется ФЧХ $\varphi(\pi/2) = -\arg K(\pi/2) = \pi/6$. ►

Результат примера 4.15 обобщается очевидным образом. Если на вход ЛДФ подается дискретное гармоническое воздействие вида $x(n) = |A| \cos(\omega n - \gamma) = \operatorname{Re}\{A e^{i\omega n}\}$, где $A = |A| e^{-i\gamma}$ — комплексная амплитуда колебания, то по известным АЧХ $|K(\omega)|$ и ФЧХ $\varphi(\omega)$ отклик ЛДФ в установившемся режиме ($n \gg 1$) находится как

$$y(n) = \operatorname{Re}\{A e^{i\omega n} |K(\omega)| e^{-i\varphi(\omega)}\} = |A| |K(\omega)| \cos(\omega n - \gamma - \varphi(\omega)). \quad (4.13)$$

Использование формулы (4.13) возможно лишь для достаточно больших значений индекса n , когда можно полагать (см. (4.12)), что

$$K_n(\omega) = \sum_{k=0}^n h(k) e^{-i\omega k} \approx K(\omega) = \sum_{k=0}^{\infty} h(k) e^{-i\omega k}.$$

Для КИХ-фильтров с ИХ $h(k) = \{h(0), h(1), \dots, h(N), 0, 0, \dots\}$ точное равенство $K_n(\omega) = K(\omega)$ получаем для значений $n \geq N$. Для устойчивых БИХ-фильтров $h(k) \xrightarrow{k \rightarrow \infty} 0$, поэтому ИХ с некоторым приближением также можно считать конечной, выбрав достаточно большое число N : $K_N(\omega) \approx K(\omega)$. Для неустойчивых фильтров формула (4.13) неприменима.

Таким образом, физический смысл частотной характеристики — комплексный коэффициент передачи гармонических колебаний.

Легко видеть, что ЧХ $K(\omega) = \sum_{k=0}^{\infty} h(k) e^{-i\omega k}$ имеет период 2π .

Учитывая, что $h(n) = 0$ при $n < 0$ и переходя к линейной частоте $v = \omega/2\pi$, частотную характеристику можно записать:

$$K_f(v) = \sum_{k=-\infty}^{\infty} h(k) e^{-2\pi i v k}.$$

Получили, что ЧХ $K_f(v)$ представляет собой частотный спектр дискретного сигнала — импульсной характеристики $h(n)$ для единичного интервала дискретизации $\Delta t = 1$ (см. раздел 3.4 и формулу (3.9)). В частотной области частота сигнала $F_{\max} = \frac{1}{2\Delta t}$ (см. (3.7)) соответствует нормированной частоте $v = 1/2$ ($\omega = \pi$), а период ЧХ (т.е. спектра ИХ), в соответствии со свойством периодичности спектра дискретного сигнала (см. раздел 3.4) в частотной области будет единичным для $K_f(v)$ и равным 2π для $K(\omega)$.

Фильтру нижних частот, например, соответствует, приведенный на рис. 4.10 вид графика нормированной по оси абсцисс АЧХ $|K(\omega)|$.

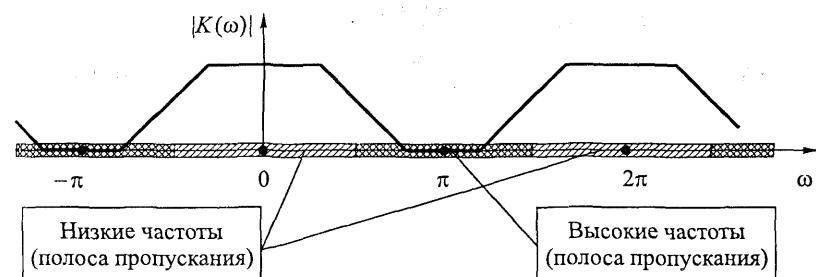


Рис. 4.10. Вид АЧХ дискретного фильтра нижних частот

Пример 4.16. Для фильтра, заданного передаточной функцией $H(z) = \frac{c}{2} + cz^{-1} + \frac{c}{2}z^{-2}$, $c > 0$, найти АЧХ и ФЧХ, определить

отклик $y(n)$ в установившемся режиме на входное воздействие $x(n) = \sin\left(\frac{\pi}{3}n - \frac{\pi}{6}\right)$ и изобразить структурную схему фильтра.

◀ ЧХ фильтра имеет вид

$$K(\omega) = e^{-i\omega} \left(\frac{c}{2} e^{i\omega} + c + \frac{c}{2} e^{-i\omega} \right) = e^{-i\omega} c(1 + \cos \omega).$$

Для $c > 0$ график АЧХ $|K(\omega)| = c(1 + \cos \omega) = 2c \cos^2 \frac{\omega}{2}$ приведен на рис. 4.11. Данный ЛДФ является фильтром нижних частот. График ФЧХ $\varphi(\omega) = -\arg K(\omega) = \arg e^{i\omega}$ постройте самостоятельно, помня, что по определению $\varphi(v) \in (-\pi; \pi]$ (см. раздел 2.1).

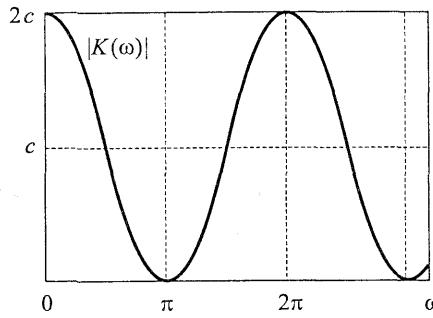


Рис. 4.11

В установившемся режиме отклик фильтра на гармоническое входное воздействие $x(n) = \text{Im } e^{i\left(\frac{\pi}{3}n - \frac{\pi}{6}\right)}$ с циклической частотой $\omega = \pi/3$ находим как

$$\begin{aligned} y(n) &= \text{Im} \left(K\left(\frac{\pi}{3}\right) e^{i\left(\frac{\pi}{3}n - \frac{\pi}{6}\right)} \right) = \text{Im} \left(\frac{3}{2} c e^{i\left(-\frac{\pi}{3} + \frac{\pi}{3}n - \frac{\pi}{6}\right)} \right) = \\ &= \frac{3}{2} c \sin\left(\frac{\pi n}{3} - \frac{\pi}{2}\right) = -\frac{3c}{2} \cos\frac{\pi n}{3}. \end{aligned}$$

Структурную схему фильтра изобразите самостоятельно. ►

■ 4.6. Синтез КИХ-фильтров по частотной характеристики

На практике часто возникает необходимость поиска таких параметров фильтра $\{a_n\}$, $\{b_k\}$, которые с заданной точностью обеспечивают требуемую ЧХ. Соответствующая задача *синтеза фильтра*, естественно, должна решаться так, чтобы число элементов задержки и умножителей в структуре фильтра было минимальным. Обычно это требует использования рекурсивных схем, однако методика синтеза таких фильтров более сложна.

Рассмотрим синтез фильтра по заданной АЧХ $|K_0(\omega)|$ на примере нерекурсивного *симметричного* фильтра, параметры которого $\{b_n = h(n)\}_{n=0}^{2N}$ таковы, что $b_{N-k} = b_{N+k}$, $k = 1, \dots, N$. Целое число $2N$ назовем *порядком* фильтра¹. Для ЧХ симметричного КИХ-фильтра получаем

$$\begin{aligned} K(\omega) &= \sum_{k=0}^{2N} b_k e^{-i\omega k} = e^{-i\omega N} \sum_{k=-N}^N b_{N+k} e^{-i\omega k} = \\ &= e^{-i\omega N} \left(b_N + \sum_{k=1}^N (b_{N+k} e^{-i\omega k} + b_{N-k} e^{i\omega k}) \right), \end{aligned}$$

тогда окончательно:

$$K(\omega) = e^{-i\omega N} \left(b_N + 2 \sum_{k=1}^N b_{N+k} \cos \omega k \right). \quad (4.14)$$

Коэффициенты фильтра $\{b_{N+k}\}_{k=0}^N$ должны быть такими, чтобы ЧХ синтезированного фильтра $K(\omega)$ хорошо аппроксимировала требуемую ЧХ $K_0(\omega)$. Обычно требуется обеспечить наилучшее равномерное приближение АЧХ: $\max_{\omega \in [0; \pi]} \|K(\omega) - K_0(\omega)\| \rightarrow \min$.

Поскольку для фильтров с вещественными коэффициентами $|K(\omega)| = |K(-\omega)|$, то с учетом 2π -периодичности ЧХ для определения АЧХ достаточно задать последнюю на отрезке $\omega \in [0; \pi]$.

¹ Для нерекурсивных фильтров порядок фильтра M обычно называют порядком полинома Лорана, соответствующего передаточной функции $H(z) = \sum_{k=0}^M b_k z^{-k}$.

Для нахождения коэффициентов фильтра $\{b_{N+k}\}_{k=0}^N$ приравняем значения требуемой АЧХ $|K_0(\omega)|$ значениям АЧХ $|K(\omega)|$ синтезируемого фильтра в некоторых точках $\omega_j \in [0; \pi]$, $j = 0, \dots, N$.

Искомые коэффициенты $\{b_{N+k}\}_{k=0}^N$ могут быть найдены (см. (4.14)) из решения следующей системы линейных уравнений:

$$b_N + 2 \sum_{k=1}^N b_{N+k} \cos \omega_j k = |K_0(\omega_j)|, \quad j = 0, \dots, N. \quad (4.15)$$

Тогда полученное из формулы (4.14) выражение

$$|K(\omega)| = \left| b_N + 2 \sum_{k=1}^N b_{N+k} \cos \omega k \right| \quad (4.16)$$

можно рассматривать как интерполяционную формулу для требуемой АЧХ $|K_0(\omega)|$, а точки ω_j задают узлы интерполяции. Вопрос оптимального расположения узлов представляет собой отдельную задачу. Мы рассмотрим простейший случай, когда выбирается равномерное расположение точек $\omega_j \in [0; \pi]$.

Пример 4.17. По идеальной АЧХ

$$|K_0(\omega)| = \begin{cases} 1, & \omega \in [\pi/2; 3\pi/2]; \\ 0, & \omega \in [0; \pi/2) \cup [3\pi/2; 2\pi] \end{cases}$$

фильтра верхних частот синтезировать нерекурсивный симметричный фильтр порядков $2N = 10$, $2N = 22$ и построить соответствующие АЧХ.

◀ Выберем $\omega_j = \frac{\pi(j+0.5)}{N+1}$, $j = 0, \dots, N$. Получаем из уравнений (4.15) систему

$$b_N + 2 \sum_{k=1}^N b_{N+k} \cos \frac{\pi k(j+0.5)}{N+1} = \left| K_0 \left(\frac{\pi(j+0.5)}{N+1} \right) \right|, \quad j = 0, \dots, N.$$

Можно показать (на основании формул (6.10) и (6.12), которые будут рассмотрены позднее), что решение полученной системы имеет вид:

$$b_{N+k} = \frac{1}{N+1} \sum_{j=0}^N \left| K_0 \left(\frac{\pi(j+0.5)}{N+1} \right) \right| \cos \frac{\pi k(j+0.5)}{N+1}, \quad k = 0, \dots, N. \quad (4.17)$$

В нашем случае коэффициенты фильтра

$$b_{N+k} = \frac{1}{N+1} \sum_{j=\frac{N+1}{2}}^N \cos \frac{\pi k(j+0.5)}{N+1},$$

по ним находим АЧХ (4.16). На рис. 4.12 приведены графики АЧХ фильтров для случаев $N = 5$ (рис. 4.12, *a*) и $N = 11$ (рис. 4.12, *b*). ►

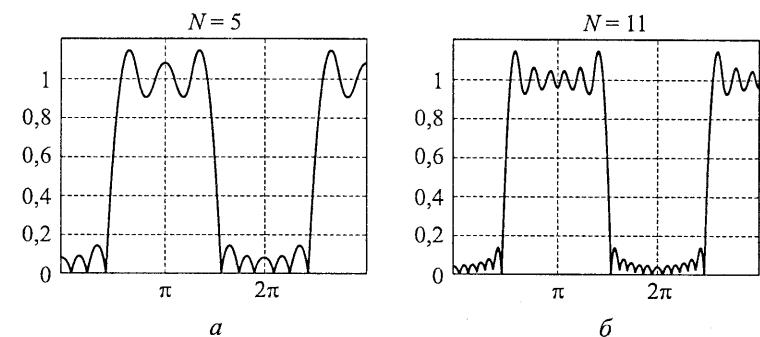


Рис. 4.12

Результаты примера 4.17 показывают, что увеличение порядка фильтра не улучшило аппроксимацию идеальной АЧХ $|K_0(\omega)|$ при помощи реальной $|K(\omega)|$: максимальное отклонение реальной АЧХ, наблюдаемое вблизи точки разрыва идеальной АЧХ $\omega = \pi/2$, осталось на том же уровне. Аппроксимация разрывной функции конечной тригонометрической суммой вида (4.14) всегда будет вызывать «всплески» вблизи точек разрыва, подобно наблюдаемым в примере 4.17. Это явление хорошо известно и носит название *эффекта Гиббса*.

Для более точного приближения тригонометрическими рядами аппроксимируемая функция должна быть как можно более гладкой, т. е. быть непрерывной и по возможности иметь непрерывные производные до как можно более высокого порядка. На практике при формировании требований к АЧХ задают ограничения на ее неравномерность в полосе пропускания и в полосе

подавления (ε_0 и ε_1 соответственно), а также определяют допустимый «переходный» участок частот, который разделяет эти полосы (рис. 4.13).

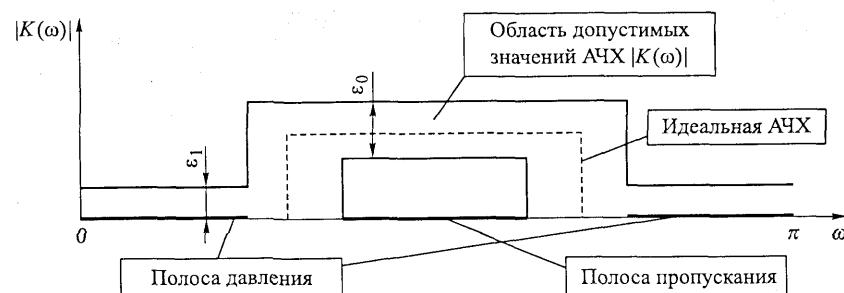


Рис. 4.13

Требуемая АЧХ синтезируемого фильтра задается уже в виде некоторого допустимого «коридора», в котором должен лежать график $|K(\omega)|$. Тогда можно подобрать более гладкую АЧХ $|K_0(\omega)|$ для ее последующей аппроксимации функцией $|K(\omega)|$ (4.14). Например, для узкополосного фильтра с полосой пропускания $\omega \in [\omega_0 - \delta; \omega_0 + \delta]$, где значение δ мало, удобно выбирать $|K_0(\omega)| = \exp(-c(\omega - \omega_0)^2)$.

Однако конечная цель синтеза фильтра состоит не в поиске приемлемой функции $|K_0(\omega)|$, а в том, чтобы найти такие коэффициенты $\{b_k\}$ фильтра минимально возможного порядка, чтобы реальная АЧХ $|K(\omega)|$ не выходила за пределы заданного «коридора». Один из классических подходов к решению данной задачи состоит в следующем¹. Задавшись некоторым порядком фильтра, равномерно расположим в частотной области узлы $\omega_j \in [0; \pi]$, $j = 0, \dots, N$. Для тех значений ω_j , которые попали в полосу подавления, приравняем значения частотной характеристики нулю; для узлов, попавших в полосу пропускания, приравняем АЧХ константе — требуемому коэффициенту передачи. Значения АЧХ

для узлов, попавших в «переходную» область между полосами подавления и пропускания, считаем свободными переменными и подбираем их так, чтобы в области пропускания и в области подавления реальная АЧХ была наиболее близка к требуемой. Если точность полученной аппроксимации оказалась избыточной, рассматриваем данную задачу для фильтра меньшего порядка, а если обеспечить требуемую аппроксимацию не удалось, то повышаем порядок синтезируемого фильтра. Выбор значений АЧХ в «переходной» области сводится к некоторой оптимизационной задаче, решаемой численными методами.

Не останавливаясь подробно на методах решения соответствующей задачи оптимизации, проиллюстрируем возможность формирования требуемой АЧХ на несложном примере.

Пример 4.18. В условиях примера 4.17 рассмотреть возможность приближения идеальной АЧХ при отсутствии ограничений на значения $|K_0(\omega)|$ в двух ближайших к частоте $\omega = \pi/2$ точках интерполяции.

◀ Пусть N — нечетное число. Воспользуемся формулой (4.17), положив $\left|K_0\left(\frac{\pi(j+0,5)}{N+1}\right)\right| = 0$ (полоса подавления) для $j = 0, \dots, \frac{N-3}{2}$ и $\left|K_0\left(\frac{\pi(j+0,5)}{N+1}\right)\right| = 1$ (полоса пропускания) для $j = \frac{N+3}{2}, \dots, N$.

Если выбрать значения АЧХ в точках $\omega_{\frac{N-1}{2}} = \frac{\pi}{2} \left(\frac{N}{N+1} \right)$ и $\omega_{\frac{N+1}{2}} = \frac{\pi}{2} \left(\frac{N+2}{N+1} \right)$ равными соответственно $\left|K_0\left(\omega_{\frac{N-1}{2}}\right)\right| = 0,16$ и $\left|K_0\left(\omega_{\frac{N+1}{2}}\right)\right| = 0,84$, то для значений $N = 5$ и $N = 11$ получим АЧХ $|K(\omega)|$, графики которых приведены на рис. 4.14 (сравните с АЧХ в примере 4.17). ►

¹ Подробнее о синтезе фильтров см.: Рабинер Л., Голд Б. Теория и применение цифровой обработки сигналов. М., 1978.

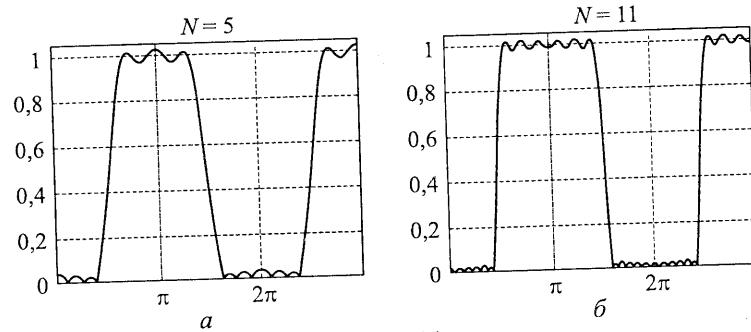


Рис. 4.14

Находя параметры нерекурсивного фильтра из решения системы (4.15), в области полосы пропускания фильтра (где $K(\omega) \neq 0$) в соответствии с (4.14) мы получаем: $K(\omega) = \exp(-i\omega N)|K(\omega)|$. Отсюда ФЧХ $\varphi(\omega) = \omega N + 2\pi m$, где целое число m выбирается из условия $\varphi(\omega) \in (-\pi; \pi]$. ФЧХ вида $\varphi(\omega) = \alpha\omega + 2\pi m$ ($\alpha \geq 0$, $m \in \mathbb{Z}$) называется *линейной*. Такая ФЧХ гарантирует, что гармонические колебания $x(n) = A \cos(\omega n - \gamma)$ любой частоты ω при прохождении через фильтр имеют одинаковую временную задержку α , и на выходе системы получается сигнал

$$y(n) = A|K(\omega)| \cos(\omega n - \gamma - \alpha\omega - 2\pi m) = |K(\omega)|x(n - \alpha).$$

Различия наблюдаются только в амплитуде колебаний, изменение которой определяется АЧХ. Свойство линейности ФЧХ очень часто является необходимым требованием, предъявляемым к синтезируемому фильтру.

■ 4.7. Нахождение отклика фильтра с использованием БПФ

Данный метод основан на свойстве 5° ДПФ (см. раздел 3.5). Пусть M — длина импульсной характеристики фильтра, т. е. $h_{M-1} \neq 0$ и $h_n = 0$ при $n \geq M$ (для БИХ-фильтра данное условие выполняется приближенно), а обрабатываемый сигнал $x(n)$ предполагается конечной последовательностью длины L : $h_{L-1} \neq 0$ и $x(n) = 0$ при $n \geq L$. Тогда отклик фильтра (см. (4.9))

$$y(n) = \sum_{m=0}^{M-1} h_m x(n-m) = \sum_{m=0}^{L-1} h_{n-m} x(m) \quad (4.18)$$

будет иметь длительность $M+L-1$, так как для $n \geq M+L-1$ приведенная формула свертки дает реакцию фильтра $y(n) = 0$, а для $n = M+L-2$ в общем случае $y(M+L-2) = h_{M-1}x(L-1) \neq 0$.

Добавлением нулевых отсчетов сформируем векторы одинаковой размерности $2N \geq \max(2L, 2M)$ (обычно $N = 2^n$):

$$\mathbf{H} = \underbrace{(h_0, h_1, \dots, h_{M-1}, 0, 0, \dots, 0)}_{2N \text{ элементов}}, \quad \mathbf{X} = \underbrace{(x(0), \dots, x(L-1), 0, 0, \dots, 0)}_{2N \text{ элементов}}$$

Затем над указанными векторами проведем обработку, аналогичную рассмотренной в примере 3.10: БПФ $\mathbf{H} \rightarrow \hat{\mathbf{H}}$ и $\mathbf{X} \rightarrow \hat{\mathbf{X}}$, покомпонентное перемножение полученных дискретных спектров $\hat{\mathbf{Y}} = \sqrt{2N} (\hat{h}_0 \hat{x}_0, \dots, \hat{h}_{2N-1} \hat{x}_{2N-1})$, обратное БПФ $\hat{\mathbf{Y}} \rightarrow \mathbf{Y}$. В полученном векторе размерности $2N$ первые $M+L-1$ компонент представляют собой отклик фильтра (4.18), а остальные компоненты будут равны нулю:

$$\mathbf{Y} = \left(\underbrace{y(0), y(1), \dots, y(L-1)}_L, \underbrace{y(L), \dots, y(L+M-2)}_{M-1}, 0, 0, \dots, 0 \right)$$

Использование описанной процедуры в вычислительном плане может быть более эффективным, чем непосредственная реализация формулы свертки. По сравнению с оценками из примера 3.8 выигрыш может быть более существенным. Действительно, фильтр можно задать не вектором ИХ (h_0, \dots, h_{M-1}) , а сразу от-

счетами ЧХ $K_f(v) = \sum_{k=0}^{M-1} h_k e^{-2\pi i v k}$, взятыми в точках $v_j = \frac{j}{2N}$,
 $j = 0, \dots, 2N-1$:

$$K_f(v_j) = \sum_{k=0}^{M-1} h_k e^{-\frac{2\pi i}{2N} j k} = \sum_{k=0}^{2N-1} h_k e^{-\frac{2\pi i}{2N} j k}$$

(здесь учтено, что $h_n = 0$ при $n \geq M$). Правая часть последнего равенства с точностью до нормировочного множителя представ-

ляет собой ДПФ импульсной характеристики

$$\mathbf{H} = (\underline{h_0, h_1, \dots, h_{M-1}, 0, 0, \dots, 0})_{\text{2N элементов}}$$

$$\hat{h}_j = \frac{1}{\sqrt{2N}} \sum_{k=0}^{2N-1} h_k e^{-\frac{2\pi i}{2N} jk}, \quad j = 0, 1, \dots, 2N-1.$$

Поэтому реально вычисление БПФ $\mathbf{H} \rightarrow \hat{\mathbf{H}}$ чаще всего не требуется, так как вектор $\hat{\mathbf{H}}$ задан (или вычислен) заранее и представляет собой отсчеты ЧХ фильтра,

$$\hat{\mathbf{H}} = (\hat{h}_0, \dots, \hat{h}_{2N-1}) = \frac{1}{\sqrt{2N}} (K_f(v_0), \dots, K_f(v_{2N-1})).$$

Кроме того, при фильтрации дискретных сигналов полосовыми фильтрами некоторые операции покомпонентного перемножения ($\hat{h}_0 \hat{x}_0, \dots, \hat{h}_{2N-1} \hat{x}_{2N-1}$) часто представляют собой умножения на числа $\hat{h}_j = 0$ (в полосе подавления) или $\hat{h}_j = 1$ (в полосе пропускания), что реально не требует выполнения арифметических действий.

Вычислительная эффективность использования БПФ тем выше, чем больше размерность используемого преобразования. Однако чаще всего обработать весь дискретный сигнал целиком невозможно из-за слишком большой длины последовательности $x(n)$. При обработке в реальном масштабе времени большая размерность используемого БПФ может вызвать также недопустимую задержку, связанную с необходимостью накопления исходных данных. Если выполнить БПФ сразу для всего сигнала не представляется возможным, последовательность $x(n)$ обрабатывается побочно. Выбирая из входного сигнала по L отсчетов и формируя блоки $\mathbf{X}_j = (x(jL), \dots, x((j+1)L-1), 0, 0, \dots, 0)$ размерности $2N$, $j = 0, 1, \dots$, далее последовательно обрабатывают каждый из них по описанной выше схеме. Входной сигнал можно представить в виде $x(n) = \sum_j \mathbf{X}_j$, и в силу линейности фильтра его отклик также представляет собой сумму

$$y(n) = \sum_j \mathbf{Y}_j, \text{ где}$$

$$\mathbf{Y}_j = \left(\underbrace{y_j(jL), \dots, y_j((j+1)L-1)}_L, \underbrace{y_j((j+1)L), \dots, y_j((j+1)L+M-2)}_{M-1}, 0, \dots, 0 \right)$$

— реакция фильтра на входное воздействие \mathbf{X}_j .

Таким образом, хотя входная последовательность разделена на непересекающиеся блоки \mathbf{X}_j , выходные блоки \mathbf{Y}_j перекрываются (рис. 4.15). Задав исходно отклик фильтра нулевой последовательностью $y(n) = 0$, далее необходимо прибавлять компоненты очередного полученного блока \mathbf{Y}_j ($j = 0, 1, \dots$) к тем элементам выходной последовательности $y(n)$, которые были сформированы ранее.

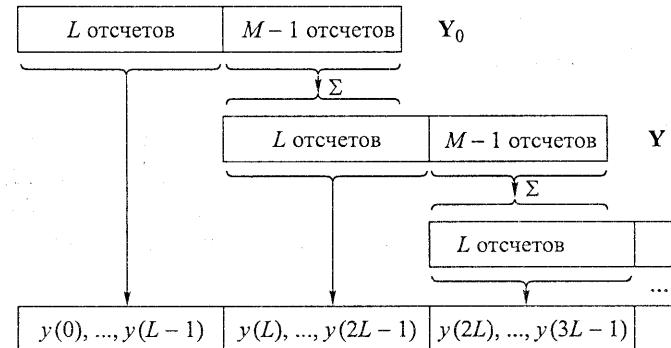


Рис. 4.15

■ 4.8. Согласованный дискретный фильтр

Пусть $u(n) = \{u(0), u(1), \dots, u(N-1)\}$ — заданный эталонный сигнал конечной длительности, $X_{\text{ш}}(n)$ — отсчеты сигнала, представляющего собой белый шум с дисперсией σ_X^2 и нулевым математическим ожиданием $m_X = 0$ (см. раздел 2.4). На вход фильтра подается случайный сигнал

$$X(n) = u(n) + X_{\text{ш}}(n).$$

Необходимо синтезировать такой линейный дискретный фильтр, чтобы получить на его выходе отклик $Y(n) = v(n) + Y_{\text{ш}}(n)$, для которого в момент дискретного времени $n = N - 1$ обеспечивается максимальное отношение мощностей полезного сигнала и шума:

$$k_{\text{ш}} = v^2(N-1) / M(Y_{\text{ш}}^2(N-1)), \quad (4.19)$$

где полезный и шумовой сигналы определяются по формуле свертки:

$$v(N-1) = \sum_{n=0}^{N-1} u(N-1-n)h(n), \quad (4.20)$$

$$Y_{\text{ш}}(N-1) = \sum_{n=0}^{N-1} X_{\text{ш}}(N-1-n)h(n). \quad (4.21)$$

Соответствующий оптимальный фильтр называется *согласованным* с сигналом $u(n) = \{u(0), u(1), \dots, u(N-1)\}$.

Найдем ИХ согласованного фильтра. Так как полезный входной сигнал $u(n)$ имеет длительность N , то, как видно из формул (4.20) и (4.21), при построении согласованного фильтра достаточно ограничиться рассмотрением КИХ-фильтров с ИХ $h(n) = \{h(0), h(1), \dots, h(N-1)\}$.

Для шумовой составляющей выходного сигнала (4.21), в силу некоррелированности и нулевого математического ожидания отсчетов белого шума (см. раздел 2.4), имеем:

$$\begin{aligned} M(Y_{\text{ш}}^2(N-1)) &= \sum_{n=0}^{N-1} \sum_{j=0}^{N-1} M(X_{\text{ш}}(N-1-n)X_{\text{ш}}(N-1-j))h(n)h(j) = \\ &= \sum_{n=0}^{N-1} M(X_{\text{ш}}^2(N-1-n))h^2(n) = \sigma_X^2 \sum_{n=0}^{N-1} h^2(n). \end{aligned} \quad (4.22)$$

Подставим полученное выражение (4.22) и соотношение (4.20) в формулу (4.19):

$$k_{\text{ш}} = \frac{\left(\sum_{n=0}^{N-1} u(N-1-n)h(n) \right)^2}{\sigma_X^2 \sum_{n=0}^{N-1} h^2(n)}. \quad (4.23)$$

Необходимо найти такую оптимальную ИХ

$$h(n) = \{h(0), h(1), \dots, h(N-1)\},$$

которая обращает в максимум выражение (4.23). Легко видеть, что для различных значений множителя α семейство ИХ $\{\alpha h(n)\}$ дает одно и то же значение $k_{\text{ш}}$ (4.23). Поэтому, не теряя общности рассуждений, можно потребовать от искомой оптимальной ИХ

выполнения условия $\sum_{n=0}^{N-1} h^2(n) = 1$. Тогда поиск безусловного максимума функции (4.23) сводится к задаче об условном максимуме выражения, стоящего в числителе дроби (4.23), т. е. значения $v^2(N-1)$ (см. (4.20)). При этом, поскольку оптимальность ИХ сохраняется при смене знака (умножении ИХ на величину $\alpha = -1$), можно потребовать, чтобы максимум значения $v^2(N-1)$ достигался для положительного знака величины $v(N-1)$ (4.20).

Окончательно получаем следующую задачу условной минимизации:

$$F(h(0), h(1), \dots, h(N-1)) = v(N-1) = \sum_{n=0}^{N-1} u(N-1-n)h(n) \rightarrow \max,$$

$$\text{при условии } \sum_{n=0}^{N-1} h^2(n) = 1.$$

Решая данную задачу методом неопределенных множителей Лагранжа, находим, что оптимальная ИХ, обращающая в максимум отношение сигнал/шум (4.19) или (4.23), имеет вид:

$$h(j) = \alpha u(N-1-j), \quad j = 0, 1, \dots, N-1, \quad (4.24)$$

т.е.

$$\alpha = \left(\sum_{k=0}^{N-1} u^2(k) \right)^{-1/2}. \quad (4.25)$$

Полученная ИХ (4.24) называется *зеркальной* к сигналу $u(n) = \{u(0), u(1), \dots, u(N-1)\}$. В соответствии со сделанными выше замечаниями параметр $\alpha \neq 0$ в формуле (4.24) можно выбрать отличным от (4.25). Несложно видеть, что при $\alpha > 0$ в равенстве (4.20) $v(N-1) > 0$, а для $\alpha < 0$ имеем $v(N-1) < 0$.

Задание. Определите самостоятельно, как связаны дискретный спектр (3.9) эталонного сигнала и частотная характеристика согласованного фильтра (последнюю запишите для линейной частоты v).

Если выбрать для согласованного фильтра с ИХ (4.24) значение масштабирующего множителя α в соответствии с (4.25), то, как это следует из (4.24) и (4.22), дисперсия (мощность) шума на выходе согласованного фильтра будет такой же, как и на входе: $\sigma_Y^2 = M(Y^2) = M(X_{\text{ш}}^2) = \sigma_X^2$, а мощность полезного сигнала в момент дискретного времени $n = N - 1$ (см. (4.20), (4.24), (4.25))

$$v^2(N-1) = 1/\alpha^2 = \sum_{k=0}^{N-1} u^2(k). \quad (4.26)$$

Данную величину, представляющую собой энергию эталонного сигнала, обозначим E_u . Заметим, что уровень полезного сигнала $v(n)$ на выходе фильтра в любой момент наблюдения $n = J$ не превосходит значения $v(N-1) = 1/\alpha = \sqrt{E_u}$. Действительно, используя (4.20), (4.24), (4.25), получим:

$$\begin{aligned} v(J) &= \sum_{k=0}^{N-1} h(k)u(J-k) = \sum_{k=-\infty}^{\infty} h(k)u(J-k) = \\ &= \alpha \sum_{k=-\infty}^{\infty} u(N-1-k)u(J-k) = \alpha \sum_{k=-\infty}^{\infty} u(N-1-J+k)u(k) = \\ &= \frac{\alpha}{2} \left(- \sum_{k=-\infty}^{\infty} (u(N-1-J+k) - u(k))^2 + \underbrace{\sum_{k=-\infty}^{\infty} u^2(N-1-J+k)}_{1/\alpha^2} + \underbrace{\sum_{k=-\infty}^{\infty} u^2(k)}_{1/\alpha^2} \right) \\ &= \frac{1}{\alpha} - \frac{\alpha}{2} \sum_{k=-\infty}^{\infty} (u(N-1-J+k) - u(k))^2 \leq \frac{1}{\alpha} = \sqrt{E_u}. \end{aligned} \quad (4.27)$$

Согласованные фильтры применяются для обнаружения заданных сигналов на фоне помех в различных практических приложениях: радиолокации, распознавании образов и др.

Рассмотрим практическую задачу, для решения которой используется согласованный фильтр. Эта задача состоит в том, чтобы определить, подан ли на вход фильтра эталонный сигнал $u(n)$, зашумленный помехами (белым шумом): $X(n) = u(n) + X_{\text{ш}}(n)$, или же входная последовательность не содержит эталона: $X(n) = X_{\text{ш}}(n)$. Так как $M(Y_{\text{ш}}(n)) = 0$, то математическое ожидание для мощности выходного сигнала $Y(n) = v(n) + Y_{\text{ш}}(n)$ при выполнении условий (4.24), (4.25)

$$M(Y^2(n)) = M(v^2(n) + 2v(n)Y_{\text{ш}}(n) + Y_{\text{ш}}^2(n)) = v^2(n) + \sigma_X^2. \quad (4.28)$$

Максимальное значение данной величины (см. (4.27)) достигается в момент дискретного времени $n = N - 1$, причем

$$M(Y^2(N-1)) = E_u + \sigma_X^2,$$

где E_u находится из (4.26). При отсутствии эталонного сигнала во входной последовательности, когда $X(n) = X_{\text{ш}}(n)$, получим:

$$M(Y^2(n)) = M(Y_{\text{ш}}^2(n)) = \sigma_X^2.$$

При использовании согласованного фильтра для определения факта наличия эталонного сигнала во входной последовательности $X(n)$ выбирается некоторый пороговый уровень T выходного сигнала $Y(n)$. Если $Y(N-1) > T$, то принимается решение о наличии эталона $u(n)$ во входном сигнале; в противном случае считаем, что эталон отсутствует. На выбор порога T оказывают влияние два противоречящих друг другу фактора: с одной стороны, высокий порог снижает вероятность ложного «срабатывания» детектора эталонного сигнала вследствие действия шума, с другой стороны, высокий порог увеличивает вероятность пропуска (исобнаружения) эталона при его появлении. Очевидно, что для снижения вероятностей различного рода ошибок энергия эталонного сигнала E_u (4.26) должна быть по возможности большей и превосходить мощность шума σ_X^2 . Обычно выбирают порог $T \in (\sigma_X; \sqrt{E_u})$.

Другая возникающая на практике задача состоит в том, чтобы в зашумленной входной последовательности

$$X(n) = \beta u(n - M) + X_{\text{ш}}(n) \quad (4.29)$$

с неизвестными уровнем $\beta > 0$ и моментом M начала эталонного сигнала определить «местоположение» эталонного сигнала во входной последовательности, т. е. значение числа M . В силу линейности фильтра и инвариантности его параметров во времени математическое ожидание мощности отклика согласованного фильтра на воздействие (4.29) находим по аналогии с (4.28):

$$\mathbb{M}(Y^2(n)) = \beta^2 v^2(n - M) + \sigma_X^2.$$

Максимальное значение данная величина принимает в момент n дискретного времени, который находится из условия $n - M = N - 1$:

$$\mathbb{M}(Y^2(M + N - 1)) = \beta^2 E_u + \sigma_X^2,$$

где E_u представляет собой энергию эталона (4.26). Таким образом, для определения числа M в выходной последовательности необходимо найти момент J дискретного времени с максимальным значением отклика: $Y(J) = \max_{j \geq N-1} Y(j)$, и положить $M = J - N + 1$. Как

уже отмечалось, для повышения достоверности определения значения M желательно, чтобы энергия $\beta^2 E_u$ полезного сигнала, поступающего на вход фильтра, была как можно большей.

Обе описанные задачи могут ставиться в комплексе, когда необходимо определить, содержится ли во входной последовательности эталонный сигнал с неизвестными временем начала появления M и уровнем β (и если да, то каково значение M), или же на вход фильтра поступает только белый шум. Тогда в выходной последовательности сначала выбирается отсчет $Y(J) = \max_{j \geq N-1} Y(j)$,

и если для некоторого порога T выполняется условие $Y(J) > T$, то принимается решение о том, что эталонный сигнал присутствует, а его первый отсчет приходится на момент дискретного времени $M = J - N + 1$.

Глава 5

Основы прикладной теории информации

■ 5.1. Дискретный источник сообщений без памяти, количество информации. Энтропия

Рассмотрим модель *дискретного источника сообщений* X , который изменяет свое состояние в заданные моменты времени, причем число возможных состояний конечно и равно N . Будем считать, что случайные переходы источника X из одного состояния в другое не зависят от текущего и предыдущих состояний (т. е. у источника сообщений «нет памяти»), а вероятности перехода в состояния x_j ($j = 1, \dots, N$) $p_j = P\{X = x_j\}$ заданы, $\sum_{j=1}^N p_j = 1$. Таким образом, если источник X находится в состоянии x_j , то с вероятностью p_j в очередной момент перехода состояние может остаться тем же, но при этом мы все равно будем говорить о переходе источника из состояния x_j в x_j .

Состояние	x_1	...	x_N
Вероятность	p_1	...	p_N

Состояние x_j можно понимать как реализацию случайной величины дискретного типа, или как случайный выбор символа из некоторого N -символьного алфавита $\{x_1, \dots, x_N\}$. Последовательность символов, создаваемых источником, представляет собой некоторое сообщение. Зададимся вопросом: какую информацию несет в себе событие, состоящее в переходе источника X в новое состояние? Как измерить количество этой информации, приходящееся на один символ сообщения, т. е. как ввести меру для количества информации, создаваемой источником X при переходе в очередное состояние? Начнем рассуждения с частного случая дискретного источника сообщений, когда все состояния являются равновероятными, $\forall k = 1, \dots, N$: $p_k = 1/N$.

Дискретный источник с равновероятными состояниями

Естественно считать, что информация, которую несет в себе один символ сообщения, т. е. один переход источника в новое состояние, тем больше, чем больше число состояний N . Например, если связать с символом сообщения результат вытягивания одной карты из колоды, состоящей из 32 карт, то при $N = 2$ один символ может нести сообщение о цвете масти, а при $N = 32$ — полную информацию о вытянутой карте. Таким образом, мера количества информации $I_X = I_X(N)$ должна представлять собой монотонно возрастающую функцию количества состояний N . Причем если у источника всего одно возможное состояние, то такой источник не производит информацию, поэтому необходимо выполнение условия $I_X(1) = 0$.

Вводя меру количества информации как некоторую монотонную функцию, необходимо учесть также следующее. Если рассмотреть источник сообщений Z , представляющий собой объединение (ансамбль) двух независимых источников X (с N равновероятными состояниями) и Y (с M равновероятными состояниями), то число равновероятных символов, описывающих состояния источника Z , будет равным MN . При этом количество информации, создаваемой источником Z , равно количеству информации, создаваемой источниками X и Y . Поэтому мера количества информации, рассматриваемая как функция числа равновероятных состояний, должна отвечать условию: $I_Z(NM) = I_X(N) + I_Y(M)$.

Указанным требованиям отвечает логарифмическая функция. В 1928 г. Хартли предложил следующую меру для количества информации, производимой источником сообщений без памяти X , который имеет N равновероятных состояний:

$$I_X = I_X(N) = \log N = -\log p, \quad (5.1)$$

где $p = 1/N$ — вероятность перехода источника в каждое из состояний. Основание логарифма при этом не имеет принципиального значения и определяет лишь масштаб, т. е. то количество информации, которое связано с одной единицей измерения. Например, если основание логарифма двоичное, то единицы измерения называют *битами*. Один бит — это такое количество информации, которое при очередном переходе создает источник сообщений, имеющий два равновероятных состояния.

Пример 5.1. Из колоды из 32 карт вытянута одна. Сколько вопросов, требующих ответа «да» или «нет», надо задать, чтобы определить, какая карта вытянута?

◀ Фактически, необходимо ответить, какое количество информации связано с источником, который имеет 32 равновероятных состояния, в единицах измерения источника с двумя («да», «нет») состояниями. Поэтому число вопросов $I = \log_2 32 = 5$. Первый вопрос можно задать, например, такой: «Масть красная?» Затем, в зависимости от результатов ответа на первый вопрос: «Пики?», или: «Черви?». Затем: «Карта — картинка (валет, дама, король, туз)?» Оставшиеся вопросы очевидны. ►

Мера Хартли не получила широкого распространения, поскольку применима лишь к весьма узкому классу источников сообщений, для которых состояния равновероятны.

Дискретный источник с неравновероятными состояниями

Рассмотрим два источника, X и Y , имеющих по два возможных состояния, т. е. алфавит каждого из которых состоит из двух символов. Распределение вероятностей приведено в табл.

x_1	x_2	y_1	y_2
0, 5	0, 5	0, 001	0, 999

Обобщая меру Хартли, будем исходить из следующего: мера информации должна отражать априорную неопределенность состояния источника, поэтому в приведенном примере источник X имеет максимальную неопределенность, поскольку его состояния равновероятны. У источника Y состояние практически предопределено, и, не имея никакой дополнительной информации, мы почти наверняка не ошибемся, сказав, что источник находится в состоянии y_2 . Таким образом, для вводимой меры количества информации в данном случае мы должны получить $I(X) > I(Y)$. Обобщая меру Хартли, естественно потребовать от вводимой меры количества информации и для источника с $N > 2$ числом состояний аналогичного свойства: максимальное значение (при занятом N) количества информации принимает в случае равной вероятности состояний $p = 1/N$, т. е. $I_{\max} = \log N$. Если же какое-то состояние источника X принимает с единичной вероятностью (т. е.

у источника одно состояние), то количество информации должно быть равным нулю: $I_{\min} = 0$. Для ансамбля Z из независимых источников сообщений X и Y как и прежде необходимо потребовать, чтобы вводимая мера количества информации отвечала условию: $I(Z) = I(X)+I(Y)$.

В качестве меры количества информации (априорной неопределенности) для дискретного источника сообщений без памяти Шеноном в конце 40-х годов прошлого века предложена следующая величина:

$$H(X) = -\sum_{k=1}^N p_k \log p_k, \quad (5.2)$$

где $p_k = P\{X = x_k\}$ — вероятность перехода источника в состояние x_k . Величина (5.2) носит название *энтропии*. Основание логарифма здесь по-прежнему не имеет принципиального значения и так же, как и для меры Хартли, определяет количество информации, приходящейся на одну единицу измерения. В цифровой обработке сигналов традиционно выбирается двоичное основание, т. е. количество информации измеряется в битах, тогда величину (5.2) называют *двоичной энтропией*.

Пример 5.2. Найти двоичную энтропию источника X , создающего сообщения из четырех символов, вероятности появления которых заданы.

x_k	a	\bar{b}	\bar{c}	\bar{d}
p_k	0,5	0,25	0,125	0,125

$$\blacktriangleleft H(X) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{8} \log_2 \frac{1}{8} - \frac{1}{8} \log_2 \frac{1}{8} = \frac{14}{8} = 1,75 \text{ бита.} \blacktriangleright$$

Трактовать величину $H(X)$ из (5.2) можно как *среднее количество информации*, приходящееся на один символ источника сообщения. Действительно, символы сообщения несут разную информацию: если вероятность появления символа велика, он несет меньшую информацию, а если вероятность мала, то большую. Простейший пример, приведенный выше — источник сообщений Y , имеющий два состояния с вероятностями $p_1 = 0,001$ и $p_2 = 0,999$. Символ y_2 почти не несет информации, его появление

практически предопределено. А вот появление символа y_1 — явление редкое. Пусть, например, символ y_2 означает, что температура воздуха (в градусах Цельсия) положительна, а y_1 означает отрицательную температуру. Источник сообщений — висящий за окном термометр. Появление какого символа, на ваш взгляд, содержит больше информации, если наблюдение ведется жарким летом?...

Считая, что информация, которую несет символ, имеющий вероятность появления p_k , определяется по формуле Хартли (5.1) как $I_k = -\log p_k$, мы и находим среднюю информативность символа в

$$\text{виде математического ожидания } H(X) = \bar{I} = \sum_{k=1}^N p_k I_k, \text{ что эквивалентно выражению (5.2).}$$

Можно трактовать эту величину как характеристику априорной неопределенности дискретного источника информации или, что то же самое, как среднее количество информации, связанное с одним символом создаваемого источником дискретного сообщения.

Отметим, что если какая-то вероятность $p_k = 0$, то соответствующее слагаемое в сумме (5.2) нужно понимать в смысле предельного значения: $\lim_{p \rightarrow 0} p \log p = 0$. Легко видеть также, что в случае равновероятных состояний мера (5.2) приобретает вид (5.1).

Укажем на важнейшие свойства энтропии (5.2).

1°. Энтропия является неотрицательной величиной. Энтропия минимальна и равна нулю тогда, когда $\exists p_k = 1$.

◀ С учетом сделанного замечания о нулевых вероятностях, каждое слагаемое $-p_k \log p_k$ в сумме (5.2) неотрицательно, поэтому неотрицательной является и вся сумма. Таким образом, нулевое значение в (5.2) получаем при равенстве нулю всех слагаемых. Так как $\sum_{k=1}^N p_k = 1$, то это возможно тогда и только тогда, когда $\exists k: p_k = 1$. ►

2°. Для независимых источников сообщений X и Y энтропия их ансамбля

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

◀ Ансамбль источников описывается вероятностями состояний: $p_{n,m} = P\{X = x_n, Y = y_m\}$, $n = 1, \dots, N$; $m = 1, \dots, M$.

x_1	...	x_N
$p_{X,1}$...	$p_{X,N}$

y_1	...	y_M
$p_{Y,1}$...	$p_{Y,M}$

Энтропия ансамбля равна

$$H(X, Y) = - \sum_{n=1}^N \sum_{m=1}^M p_{n,m} \log p_{n,m}.$$

Для независимых источников $p_{n,m} = p(x_n)p(y_m)$, поэтому энтропия

$$\begin{aligned} H(X, Y) &= - \sum_{n=1}^N \sum_{m=1}^M p(x_n)p(y_m)(\log p(x_n) + \log p(y_m)) = \\ &= - \underbrace{\sum_{n=1}^N p(x_n) \log p(x_n)}_{H(X)} \underbrace{\sum_{m=1}^M p(y_m)}_{1} + \underbrace{\sum_{n=1}^N p(x_n)}_{1} \underbrace{\left(- \sum_{m=1}^M p(y_m) \log p(y_m) \right)}_{H(Y)} = \\ &= H(X) + H(Y). \end{aligned}$$

3°. При фиксированном значении N энтропия принимает максимальное значение $H(X) = \log N$ в случае, когда $p_k = 1/N$, $k = 1, \dots, N$.

◀ Найдем максимум функции $H(p_1, \dots, p_N) = - \sum_{k=1}^N p_k \log p_k$

при условиях $\sum_{k=1}^N p_k = 1$, $0 \leq p_k \leq 1$. Для этого выразим

$p_N = 1 - \sum_{k=1}^{N-1} p_k$, и рассмотрим задачу безусловной минимизации функции

$$H(p_1, \dots, p_{N-1}) = - \sum_{k=1}^{N-1} p_k \log p_k - \left(1 - \sum_{k=1}^{N-1} p_k\right) \log \left(1 - \sum_{k=1}^{N-1} p_k\right).$$

Необходимым условием безусловного экстремума является равенство нулю частных производных. Полагая для упрощения записи, что основание логарифма является натуральным, имеем $\forall j = 1, \dots, N-1$:

$$\frac{\partial H(p_1, \dots, p_{N-1})}{\partial p_j} = -\log p_j + \log \left(1 - \sum_{k=1}^{N-1} p_k\right) = 0.$$

Отсюда $\forall j = 1, \dots, N-1$: $p_j = 1 - \sum_{k=1}^{N-1} p_k = p_N$, т. е. все вероятности

одинаковы, $p_j = 1/N$. Достаточным условием максимума функции является отрицательный знак ее второго дифференциала в найденной критической точке, что соответствует отрицательной определенности матрицы вторых производных (матрицы Гессе)

$\left\{ \frac{\partial^2 H(p_1, \dots, p_{N-1})}{\partial p_i \partial p_j} \right\}_{i,j=1}^{N-1}$. В соответствии с критерием Сильвестра

матрица размерности $(N-1) \times (N-1)$ отрицательно определена тогда и только тогда, когда угловые миноры $|\Delta_j|$ (определители подматриц, составленных из первых j строк и j столбцов исходной матрицы) являются знакочередующимися, а именно: $\text{sign} |\Delta_j| = (-1)^j$, $j = 1, \dots, N-1$. Найдем вторые производные функции $H(p_1, \dots, p_{N-1})$ в критической точке $p_1 = \dots = p_{N-1} = 1/N$. Несложно убедиться, что

$$\left. \frac{\partial^2 H(p_1, \dots, p_{N-1})}{\partial p_j \partial p_k} \right|_{p_1 = \dots = p_{N-1} = 1/N} = \begin{cases} -N, & j \neq k; \\ -2N, & j = k. \end{cases}$$

Подсчитаем j -й угловой минор матрицы Гессе; $\forall j = 1, \dots, N-1$:

$$|\Delta_j| = \det \begin{pmatrix} -2N & -N & \dots & -N \\ -N & -2N & \dots & -N \\ \dots & \dots & \dots & \dots \\ -N & -N & \dots & -2N \end{pmatrix} = (-N)^j \begin{vmatrix} 2 & 1 & \dots & 1 \\ 1 & 2 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 2 \end{vmatrix} =$$

Последнюю строку вычтем из всех остальных

$$=(-N)^j \begin{vmatrix} 1 & 0 & \dots & -1 \\ 0 & 1 & \dots & -1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 2 \end{vmatrix} = (-N)^j \begin{vmatrix} 1 & 0 & \dots & -1 \\ 0 & 1 & \dots & -1 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1+j \end{vmatrix} = (-N)^j(j+1),$$

Из последней строки вычтем
сумму всех предыдущих

поскольку определитель треугольной матрицы равен произведению диагональных элементов. Видим, что условие $\operatorname{sign}|\Delta_j| = (-1)^j$ выполняется, поэтому максимум энтропии (5.2) достигается при равенстве вероятностей состояния дискретного источника: $H_{\max}(p_1, \dots, p_N) = H(1/N, \dots, 1/N) = \log N$. ►

4°. Энтропия характеризует неопределенность выбора источником состояния, но не саму содержательную сущность этого выбора.

Поясним последнее свойство. Например, если лекарство A в 90 % случаев вызывает улучшение состояния больного, а в 10 % случаев не оказывает эффекта, то неопределенность (энтропия) его действия такая же, как и у «лекарства» B , которое в 90 % случаев не вызывает изменения состояния больного, а в 10 % случаев даже влечет его ухудшение:

$$H(A) = H(B) = -0,1 \log_2 0,1 - 0,9 \log_2 0,9 \approx 0,47 \text{ бита.}$$

При этом сущность результатов действия лекарств совершенно разная.

■ 5.2. Основные теоремы о кодировании источника без памяти

Прежде чем рассматривать вопросы построения двоичных кодов, отметим, что термин «бит» часто используется для обозначения двух близких, но, вообще говоря, незэквивалентных понятий. В разделе 5.1 мы определили бит как единицу измерения количества информации. Другое значение термина «бит» — двоичный символ. Обычно, если нет явного указания на то, в каком смысле понимается термин «бит», имеется в виду второе значение. Двоичный символ несет количество информации равное одни.

ному биту тогда и только тогда, когда символ создан источником сообщений, имеющим два равновероятных состояния.

Рассмотрим источник X , имеющий N состояний x_k с заданными вероятностями p_k , $k = 1, \dots, N$. Каждому состоянию источника можно непосредственно (фактически, записав в двоичном представлении номер состояния k) поставить в соответствие двоичный код из R битов, где целое число R находится из условия: $N \leq 2^R$, т. е. для представления алфавита $\{x_k\}_{k=1}^N$ исходного сообщения, порожденного источником X , можно использовать другой «алфавит кодирования», в данном случае $\{0, 1\}$, что представляет для нас особый интерес как основной способ представления информации в цифровых электронных устройствах. Поэтому и для энтропии (5.2) также будем использовать далее двоичные единицы измерения (биты), выбрав основание логарифма равное двум.

Если двоичная энтропия источника $H(X)$ меньше величины R , находимой из условия $N \leq 2^R$, то можно предложить более эффективные способы двоичного кодирования состояний источника X .

Пример 5.3. Поставим в соответствие символам источника сообщений из примера 5.2 битовые последовательности («слова») следующим образом: $a \leftrightarrow 0$, $b \leftrightarrow 10$, $c \leftrightarrow 110$, $d \leftrightarrow 111$. Покажем, что любое сообщение источника, закодированное в двоичном виде, является однозначно декодируемым, а среднее число битов, приходящееся на один символ, совпадает со значением двоичной энтропии источника.

◀ Действительно, так как более короткие битовые слова не являются началом более длинных, то любая последовательность битов однозначно декодируется, например: $\underline{\hspace{1cm}} \underline{\hspace{1cm}} \underline{\hspace{1cm}} \underline{\hspace{1cm}} \underline{\hspace{1cm}}$ или $\underline{\hspace{1cm}} \underline{\hspace{1cm}} \underline{\hspace{1cm}} \underline{\hspace{1cm}} \underline{\hspace{1cm}}$

(обозначая R_k число битов в коде символа x_k , находим среднее количество битов, приходящееся на кодирование одного символа:

$$R = \sum_{k=1}^4 p_k R_k = 0,5 \cdot 1 + 0,25 \cdot 2 + 0,125 \cdot 3 + 0,125 \cdot 3 = 1,75 \text{ бита.}$$

Что совпадает со значением двоичной энтропии $H(X)$. Как видим, по сравнению с непосредственным кодированием, при котором для каждого символа требуется код длиной 2 бита, получаем бо-

лее эффективный способ представления сообщений источника X в двоичном виде. ►

Преобразование способа представления данных, приводящее к повышению эффективности кодирования, часто называют *сжатием* или *компрессией* данных.

Среднее число битов, необходимое для кодирования одного символа сообщения, будем называть также *битовыми затратами*. Коды переменной длины, у которых короткие битовые слова не являются началом более длинных (декодируемые коды), называются *префиксными*. Префиксному двоичному коду можно поставить во взаимно однозначное соответствие граф, имеющий вид дерева. Например, префиксному коду из примера 5.3 соответствует граф, приведенный на рис. 5.1.

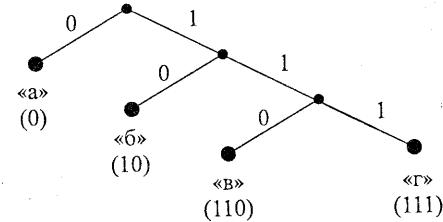


Рис. 5.1. Кодовое дерево, которое определяет префиксный код, приведенный в примере 5.3

Двоичные коды символов формируются последовательным приписыванием битов, соответствующих ребрам графа, которые нужно пройти от вершины дерева до окончного узла («листа»), соответствующего определенному символу заданного алфавита.

В примере 5.3 битовые затраты совпали со значением энтропии источника. Оказывается, величина двоичной энтропии является минимально возможным значением для битовых затрат, какие бы способы кодирования не использовались. Прежде чем доказывать соответствующую теорему, сформулируем ряд вспомогательных утверждений.

Лемма 5.1. Если целые числа $I_j \geq 0$, $j = 1, \dots, n$ таковы, что выполняется неравенство

$$\sum_{j=1}^{n-1} I_j 2^{-j} \leq 1, \quad (5.3)$$

то верно также и неравенство

$$\sum_{j=1}^{n-1} I'_j 2^{-j} \leq 1,$$

где для $j = 1, \dots, n-2$ $I'_j = I_j$, а

$$I'_{n-1} = \begin{cases} I_{n-1} + \frac{I_n}{2}, & \text{если } I_n \text{ — четное;} \\ I_{n-1} + \frac{I_n + 1}{2}, & \text{если } I_n \text{ — нечетное.} \end{cases}$$

◀ Для случая четного числа I_n имеем:

$$\sum_{j=1}^{n-1} I'_j 2^{-j} = \sum_{j=1}^{n-2} I_j 2^{-j} + \left(I_{n-1} + \frac{I_n}{2} \right) 2^{-(n-1)} = \sum_{j=1}^n I_j 2^{-j} \leq 1.$$

Пусть теперь I_n — нечетное число. Умножим неравенство (5.3) на 2^n :

$$\underbrace{\sum_{j=1}^{n-1} I_j 2^{n-j}}_{\text{Четное число}} + \underbrace{\frac{I_n}{2}}_{\substack{\text{Нечетное} \\ \text{число}}} \leq \underbrace{2^n}_{\substack{\text{Четное} \\ \text{число}}}.$$

В полученном неравенстве все слагаемые, стоящие под знаком суммы, четные, поэтому четна и вся сумма. Тогда все выражение, стоящее в левой части неравенства, есть нечетное число, а прибавление единицы в левую часть неравенства не изменит его истинности:

$$\sum_{j=1}^{n-1} I_j 2^{n-j} + I_n + 1 \leq 2^n.$$

Поделив это неравенство на 2^n , получим утверждение леммы для случая нечетного числа I_n . ►

Лемма 5.2 (неравенство Крафта). Для существования однозначно декодируемого двоичного кода $\{(x_k, R_k)\}_{k=1}^N$, содержащего N кодовых слов с длинами R_1, \dots, R_N , необходимо и достаточно, чтобы выполнялось неравенство Крафта:

$$\sum_{k=1}^N 2^{-R_k} \leq 1. \quad (5.4)$$

◀ Необходимость. Обозначим $n = \max_{k=1,\dots,N} R_k$ (максимальная

длина кода одного символа). Обобщая результаты примера 5.3, схему формирования префиксного кода символов в общем случае можно представить в виде графа, пример которого для случая $n = 4$ и $N = 11$ приведен на рис. 5.2 (пунктиром обозначены ребра, которые потенциально можно было бы использовать при построении кода без изменения величины n). Для построения кода каждому ребру, выходящему из узла, необходимо присвоить бит кода (например, ноль для левого ребра и единицу для правого).

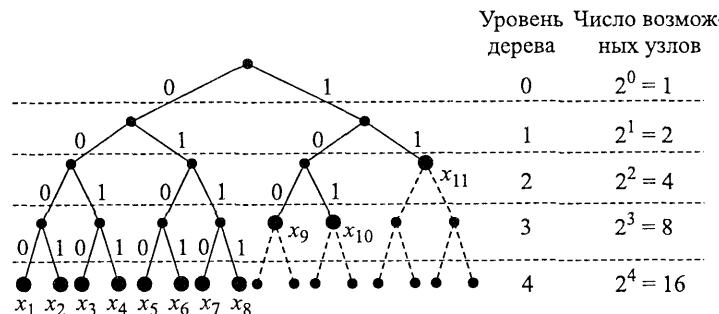


Рис. 5.2. Пример кодового графа для $n = 4$ и $N = 11$

Из рис. 5.2 видно, что каждый узел i -го уровня ($i = 0, \dots, n$) может породить ровно 2^{n-i} узлов последнего уровня n . Листья дерева (окончательные узлы) соответствуют символам кодируемого источника $\{x_k\}_{k=1}^N$, причем уровень соответствующего узла в дереве равен длине битового кода R_k . Подсчитывая общее число возможных узлов последнего уровня, которые могут быть порождены узлами-листьями, получаем

$$\sum_{k=1}^N 2^{n-R_k} = 2^n.$$

Заметим, что требование префиксности кода не будет нарушено, если в графе, задающем правило построения кода, допус-

тить наличие узлов, имеющих одного потомка. Так, свойство декодируемости кода сохранится, если приведенный на рис. 5.2 график преобразовать, например, к виду, изображенному на рис. 5.3. Очевидно, что средняя длина кода одного символа при этом увеличится, а просуммировав общее количество возможных потомков последнего уровня, порождаемых узлами-листьями, мы получим величину меньшую, чем 2^n . Поэтому в общем случае для префиксного кода верно соотношение

$$\sum_{k=1}^N 2^{n-R_k} \leq 2^n,$$

из которого делением на 2^n получаем неравенство (5.4).

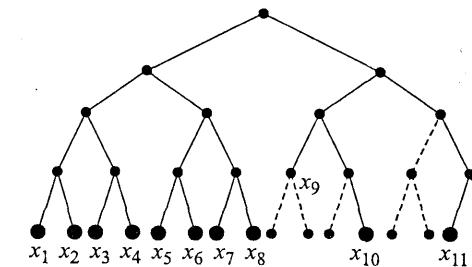


Рис. 5.3

Приведенные рассуждения можно провести аналогичным образом для любого числа n . Взяв произвольный префиксный код и построив порождающий его график, мы сможем убедиться, что неравенство Крафта (5.4) выполняется.

Достаточность. Нужно показать, что если для некоторого набора символов $X = \{x_k\}_{k=1}^N$ заданы длины их кодов $\{R_k\}_{k=1}^N$, удовлетворяющие условию (5.4), то соответствующие декодируемые коды существуют. Для доказательства данного утверждения опишем процедуру построения таких кодов.

Вновь обозначим $n = \max_{k=1,\dots,N} R_k$. Пусть, для определенности, $R_N \leq R_{N-1} \leq \dots \leq R_1 = n$. Дерево T , определяющее двоичный код, будем строить по «каркасу», приведенному на рис. 5.4.

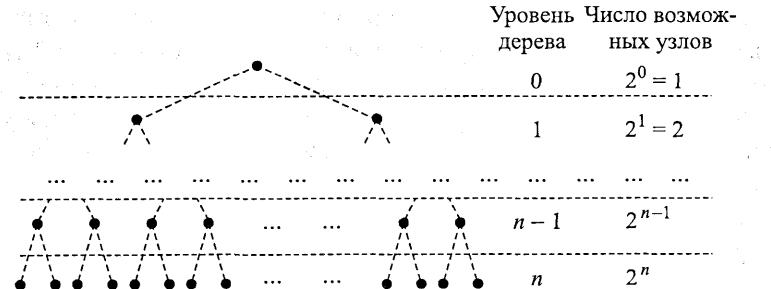


Рис. 5.4. Каркас (шаблон) графа для построения кодового дерева

Обозначим $I_j = |\{x_k \in X | R_k = j\}|$ — число элементов алфавита X , которым должны быть приписаны битовые коды длины j ($j = 1, \dots, n$). Тогда на кодовом дереве T , которое необходимо построить, на j -м уровне должно быть I_j узлов-листьев, соответствующих некоторым символам алфавита X , а неравенство Крафта (5.4) можно переписать в виде (5.3).

Из неравенства Крафта в форме (5.3) следует, что число узлов-листьев на последнем уровне n удовлетворяет неравенству $I_n \leq 2^n$, поэтому для каждого листа последнего уровня найдутся позиции в каркасе (рис. 5.4). Заполним их символами $\{x_k \in X | R_k = n\}$ без пропусков, начиная с крайней левой позиции.

Начнем построение кодового дерева T снизу, объединяя в пары узлы-листья последнего уровня n и ставя в соответствие каждой такой паре узел графа на уровне $n-1$ (рис. 5.5). Если число I_n — четное, то листья уровня n добавят в граф $\tilde{I}_{n-1} = I_n / 2$ узлов на уровень $n-1$, а если I_n — нечетное, то $\tilde{I}_{n-1} = (I_n + 1) / 2$ узлов.

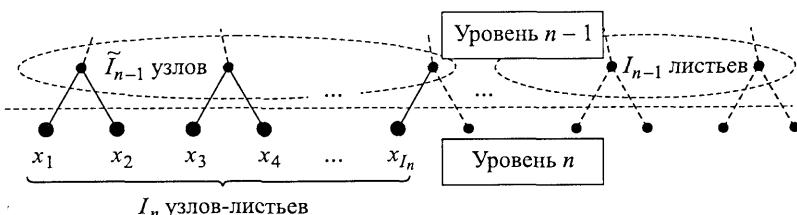


Рис. 5.5. Процедура, сводящая процесс построения n -уровневого кодового дерева к построению $(n-1)$ -уровневого дерева (число I_n выбрано нечетным)

Построение кодового дерева продолжаем далее, рассматривая уже в качестве начального новое дерево T' из $n-1$ уровней, для которого заданы величины $I'_j = I_j$ ($j = 1, \dots, n-2$) и $I'_{n-1} = I_{n-1} + \tilde{I}_{n-1}$. Для дерева T' (см. лемму 5.1) вновь выполняется неравенство Крафта:

$$\sum_{j=1}^{n-1} I'_j 2^{-j} \leq 1,$$

поэтому количество узлов в дереве T (или T') на уровне $n-1$ не превысит максимально допустимого, $I'_{n-1} \leq 2^{n-1}$, и построение кодового дерева можно продолжить. Описанный процесс построения дерева продолжается рекуррентно до тех пор, пока не будет достигнута вершина кодового дерева, которое в итоге будет представлять собой некоторое подмножество каркаса, изображенного на рис. 5.4. При этом лемма 5.1 гарантирует, что число формируемых узлов на каждом слое не будет превышать максимально возможного количества. ►

Докажем теперь, что средние битовые затраты на кодирование одного символа сообщения источника без памяти не могут быть меньше его двоичной энтропии.

Теорема 5.1. Для любого дискретного источника сообщений без памяти X с заданным распределением вероятностей $\{(x_k, p_k)\}_{k=1}^N$ и любого декодируемого битового кода $\{(x_k, R_k)\}_{k=1}^N$ выполняется неравенство

$$R = \sum_{k=1}^N p_k R_k \geq -\sum_{k=1}^N p_k \log_2 p_k = H(X).$$

◀ Учитывая, что $\ln x \leq x-1$, $\sum_{k=1}^N p_k = 1$, $\sum_{k=1}^N 2^{-R_k} \leq 1$ (см. (5.4)),

имеем:

$$-\sum_{k=1}^N p_k \ln p_k - \sum_{k=1}^N p_k R_k \ln 2 = \sum_{k=1}^N p_k \ln \frac{2^{-R_k}}{p_k} \leq \sum_{k=1}^N p_k \left(\frac{2^{-R_k}}{p_k} - 1 \right) \leq 0.$$

Поделив полученное неравенство на величину $\ln 2$, получим утверждение теоремы. ►

При доказательстве леммы 5.2 мы, фактически, описали процедуру построения префиксного кода по заданным длинам кодов символов $\{R_k\}_{k=1}^N$, удовлетворяющим неравенству Крафта (5.4). Следующая теорема определяет правило выбора величин $\{R_k\}_{k=1}^N$ и верхнюю оценку для средних битовых затрат соответствующего кода.

Теорема 5.2. Для любого дискретного источника без памяти существует такой декодируемый код, что средние битовые затраты

$$R = \sum_{k=1}^N p_k R_k < -\sum_{k=1}^N p_k \log_2 p_k + 1 = H(X) + 1.$$

◀ Запишем неравенство Крафта (5.4) в виде

$$\sum_{k=1}^N 2^{-R_k} \leq \sum_{k=1}^N p_k = 1.$$

Тогда по схеме, описанной при доказательстве леммы 5.2, мы сможем построить префиксный код, выбрав наименьшее целое число R_k из условия $2^{-R_k} \leq p_k$, эквивалентного требованию $R_k \geq -\log_2 p_k$. Тогда

$$R_k - 1 < -\log_2 p_k, \quad k = 1, 2, \dots, N.$$

Умножив неравенства на p_k и просуммировав для $k = 1, \dots, N$, получаем

$$\sum_{k=1}^N p_k R_k - \underbrace{\sum_{k=1}^N p_k}_{1} < -\sum_{k=1}^N p_k \log_2 p_k,$$

откуда следует утверждение теоремы. ►

Следствие из теорем 5.1 и 5.2. Выбирая для заданного источника $\{(x_k, p_k)\}_{k=1}^N$ величины R_k как наименьшие целые числа, от-

вечающие условию $R_k \geq -\log_2 p_k$, можно построить префиксный код со средними битовыми затратами $R = \sum_{k=1}^N p_k R_k$, которые удовлетворяют неравенствам

$$H(X) \leq R < H(X) + 1, \quad (5.5)$$

где $H(X)$ — двоичная энтропия (2) источника X .

Доказанные теоремы позволяют сформулировать следующее основополагающее утверждение теории информации.

Теорема 5.3 (Шэннона). При бесконечной длительности сообщение источника без памяти X с двоичной энтропией $H(X)$ можно закодировать так, что средние битовые затраты R , приходящиеся на один символ сообщения, $\forall \varepsilon > 0$ удовлетворяют неравенствам $H(X) \leq R < H(X) + \varepsilon$.

◀ Рассмотрим блок из M символов x_{k_1}, \dots, x_{k_M} , созданных дискретным источником сообщений без памяти с двоичной энтропией $H(X)$. Так как источник не имеет памяти, блок x_{k_1}, \dots, x_{k_M} может рассматриваться как реализация состояния ансамбля из M независимых источников, каждый из которых имеет энтропию $H(X)$. Поэтому на основании свойства 2° энтропии (5.2) получим:

$$H(X_1, \dots, X_M) = \sum_{m=1}^M H(X_m) = M \cdot H(X).$$

Блок x_{k_1}, \dots, x_{k_M} в соответствии с теоремами 5.1 и 5.2 можно закодировать со средними битовыми затратами R_M , приходящими на M символов:

$$M \cdot H(X) \leq R_M < M \cdot H(X) + 1.$$

Отсюда

$$H(X) \leq \frac{R_M}{M} < H(X) + \frac{1}{M}$$

и $\forall \varepsilon > 0$ можно выбрать такой размер блока M , что средние битовые затраты на кодирование одного символа $R = R_M/M$ будут удовлетворять условию $H(X) \leq R < H(X) + \varepsilon$. ►

Таким образом, двоичная энтропия источника (5.2) является точной оценкой снизу для минимально возможных битовых затрат R , приходящихся в среднем на кодирование одного символа. Невозможно построить такой декодируемый двоичный код, для которого $R < H(X)$. Коды, обеспечивающие минимально возможные битовые затраты, называются *эффективными*. Эффективное кодирование сообщений, основанное на учете статистики источника, называют *статистическим* (или, в зарубежной литературе, *энтропийным*) кодированием.

■ 5.3. Эффективное кодирование дискретного источника без памяти по методам Шэннона — Фано и Хаффмана

Теорема Шэннона определяет теоретически возможный предел для эффективности кодирования. В примере 5.3 для заданного источника был приведен код, обладающий максимальной эффективностью, однако в общем случае достижение максимальной эффективности возможно лишь теоретически, при кодировании сообщений из бесконечного числа символов.

После двоичного кодирования получаемую последовательность битов можно рассматривать как сообщение, созданное некоторым «приведенным» источником B , имеющим два состояния (рис. 5.6). При этом, очевидно, информация, содержащаяся в битовой последовательности, в точности та же, которая была создана исходным источником X .

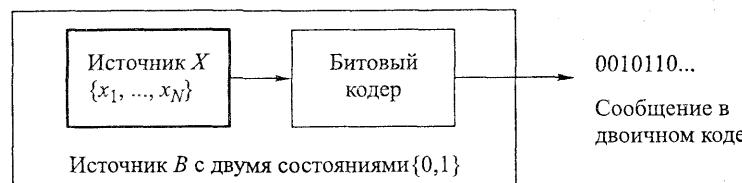


Рис. 5.6

При таком рассмотрении процесса двоичного кодирования из теоремы Шэннона следует важное следствие: появление нулевых и единичных битов при эффективном двоичном кодировании являет-

ся равновероятным, т. е. $p_0 = p_1 = 1/2$. Действительно, изменения состояний источника B кодируются с битовыми затратами $R = 1$ (бит). При этом количество информации $H(B) = -p_0 \log_2 p_0 - p_1 \log_2 p_1$, создаваемое источником, максимально и равно 1 биту, когда состояния источника B равновероятны. Если же $p_0 \neq p_1$, т. е. $H(B) < 1$ (бит), то в соответствии с теоремой Шэннона существует способ повышения эффективности кодирования, при котором среднее число двоичных символов R , приходящееся на кодирование очередного состояния источника B , может быть уменьшено до величины $R < H(B) + \varepsilon < 1$ (бит), где ε — как угодно малое число.

Способы построения эффективных двоичных кодов основаны на необходимости получения равных вероятностей появления нулевого и единичного битов в двоичном представлении информации. Практически не отличающиеся друг от друга методы построения эффективных кодов первоначально были предложены независимо друг от друга Шэнноном и Фано.

Метод построения эффективных кодов по Шэннону — Фано

1. Упорядочиваем таблицу символов (состояний источника) $\{x_1, \dots, x_N\}$ в порядке невозрастания вероятностей: $p_1 \geq \dots \geq p_N$. Полученная таблица является первым блоком символов для последующего разбиения.

2. Разбиваем текущий блок символов на две части (два подблока) так, чтобы вероятности попадания символа сообщения в подблоки были как можно ближе друг к другу.

3. Приписываем очередной бит в коды символов подблоков: «0» — для первого подблока и «1» — для второго подблока. Далее подблоки рассматриваем независимо друг от друга как новые блоки в разбиении таблицы символов.

4. Если в разбиении таблицы символов присутствует блок, содержащий более одного символа, то повторяем с этим блоком шаги 2 и 3, иначе заканчиваем процедуру построения двоичных кодов.

Пример 5.4. Используя методику Шэннона — Фано, построить эффективные коды для источника сообщений Z с приведенным распределением вероятностей появления символов.

z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8
0, 22	0, 20	0, 16	0, 16	0, 10	0, 10	0, 04	0, 02

◀ Последовательность разбиений таблицы символов и соответствующая схема построения кода приведены ниже. Битовые затраты на кодирование одного символа $R = \sum_{k=1}^8 p_k R_k = 2,84$ бита,

$$\text{где } R_k \text{ — число битов в коде символа } z_k.$$

z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8
0, 22	0, 20	0, 16	0, 16	0, 10	0, 10	0, 04	0, 02
0		1					
00	01		10		11		
	010	011	100	101	110	111	
					1110	1111	

Двоичная энтропия источника в данном случае $H(Z) \approx 2,76$ бита, т. е. полученный код обладает избыточностью $R - H \approx 0,08$ бита на символ. Заметим, что методика Шэннона — Фано обладает некоторой неоднозначностью построения кода: она не регламентирует, как следует производить разбиение блока на подблоки в том случае, когда возможны два варианта. Так, для данного примера можно построить и другую таблицу кодов, приведенную ниже.

z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8
0, 22	0, 20	0, 16	0, 16	0, 10	0, 10	0, 04	0, 02
0		1					
00	01	10		11			
		100	101	110	111		
				1110	1111	11110	11111

Для данного разбиения получим битовые затраты $R = 2,80$ бита и вдвое меньшую избыточность кода: $R - H \approx 0,04$ бита на символ. ►

Как видим, некоторая избыточность в полученных в примере 5.4 кодах все-таки осталась. Из теоремы Шэннона следует, что избыточность кодов, т. е. величину $\Delta = R - H$, можно сделать как угодно малой. Для этого символы сообщения надо кодировать не

по одному, а объединять их в серии (по два, три и т.д.), следуя идею, приведенной в доказательстве теоремы 5.3.

Пример 5.5. Построить эффективные коды для источника сообщений Z , имеющего вероятности появления символов: $p(z_1) = 0,9$ и $p(z_2) = 0,1$.

◀ Энтропия источника

$$H(Z) = -0,1 \log_2 0,1 - 0,9 \log_2 0,9 \approx 0,47 \text{ бита.}$$

Посимвольное кодирование источника дает единственно возможный код, когда одному символу приписывается нулевой, а другому — единичный бит, что дает битовые затраты $R = 1$ бит/символ и более чем стопроцентную избыточность кода:

$$(R-H)/H \approx 0,53/0,47 \approx 1,13.$$

Для того чтобы повысить эффективность кодирования, объединим последовательно выдаваемые источником Z двоичные символы в пары. В качестве исходного для построения кода рассматриваем новый алфавит с четырьмя символами. Соответствующее распределение вероятностей и полученные по методу Шэннона — Фано коды приведены в таблице.

Z	$(z_1 z_1)$	$(z_1 z_2)$	$(z_2 z_1)$	$(z_2 z_2)$
p	0,81	0,09	0,09	0,01
Код	0	10	110	111

Битовые затраты на кодирование пары символов составляют величину

$$2R = 0,81 \cdot 1 + 0,09 \cdot 2 + 0,09 \cdot 3 + 0,01 \cdot 3 = 1,29 \text{ бита.}$$

То есть битовые затраты в пересчете на один символ сообщения источника Z уменьшились и составили величину $R = 0,645$ бита, сократив избыточность кода со 113 % до 37 %. Кодируя сообщение источника Z тройками символов, т. е. рассматривая при построении кодов Шэннона — Фано алфавит из 8 символов, можно дополнительно повысить эффективность кодирования. (Проверьте это самостоятельно.) ►

Следуя общей схеме объединения символов в пары, тройки и т. д., теоретически можно как угодно приблизить величину бито-

вых затрат на кодирование одного символа к значению двоичной энтропии, для самого общего случая источника сообщений с произвольными объемом (числом символов) алфавита и распределением вероятностей. Однако на практике это не всегда целесообразно и технически возможно: например, если источник имеет 256 состояний, то при рассмотрении пар символов потребуется составлять таблицу кодов уже для $256 \times 256 = 65536$ символов, а снижение избыточности кодирования при этом может быть крайне незначительным. В любом случае, невозможно получить битовые затраты меньшие, чем теоретический предел, который определяется величиной двоичной энтропии источника.

Как отмечалось при рассмотрении примера 5.4, методика Шэннона — Фано имеет некоторую неоднозначность, которая может приводить к различным по эффективности кодам. Этой неоднозначности лишен получивший значительно более широкое распространение метод построения эффективных кодов, предложенный Хаффманом. Коды Хаффмана нередко оказываются эффективнее кодов Шэннона — Фано.

Метод построения эффективных кодов по Хаффману

1. Построение дерева Хаффмана.

1.1. Упорядочиваем текущую таблицу символов в порядке неввозрастания вероятностей: $p_1 \geq \dots \geq p_k \geq \dots$

1.2. Два последних символа, имеющих наименьшие вероятности появления, объединяем в один новый символ, которому приписываем суммарную вероятность объединенных символов. Если в полученном алфавите имеется более одного символа, то переходим на шаг 1.1, иначе — на шаг 2.

Комментарий. Вершина дерева Хаффмана — это единственный символ окончательного алфавита, листья — символы исходного алфавита, прочие узлы — символы промежуточных алфавитов, полученные в результате слияния других символов.

2. Построение битового кода. Из каждого узла дерева Хаффмана (за исключением листьев) строим по два ребра; приписываем одному ребру бит «0», другому — «1». Код каждого символа исходного алфавита получается в результате последовательного дописывания битов, соответствующих ребрам графа, которые необходимо пройти от вершины дерева до листа-символа.

Пример 5.6. Построить по методике Хаффмана эффективный двоичный код для источника, имеющего 5 равновероятных состояний.

◀ Для определенности положим, что символами исходного алфавита являются буквы $\{A, B, C, D, E\}$. Процедуру построения дерева Хаффмана (шаг 1 алгоритма) тогда можно изобразить в виде схемы, представленной на рис. 5.7.

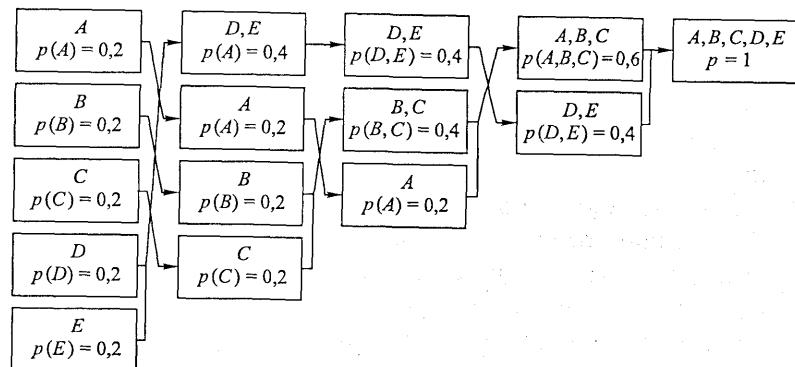


Рис. 5.7

Построение соответствующего двоичного кода по дереву Хаффмана (выполнение шага 2 алгоритма) отражает рис. 5.8. Двоичная энтропия источника сообщений в данном случае $H = \log_2 5 \approx 2,32$ бита, а полученные битовые затраты $R = 2,4$ бита (избыточность кода — около 3 %). ►

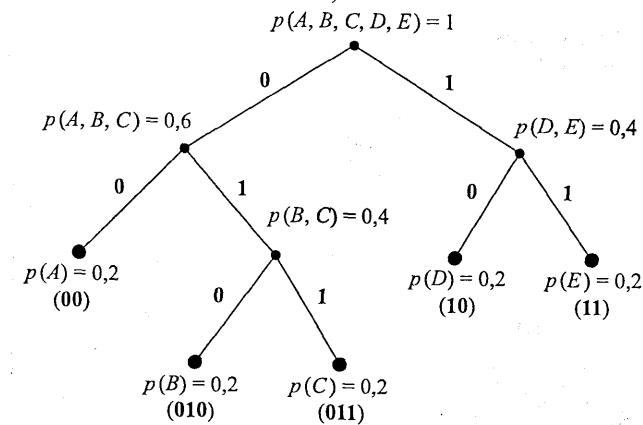


Рис. 5.8

Задание. Постройте самостоятельно коды Хаффмана для источника сообщений, рассмотренного в примере 5.4. Убедитесь в том, что битовые затраты составляют 2,80 бита на символ, т. е. соответствуют лучшему из рассмотренных в примере 5.4 вариантов построения кода Шэннона — Фано.

■ 5.4. Кодирование длин серий

Если энтропия источника сообщений мала (т. е. имеется состояние, в котором источник находится с вероятностью, близкой к единице), то вместо кодирования сообщений по Хаффману или Шэннону — Фано с объединением символов в блоки (см. пример 5.5) можно предложить другой простой в реализации метод, который во многих случаях дает незначительную избыточность кода, — метод кодирования длин серий (КДС). В этом методе последовательность символов сообщения также разбивается на блоки, но эти блоки имеют переменное число символов, а кодируются кодами фиксированной битовой длины b .

Рассмотрим метод КДС на примере двоичного источника $Z = \{z_1, z_2\}$. Обозначим вероятности

$$P\{Z = z_1\} = p, \quad P\{Z = z_2\} = 1 - p = q$$

и положим для определенности $p > q$. Введем в рассмотрение следующую систему из $M = 2^b$ блоков, состоящих из разного числа символов сообщения:

Блок символов	Количество символов в блоке	Вероятность появления блока
$S_0 = z_2$	1	$1 - p$
$S_1 = z_1 z_2$	2	$p(1 - p)$
$S_2 = z_1 z_1 z_2$	3	$p^2(1 - p)$
...
$S_{M-2} = z_1 \dots z_1 z_2$	$M-1$	$p^{M-2}(1 - p)$
$S_{M-1} = z_1 \dots z_1 z_1$	$M-1$	p^{M-1}

Каждый блок данной системы содержит на один символ z_1 больше предыдущего. Любое сообщение из последовательности символов, созданных источником Z , можно однозначно преобразовать в последовательность блоков из системы $\{S_k\}_{k=0}^{M-1}$. Битовый код сообщения будет представлять собой последовательно записанные (в двоичном представлении из b битов) номера соответствующих блоков. Фактически, кодируется длина серии символов z_1 , разделяющих в сообщении появление символа z_2 , что и дало название описанному методу. Реализация метода КДС весьма проста, так как для стороны кодера сводится к простому подсчету числа повторений символов, а для декодера — к формированию серий повторяющихся символов.

Оценим эффективность применения метода КДС для кодирования двоичного источника. Найдем среднее число символов сообщения, которое приходится на блок, кодируемый b битами. Математическое ожидание для числа символов, которые содержатся в одном блоке системы $\{S_k\}_{k=0}^{M-1}$, находим по данным из приведенной выше таблицы:

$$\begin{aligned} K &= (1 - p) + 2p(1 - p) + 3p^2(1 - p) + \dots + (M - 1)p^{M-2}(1 - p) + (M - 1)p^{M-1} = \\ &= (1 - p) \sum_{k=0}^{M-1} kp^{k-1} + (M - 1)p^{M-1} = (1 - p) \frac{d}{dp} \left(\sum_{k=0}^{M-1} p^k \right) + (M - 1)p^{M-1} = \\ &= (1 - p) \frac{d}{dp} \left(\frac{p^M - 1}{p - 1} \right) + (M - 1)p^{M-1} = \frac{1 - p^{M-1}}{1 - p}. \end{aligned}$$

Учитывая, что $M = 2^b$, находим среднее число битов $R = b/K$, затраченных на кодирование одного символа сообщения:

$$R = \frac{(1 - p)b}{1 - p^{2^b-1}}. \quad (5.6)$$

На рис. 5.9 приведены графики функций $R = R(b, p)$ для некоторых значений b , а также для сравнения указаны минимально возможные битовые затраты, определяемые значением энтропии двоичного источника $H(p) = -p \log_2 p - (1 - p) \log_2 (1 - p)$.

Разрядность b представления номера блока является единственным параметром метода КДС. Как видно из графиков, приве-

денных на рис. 5.9, параметр b следует выбирать в зависимости от значения вероятности p , которая определяет модель двоичного источника без памяти. При надлежащем выборе b избыточность кода, полученного по методу КДС, будет незначительной.

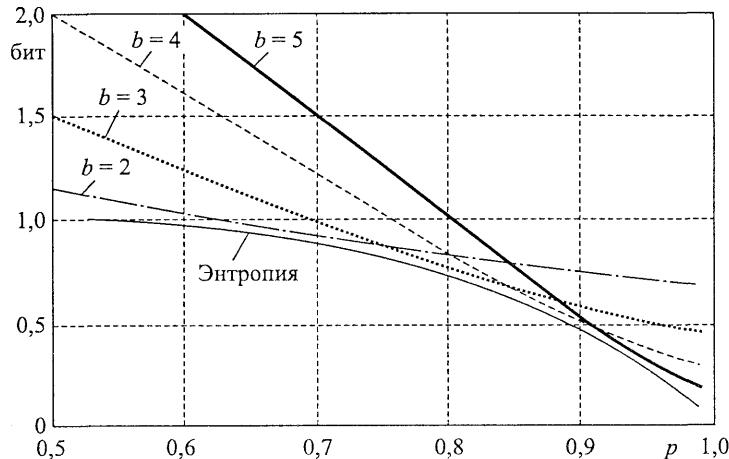


Рис. 5.9. Графики функции (5.6) и энтропии двоичного источника

Пример 5.7. Для источника, рассмотренного в примере 5.5, найти оптимальное число блоков для использования метода КДС и оценить избыточность кода.

◀ Из графиков (см. рис. 5.9) находим оптимальное значение разрядности для представления номера блоков: $b = 4$. Отсюда получаем число блоков для метода КДС: $M = 2^4 = 16$. Энтропия источника найдена в примере 5.5: $H \approx 0,47$ бита. Битовые затраты находим по формуле (5.6): $R \approx 0,504$ бита. Получаем избыточность кода для метода КДС:

$$\frac{R - H}{H} \times 100 \% \approx 7 \%.$$

■ 5.5. Арифметическое кодирование

Рассмотрим метод кодирования, который был предложен в 70-х годах прошлого века и представляет собой альтернативу ме-

тодам Шэннона — Фано и Хаффмана. Суть его состоит в том, что последовательности символов $\{x(1), x(2), \dots, x(M)\}$, созданных дискретным источником сообщений X с известным алфавитом и вероятностями появления символов $\{(x_k, p_k)\}_{k=1}^N$, ставится в соответствие некоторое число, однозначно задающее данную последовательность. Идею метода проиллюстрируем на примере.

Пример 5.8. Рассмотрим источник, алфавит которого и соответствующие вероятности появления символов заданы следующей таблицей:

A	B	C	D
0,5	0,2	0,2	0,1

Найдем число, соответствующее сообщению «DABAC».

◀ Сначала по таблице распределения вероятностей необходимо построить *интервалы вероятностей*¹ $\Delta_k = [L_{x_k}; U_{x_k})$ для каждого символа x_k алфавита $\{x_k\}_{k=1}^N$ по правилу: $U_{x_k} = \sum_{j=1}^k p_j$,

$L_{x_k} = U_{x_k} - p_k$. То есть интервал $[0; 1)$ разбивается на N непересекающихся интервалов с длинами, равными вероятностям появления соответствующих символов: $|\Delta_k| = U_{x_k} - L_{x_k} = p_k$. В нашем случае имеем четыре интервала (см. табл.).

Символ	A	B	C	D
Интервал вероятности	[0, 0; 0, 5)	[0, 5; 0, 7)	[0, 7; 0, 9)	[0, 9; 1, 0)

Далее устанавливаются границы начального *текущего интервала* $[L; U]$: $L = 0$ (нижняя граница), $U = 1$ (верхняя граница).

Затем начинается последовательная обработка символов. Получая очередной символ x_k , кодер производит разбиение текущего интервала пропорционально начальным интервалам вероятностей символов и в качестве следующего текущего интервала выбирает тот, которому соответствует символ x_k . При этом каждый

¹ При описании метода арифметического кодирования полуинтервалы вида $[a; b)$ традиционно называют интервалами.

следующий текущий интервал вложен в предыдущий. Так, для нашего случая кодирование сообщения «DABAC» сводится к процедуре построения последовательности текущих интервалов, показанной на рис. 5.10. Получили окончательный интервал $[0, 9285; 0, 9295]$. Из этого интервала можно взять любое число, например $B = 0, 929$ (как содержащее наименьшее число цифр), и восстановить по нему исходную последовательность из 5 символов «DABAC».

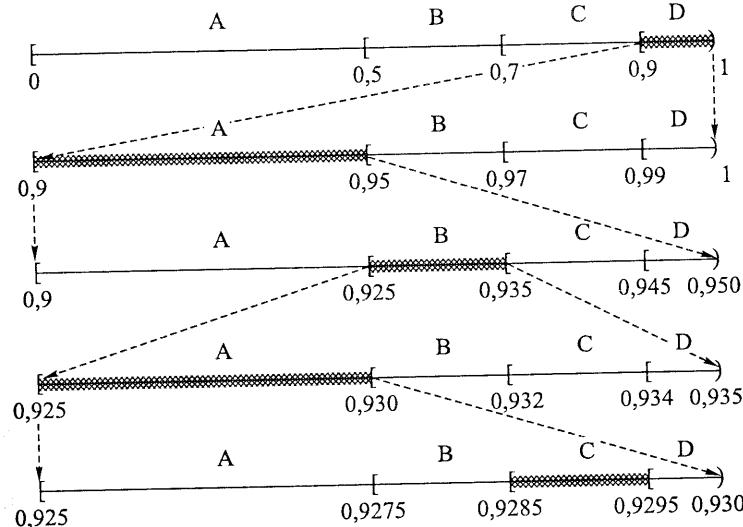


Рис. 5.10

Для восстановления последовательности на стороне декодера должно быть известно распределение вероятностей символов, т. е. разбиение на интервалы вероятностей. Изначально выбираем текущий интервал $[0; 1)$. Далее разбиваем текущий интервал пропорционально интервалам вероятностей символов, проверяем, в интервал какого символа попадает число B . Так находим очередной декодированный символ и выбираем его интервал вероятности в качестве следующего текущего интервала. И т. д. ►

Заметим, что появление того или иного символа можно отождествить с выбором наудачу числа из текущего интервала. Попадание числа в интервал вероятности некоторого символа означает

событие, состоящее в появлении этого символа. Подобная трактовка хорошо известна как «геометрическая вероятность».

Дадим теперь формализованное описание процедуры нахождения числа B , которое ставится в соответствие последовательности $\{x(1), x(2), \dots, x(M)\}$, созданной источником с алфавитом $\{x_k\}_{k=1}^N$, для которого задано (известно) разбиение на интервалы вероятностей $\Delta_k = [L_{x_k}; U_{x_k})$, $k = 1, \dots, N$.

Шаг 1. Установить границы текущего интервала:

$$L = 0, U = 1. \quad \text{Установить счетчик символов } j = 1.$$

Шаг 2. Определить новые границы текущего интервала:

$$W = U - L,$$

$$U = L + W \times U_{x(j)},$$

$$L = L + W \times L_{x(j)}.$$

Шаг 3. Увеличить счетчик символов: $j = j + 1$. Если $j > M$, то перейти на шаг 4, иначе — перейти на шаг 2.

Шаг 4. Выдать число $B \in [L; U)$, содержащее наименьшее число цифр.

Наличие арифметических операций на шаге 2 послужило причиной того, что описанный метод получил название *арифметического кодирования*.

Декодирование сообщения из M символов, которому соответствует число B , можно описать в виде следующей процедуры, также основанной на выполнении арифметических операций.

Шаг 1. Ввести число B . Установить счетчик символов $j = 1$.

Шаг 2. Найти интервал Δ_k , в который попадает число

$$B \in \Delta_k. \quad \text{Выдать символ } x(j) = x_k.$$

Шаг 3. Привести текущий интервал $[L_{x_k}; U_{x_k})$ к интервалу $[0; 1)$:

$$B = \frac{B - L_{x_k}}{U_{x_k} - L_{x_k}}.$$

Шаг 4. Увеличить счетчик символов: $j = j + 1$. Если $j \leq M$, то перейти на шаг 2, иначе — закончить работу.

На шаге 3 осуществляется своего рода «проектирование» текущего интервала $[L_{x_k}; U_{x_k})$ в интервал $[0; 1)$ с соответствующим изменением значения кодового числа $B \rightarrow B'$ (рис. 5.11).

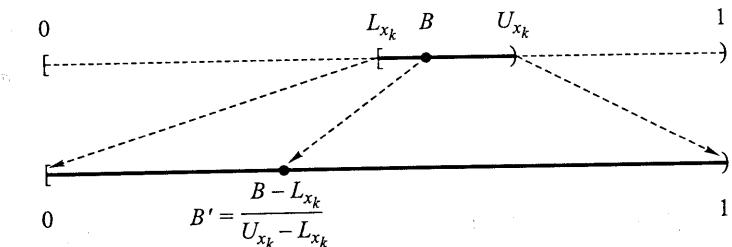


Рис. 5.11

Поскольку в компьютере используется двоичное представление чисел, то выходным битовым кодом арифметического кодера является двоичная запись числа B . Поэтому на шаге 4 в процедуре кодирования следует выбирать число, которое можно записать, наименьшим числом двоичных цифр $\{0; 1\}$.

Пример 5.9. В условиях примера 5.8 определить кодовое число, содержащее наименьшее число двоичных цифр.

◀ Границы текущего интервала, полученные после окончания рассмотренной в примере 5.8 процедуры кодирования, представим в двоичном виде. Для произвольного числа $T \in [0; 1)$ двоичные цифры в представлении $T = 0, t_1 t_2 \dots t_m \dots_{(2)} = \sum_{m=1}^{\infty} t_m 2^{-m}$ находятся по формуле

$$t_m \equiv [T \cdot 2^m] \pmod{2}, \quad t_m \in \{0; 1\},$$

используя которую получим:

$$U = 0,9295_{(10)} = 0,11101101111\dots_{(2)},$$

$$L = 0,9285_{(10)} = 0,111011011011\dots_{(2)}.$$

Отсюда находим число $B \in [L; U)$ с минимальным числом двоичных цифр: $B = 0,1110110111_{(2)} = 0,9287109375_{(10)}$. ►

После окончания процедуры арифметического кодирования M символов первые биты в двоичной записи чисел $L = 0, l_1 l_2 \dots_{(2)}$ и $U = 0, u_1 u_2 \dots_{(2)}$ совпадают. Пусть количество первых совпадших разрядов $\{b_j\}$ равно K : $u_j = l_j = b_j$, $j = 1, \dots, K$. Тогда, поскольку $L < U$, первые несовпадающие биты: $l_{K+1} = 0$, $u_{K+1} = 1$.

Можем записать:

$$\begin{aligned} U = 0, b_1 b_2 \dots b_K 1 u_{K+2} u_{K+3} \dots_{(2)} &= \sum_{k=1}^K b_k 2^{-k} + \underbrace{\frac{1}{2^{K+1}}}_{\leqslant 1/2^{K+1}} + \sum_{k=K+2}^{\infty} u_k 2^{-k}; \\ L = 0, b_1 b_2 \dots b_K 0 l_{K+2} l_{K+3} \dots_{(2)} &= \sum_{k=1}^K b_k 2^{-k} + \underbrace{\sum_{k=K+2}^{\infty} l_k 2^{-k}}_{\leqslant 1/2^{K+1}}. \end{aligned} \quad (5.7)$$

Считаем, что в представлении (5.7) для верхней границы среди двоичных коэффициентов $\{u_k\}_{k=K+2}^{\infty}$ есть ненулевые, в противном случае заменим представление $U = 0, b_1 \dots b_K 1_{(2)}$ на $U = 0, b_1 \dots b_K 01111\dots_{(2)}$, эквивалентность которого следует из

формулы для суммы геометрической прогрессии: $\sum_{k=K+2}^{\infty} 2^{-k} = \frac{1}{2^{K+1}}$.

Очевидно, что число совпадших старших разрядов в двоичных представлениях (5.7) тем больше, чем ближе друг к другу границы L и U , которые сближаются тем быстрее, чем меньше вероятность появления кодируемых символов. При этом длина текущего интервала (см. (5.7)), равная вероятности реализации серии $\{x(1), \dots, x(M)\}$:

$$P(x(1), x(2), \dots, x(M)) = \prod_{m=1}^M P(x(m)) = U - L \leqslant \frac{1}{2^{K+1}} + \frac{1}{2^{K+1}} = \frac{1}{2^K},$$

откуда

$$K = K(x(1), \dots, x(M)) \leqslant -\log_2 P(x(1), \dots, x(M)).$$

Усредняя последнее неравенство по всем возможным реализациям серии $\{x(1), \dots, x(M)\}$, получаем математическое ожидание

количества совпавших старших разрядов в (5.7) после кодирования M символов:

$$\begin{aligned}\bar{K} &= \sum_{k_1=1}^N \dots \sum_{k_M=1}^N P(x_{k_1}, \dots, x_{k_M}) K(x_{k_1}, \dots, x_{k_M}) \leqslant \\ &\leqslant -\sum_{k_1=1}^N \dots \sum_{k_M=1}^N P(x_{k_1}, \dots, x_{k_M}) \log_2 P(x_{k_1}, \dots, x_{k_M}) = M \cdot H(X), \quad (5.8)\end{aligned}$$

поскольку M символов, созданных источником без памяти, могут рассматриваться как символы, созданные M независимыми источниками, каждый из которых имеет энтропию $H(X)$. Это позволяет использовать свойство 2° энтропии.

Так как среди коэффициентов $\{u_{K+2}, u_{K+3}, \dots\}$ существует хотя бы один, отличный от нуля, то в качестве кодового числа можно выбрать $B = 0.b_1 b_2 \dots b_K 1_{(2)}$, что гарантирует выполнение требования $L \leq B < U$. Получаем, что двоичное представление кодового числа всегда заканчивается единичным битом, т. е. последний можно не передавать, а при декодировании всегда сначала добавлять единичный бит в конец кодовой последовательности. Тогда кодом сообщения будет являться последовательность из K битов: $\{b_1, b_2, \dots, b_K\}$, а для средних битовых затрат на кодирование одного символа сообщения $R = \bar{K}/M$ получим (см. (5.8)):

$$R \leq H(X).$$

В теореме 5.1 мы показали, что битовые затраты на кодирование одного символа сообщения не могут быть меньше энтропии источника. Поэтому на основании полученного неравенства можно сделать предположение, что арифметическое кодирование всегда обеспечивает максимально возможную эффективность кодирования, когда $H(X) = R$. Однако это предположение неверно. Дело в том, что в описанной выше схеме кодирования кодовое число (и, соответственно, выходной битовый код) не несет информации о том, из какого количества символов M состоит сообщение. По одному и тому же кодовому числу для заданного источника и соответствующего разбиения интервала вероятностей можно восстановить бесконечное число различных сооб-

щений, отличающихся количеством символов ($M = 1, 2, 3, \dots$), при этом короткие сообщения будут являться началом более длинных. Так, для источника из примера 5.8 по последовательности битов «111011011», соответствующей кодовому числу $B = 0.1110110111_{(2)} = 0.9287109375_{(10)}$ (см. пример 5.9), можно восстановить сообщения: D, DA, DAB, DABA, DABAC, DABACA, ... Нельзя указать, какое именно сообщение было закодировано. Соответственно, описанный способ арифметического кодирования является неоднозначным.

На практике к алфавиту источника добавляют еще один специальный символ признака конца сообщения, при этом для битовых затрат на кодирование одного символа может быть получена оценка¹:

$$H(X) + \frac{1}{M} \leq R \leq H(X) + \frac{2}{M}. \quad (5.9)$$

Оценка (5.9) показывает, что с увеличением числа обработанных символов M битовые затраты на кодирование одного символа сообщения приближаются к теоретическому минимуму. Арифметическое кодирование не требует специальных ухищрений при кодировании источников с малой энтропией, при этом не встречает тех технических проблем увеличения (в геометрической прогрессии) размерности таблиц кодов при объединении символов сообщения в блоки, которые возникают при использовании методик Хаффмана или Шэннона — Фано (см. пример 5.5). В результате, в большинстве случаев использование арифметического кодирования приводит к более эффективному сжатию данных. За это приходится расплачиваться большей вычислительной сложностью реализации метода.

Приведенные выше рассуждения были призваны проиллюстрировать идею метода арифметического кодирования и относились к гипотетическому вычислительному устройству, которое оперирует бесконечной разрядностью двоичного представления чисел. Выполнение арифметического кодирования на реальных вычислительных устройствах обладает спецификой, связанной с

¹ Следует из неравенства (1.1), приведенного в монографии: Сэломон Д. Сжатие данных, изображений и звука. М., 2004.

конечной разрядностью представления чисел. Так, верхнюю границу U представляем с помощью целого r -битового числа $U_r \in \{0, \dots, 2^r - 1\}$ в виде $U = (U_r + 1) \cdot 2^{-r}$. Изначально, для $U = 1$, задаем значение $U_r = 2^r - 1 = 111\dots 1_{(2)}$. Нижнюю границу и другие промежуточные числа $T \in [0; 1]$ представляем как $T = T_r \cdot 2^{-r}$, где целое число $T_r \in \{0, \dots, 2^r - 1\}$. Например, шаг 2 процедуры кодирования (определение границ текущего интервала) при обработке j -го символа сообщения будет выглядеть тогда (в целочисленной арифметике!) следующим образом:

$$\begin{aligned} W_r &= U_r - L_r + 1, \\ U_r &= L_r + (W_r \cdot U_{x(j),r}) 2^{-r} - 1, \\ L_r &= L_r + (W_r \cdot L_{x(j),r}) 2^{-r}. \end{aligned}$$

Каждый раз после выполнения указанных вычислений проверяются на совпадение старшие биты чисел U_r и L_r . Совпадшие биты (вновь обозначим их число через K) уже не могут измениться в процессе работы алгоритма кодирования и выводятся кодером в выходной поток данных. Затем выполняется арифметический сдвиг регистров U_r и L_r влево на K позиций, а освободившиеся K младших разрядов заполняются единицами для регистра U_r и нулями для регистра L_r .

При реализации алгоритма арифметического кодирования в целочисленной арифметике с конечной разрядностью r представления чисел существует также опасность возникновения ситуации, когда верхняя и нижняя границы недопустимо сближаются: $U_r = \underbrace{100\dots 00}_{r \text{ разрядов}}, L_r = \underbrace{011\dots 11}_{r \text{ разрядов}}$. В этом случае алгоритм зациклится,

так как дальнейшее уменьшение текущего интервала (далее сближение границ) невозможно. Для избежания подобной ситуации существуют специальные меры технического характера, описание которых мы опустим.

■ 5.6. Условная энтропия

Вводя понятие энтропии (5.2) как меры априорной неопределенности состояния источника сообщений, мы требовали, чтобы для ансамбля независимых источников X и Y выполнялось условие $H(X, Y) = H(X) + H(Y)$. Если же источники *зависимы*, то информация о реализации того или иного состояния одного источника позволяет уменьшить априорную неопределенность состояния другого; интуитивно понятно, что это должно означать следующее: $H(X, Y) < H(X) + H(Y)$, и в общем случае

$$H(X, Y) \leq H(X) + H(Y). \quad (5.10)$$

С учетом (5.2) неравенство (5.10) принимает вид:

$$\underbrace{-\sum_i \sum_j p(x_i, y_j) \log p(x_i, y_j)}_{H(X, Y)} \leq \underbrace{-\sum_i p(x_i) \log p(x_i)}_{H(X)} + \underbrace{-\sum_j p(y_j) \log p(y_j)}_{H(Y)}$$

(здесь суммирование производится «по всем возможным значениям», т. е. пределы суммирования определяются способом нумерации и числом возможных состояний источников сообщений). Справедливость неравенства (5.10) мы докажем несколько позже (см. теорему 5.4).

Ансамбль двух источников сообщений можно описать, задав закон совместного распределения вероятностей всевозможных пар состояний (x_i, y_j) , например, в виде таблицы распределения вероятностей $\{p_{i,j} = P(X = x_i, Y = y_j)\}$. Удобно трактовать пары (x_i, y_j) как реализации случайного вектора (X, Y) , компоненты которого являются случайными величинами дискретного типа. Дискретные случайные величины X, Y являются независимыми, если $\forall i, j$:

$$\underbrace{P(X = x_i, Y = y_j)}_{p_{i,j} = p(x_i, y_j)} = \underbrace{P(X = x_i)}_{p(x_i)} \underbrace{P(Y = y_j)}_{p(y_j)}.$$

В общем же случае, когда случайные величины зависимы, вероятность реализации случайного вектора выражается через условные вероятности:

$$p(x_i, y_j) = p(x_i)p(y_j|x_i) = p(y_j)p(x_i|y_j),$$

причем $\exists i, j: p(x_i, y_j) \neq p(x_i)p(y_j)$. Тогда энтропию ансамбля источников можно представить в следующем виде:

$$\begin{aligned} H(X, Y) &= -\sum_{n,m} p(x_n, y_m) \log p(x_n, y_m) = \\ &= -\sum_{n,m} p(x_n)p(y_m|x_n) \log(p(x_n)p(y_m|x_n)) = \\ &= -\underbrace{\sum_n p(x_n) \log p(x_n)}_{H(X)} \underbrace{\sum_m p(y_m|x_n)}_1 + \\ &\quad + \sum_n p(x_n) \underbrace{\left(-\sum_m p(y_m|x_n) \log p(y_m|x_n) \right)}_{H(Y|x_n)}. \end{aligned}$$

Таким образом,

$$H(X, Y) = H(X) + \sum_n p(x_n) H(Y|x_n),$$

где выражение

$$H(Y|x_n) = -\sum_m p(y_m|x_n) \log p(y_m|x_n) \quad (5.11)$$

называется *частной условной энтропией* источника Y относительно состояния x_n источника X . Частная условная энтропия (5.11) характеризует априорную неопределенность источника Y при условии, что состояние источника $X = x_n$ известно. Усредняя величину (5.11) по всем возможным состояниям источника X , получим *условную энтропию* источника Y относительно источника X :

$$H(Y|X) = \sum_n p(x_n) H(Y|x_n). \quad (5.12)$$

Аналогичное выражение можно записать и для условной энтропии источника X относительно источника Y .

С использованием понятия условной энтропии можно записать общее соотношение для энтропии ансамбля двух, вообще говоря, зависимых источников сообщений X и Y :

$$H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y). \quad (5.13)$$

Для обоснования неравенства (5.10) осталось доказать важное свойство условной энтропии (5.12), которое мы сформулируем в виде следующей теоремы.

Теорема 5.4. Для ансамбля любых дискретных источников (X, Y) условная энтропия каждого источника не превосходит его безусловной энтропии:

$$H(X|Y) \leq H(X), \quad H(Y|X) \leq H(Y).$$

◀ Не теряя общности рассуждений, положим, что в формулах (5.11)–(5.13) используется натуральное основание логарифма, тогда $\forall x > 0$ справедлива оценка $\log x \leq x - 1$, используя которую получим:

$$\begin{aligned} H(X|Y) - H(X) &= -\sum_m p(y_m) \sum_n p(x_n|y_m) \log p(x_n|y_m) + \\ &\quad + \sum_n p(x_n) \log p(x_n) \underbrace{\sum_m p(y_m|x_n)}_1 = \sum_{n,m} p(x_n, y_m) \log \frac{p(x_n)}{p(x_n|y_m)} \leq \\ &\leq \sum_{n,m} p(x_n, y_m) \left(\frac{p(x_n)}{p(x_n|y_m)} - 1 \right) = \\ &= \underbrace{\sum_n p(x_n)}_1 \underbrace{\sum_m p(y_m)}_1 - \underbrace{\sum_{n,m} p(x_n, y_m)}_1 = 0. \end{aligned}$$

Отсюда $H(X|Y) \leq H(X)$. Аналогично, $H(Y|X) \leq H(Y)$. ►

Если источники независимы, то условная энтропия совпадает с безусловной, $H(Y|X) = H(Y)$ и $H(X|Y) = H(X)$, что следует непосредственно из формул (5.11) и (5.12), так как в этом случае условные вероятности в формуле (5.11) совпадают с безусловными.

Итак, если между источниками существует статистическая зависимость, то знание о реализации состояния одного источника может только уменьшить неопределенность другого.

Пример 5.10. Показать, что если состояния источника Y связаны функциональной зависимостью ϕ с состояниями источника X , то условная энтропия $H(Y|X) = 0$ и $H(X, Y) = H(X)$.

◀ Так как $p(y_k|x_m) = P\{Y = y_k | X = x_m\} = \begin{cases} 1, & y_k = \phi(x_m); \\ 0, & y_k \neq \phi(x_m), \end{cases}$ то

частная условная энтропия $\forall m$

$$H(Y|x_m) = -\sum_k p(y_k|x_m) \log p(y_k|x_m) = 0,$$

и, следовательно, $H(Y|X) = 0$. Отсюда, в силу соотношения (5.13), $H(X, Y) = H(X)$. ►

Если приведенные рассуждения обобщить на ансамбль из n дискретных источников сообщений (X_1, \dots, X_n) , то для энтропии ансамбля можно получить следующее выражение:

$$H(X_1, X_2, \dots, X_n) = H(X_1) + H(X_2 | X_1) + \\ + H(X_3 | X_1 X_2) + \dots + H(X_n | X_1 X_2 \dots X_{n-1}),$$

где

$$H(X_j | X_1 X_2 \dots X_{j-1}) = \sum_{k_1} \dots \sum_{k_{j-1}} p(x_{k_1}, \dots, x_{k_{j-1}}) H(X_j | x_{k_1}, \dots, x_{k_{j-1}}) \quad (5.14)$$

— условная энтропия источника X_j относительно ансамбля (X_1, \dots, X_{j-1}) ,

$$H(X_j | x_{k_1}, \dots, x_{k_{j-1}}) = -\sum_{k_j} p(x_{k_j} | x_{k_1}, \dots, x_{k_{j-1}}) \log p(x_{k_j} | x_{k_1}, \dots, x_{k_{j-1}}) \quad (5.15)$$

— частная условная энтропия источника X_j относительно реализации ансамбля $(X_1 = x_{k_1}, X_2 = x_{k_2}, \dots, X_{j-1} = x_{k_{j-1}})$.

Можно показать (доказательство аналогично приведенному для теоремы 5.4), что для условной энтропии (5.14) вновь верна оценка

$$H(X_j | X_1 X_2 \dots X_{j-1}) \leq H(X_j).$$

Отсюда

$$H(X_1, X_2, \dots, X_n) \leq \sum_{j=1}^n H(X_j)$$

(равенство выполняется только для ансамбля независимых источников).

Пример 5.11. Для ансамбля источников U и V найти по заданному закону совместного распределения вероятностей двоичную условную энтропию $H(V|U)$.

$u_i \backslash v_j$	v_1	v_2	v_3
u_1	0,4	0,1	0
u_2	0	0,2	0,3

◀ Для источника U безусловные вероятности состояний $p(u_1) = p(u_2) = 0,5$. Соответственно, безусловная энтропия $H(U) = 1$ бит. Энтропия ансамбля (U, V)

$$H(UV) = -\sum_i \sum_j p_{i,j} \log_2 p_{i,j} = \\ = -0,4 \log_2 0,4 - 0,1 \log_2 0,1 - 0,2 \log_2 0,2 - 0,3 \log_2 0,3 \approx 1,85 \text{ бита.}$$

Отсюда, в соответствии с (5.13), условная энтропия равна

$$H(V|U) = H(UV) - H(U) \approx 0,85 \text{ бита.}$$

Заметим, что безусловная энтропия источника V почти в два раза выше:

$$H(V) = -0,4 \log_2 0,4 - 0,3 \log_2 0,3 - 0,3 \log_2 0,3 \approx 1,57 \text{ бита.} \blacktriangleright$$

■ 5.7. Кодирование дискретного источника с памятью

Используя понятие условной энтропии, рассмотрим источник сообщений «с памятью», для которого вероятности выбора очередного состояния зависят от предыстории прошлых состояний. Другими словами, имеется вероятностная зависимость между оче-

редным символом, создаваемым источником сообщения, и реализациами предыдущих символов (состояний). При этом мы вновь рассматриваем дискретный источник X , число состояний которого конечно и описывается некоторым алфавитом $\{x_1, \dots, x_N\}$.

В качестве модели источника сообщений рассмотрим источник с конечной памятью. Это означает, что на распределение вероятностей для текущего перехода источника в очередное состояние оказывает влияние лишь конечное число M предыдущих состояний, т. е. вероятности перехода задаются *условными* законами распределения. Обозначим m -ое состояние источника при априорном рассмотрении (нереализованное состояние) через $X(m)$, а при апостериорном рассмотрении (реализованное состояние) — через $x(m)$. Тогда для описания вероятностей перехода источника с памятью необходимо задать условные распределения

$$P(X(m)=x(m)|X(m-1)=x(m-1), \dots, X(m-M)=x(m-M)),$$

где $x(m-j) \in \{x_1, \dots, x_N\}$, $j = 0, \dots, M$. Будем рассматривать *стационарный* источник, для которого указанные условные распределения не зависят от индекса m , и будем использовать запись

$$p(x_{k_0}|x_{k_1}, \dots, x_{k_M}) = P(X(m)=x_{k_0}|x(m-1)=x_{k_1}, \dots, x(m-M)=x_{k_M}), \quad (5.16)$$

где $k_0, k_1, \dots, k_M \in \{1, \dots, N\}$ и $x_{k_j} \in \{x_1, \dots, x_N\}$, $j = 0, \dots, M$. Если символы алфавита источника представить как числа, то рассматриваемый источник сообщений описывается стационарным случайным процессом с дискретным временем и дискретными состояниями (значениями).

Определение. Стационарный в широком смысле случайный процесс $X(n)$ с дискретным временем и дискретными состояниями, условный закон распределения которого зависит только от M предыдущих значений $x(n-1), \dots, x(n-M)$, называется *дискретным марковским процессом порядка M* , или *сложной цепью Маркова порядка M* . Если $M=1$, то цепь Маркова называют простой.

Таким образом, если символы алфавита рассматривать как числа, то сообщение стационарного источника X с памятью размера M представляет собой дискретный марковский процесс M -го порядка.

Для того чтобы полностью описать математическую модель источника с конечной памятью, необходимо задать закон *совместного* распределения вероятностей, т. е. $(M+1)$ -мерную таблицу:

$$p(x_{k_0}, x_{k_1}, \dots, x_{k_M}) = P(X(m)=x_{k_0}, X(m-1)=x_{k_1}, \dots, X(m-M)=x_{k_M}), \\ k_j = 1, \dots, N; \quad j = 0, \dots, M. \quad (5.17)$$

(Очевидно, что число элементов в этой таблице N^{M+1} .) Тогда по совместному закону распределения (5.17) условный закон (5.16) может быть найден следующим образом:

$$p(x_{k_0}|x_{k_1}, \dots, x_{k_M}) = \frac{p(x_{k_0}, x_{k_1}, \dots, x_{k_M})}{p(x_{k_1}, \dots, x_{k_M})},$$

где

$$p(x_{k_1}, \dots, x_{k_M}) = P(X(m-1)=x_{k_1}, \dots, X(m-M)=x_{k_M}) = \\ = \sum_{k_0=1}^N p(x_{k_0}, \dots, x_{k_M}).$$

Если известна реализованная предыстория состояний источника $(x(m-1)=x_{k_1}, \dots, x(m-M)=x_{k_M})$, то известен *частный* закон распределения вероятностей перехода в очередное состояние $X(m)$ и неопределенность перехода в новое состояние можно оценить по частной условной энтропии (сравните с (5.15)):

$$H(X(m)|x(m-1)=x_{k_1}, \dots, x(m-M)=x_{k_M}) = \\ = H(X|x_{k_1}, \dots, x_{k_M}) = - \sum_{k_0=1}^N p(x_{k_0}|x_{k_1}, \dots, x_{k_M}) \log p(x_{k_0}|x_{k_1}, \dots, x_{k_M}). \quad (5.18)$$

При этом каждому возможному набору из N^M предысторий источника соответствует, вообще говоря, своя таблица условного распределения вероятностей (5.16), на которой должно основываться построение процедуры эффективного кодирования очередного символа (состояния источника) с битовыми затратами, минимальное значение которых оценивается по величине (5.18) (с двоичным основанием логарифма). Среднее количество ин-

формации, связанное с очередным переходом источника в новое состояние и выдачей очередного символа сообщения, оценивается по условной энтропии (сравните с (5.14)):

$$\begin{aligned} H_M(X) &= H(X(m)|X(m-1), \dots, X(m-M)) = \\ &= \sum_{k_1=1}^N \dots \sum_{k_M=1}^N p(x_{k_1}, \dots, x_{k_M}) H(X|x_{k_1}, \dots, x_{k_M}). \end{aligned} \quad (5.19)$$

Условная энтропия (5.19) характеризует неопределенность источника с памятью и носит также название «условной энтропии M -го порядка». В частном случае для источника информации без памяти, когда $p(x_{k_0}|x_{k_1}, \dots, x_{k_M}) = P(X(m)=x_{k_0}) = p_{k_0}$ (см. (5.16)), величина (5.19) совпадает с введенной ранее безусловной энтропией (5.2), так как для любой предыстории получаем в формуле (5.18):

$$H(X|x_{k_1}, \dots, x_{k_M}) = -\sum_{k_0=1}^N p_{k_0} \log p_{k_0} = H(X).$$

Таким образом, энтропия (5.19) является обобщением понятия безусловной энтропии (или условной энтропии нулевого порядка) $H_0(X) = H(X)$ из (5.2).

Для условной энтропии (5.19) верна теорема 5.3 (Шэннона), т. е. значение $H_M(X)$ (5.19) является точной оценкой минимально возможных битовых затрат, приходящихся в среднем на кодирование одного символа сообщения источника с памятью. Схема построения эффективных битовых кодов для источника с памятью размера M основана на тех же общих идеях, которые изложены ранее в доказательствах леммы 5.2 и теорем 5.1–5.3.

Если алфавит источника состоит из N символов, то при кодировании очередного символа надо учитывать реализации M предшествующих состояний, т. е. всего N^M возможных вариантов, каждому из которых соответствует своя таблица условных распределений вероятностей (5.16). Это требует N^M таблиц эффективных кодов из N элементов каждая, что эквивалентно массиву данных из N^{M+1} ячеек. Например, если в алфавите 256 символов, то влияние только одного предшествующего состояния ($M = 1$) на очередной переход источника в общем случае потребует наличия таблиц кодов из $256^2 = 65536$ ячеек. Если не учиты-

вать предшествующие состояния при кодировании очередного символа, то полученные коды могут оказаться неэффективными и обладать значительной избыточностью, характеризуемой большим различием значений безусловной энтропии источника (5.2) и условной энтропии (5.19).

Пример 5.12. Пусть последовательность символов сообщения $X(m)$ создается источником, для которого вероятности перехода в новое состояние зависят только от текущего состояния, а совместный закон распределения вероятностей

$$P_{m_0, m_1} = P(X(m)=x_{m_0}, X(m-1)=x_{m_1})$$

задан приведенной таблицей.

		$X(m-1)$	x_1	x_2	x_3
		$X(m)$			
x_1	x_1	1/8	1/8	1/4	
	x_2	1/8	1/8	0	
	x_3	1/4	0	0	

Построить эффективный двоичный код, если предыдущее состояние источника:

- a) не учитывается,
- b) учитывается.

◀ a) Если нет возможности учитывать предыдущее состояние источника, то эффективные коды строим по безусловному распределению вероятностей. Используя методику Шэннона — Фано или Хаффмана, получим среднюю длину битового кода $R = 1,5$ бита, что совпадает с безусловной двоичной энтропией источника.

X	x_1	x_2	x_3
P	1/2	1/4	1/4
Двоичный код	1	01	00

б) Возможны три варианта выбора статистической модели условного распределения вероятностей для символа $X(m)$ по реализации символа $X(m-1)$.

- $X(m-1) = x_1$. Тогда условное распределение для символа $X(m)$ задается приведенной ниже таблицей вероятностей. Частная

условная энтропия (5.18) источника совпадает со средними битовыми затратами полученного кода:

$$H(X(m) | X(m-1) = x_1) = R(X(m) | X(m-1) = x_1) = 1,5 \text{ бита.}$$

$X(m)$	x_1	x_2	x_3
$p(x_k x_1)$	1/4	1/4	1/2
Двоичный код	11	10	0

• $X(m-1) = x_2$. Частная условная энтропия (5.18) совпадает со средними битовыми затратами полученного кода:

$$H(X(m) | X(m-1) = x_2) = R(X(m) | X(m-1) = x_2) = 1 \text{ бит.}$$

$X(m)$	x_1	x_2	x_3
$p(x_k x_2)$	1/2	1/2	0
Двоичный код	1	0	—

• $X(m-1) = x_3$. В этом случае очередное состояние источника полностью определено, кодирование состояния $X(m) = x_1$ не требуется:

$$H(X(m) | X(m-1) = x_3) = R(X(m) | X(m-1) = x_3) = 0 \text{ битов.}$$

В данном примере при использовании информации о предыдущем состоянии источника мы получили минимально возможную среднюю длину битового кода одного символа, равную условной энтропии источника:

$$\begin{aligned} H_1(X) &= H(X(m) | X(m-1)) = \sum_{k=1}^3 p(x_k) H(X(m) | X(m-1) = x_k) = \\ &= R(X(m) | X(m-1)) = \sum_{k=1}^3 p(x_k) R(X(m) | X(m-1) = x_k) = \\ &= 0,5 \times 1,5 + 0,25 \times 1 + 0,25 \times 0 = 1 \text{ бит.} \end{aligned} \quad \blacktriangleright$$

Следующий пример также иллюстрирует общее утверждение о том, что $H_1(X) \leq H_0(X)$, а учет зависимостей между символами сообщения позволяет кодировать их более эффективно.

Пример 5.13. Пусть последовательность символов сообщения $X(m)$ создается источником, для которого вероятности перехода в новое состояние зависят только от текущего состояния, а совместный закон распределения

$$P_{m_0, m_1} = P(X(m) = x_{m_0}, X(m-1) = x_{m_1}),$$

$x_0 = -1, x_1 = 0, x_2 = 1$, задан приведенной таблицей.

$X(m)$	$X(m-1)$	-1	0	1
	-1	3/16	3/16	2/16
	0	4/16	0	0
	1	1/16	1/16	2/16

Оценить минимальные битовые затраты R , которые необходимы для эффективного кодирования символа сообщения, если предыдущее состояние источника:

- a) не учитывается,
- b) учитывается.

◀ Оцените битовые затраты, самостоятельно подсчитав величину безусловной и условной энтропии данного источника. ►

Для кодирования источника с памятью можно использовать и арифметическое кодирование. В этом случае (см. раздел 5.5) разбиение текущего интервала на интервалы вероятностей символов строится по условному распределению вероятностей, соответствующему конкретной предыстории. Так же, как и для источника без памяти, арифметическое кодирование обычно показывает более высокие результаты сжатия, но влечет замедление обработки из-за большей сложности реализации метода.

В общем случае $H_0(X) \geq H_1(X) \geq \dots \geq H_M(X) \geq \dots$, что вытекает из следующей теоремы.

Теорема 5.5. Для условной энтропии (5.19) любого стационарного источника сообщений X с алфавитом $\{x_1, \dots, x_N\}$ $\forall K = 1, 2, \dots$ справедливо неравенство

$$H_{K-1}(X) \geq H_K(X).$$

◀ Схема доказательства та же, что и для теоремы 5.4:

$$\begin{aligned} H_K(X) - H_{K-1}(X) &= \\ &= - \sum_{k_1=1}^N \dots \sum_{k_K=1}^N p(x_{k_1}, \dots, x_{k_K}) \sum_{k_0=1}^N p(x_{k_0} | x_{k_1}, \dots, x_{k_K}) \log p(x_{k_0} | x_{k_1}, \dots, x_{k_K}) + \end{aligned}$$

$$\begin{aligned}
& + \sum_{k_1=1}^N \dots \sum_{k_{K-1}=1}^N p(x_{k_1}, \dots, x_{k_{K-1}}) \sum_{k_0=1}^N p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}}) \log p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}}) \times \\
& \times \underbrace{\sum_{k_K=1}^N p(x_{k_K} | x_{k_0}, \dots, x_{k_{K-1}})}_1 = - \sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K}) \log p(x_{k_0} | x_{k_1}, \dots, x_{k_K}) + \\
& + \sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K}) \log p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}}) = \\
& = \sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K}) \log \frac{p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}})}{p(x_{k_0} | x_{k_1}, \dots, x_{k_K})} \leq \\
& \leq \sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K}) \left(\frac{p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}})}{p(x_{k_0} | x_{k_1}, \dots, x_{k_K})} - 1 \right) = \\
& = \underbrace{\sum_{k_1=1}^N \dots \sum_{k_K=1}^N p(x_{k_1}, \dots, x_{k_K})}_1 \underbrace{\sum_{k_0=1}^N p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}})}_1 - \underbrace{\sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K})}_1 = 0,
\end{aligned}$$

откуда следует утверждение данной теоремы. ▶

Следствие. Несложно показать, что $\forall K H_K(X) \geq 0$. Поэтому для любого стационарного источника сообщений X с алфавитом $\{x_1, \dots, x_N\}$ $\exists \lim_{K \rightarrow \infty} H_K(X) = H_\infty(X)$, где $H_K(X)$ определяется по формуле (5.19).

Таким образом, чем выше используемый для описания произвольного стационарного источника X порядок условной энтропии $H_K(X)$, тем ближе будет оценка априорной неопределенности источника к минимально возможному значению, которое определяется величиной $H_\infty(X)$ (для источника с конечной памятью M выбирать порядок $K > M$, очевидно, не имеет смысла). Однако увеличение порядка энтропии K влечет за собой и увеличение (в геометрической прогрессии) размерности массива совместного распределения вероятностей (5.17), описывающего модель источника с памятью. Построение оптимальных кодов, учитывающих межсимвольные зависимости (память источника), также, к сожалению, обычно упирается на практике в ту же проблему размерности таблиц распределения вероятностей.

■ 5.8. Статистическое моделирование источника

В рассмотренных нами методах кодирования источников сообщений предполагалось, что распределение вероятностей $\{p(x_k)\}_{k=1}^N$ (для источника без памяти) или $\{p(x_0 | x_1, \dots, x_M)\}_{k_0, \dots, k_M=1}^N$ (для источника с памятью) считалось заданным. Вместе с тем, вероятностное описание (статистическая модель) источника известно не всегда. В таких случаях построение модели можно осуществить одновременно с кодированием.

Покажем, как осуществляется статистическое моделирование, на примере источника без памяти. В качестве начального распределения вероятностей источника можно выбрать равномерное: $p_k = 1/N$, $k = 1, \dots, N$, а затем, закодировав первый символ сообщения $x(1) = x_j$, надо внести изменения в модель, повысив вероятность символа x_j : $p_j = 2/(N+1)$, $\forall k \neq j p_k = 1/(N+1)$.

Декодер, начиная работать с тем же распределением вероятностей, что и кодер, после декодирования первого символа сообщения повышает в модели источника соответствующую вероятность так же, как и кодер. Такая адаптация статистических моделей производится после каждого кодирования/декодирования очередного символа. Фактически, в качестве статистической модели при этом используется гистограмма частот, полученная по выборке из обработанных символов.

По мере обработки (кодирования или декодирования) сообщения объем выборки растет, и полученная по накопленной выборке гистограмма все более точно описывает исходно неизвестное распределение вероятностей стационарного источника. Очевидно, что для адекватного описания распределения вероятностей объем выборки должен намного превышать количество ячеек в гистограмме, т. е. число символов N в алфавите источника. Поэтому чем меньше символов в алфавите, тем быстрее «настраивается» модель источника.

Аналогичный подход к построению модели источника по накопленной гистограмме используется и при моделировании источника с памятью. Вместо одной гистограммы используется N^M гистограмм (по количеству предысторий), где M — величина па-

мяти источника, т. е. порядок условной энтропии $H_M(X)$, которая используется для его описания. Общее количество ячеек в этих гистограммах равно N^{M+1} . Для обеспечения точности статистической модели объем выборки, по которой она построена, вновь должен быть намного больше числа ячеек гистограмм, т. е. величины N^{M+1} . По этой причине наличие памяти источника не только вызывает проблемы с размерностью модели, но и существенно замедляет скорость ее адаптации к статистике реального источника.

Отметим, что при использовании метода арифметического кодирования каждый обработанный символ сразу же влечет изменение параметров алгоритма, которые определяются интервалами вероятностей символов (см. раздел 5.5). Длительность процедуры построения эффективных префиксных кодов в методах Хаффмана или Шэннона — Фано реально не позволяет производить изменение таблиц кодов после каждого обработанного символа, и обновление таблиц кодов в этом случае производится «время от времени», т. е. после обработки очередного блока из некоторого числа символов.

■ 5.9. Неопределенность непрерывного источника сообщений. Дифференциальная энтропия

Рассмотрим возможность применения понятия энтропии к случайной величине непрерывного типа X , соответствующей непрерывному источнику информации. Будем считать, что X принимает значения из полуинтервала $x \in [0, 1]$ и описывается заданной непрерывной функцией плотности распределения вероятностей $f(x)$, $x \in [0, 1]$, $\int_0^1 f(x)dx = 1$. Разобъем полуинтервал $[0, 1]$ на N равных частей $\Delta_j = [jq, (j+1)q)$ длины $q = 1/N$, $j = 0, 1, \dots, N - 1$. Случайной величине X , попавшей в полуинтервал Δ_j , поставим в соответствие значение $x_j = q(j + \frac{1}{2})$. Получим из непрерывной величины X равномерно проквантованную с шагом q дискретную случайную величину \hat{X} , энтропия (5.2) которой, в соответствии с теоремой о среднем для интегралов, имеет вид:

$$\begin{aligned} H(\hat{X}) &= -\sum_{j=0}^{N-1} \int_{\Delta_j} f(x)dx \log \int_{\Delta_j} f(x)dx = -\sum_{j=0}^{N-1} f(\theta_j)q \log(f(\theta_j)q) = \\ &= -\sum_{j=0}^{N-1} f(\theta_j)q \log f(\theta_j) - \sum_{j=0}^{N-1} \int_{\Delta_j} f(x)dx \log q = \\ &= -\sum_{j=0}^{N-1} f(\theta_j) \log f(\theta_j)q - \log q, \end{aligned} \quad (5.20)$$

где некоторые точки $\theta_j \in \Delta_j$ таковы, что $\int_{\Delta_j} f(x)dx = f(\theta_j)q$. При этом математическое ожидание квадрата ошибки квантования $\varepsilon^2 = M((X - \hat{X})^2)$:

$$\varepsilon^2 = \sum_{j=0}^{N-1} \int_{\Delta_j} \underbrace{(x - (j + \frac{1}{2})q)^2}_{\geq 0} f(x)dx = \sum_{j=0}^{N-1} f(\xi_j) \int_{\Delta_j} (x - (j + \frac{1}{2})q)^2 dx,$$

где ξ_j — некоторая точка, которая для непрерывной функции $f(x)$ найдется на интервале Δ_j . После интегрирования получим окончательно:

$$\varepsilon^2 = \frac{q^2}{12} \sum_{j=0}^{N-1} f(\xi_j)q. \quad (5.21)$$

Увеличивая число разбиений N , получим для энтропии (5.20) в пределе:

$$\lim_{N \rightarrow \infty} H(\hat{X}) = \lim_{q \rightarrow 0} H(\hat{X}) = -\int_0^1 f(x) \log f(x)dx - \lim_{q \rightarrow 0} \log q = +\infty \quad (5.22)$$

(если интеграл в (5.22) существует, т. е. интегральная сумма в (5.20) сходится). При этом для ошибки (5.21) справедливо:

$$\lim_{q \rightarrow 0} \frac{\varepsilon^2}{q^2/12} = \lim_{q \rightarrow 0} \sum_{j=0}^{N-1} f(\xi_j)q = \int_0^1 f(x)dx = 1.$$

Таким образом, при стремлении дисcreteta (шага) равномерного квантования q к нулю ошибка квантования (5.21) есть бесконечно

малая величина, $\varepsilon^2 \sim q^2/12$ для любой функции $f(x)$, т. е. *независимо от закона распределения*.

Предел (5.22) для энтропии (5.20) не существует — число битов, необходимое для кодирования дискретной величины \hat{X} , с уменьшением шага квантования q непрерывной величины X бесконечно возрастает, $H(\hat{X}) \sim -\log_2 q$. Однако если имеются две случайные величины X и Y , определяемые на полуинтервале $t \in [0, 1]$ функциями плотности распределения вероятностей $f(t)$ и $g(t)$ соответственно, то при бесконечно малом шаге разбиения интервала $[0, 1]$ предел разности для значений энтропии (5.20) дискретных величин \hat{X} и \hat{Y} можно записать в виде:

$$\lim_{q \rightarrow 0} (H(\hat{X}) - H(\hat{Y})) = - \int_0^1 f(t) \log f(t) dt + \int_0^1 g(t) \log g(t) dt \quad (5.23)$$

(если интегралы существуют). При этом ошибки квантования (5.21) для X и Y будут эквивалентными бесконечно малыми величинами: $\varepsilon^2(\hat{X}) \sim \varepsilon^2(\hat{Y}) \sim q^2/12$.

Проведенные рассуждения можно обобщить и на случай, когда функция плотности распределения вероятностей $f(x)$ задана на всей числовой оси и может принимать ненулевые значения не только на полуинтервале $x \in [0, 1]$. Для непрерывной случайной величины X с функцией плотности распределения вероятностей $f(x)$ величина

$$H(X) = - \int_{-\infty}^{+\infty} f(x) \log f(x) dx \quad (5.24)$$

есть, по определению, энтропия или *дифференциальная энтропия*.

Для непрерывного источника информации X дифференциальная энтропия является мерой априорной неопределенности его состояния. Смыслоное содержание понятия «дифференциальная энтропия» следующее. Если для некоторых непрерывных источников информации X, Y выполняется неравенство $H(X) < H(Y)$, то при достаточно малом шаге квантования q ошибки квантования соответствующих случайных величин эквивалентны, $\varepsilon^2(\hat{X}) \sim \varepsilon^2(\hat{Y}) \sim q^2/12$, а для кодирования величины \hat{X} потребу-

ется примерно на $H(Y) - H(X)$ меньшее число битов, чем необходимо для кодирования \hat{Y} — это следует из (5.23). Оценку же собственно битовых затрат по величине энтропии (5.24) получить нельзя.

На непрерывный источник информации можно распространить и понятие условной энтропии. Так, если X, Y — непрерывные источники информации (или случайные величины непрерывного типа), то частной условной энтропией источника Y , соответствующей реализованному состоянию x источника X , назовем величину

$$H(Y|x) = - \int_{-\infty}^{+\infty} f_{Y|X}(x, y) \log f_{Y|X}(x, y) dy,$$

где $f_{Y|X}(x, y)$ — условная плотность распределения вероятностей, определяемая по совместной плотности распределения $f(x, y)$:

$$f_{Y|X}(x, y) = f(x, y) / f_X(x);$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy \quad \text{— безусловная плотность распределения}$$

случайной величины X . Условной энтропией источника Y по X называем величину

$$H(Y|X) = \int_{-\infty}^{+\infty} H(Y|x) f_X(x) dx.$$

Для непрерывных источников остаются в силе соотношения

$$\begin{aligned} H(X, Y) &= - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \log f(x, y) dx dy = \\ &= H(X) + H(Y|X) = H(Y) + H(Y|X) \end{aligned}$$

и $H(X|Y) \leq H(X)$, $H(Y|X) \leq H(Y)$.

■ 5.10. Словарные методы кодирования

Рассмотренные нами ранее методы эффективного кодирования опираются на вероятностную модель сообщения, описываемую простой или сложной цепью Маркова. При этом мы указывали (см. разделы 5.7, 5.8) на значительные сложности, которые возникают на практике при статистическом моделировании и кодировании источников с памятью. В литературных текстах, например, возможно появление полностью повторяющихся слов и фраз из десятков символов, поэтому использование марковской модели для описания источника сообщений и реализация процедуры эффективного кодирования становятся здесь просто невозможными.

Альтернативный подход к кодированию сообщений, для которых характерно появление повторяющихся последовательностей символов, основан на идеях, опубликованных в 1977 г. Лемпелем (Lempel) и Зивом (Ziv). Они предложили первый *словарный* метод кодирования данных, названный по первым буквам фамилий авторов и году его опубликования LZ77.

Метод LZ77

В методе LZ77 по последовательности кодируемых символов перемещается окно (рис. 5.12), состоящее из двух частей: словаря, в который помещаются уже закодированные символы сообщения, и опережающего буфера. Обычно словарь содержит несколько тысяч символов, а буфер — несколько десятков.

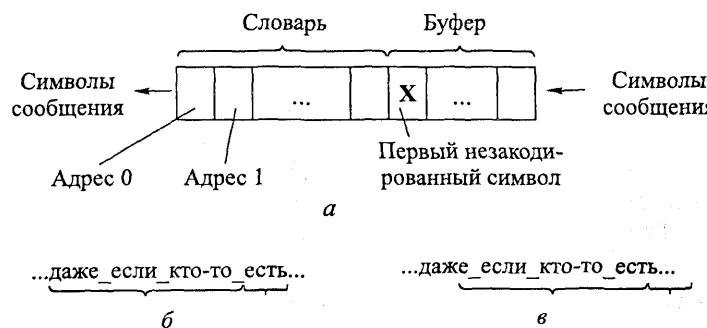


Рис. 5.12

Рассматривая первый незакодированный символ сообщения (первый символ в буфере) как начало символьной строки, кодер ищет в словаре совпадающую с ней строку максимально возможной длины. Если такая строка в словаре найдена, то на выход кодера выдается единица записи, состоящая из трех информационных полей (A, L, c), где A — адрес начала строки в словаре, L — ее длина, c — символ, следующий в буфере сразу за найденной строкой. После этого окно (словарь и буфер) сдвигается по последовательности сообщения на $L + 1$ позицию, с соответствующим сдвигом адресов символов в словаре. Если строка в словаре не была найдена, то единица записи выдается в виде $(0, 0, c)$, где c — первый символ буфера, а окно, соответственно, смешается на одну позицию. Например, если окно кодера находится в положении, приведенном на рис. 5.12, б, то кодер обнаружит в словаре по адресу 5 совпадающую с началом буфера строку из двух символов «ес» и выдаст запись $(5, 2, \langle t \rangle)$. После этого скользящее окно переместится по сообщению на 3 символа (рис. 5.12, в). Следующей записью, которую выдаст кодер, будет тройка $(0, 0, \langle y \rangle)$, а окно переместится затем на одну позицию.

Поиск подстрок в словаре представляет собой вычислительно сложную процедуру. Кроме того, в том случае, когда сообщение содержит малое количество повторяющихся серий символов, кодер LZ77 работает неэффективно. Например, если взять объем словаря в 4096 символов (12-битная адресация), а объем буфера — 16 символов (для кодирования длины подстроки требуется 4 бита), то при 256 символах алфавита сообщения появление символа c , отсутствующего в словаре, потребует выдачу 24 битов для записи $(0, 0, c)$ вместо 8 битов, которые требовались при непосредственном представлении символа c .

Метод LZSS

Метод LZSS является развитием LZ77 и был предложен в 1982 г. Сторером (Storer) и Шиманским (Szimansky). В методе LZSS особое внимание было уделено организации структуры словаря в виде лексикографически упорядоченного дерева для ускорения алгоритма поиска строк (на деталях реализации структуры словаря и алгоритма поиска мы останавливаться не будем). Кодер

LZSS выдает записи двух видов: (Адрес, Длина) — в том случае, если строка из буфера найдена в словаре; или же непосредственно код символа. Для того чтобы различать тип записи, перед ней добавляется один бит-признак (префикс записи). Такой способ кодирования позволяет получить заметный выигрыш в сжатии по сравнению с методом LZ77.

Методы LZ77 и LZSS основаны на предположении, что повторяющиеся последовательности символов находятся в сообщении недалеко друг от друга, т. е. при очередном повторении строки ее еще можно найти в постоянно обновляемом словаре. Может случиться, что повторившаяся строка находится уже вне скользящего по сообщению словаря. Желательно было бы сохранять часто встречавшиеся строки символов вне зависимости от того, как давно они встретились. Естественным выходом здесь кажется увеличение объема скользящего словаря, но это означает как увеличение времени поиска в словаре, так и большие битовые затраты на кодирование адреса строки в словаре, что, наоборот, может привести к ухудшению сжатия данных. То же самое касается и увеличения числа символов в упреждающем буфере (усложнение поиска в словаре, увеличение битовых затрат на кодирование длины строки). На устранение указанных проблем направлен другой метод словарного кодирования — LZ78, который Лемпел и Зив предложили в 1978 г.

Метод LZ78

Если в методе LZ77 словарь представлял собой множество подстрок символьной строки конечной длины — последней части кодируемого сообщения, то в LZ78 словарь $\mathbf{P} = \{P(i)\}_i$ представляет собой потенциально бесконечный массив из строк («фраз») $P(i)$ различной длины, которые встретились при обработке сообщения. В начале работы и у кодера, и у декодера массив словаря состоит из одного элемента — пустой строки. Словарь пополняется в процессе обработки сообщения.

Рассмотрим работу кодера. Считывая очередной незакодированный символ сообщения c из входного потока, кодер присоединяет его к концу текущей строки S посредством операции конкатенации: $S = S + c$ (в начале работы строка S пустая). До тех пор

пока текущая строка соответствует какой-либо фразе из словаря, т. е. $\exists i: S = P(i), P(i) \in \mathbf{P}$, процесс присоединения очередного считанного символа сообщения к строке S продолжается. В какой-то момент присоединение символа дает строку, отсутствующую в словаре, т. е. $S = P(i) + c: S \notin \mathbf{P}, P(i) \in \mathbf{P}$.

Тогда кодер:

1. Выдает в выходной поток запись, содержащую два поля: (i, c) , где i — номер последней найденной фразы $P(i)$ в массиве-словаре $\mathbf{P} = \{P(i)\}_i$, c — символ, присоединение которого дало текущую строку S , отсутствующую в словаре.
2. Строку $S = P(i) + c$ добавляет в словарь в качестве нового (очередного) элемента массива: $\mathbf{P} = \mathbf{P} + S$.
3. Текущую строку S устанавливает в пустую.

Затем кодер продолжает работу по описанному алгоритму, считывая очередной символ сообщения, присоединяя его к текущей строке и проверяя наличие этой строки в словаре.

Разумеется, декодер использует те же правила построения словаря.

Поскольку размер словаря \mathbf{P} в методе LZ78 ничем не ограничен, то при работе алгоритма кодирования может возникнуть переполнение — нехватка памяти для хранения словаря или нехватка производительности вычислительной системы для реализации поиска в словаре. Оригинальный метод LZ78 не определяет, что необходимо делать в случае переполнения словаря. Поэтому метод LZ78 является скорее теоретической конструкцией. Для ее использования на практике необходимо определиться с тем, как избежать переполнения словаря.

Укажем на некоторые возможные пути решения этой проблемы.

1. «Замораживание» словаря. Как только объем словаря достиг максимально возможного значения, он перестает пополняться, становится статическим. Если распределение повторяющихся фраз в сообщении достаточно однородно, то сформированный по начальным данным словарь будет эффективным и для обработки всего сообщения.

2. Сброс словаря в начальное состояние (единственная пустая строка) при переполнении. При этом кодируемое сообщение, по сути дела, разбивается на блоки символов, для каждого из которых в процессе кодирования формируется свой словарь. Такой

подход более предпочтителен, если состав повторяющихся фраз в сообщении меняется от блока к блоку. Как и в LZ77, здесь неявно предполагается, что похожие данные будут группироваться в сообщении близко друг к другу.

3. Когда словарь заполнился, удалить из него некоторые самые старые строки. К сожалению, не существует универсального алгоритма, который определял бы, какие строки и в каком количестве удалять.

Метод LZW

Этот, возможно наиболее популярный на сегодняшний день словарный метод кодирования, представляющий собой модификацию LZ78, был предложен Уэлчем (Welch) в 1984 г.

Основное отличие метода LZW от LZ78 состоит в том, что кодер LZW выдает только ссылки на словарь P , который при инициализации перед началом работы заполняется всеми возможными фразами длиной в один символ, т. е. всеми символами алфавита. Когда в процессе работы кодера присоединение очередного символа c дает текущую строку $S = P(i) + c$, отсутствующую в словаре ($S \notin P$, $P(i) \in P$), кодер LZW обрабатывает эту ситуацию следующим образом:

1. Выдает в выходной поток индекс i последней найденной фразы $P(i)$ из массива-словаря $P = \{P(i)\}_i$.
2. Строку $S = P(i) + c$ запоминает для последующего добавления в словарь: $St = S$.
3. Присваивает текущей строке значение $S = c$.

Строка St будет добавлена в словарь P сразу после того, как в следующий раз кодер выдаст (а декодер прочитает) очередной индекс k некоторой фразы $P(k)$. Это сделано для того, чтобы и кодер, и декодер работали с одинаковыми словарями, пополняя их синхронно. До момента выбора следующей фразы $P(k)$ из словаря декодер не имеет информации о том, какой символ (он является первым в $P(k)$ символом) надо приписать к строке $P(i)$ для формирования фразы St , подлежащей добавлению в словарь.

Так же, как и для LZ78, ключевым для метода LZW является предотвращение переполнения словаря. Для этого используются уже описанные выше подходы. Отметим, что задание правил для «чистки» словаря (удаления фраз) во многом определяет эффективность кодирования данных и по-прежнему остается актуальной задачей, которая оставляет значительный простор для исследований.

Мы рассмотрели общие идеи словарных методов кодирования, не вдаваясь в подробности их реализации. Вместе с тем, способы организации структуры словаря и реализации поиска в этой структуре являются важнейшим звеном алгоритмов сжатия данных, так как в значительной мере определяют их быстродействие. Например, структуру словаря для метода LZW целесообразно представлять не в виде массива строк, а в виде дерева-графа, узлами которого являются символы, а фразы словаря формируются при проходе ветвей графа. Заинтересованного читателя мы отсылаем к дополнительной литературе¹.

В описанных в данном разделе методах кодирования словарь играл роль модели сообщения. Поскольку разные фразы словаря встречаются в сообщениях с разной частотой, то само собой напрашивается использование статистического кодирования для адресов фраз. Данный подход, заключающийся в одновременном использовании идей словарных методов и статистического кодирования, и является тем ядром, на основе которого построены практически все широко распространенные компьютерные программы архивации (arj, zip, rar и др.).

¹ Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. М., 2002

Глава 6

Теоретические основы применения ортогональных преобразований для представления дискретных сигналов

■ 6.1. Корреляция как мера статистической зависимости данных. Преобразование Карунена — Лоэва

Пусть $\mathbf{X} = (X_0, \dots, X_{N-1})^T$ — случайный вектор, состоящий из временных отсчетов дискретного сигнала. Для реальных физических процессов соседние отсчеты вектора обычно имеют близкие значения, т. е. между компонентами вектора имеется статистическая зависимость. В терминах главы 5 дискретный сигнал можно рассматривать как сообщение, созданное непрерывным источником информации с памятью (если отсчеты проквантованы, то дискретным источником с памятью).

Зависимость между случайными величинами X и Y можно характеризовать ковариацией: $\text{cov}(X, Y) = M[(X - m_X)(Y - m_Y)]$, где m_X, m_Y — математические ожидания случайных величин X и Y соответственно. Если $\text{cov}(X, Y) \neq 0$, то говорят, что случайные величины коррелированы. Из коррелированности случайных величин следует их зависимость. Обратное неверно. Например, пусть $X \sim R(-1; 1)$ (равномерное распределение), а $Y = X^2$. Несмотря на наличие функциональной зависимости между случайными величинами X и Y имеем: $\text{cov}(X, Y) = 0$ (убедитесь!), т. е. X и Y некоррелированы.

Однако на практике о наличии или отсутствии статистических зависимостей часто судят именно по результатам изучения корреляционных связей. При этом в качестве меры зависимости случайных величин X, Y удобно использовать *коэффициент корреляции* $\rho(X, Y) = \text{cov}(X, Y) / \sqrt{D_X D_Y}$, где D_X, D_Y — их дисперсии. Коэффициент корреляции может принимать значения из

диапазона $\rho \in [-1; 1]$, чем ближе к единице его абсолютная величина, тем сильнее зависимость между случайными величинами. При $|\rho| = 1$ случайные величины жестко связаны линейной зависимостью вида $Y = aX + b$, и при известных численных значениях параметров a, b реализация одной случайной величины позволяет указать точное значение другой. Таким образом, в первом приближении наличие или отсутствие зависимости между случайными величинами характеризуется их коррелированностью. В данной главе понятия коррелированности и зависимости случайных величин мы отождествим, т. е. ограничим изучение статистических зависимостей рамками корреляционной теории.

В корреляционной теории для описания математической модели вектора данных $\mathbf{X} = (X_0, \dots, X_{N-1})^T$ необходимо задать вектор математических ожиданий компонент $\mathbf{m}_X = (m_{X_0}, \dots, m_{X_{N-1}})^T$ и ковариационную матрицу $\mathbf{K}_X = (\text{cov}(X_k, X_j))_{k,j=0}^{N-1}$. Напомним, что ковариационная матрица всегда симметрическая, положительно определена, а ее диагональные элементы представляют собой дисперсии компонент случайного вектора \mathbf{X} . Считая \mathbf{X}, \mathbf{m}_X векторами-столбцами, ковариационную матрицу можно записать как $\mathbf{K}_X = M[(\mathbf{X} - \mathbf{m}_X)(\mathbf{X} - \mathbf{m}_X)^T]$.

Так как для реальных дискретных сигналов в большинстве случаев соседние отсчеты имеют близкие значения, то корреляция между близко расположенным на временной оси отсчетами положительна, $\rho(X_k, X_{k+j}) > 0$, причем коэффициент корреляции тем больше, чем меньше величина $|j|$. Вследствие наличия межкомпонентных связей в векторе \mathbf{X} независимое покомпонентное статистическое кодирование компонент (отсчетов дискретного сигнала), следующее за их квантованием, порождает избыточные, неэффективные коды (см. раздел 5.7 и пример 5.12).

Учет статистических зависимостей при построении эффективных кодов технически сложен из-за больших размерностей моделей, описывающих совместные распределения вероятностей компонент вектора \mathbf{X} . По этой причине один из методов, предложенных для повышения эффективности кодирования дискретных сигналов, представляет собой *кодирование с предсказанием сигна-*

ла и заключается в следующем. По уже каким-то образом закодированным проквантованным отсчетам дискретного сигнала $\hat{x}_{k-m}, \dots, \hat{x}_{k-1}$ составляется прогноз $p_k = p(\hat{x}_{k-m}, \dots, \hat{x}_{k-1})$ для отсчета x_k в виде экстраполирующего полинома, линейной регрессионной функции и др. Затем квантованию и кодированию подвергается ошибка прогноза $\Delta_k = x_k - p_k$. В простейшем случае используется экстраполирующий полином нулевого порядка; тогда прогноз имеет вид $p_k = x_{k-1}$, т. е. квантованию и кодированию подвергаются первые разности $\Delta_k = x_k - x_{k-1}$ дискретного сигнала — этот метод кодирования называют *дифференциальной импульсно-кодовой модуляцией* (ДИКМ). Естественно ожидать повышения эффективности кодирования при построении прогноза по большему числу предыдущих отсчетов, однако это влечет за собой и усложнение вычислений.

Более общий подход, который используется для повышения эффективности кодирования дискретных сигналов, состоит в предварительной обработке исходных данных с помощью обратного преобразования F , переводящего вектор \mathbf{X} в некоторый вектор $\mathbf{Y} = F\{\mathbf{X}\}$, в котором зависимость между компонентами ослаблена (в рамках корреляционного подхода это означает, что компоненты менее коррелированы). Тогда независимое покомпонентное кодирование вектора \mathbf{Y} , а не вектора \mathbf{X} , становится более обоснованным. Вектор \mathbf{Y} будем называть вектором трансформант (от англ. *transform* — преобразование).

Положим, что для вектора данных \mathbf{X} известны ковариационная матрица \mathbf{K}_x и вектор математических ожиданий компонент $\mathbf{m}_x = M(\mathbf{X}) = (m_{X_0}, \dots, m_{X_{N-1}})^T$. Пусть вектор \mathbf{Y} получен из вектора \mathbf{X} в результате некоторого линейного преобразования, которое определяется матрицей \mathbf{W} : $\mathbf{Y} = \mathbf{WX}$. Наша задача состоит в том, чтобы ослабить (в идеале — исключить) межкомпонентную зависимость данных в векторе \mathbf{Y} . Поскольку статистическую зависимость мы отождествили с коррелированностью, это означает, что необходимо добиться того, чтобы ковариационная матрица вектора \mathbf{Y} приняла диагональный вид. Возможно ли это? Для ответа на данный вопрос найдем сначала выражение для ковариационной матрицы вектора \mathbf{Y} .

В силу линейности операции взятия математического ожидания $\mathbf{m}_y = (m_{y_0}, \dots, m_{y_{N-1}})^T = M(\mathbf{Y}) = M(\mathbf{WX}) = \mathbf{WM}(\mathbf{X}) = \mathbf{Wm}_x$.

Поэтому для ковариационной матрицы имеем:

$$\begin{aligned} \mathbf{K}_y &= M[(\mathbf{Y} - \mathbf{m}_y)(\mathbf{Y} - \mathbf{m}_y)^T] = M[(\mathbf{WX} - \mathbf{Wm}_x)(\mathbf{WX} - \mathbf{Wm}_x)^T] = \\ &= M[\mathbf{W}(\mathbf{X} - \mathbf{m}_x)(\mathbf{X} - \mathbf{m}_x)^T \mathbf{W}^T] = \mathbf{W} \cdot M[(\mathbf{X} - \mathbf{m}_x)(\mathbf{X} - \mathbf{m}_x)^T] \mathbf{W}^T, \end{aligned}$$

откуда

$$\mathbf{K}_y = \mathbf{WK}_x \mathbf{W}^T. \quad (6.1)$$

Напомним два утверждения, известных из линейной алгебры.

1. Пусть матрица \mathbf{A} размерности $n \times n$ обладает линейно независимыми собственными векторами $\mathbf{r}_1, \dots, \mathbf{r}_n$. Тогда матрица $\mathbf{B} = \mathbf{C}^{-1} \mathbf{AC}$ будет иметь диагональный вид, если матрица \mathbf{C} составлена из собственных векторов-столбцов \mathbf{A} : $\mathbf{C} = (\mathbf{r}_1 \mathbf{r}_2 \dots \mathbf{r}_n)$.

2. Если, к тому же, матрица \mathbf{A} — симметрическая, то матрица \mathbf{C} из нормированных ($\|\mathbf{r}_k\|_E = 1$, $k = 1, \dots, n$) собственных векторов — ортогональна и $\mathbf{C}^{-1} = \mathbf{C}^T$.

Из приведенных утверждений с учетом формулы (6.1) следует, что ковариационная матрица \mathbf{K}_y вектора $\mathbf{Y} = \mathbf{WX}$ будет иметь диагональный вид, если матрицу \mathbf{W} составить из транспонированных собственных векторов (векторов-строк) ковариационной матрицы \mathbf{K}_x . При этом матрица \mathbf{W} будет ортогональной в силу симметричности \mathbf{K}_x . Данное (ортогональное!) преобразование, приводящее матрицу \mathbf{K}_y к диагональному виду, называется *преобразованием Карунена — Лоэва*.

Таким образом, по заданной ковариационной матрице вектора \mathbf{X} с помощью ортогонального преобразования Карунена — Лоэва можно получить вектор $\mathbf{Y} = \mathbf{WX}$, компоненты которого полностью не коррелированы. В рамках корреляционной модели мы считаем компоненты вектора \mathbf{Y} независимыми, поэтому становится обоснованной и независимая покомпонентная обработка (квантование и статистическое кодирование) вектора.

Однако возможности практического использования преобразования Карунена — Лоэва существенно ограничиваются следующими факторами. Прежде всего, это преобразование является неуниверсальным: матрица преобразования \mathbf{W} определяется конкретной ковариационной матрицей \mathbf{K}_x . Применяя для обрабатываемого сигнала модель случайного процесса, не всегда можно предполагать его стационарность, т. е. корреляционные зависи-

ности между отсчетами дискретного сигнала непостоянны.

В этих случаях различные выборки из дискретного сигнала (векторы данных \mathbf{X}) могут иметь существенно различающиеся ковариационные матрицы, для которых априори не известны точные значения элементов $\{k_{k,m} = \text{cov}(X_k, X_m)\}_{k,m=0}^{N-1}$. Использование же некоторой «средней» ковариационной матрицы \mathbf{K}_x для построения преобразования Карунена — Лоэва не позволяет говорить о его оптимальности для каждого отдельного вектора \mathbf{X} — очередной выборки из дискретного сигнала. Другая проблема, связанная с использованием преобразования Карунена — Лоэва, заключается в отсутствии быстрых алгоритмов его вычисления.

■ 6.2. Эффективность использования дискретных ортогональных преобразований для кодирования коррелированных данных

Отсутствие быстрых алгоритмов вычисления и зависимость параметров оптимального преобразования Карунена — Лоэва от структуры матрицы \mathbf{K}_x вынуждают использовать на практике другие ортогональные преобразования. При этом общую схему кодирования и декодирования дискретного сигнала можно представить в следующем виде (см. рис. 6.1).

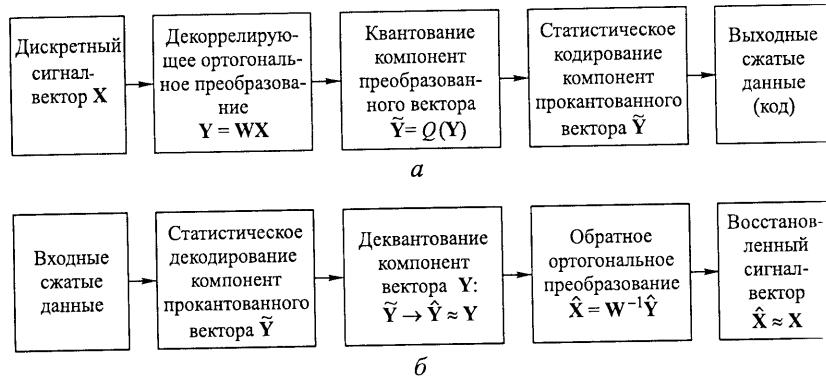


Рис. 6.1. Схема сжатия (а) и восстановления (б) коррелированных данных с использованием ортогонального преобразования

Наличие этапа квантования компонент вектора трансформант \mathbf{Y} (будем также называть этот вектор *спектром*, понимая данный термин обобщенно) вносит определенную ошибку $\|\mathbf{X} - \hat{\mathbf{X}}\|^2 = \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2$ в восстановленные данные — вектор $\hat{\mathbf{X}}$ (см. также пример 3.11), которая может быть снижена до пренебрежимо малого уровня путем квантования спектральных компонент с большим числом уровней квантования. Кодирование спектра \mathbf{Y} , а не исходного вектора \mathbf{X} , позволяет во многих случаях получить существенно меньшие битовые затраты. Критерий, по которому может быть оценена эффективность применения вещественных ортогональных преобразований для целей кодирования, основан на следующих рассуждениях.

Мерой априорной неопределенности для непрерывной случайной величины с функцией плотности распределения вероятностей $f(x)$ является дифференциальная энтропия (5.24). Пусть $f_k(m_k, \sigma_k, x)$ — функция плотности распределения вероятностей для Y_k — k -й компоненты вектора \mathbf{Y} , где m_k — математическое ожидание, σ_k — среднеквадратичное отклонение. Обозначая

$$f_k^0(x) = f_k(0, 1, x) \text{ и учитывая, что } f_k(m_k, \sigma_k, x) = \frac{1}{\sigma_k} f_k^0\left(\frac{x-m_k}{\sigma_k}\right),$$

$$\int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) dx = 1, \text{ среднюю безусловную энтропию одной}$$

компоненты вектора трансформант можно записать следующим образом:

$$\begin{aligned}
 H_{cp} &= -\frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) \log f_k(m_k, \sigma_k, x) dx = \\
 &= -\frac{1}{N} \sum_{k=0}^{N-1} \log \frac{1}{\sigma_k} \int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) dx - \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0\left(\frac{x-m_k}{\sigma_k}\right) \log f_k^0\left(\frac{x-m_k}{\sigma_k}\right) d\left(\frac{x}{\sigma_k}\right) = \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} \log \sigma_k - \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx. \quad (6.2)
 \end{aligned}$$

Чем меньше средняя энтропия (6.2), тем эффективнее будет последующее независимое кодирование компонент вектора \mathbf{Y} . Отметим, что требование ортогональности исследуемого преоб-

разования является принципиальным. Действительно, анализируя среднюю неопределенность компоненты вектора $\mathbf{Y} = \mathbf{W}\mathbf{X}$ по величине средней дифференциальной энтропии (6.2), мы (см. раздел 5.9) фактически предполагаем, что компоненты вектора \mathbf{Y} равномерно проквантованы с шагом квантования $q \rightarrow 0$. При этом квадратичная ошибка (5.21), вносимая в результате квантования в каждую компоненту вектора \mathbf{Y} : $\varepsilon^2 \sim q^2/12$, поэтому суммарная ошибка от квантования его компонент $\mathbf{Y} \xrightarrow{\text{квантование}} \hat{\mathbf{Y}}$ для любого преобразования $\|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 \sim Nq^2/12$. Эта ошибка равна ошибке, внесенной в восстановленный вектор $\hat{\mathbf{X}}$, только для ортогональных преобразований (см. лемму 3.1). Если анализируемые преобразования не являются ортогональными, то для них нельзя говорить об эквивалентности ошибки $\|\mathbf{X} - \hat{\mathbf{X}}\|^2 \sim Nq^2/12$, вносимой скалярным квантованием компонент вектора \mathbf{Y} в восстановленный вектор $\hat{\mathbf{X}}$, и сравнение декоррелирующей эффективности преобразований по критерию (6.2), основанному на дифференциальной энтропии, является некорректным.

Поскольку в рамках рассматриваемой нами корреляционной модели законы распределения компонент вектора \mathbf{X} неизвестны, то и точное определение вида функций плотности распределения $f_k^0(x)$ для компонент вектора \mathbf{Y} также невозможно. Однако компоненты вектора \mathbf{Y} представляют собой определенные взвешенные суммы компонент вектора \mathbf{X} , $y_k = \sum_{j=0}^{N-1} w_{k,j} x_j$, где весовые

коэффициенты представляют собой элементы матрицы преобразования $\mathbf{W} = \{w_{k,j}\}_{k,j=0}^{N-1}$. Поэтому можно допустить возможность применения центральной предельной теоремы и положить, что распределение каждой из компонент вектора \mathbf{Y} подчиняется нормальному закону (отличия имеют место только в параметрах распределения). Такое допущение не является строгим, однако часто применяется на практике, поэтому можно считать, что $\forall k = 0, 1, \dots, N-1$:

$$\int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx \approx \text{const.}$$

Мы будем использовать менее жесткое ограничение, а именно:

$$\frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx \approx \text{const}$$

(для любых преобразований при $N \gg 1$). Тогда среднюю энтропию (6.2) одной компоненты вектора трансформант можно записать в виде:

$$H_{cp} = \frac{1}{2N} \sum_{k=0}^{N-1} \log \sigma_k^2 + C = \frac{1}{2N} \log \prod_{k=0}^{N-1} \sigma_k^2 + C, \quad (6.3)$$

где C — некоторая константа, не зависящая от вида используемого преобразования и его размерности.

Необходимые в выражении (6.3) значения $\sigma_k^2 = \sigma_{y_k}^2$ дисперсий компонент y_k являются диагональными элементами матрицы $\mathbf{K}_Y = \mathbf{W}\mathbf{K}_X\mathbf{W}^T$ и могут быть найдены по формуле (убедитесь!)

$$\sigma_k^2 = \sum_{m=0}^{N-1} w_{k,m} \sum_{j=0}^{N-1} w_{k,j} \text{cov}(x_m, x_j).$$

Для оптимального преобразования Карунена — Лоэва расчет по формуле (6.3) упрощается, так как матрица \mathbf{K}_Y имеет диагональный вид и

$$\prod_{k=0}^{N-1} \sigma_k^2 = \det \mathbf{K}_Y = \det \left(\mathbf{W} \mathbf{K}_X \underbrace{\mathbf{W}^T}_{\mathbf{W}^{-1}} \right) = \det \mathbf{W} \det \mathbf{W}^{-1} \det \mathbf{K}_X = \det \mathbf{K}_X.$$

Поэтому для данного преобразования

$$H_{cp} = \frac{1}{2N} \log \prod_{k=0}^{N-1} \sigma_k^2 + C = \frac{1}{2N} \log \det \mathbf{K}_X + C. \quad (6.4)$$

При использовании понятия дифференциальной энтропии для анализа априорной неопределенности непрерывных случайных величин реальное содержание имеют не сами значения энтропии, а их отличия друг от друга (т. е. разности) (см. раздел 5.9). Рассматривая среднюю безусловную энтропию одного элемента вектора трансформант (6.2) как характеристику декоррелирующих свойств ортогональных преобразований, естественно в каче-

стве «точки отсчета» принять значение (6.2) для оптимального преобразования Карунена — Лоэва и ввести в рассмотрение следующий параметр:

$$\Delta H(\mathbf{W}, \mathbf{K}_x) = H_{\text{cp}}(\mathbf{W}, \mathbf{K}_x) - H_{\text{cp}}(\mathbf{W}_{\text{opt}}, \mathbf{K}_x),$$

или, с учетом формул (6.3), (6.4):

$$\Delta H(\mathbf{W}, \mathbf{K}_x) = \frac{1}{2N} \log \left(\prod_{k=0}^{N-1} \sigma_k^2 / \det \mathbf{K}_x \right). \quad (6.5)$$

Назовем данную величину *средней избыточной энтропией*. Величина (6.5) характеризует избыточность данных в векторе \mathbf{Y} , связанную с неполной декорреляцией компонент (чем больше значение (6.5), тем меньше эффективность декоррелирующего преобразования с матрицей \mathbf{W}).

Можно показать, что величина (6.5) всегда является неотрицательной и принимает нулевое значение для преобразования Карунена — Лоэва. Если использовать в (6.5) двоичное основание логарифма, то физический смысл величины (6.5) можно определить как избыточные битовые затраты, приходящиеся в среднем на одну компоненту вектора \mathbf{Y} при независимом эффективном кодировании компонент после квантования с достаточно малым шагом. Таким образом, чем меньше величина средней избыточной энтропии, тем ближе свойства используемого преобразования к оптимальным.

Лемма 6.1. Для всех ортогональных преобразований след ковариационной матрицы инвариантен:

$$\text{trace } \mathbf{K}_y = \sum_{k=0}^{N-1} \sigma_k^2 = \text{trace } \mathbf{K}_x = \sum_{k=0}^{N-1} \sigma_{x_k}^2.$$

$$\begin{aligned} \blacktriangleleft \quad & \sum_{k=0}^{N-1} \sigma_k^2 = \sum_{k=0}^{N-1} M((Y_k - m_{Y_k})^2) = M \left(\underbrace{(\mathbf{Y} - \mathbf{m}_Y)^T}_{\text{Строка}} \underbrace{(\mathbf{Y} - \mathbf{m}_Y)}_{\text{Столбец}} \right) = \\ & = M \left((\mathbf{X} - \mathbf{m}_x)^T \underbrace{\mathbf{W}^T \mathbf{W}}_{\mathbf{E}} (\mathbf{X} - \mathbf{m}_x) \right) = M((\mathbf{X} - \mathbf{m}_x)^T (\mathbf{X} - \mathbf{m}_x)) = \sum_{k=0}^{N-1} \sigma_{x_k}^2. \end{aligned}$$



Замечание. Из формулы (6.5) следует, что при фиксированной размерности N самый «худший» вариант использования ортогонального преобразования дает максимальное значение параметра

$$P = \prod_{k=0}^{N-1} \sigma_k^2. \quad (6.6)$$

Наложив на дисперсию коэффициентов-трансформант вытекающее из леммы 6.1 ограничение $\sum_{k=0}^{N-1} \sigma_k^2 = \text{const}$, с помощью метода

неопределенных множителей Лагранжа можно убедиться в том, что максимум выражения (6.6) достигается при $\sigma_0^2 = \sigma_1^2 = \dots = \sigma_{N-1}^2$ (удобнее искать максимум от логарифма выражения (6.6)). То есть наихудший случай имеет место тогда, когда суммарная дисперсия (фактически, энергия сигнала) равномерно распределена по всем коэффициентам-трансформантам. Оптимальность же применения преобразования можно понимать как его способность к концентрации большей части энергии исходного дискретного сигнала (вектора \mathbf{X}) в малом количестве коэффициентов-трансформант. В этом смысле задача выбора преобразования принимает вид задачи дискретной L_2 -аппроксимации: найти такое преобразование, чтобы для заданного набора Ω отбрасываемых коэффициентов-трансформант их суммарная дисперсия (энергия) $\sum_{k \in \Omega} \sigma_k^2$ была минимальной. Такой подход к анализу эффективно-

сти применения преобразований для кодирования данных менее универсален (по сравнению с критерием избыточной энтропии), он может дать разные результаты сравнения преобразований при различном количестве отбрасываемых коэффициентов. Тем не менее, подобное изучение свойств преобразований обосновано и описано в литературе.

■ 6.3. ДПФ в вещественной форме. Дискретное преобразование Хартли

При выводе критерия средней избыточной энтропии (6.5) мы предполагали, что ортогональное преобразование, определяемое

матрицей \mathbf{W} , вещественное. Вместе с тем, при использовании ДПФ (3.14), имеющего комплексную природу, для обработки вектора \mathbf{X} из N вещественных компонент в векторе-спектре \mathbf{Y} получим уже N , вообще говоря, комплексных коэффициентов, или $2N$ вещественных чисел. Двойную избыточность преобразованных данных можно устраниć, если воспользоваться свойством 1° ДПФ (см. раздел 3.5). Для коэффициента ДПФ y_k (3.14) введем обозначения: $a_k = \operatorname{Re} y_k$, $b_k = \operatorname{Im} y_k$. Тогда при обработке вещественного вектора \mathbf{X} получим $y_0 = a_0$, и если размерность преобразования N есть четное число, то $y_{N/2} = \bar{y}_{\frac{N-N}{2}} = \bar{y}_{N/2} = a_{N/2}$. Остальные спектральные коэффициенты в общем случае являются комплексными числами, и в силу свойства 1° ДПФ получаем для них

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N} kj} = \underbrace{\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \cos \frac{2\pi i}{N} kj}_{a_k} + i \underbrace{\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \sin \frac{2\pi i}{N} kj}_{b_k} = a_k + ib_k,$$

$$y_{N-k} = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{\frac{2\pi i}{N} kj} = \underbrace{\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \cos \frac{2\pi i}{N} kj}_{a_{N-k}} + i \underbrace{\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \sin \frac{2\pi i}{N} kj}_{b_{N-k}} = a_k - ib_k.$$

Из вектора-спектра \mathbf{Y} , полученного в результате ДПФ вещественного вектора \mathbf{X} , сформируем вектор $\tilde{\mathbf{Y}}$ по следующему правилу:

$$\tilde{y}_k = \begin{cases} a_k = y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \cos \frac{2\pi i}{N} kj & \text{для } k = 0 \text{ или } k = N/2; \\ a_k \sqrt{2} = \sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} x_j \cos \frac{2\pi i}{N} kj & \text{для } 1 \leq k < N/2; \\ -b_k \sqrt{2} = \sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} x_j \sin \frac{2\pi i}{N} kj & \text{для } N/2 < k \leq N-1, \end{cases} \quad (6.7)$$

$k = 0, 1, \dots, N-1$.

Лемма 6.2. Дискретное преобразование вещественного вектора \mathbf{X} в вектор $\tilde{\mathbf{Y}} = \tilde{\mathbf{W}}\mathbf{X}$, определяемое формулой (6.7), является ортогональным.

◀ Пусть N — четное число. Так как $b_k = -b_{N-k}$, $a_k = a_{N-k}$, то

$$\sum_{k=\frac{N}{2}+1}^{N-1} b_k^2 = \sum_{k=\frac{N}{2}+1}^{N-1} b_{N-k}^2 = \sum_{k=1}^{\frac{N}{2}-1} b_k^2, \quad \sum_{k=\frac{N}{2}+1}^{N-1} a_k^2 = \sum_{k=\frac{N}{2}+1}^{N-1} a_{N-k}^2 = \sum_{k=1}^{\frac{N}{2}-1} a_k^2.$$

Для евклидовой нормы вектора $\tilde{\mathbf{Y}}$ имеем:

$$\begin{aligned} \|\tilde{\mathbf{Y}}\|_E^2 &= \sum_{k=0}^{N-1} \tilde{y}_k^2 = |y_0|^2 + |y_{N/2}|^2 + \sum_{k=1}^{\frac{N}{2}-1} 2a_k^2 + \sum_{k=\frac{N}{2}+1}^{N-1} 2b_k^2 = \\ &= |y_0|^2 + \sum_{k=1}^{\frac{N}{2}-1} (a_k^2 + b_k^2) + |y_{N/2}|^2 + \sum_{k=\frac{N}{2}+1}^{N-1} (a_k^2 + b_k^2) = \sum_{k=0}^{N-1} |y_k|^2 = \|\mathbf{Y}\|_E^2. \end{aligned}$$

В силу унитарности ДПФ (см. раздел 3.5) для любого вектора \mathbf{X} имеем $\|\mathbf{Y}\|_E = \|\mathbf{X}\|_E$, поэтому $\|\tilde{\mathbf{Y}}\|_E = \|\mathbf{X}\|_E$, и на основании леммы 3.1 получим, что определяемое формулой (6.7) преобразование является ортогональным, т. е. $\tilde{\mathbf{W}}^T = \tilde{\mathbf{W}}^{-1}$.

Для нечетного числа N схема доказательства аналогична. ►

Преобразование (6.7) назовем ДПФ в вещественной форме, или вещественным ДПФ (ВДПФ). Другим родственным ДПФ (3.14) вещественным преобразованием является *дискретное преобразование Хартли*, которое переводит вектор $\mathbf{f} = (f_0, \dots, f_{N-1})^T$ в вектор $\mathbf{h} = (h_0, \dots, h_{N-1})^T$ по следующей формуле:

$$h_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} f_j \left(\cos \frac{2\pi kj}{N} + \sin \frac{2\pi kj}{N} \right), \quad (6.8)$$

$k = 0, 1, \dots, N-1$. Или в матричном виде:

$$\mathbf{h} = \mathbf{C} \cdot \mathbf{f},$$

$$\text{где } \mathbf{C} = \left\{ c_{k,j} = \frac{1}{\sqrt{N}} \left(\cos \frac{2\pi kj}{N} + \sin \frac{2\pi kj}{N} \right) \right\}_{k,j=0}^{N-1}.$$

Очевидно, что $\mathbf{C} = \mathbf{C}^T$.

Лемма 6.3. Дискретное преобразование Хартли (6.8) — ортогонально.

◀ Необходимо показать (см. лемму 3.1), что преобразование любого вектора \mathbf{f} по формуле (6.8) сохраняет евклидову норму: $\|\mathbf{f}\|_E = \|\mathbf{h}\|_E$.

Пусть $\mathbf{Y} = \mathbf{W}\mathbf{f}$ — ДПФ (3.14) вещественного вектора \mathbf{f} . Из формул (6.8) и (3.14) видно, что $\mathbf{h} = \mathbf{C}\cdot\mathbf{f} = \operatorname{Re} \mathbf{Y} - \operatorname{Im} \mathbf{Y}$ (отметим, что на этом свойстве основан алгоритм вычисления преобразования Хартли при помощи БПФ). То есть $h_k = a_k - b_k$ ($k = 0, \dots, N-1$), где, как и выше, $a_k = \operatorname{Re} y_k$, $b_k = \operatorname{Im} y_k$. Для нормы вектора \mathbf{h} имеем:

$$\|\mathbf{h}\|_E^2 = \sum_{k=0}^{N-1} h_k^2 = \sum_{k=0}^{N-1} (a_k - b_k)^2 = \sum_{k=0}^{N-1} \underbrace{\left(a_k^2 + b_k^2\right)}_{|y_k|^2} - 2 \sum_{k=0}^{N-1} a_k b_k = \|\mathbf{Y}\|_E^2 - 2 \sum_{k=0}^{N-1} a_k b_k.$$

Положим, что N — четное число. Так как, с учетом свойства 1° ДПФ (см. раздел 3.5), $b_k = -b_{N-k}$, $a_k = a_{N-k}$, причем $b_0 = b_{N/2} = 0$, то

$$\begin{aligned} \sum_{k=0}^{N-1} a_k b_k &= \sum_{k=1}^{\frac{N}{2}-1} a_k b_k + \sum_{k=\frac{N}{2}+1}^{N-1} a_k b_k = \sum_{k=1}^{\frac{N}{2}-1} a_k b_k - \sum_{k=\frac{N}{2}+1}^{N-1} a_{(N-k)} b_{(N-k)} = \\ &= \sum_{k=1}^{\frac{N}{2}-1} a_k b_k - \sum_{k=1}^{\frac{N}{2}-1} a_k b_k = 0, \end{aligned}$$

поэтому $\forall \mathbf{Y} \|\mathbf{h}\|_E^2 = \|\mathbf{Y}\|_E^2$.

В силу унитарности ДПФ $\|\mathbf{f}\|_E = \|\mathbf{Y}\|_E$, поэтому $\|\mathbf{f}\|_E = \|\mathbf{h}\|_E$. ▶

Следствие. Матрица обратного преобразования Хартли совпадает с матрицей прямого преобразования: $\mathbf{C} = \mathbf{C}^{-1}$.

■ 6.4. Дискретный марковский процесс первого порядка. Дискретное косинусное преобразование (ДКП)

Одной из важнейших корреляционных моделей для дискретных сигналов является модель дискретного марковского процесса

первого порядка¹. Данный процесс является стационарным, т. е. его параметры не зависят от момента наблюдения. Стационарность (в широком смысле) дискретного процесса означает, что отсчеты сигнала $\{X_k\}$ имеют одинаковые математические ожидания $M(X_k) = m_X$ и дисперсии $D(X_k) = \sigma_X^2$, а корреляция отсчетов X_k, X_j зависит только от расстояния между моментами наблюдения дискретного процесса, т. е. от величины $|k-j|$. Так, для марковского процесса первого порядка коэффициент корреляции $r(X_k, X_j) = \rho^{|k-j|}$, где ρ — коэффициент корреляции соседних отсчетов сигнала. Таким образом, если случайный вектор $\mathbf{X} = (X_0, \dots, X_{N-1})^T$ является сигналом, описываемым моделью дискретного марковского процесса первого порядка, то он имеет следующую ковариационную матрицу:

$$\mathbf{K}_x(\rho) = \sigma_X^2 \begin{pmatrix} 1 & \rho & \dots & \rho^{N-1} \\ \rho & 1 & \dots & \rho^{N-2} \\ \dots & \dots & \dots & \dots \\ \rho^{N-1} & \rho^{N-2} & \dots & 1 \end{pmatrix}. \quad (6.9)$$

Без потери общности, можно считать далее, что $D(X_k) = \sigma_X^2 = 1$. Модель (6.9) представляет особый интерес, так как часто используется на практике для описания дискретных сигналов, причем обычно параметр ρ имеет значение, близкое к единице.

Среди преобразований, имеющих быстрые алгоритмы вычислений (т. е. реализуемых $\sim N \log N$ вычислительными операциями вместо $\sim N^2$), наибольшую эффективность для кодирования сигнала, описываемого моделью марковского процесса первого порядка, показывает *дискретное косинусное преобразование* (ДКП), которое определяется формулой

$$y_k = \sqrt{\frac{2}{N}} c(k) \sum_{j=0}^{N-1} x_j \cos\left(\frac{\pi k}{N}(j + \frac{1}{2})\right), \quad k = 0, 1, \dots, N-1, \quad (6.10)$$

где $c(k) = \begin{cases} 1/\sqrt{2} & \text{при } k=0; \\ 1 & \text{при } k \neq 0. \end{cases}$

¹ Определение дискретного марковского процесса см. в разделе 5.7.

ДКП (6.10) можно записать в матричном виде, $\mathbf{Y} = \mathbf{WX}$, где

$$\mathbf{W} = \left\{ w_{k,j} = \sqrt{2/N} c(k) \cos\left(\frac{\pi k}{N}(j + \frac{1}{2})\right) \right\}_{k,j=0}^{N-1}.$$

Выясним «происхождение» формулы (6.10), для чего покажем, как связано ДКП с непрерывным преобразованием Фурье по тригонометрической системе.

Если для некоторого сигнала $f(t)$, заданного на отрезке наблюдения $t \in [0, T]$, составить ряд Фурье (2.3) с коэффициентами, найденными по формуле (2.4), то свойства сходимости полученного ряда будут определяться характером функции $\tilde{f}(t)$ — продолжения с периодом T на всю числовую ось функции $f(t)$ (рис. 6.2).

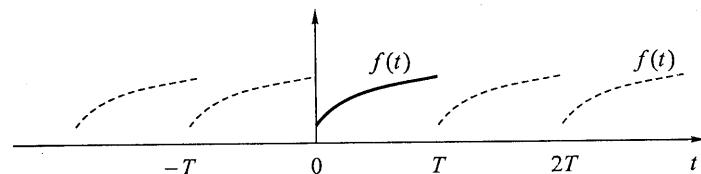


Рис. 6.2

Как отмечалось в разделе 3.5, с точностью до нормировочного множителя ДПФ представляет собой приближенное вычисление первых коэффициентов (2.4) ряда Фурье (2.3) по формуле прямоугольников:

$$c_k = \frac{1}{T} \int_0^T f(t) \exp\left(-i \frac{2\pi k}{T} t\right) dt \approx \frac{1}{T} \sum_{j=0}^{N-1} f(j\Delta t) \exp\left(-i \frac{2\pi k}{T} j\Delta t\right) \Delta t = \frac{1}{N} \sum_{j=0}^{N-1} x_j e^{-i \frac{2\pi kj}{N}},$$

где $N = T/\Delta t$, $x_j = f(j\Delta t)$. Это позволяет предположить, что чем выше скорость сходимости ряда (2.3) (т. е. чем быстрее убывают абсолютные величины коэффициентов Фурье), тем сильнее выражена локализация энергии вблизи первых коэффициентов вектора трансформант ДПФ.

Если функция $\tilde{f}(t)$ — непрерывная, то соответствующий ряд (2.3) будет сходиться быстрее по сравнению со случаем разрывной функции $\tilde{f}(t)$. Чтобы гарантировать непрерывность $\tilde{f}(t)$,

сначала дополним функцию $f(t)$ четным образом на отрезке $t \in [-T, 0]$, а затем полученную функцию (заданную уже на области $t \in [-T, T]$) продолжим на всю числовую ось с периодом $2T$ (рис. 6.3).

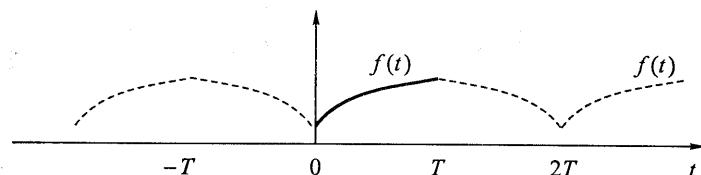


Рис. 6.3

Найдем коэффициенты (2.4) ряда Фурье (2.3) функции $\tilde{f}(t)$ периода $2T$. В силу четности $\tilde{f}(t)$ имеем:

$$c_k = \frac{1}{2T} \int_{-T}^T \tilde{f}(t) e^{-i \frac{2\pi k}{2T} t} dt = \frac{1}{2T} \int_{-T}^T \tilde{f}(t) \cos\left(\frac{\pi k}{T} t\right) dt - \underbrace{\frac{i}{2T} \int_{-T}^T \tilde{f}(t) \sin\left(\frac{\pi k}{T} t\right) dt}_0,$$

$$c_k = \frac{1}{T} \int_0^T \tilde{f}(t) \cos\left(\frac{\pi k}{T} t\right) dt. \quad (6.11)$$

Таким образом, коэффициенты Фурье являются вещественными и $c_k = c_{-k}$.

Подсчитаем приближенно значение интеграла (6.11) по квадратурной формуле прямоугольников, выбрав узлы в точках $t_j = (j + 0,5)\Delta t$, $j = 0, 1, \dots, N - 1$, где $N = T/\Delta t$:

$$c_k = \frac{1}{T} \int_0^T f(t) \cos\left(\frac{\pi k}{T} t\right) dt \approx \frac{1}{N} \sum_{j=0}^{N-1} x_j \cos\left(\frac{\pi k}{N} (j + 0,5)\right).$$

Здесь обозначено $x_j = f((j + 0,5)\Delta t)$. Полученное выражение с точностью до нормировочного множителя совпадает с формулой ДКП (6.10). Оказывается, выбор узлов квадратурной формулы прямоугольников в точках $t_j = (j + 0,5)\Delta t$ обеспечивает ортогональность ДКП.

Лемма 6.4. Матрица \mathbf{W} ДКП (6.10) ортогональна.

◀ Необходимо показать, что матрица $\mathbf{U} = \mathbf{WW}^T$ является единичной. В соответствии с определением (6.10) получаем для произвольного элемента матрицы \mathbf{U} :

$$\begin{aligned} u_{k,m} &= \sum_{j=0}^{N-1} w_{k,j} w_{m,j} = \frac{2}{N} c(k)c(m) \sum_{j=0}^{N-1} \cos\left(\frac{\pi k}{N}\left(j + \frac{1}{2}\right)\right) \cos\left(\frac{\pi m}{N}\left(j + \frac{1}{2}\right)\right) = \\ &= \frac{1}{N} c(k)c(m) \left(\underbrace{\sum_{j=0}^{N-1} \cos\left(\frac{\pi(k+m)}{N}\left(j + \frac{1}{2}\right)\right)}_{s(k+m)} + \underbrace{\sum_{j=0}^{N-1} \cos\left(\frac{\pi(k-m)}{N}\left(j + \frac{1}{2}\right)\right)}_{s(k-m)} \right), \\ &\quad k = 0, \dots, N-1; m = 0, \dots, N-1. \end{aligned}$$

Найдем значение суммы $s(l) = \sum_{j=0}^{N-1} \cos\left(\frac{\pi l}{N}\left(j + \frac{1}{2}\right)\right) = s(-l)$ для $l = 0, \dots, 2N-2$. Очевидно, что $s(0) = N$, а при $l = 1, \dots, 2N-2$ получаем по формуле для суммы геометрической прогрессии:

$$\begin{aligned} s(l) &= \operatorname{Re} \left(\sum_{j=0}^{N-1} e^{i \frac{\pi l}{N} (j + \frac{1}{2})} \right) = \operatorname{Re} \left(e^{i \frac{\pi l}{2N}} \frac{e^{i \frac{\pi l}{2}} - 1}{e^{i \frac{\pi l}{N}} - 1} \right) = \operatorname{Re} \left(e^{i \frac{\pi l}{2}} \frac{e^{i \frac{\pi l}{2}} - e^{-i \frac{\pi l}{2}}}{e^{i \frac{\pi l}{2N}} - e^{-i \frac{\pi l}{2N}}} \right) = \\ &= \operatorname{Re} \left(e^{i \frac{\pi l}{2}} \frac{\sin \frac{\pi l}{2}}{\sin \frac{\pi l}{2N}} \right) = \frac{\cos \frac{\pi l}{2} \sin \frac{\pi l}{2}}{\sin \frac{\pi l}{2N}} = \frac{\sin \pi l}{2 \sin \frac{\pi l}{2N}} = 0. \end{aligned}$$

Поэтому

$$u_{k,m} = \frac{c(k)c(m)}{N} (s(k+m) + s(k-m)) = \begin{cases} 0 & \text{при } k \neq m; \\ 2c(k)c(m) & \text{при } k = m = 0; \\ c(k)c(m) & \text{при } k = m \neq 0. \end{cases}$$

Отсюда, поскольку $c(j) = \begin{cases} 1/\sqrt{2} & \text{при } j = 0; \\ 1 & \text{при } j \neq 0, \end{cases}$ получим:

$$u_{k,m} = \begin{cases} 0 & \text{при } k \neq m; \\ 1 & \text{при } k = m. \end{cases}$$

Следствие. Обратное для ДКП (6.10) преобразование (обратное ДКП, или ОДКП) определяется формулой

$$x_j = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} c(k) y_k \cos\left(\frac{\pi k}{N}\left(j + \frac{1}{2}\right)\right), \quad j = 0, 1, \dots, N-1, \quad (6.12)$$

где $c(k) = \begin{cases} 1/\sqrt{2} & \text{при } k = 0; \\ 1 & \text{при } k \neq 0. \end{cases}$

Для иллюстрации декоррелирующих свойств ДКП приведем расчетные значения средней избыточной энтропии (6.5), соответствующие некоторым дискретным преобразованиям при обработке вектора \mathbf{X} , который состоит из 16 компонент и описывается моделью марковского процесса первого порядка с ковариационной матрицей (6.9). На рис. 6.4 отражены результаты сравнения следующих ортогональных преобразований: ВДПФ, ДКП, дискретные преобразования Уолша (ДПУ) и Хаара (ДПХ) (см. разделы 3.8, 3.9). Дискретное преобразование Хартли не отражено на графиках, поскольку оно показывает практически те же характеристики, что и ВДПФ (значение средней избыточной энтропии (6.5) для преобразования Хартли по сравнению с ВДПФ меньше в

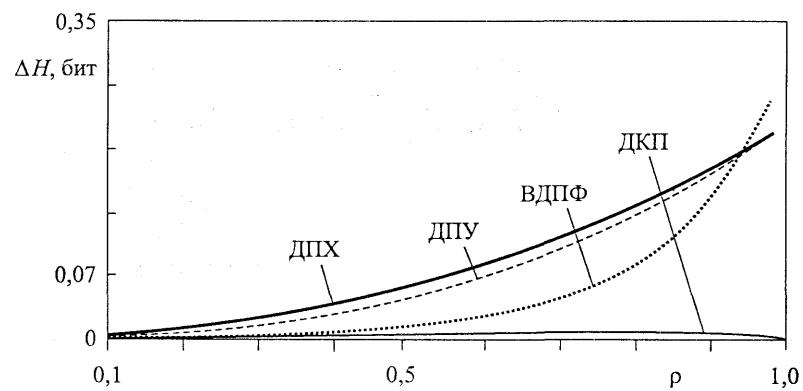


Рис. 6.4. Графики средней избыточной энтропии (6.5) при использовании некоторых ортогональных преобразований для обработки 16-компонентного вектора данных с ковариационной матрицей (6.9)

среднем на 2 %). Как видно из рис. 6.4, наиболее близкими характеристиками к оптимальному преобразованию Карунена — Лоэва обладает ДКП, причем явное преимущество ДКП над другими преобразованиями наблюдается для различных значений параметра ρ . Это объясняет широкое практическое применение ДКП во многих, в том числе стандартизованных, алгоритмах сжатия дискретных сигналов.

■ 6.5. Компрессия изображений на основе двумерного ДКП

Широкий класс линейных дискретных преобразований, используемых для цифровой обработки изображений, составляют двумерные (т. е. переводящие матрицу в матрицу) *сепарабельные* преобразования, которые получают из одномерных (переводящих вектор в вектор) следующим образом.

Определение. Пусть преобразование матрицы $\mathbf{X} = \{x_{j,m}\}$, ($j = 0, \dots, N - 1; m = 0, \dots, M - 1$) в матрицу \mathbf{Y} той же размерности производится по формуле

$$y_{k,l} = \sum_{j=0}^{N-1} w_{k,j}^{(N)} \sum_{m=0}^{M-1} w_{l,m}^{(M)} x_{j,m}, \quad (6.13)$$

$$k = 0, \dots, N - 1, \quad l = 0, \dots, M - 1,$$

где $\mathbf{W}_K = \{w_{j,n}^{(K)}\}_{j,n=0}^{K-1}$ — матрица одномерного преобразования

вида $\mathbf{U} = \mathbf{WV}$: $u_j = \sum_{n=0}^{K-1} w_{j,n}^{(K)} v_n$, $j = 0, \dots, K - 1$. Тогда двумерное преобразование (6.13) называется дискретным сепарабельным преобразованием.

Из формулы (6.13) видно, что вычисление сепарабельного преобразования сводится к выполнению одномерных дискретных преобразований размерности M вдоль строк $j = 0, \dots, N - 1$ матрицы \mathbf{X} :

$$z_{j,l} = \sum_{m=0}^{M-1} x_{j,m} w_{l,m}^{(M)}, \quad l = 0, 1, \dots, M - 1$$

(или $\mathbf{Z} = \mathbf{XW}_M^T$) и к последующим одномерным преобразованиям размерности N вдоль столбцов $l = 0, \dots, M - 1$ полученной матрицы $\mathbf{Z} = \{z_{j,l}\}$:

$$y_{k,l} = \sum_{j=0}^{N-1} w_{k,j}^{(N)} z_{j,l}, \quad k = 0, 1, \dots, N - 1$$

(или $\mathbf{Y} = \mathbf{W}_N \mathbf{Z}$). Причем сначала можно выполнить одномерные преобразования по столбцам матрицы, а потом по строкам, так как порядок суммирования в выражении (6.13) можно поменять. Из сказанного следует, что обратить двумерное сепарабельное преобразование можно, выполнив одномерные обратные преобразования сначала вдоль столбцов, а потом вдоль строк матрицы \mathbf{Y} (или наоборот).

Запишем сепарабельное преобразование (6.13) в эквивалентном матричном виде: $\mathbf{Y} = \mathbf{W}_N \mathbf{XW}_M^T$. Тогда если соответствующее одномерное преобразование является ортогональным, то обратное к (6.13) преобразование записывается как $\mathbf{X} = \mathbf{W}_N^T \mathbf{Y} \mathbf{W}_M$, а обращение формулы (6.13) принимает вид (убедитесь!):

$$x_{j,m} = \sum_{k=0}^{N-1} w_{k,j}^{(N)} \sum_{l=0}^{M-1} w_{l,m}^{(M)} y_{k,l}, \quad (6.14)$$

$$j = 0, \dots, N - 1, \quad m = 0, \dots, M - 1.$$

В соответствии с формулами (6.13), (6.14) и данными ранее определениями одномерных преобразований вводятся двумерные дискретные преобразования Фурье, Уолша, Хаара, Хартли. В частности, прямое и обратное двумерные ДКП (см. (6.10), (6.12)) задаются соответственно следующими формулами:

$$y_{k,l} = \frac{2}{\sqrt{MN}} c(k)c(l) \sum_{j=0}^{N-1} \cos\left(\frac{\pi k}{N}\left(j + \frac{1}{2}\right)\right) \sum_{m=0}^{M-1} x_{j,m} \cos\left(\frac{\pi l}{M}\left(m + \frac{1}{2}\right)\right),$$

$$k = 0, 1, \dots, N - 1, \quad l = 0, 1, \dots, M - 1, \quad (6.15)$$

$$x_{j,m} = \frac{2}{\sqrt{MN}} \sum_{k=0}^{N-1} c(k) \cos\left(\frac{\pi k}{N}\left(j + \frac{1}{2}\right)\right) \sum_{l=0}^{M-1} c(l) \cdot y_{k,l} \cos\left(\frac{\pi l}{M}\left(m + \frac{1}{2}\right)\right),$$

$$j = 0, 1, \dots, N-1, \quad m = 0, 1, \dots, M-1,$$
(6.16)

где $c(n) = \begin{cases} 1/\sqrt{2} & \text{при } n = 0; \\ 1 & \text{при } n \neq 0. \end{cases}$

Именно двумерное ДКП было положено в основу появившегося в начале 90-х годов прошлого века международного стандарта JPEG (Joint Photographic Experts Group), который определяет методы эффективного представления фотографических изображений. Метод JPEG использует общую схему сжатия данных с потерями, приведенную ранее на рис. 6.1. Несмотря на то что с 2001 года официально действует расширенная спецификация стандарта «JPEG-2000», которая допускает использование альтернативной схемы кодирования на основе вейвлет-преобразований, кодирование с применением ДКП остается основным и наиболее распространенным методом, реализованным в многочисленных программных продуктах и различной аппаратуре.

Рассмотрим кратко основной вариант схемы JPEG на примере обработки *полутонового* цифрового изображения, которое можно представить в виде матрицы, элементами которой являются значения яркости точек раstra. Точки изображения (элементы матрицы) называют также *пикселями*.

Поскольку фотографические изображения в большинстве случаев представляют собой нестационарные двумерные сигналы, для адаптации способа обработки под локально изменяющиеся характеристики изображения последнее обрабатывается небольшими блоками размерности 8×8 пикселей. При такой размерности, с одной стороны, декорелирующие свойства ДКП уже хорошо выражены, а с другой стороны, в столь малой области статистические характеристики изображения можно считать локально стационарными, с достаточной точностью подчиняющимися марковской модели сигнала (см. (6.9)). Фрагменты обрабатываются последовательно (очередной обрабатываемый блок является соседним к предыдущему), общая схема обработки (рис. 6.1) принимает вид, приведенный на рис. 6.5, и включает в себя четыре основных шага.

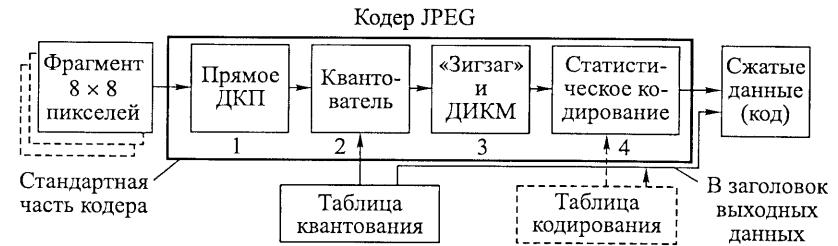


Рис. 6.5. Схема сжатия изображений по стандарту JPEG

Шаг 1. Вычисление двумерного ДКП (6.15) блока размерности 8×8 .

Шаг 2. Скалярное квантование спектра ДКП: каждый элемент спектра равномерно квантуется с округлением до целого:

$$\tilde{y}_{k,l} = \text{round}(y_{k,l}/q_{k,l}).$$

Целые числа $q_{k,l}$ из диапазона $\{0, \dots, 255\}$ определяются по таблице (матрице) квантования \mathbf{Q} (размера 8×8). Матрица квантования стандартом не регламентируется (но есть рекомендуемый JPEG набор таблиц) и передается в заголовке выходных данных. Диапазон возможных значений, которые могут принимать элементы спектра, после выполнения процедуры квантования существенно уменьшается, появляется большое число нулей.

Шаг 3. Коэффициент спектра $\tilde{y}_{0,0}$ (постоянная составляющая яркости фрагмента изображения) подвергается дифференциальной импульсно-кодовой модуляции (ДИКМ), т. е. $y_{0,0}$ заменяется значением $\Delta\tilde{y}_{0,0} = \tilde{y}_{0,0}(j) - \tilde{y}_{0,0}(j-1)$, где j — номер текущего фрагмента 8×8 . ДИКМ целесообразно применять по причине того, что средние значения яркости соседних фрагментов, как правило, имеют близкие значения. Полученный после ДИКМ коэффициента $\tilde{y}_{0,0}$ двумерный набор из 64 чисел зигзагообразно (см. рис. 6.6) считывается в одномерную последовательность, образуя промежуточный поток данных, который является входным для 4-го шага. Смысл такого зигзагообразного считывания заключается в том, чтобы сгруппировать основную часть нулевых элементов спектра, которые оказываются сосредоточенными в конце 64-элементной последовательности промежуточного потока данных.

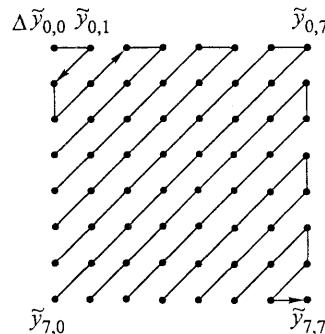


Рис. 6.6

Шаг 4.1. Удаление нулей. Нули удаляются путем преобразования промежуточного потока данных в два выходных потока. Запись в выходные потоки данных производится при появлении во входном промежуточном потоке ненулевого элемента $S \neq 0$. Первый выходной поток состоит из байтов-ключей L : в полубайте L_1 записывается число нулей, встретившихся перед элементом S после предыдущего появления ненулевого элемента¹, а в полубайте L_2 записывается число B , которое равно числу двоичных разрядов, необходимых для представления элемента S (см. таблицу). Во второй поток заносится B бит двоичного кода элемента S . Таким образом, первый поток данных — битовый, второй — битовый. После записи последнего ненулевого элемента из «змейки» в битовый поток заносится специальный символ признака конца блока ненулевых коэффициентов — EOB (end of block). Битовый поток данных дополнительно не обрабатывается, а поток байтов-ключей кодируется статистически.

Диапазон значений числа $S \neq 0$	B
$\{-1, 1\}$	1
$\{-3, -2\} \cup \{2, 3\}$	2
$\{-7, -6, -5, -4\} \cup \{4, 5, 6, 7\}$	3
$\{-15, -14, \dots, -8\} \cup \{8, 9, \dots, 15\}$	4
$\{-31, \dots, -16\} \cup \{16, \dots, 31\}$	5
\cdots	\cdots
$\{-2^{15} + 1, \dots, -2^{14}\} \cup \{2^{14}, \dots, 2^{15} - 1\}$	15

¹ Фактически, используется вариант метода КДС, описанного в разделе 5.4.

Шаг 4.2. Статистическое кодирование байтов-ключей. Для этой цели используется алгоритм Хаффмана. Таблицы кодов Хаффмана должен строить сам разработчик системы сжатия, поэтому таблицы необходимо помещать в заголовок выходных данных (имеются также рекомендуемые JPEG коды).

Второй шаг схемы сжатия JPEG вносит потерю информации. Задание таблицы квантования Q является тем инструментом, с помощью которого выбирается компромисс между сжатием информации и достоверностью ее представления: чем грубее квантование (чем больше значения шага равномерного квантования q_k), тем большая ошибка возникнет в изображении после восстановления, но тем выше степень сжатия данных.

Каждому шагу в схеме сжатия (рис. 6.5) соответствует шаг в схеме восстановления изображения (рис. 6.7), на котором выполняются обратные операции. За исключением квантования, все операции являются обратимыми. «Деквантование» представляет собой масштабирование $\hat{y}_{k,l} = \tilde{y}_{k,l} \cdot q_{k,l}$, задаваемое той же матрицей квантования Q .

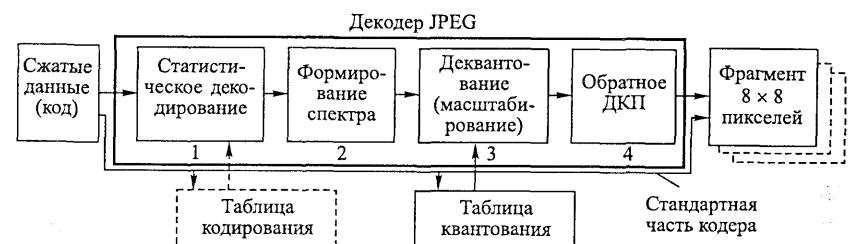


Рис. 6.7. Схема восстановления изображений по стандарту JPEG

При обработке цветных изображений каждый пиксель из цветового трехкомпонентного пространства (Красный — Red, Зеленый — Green, Синий — Blue) — сокращенно (R, G, B) — сначала переводится в представление

$$\begin{pmatrix} Y \\ Cr \\ Cb \end{pmatrix} = \begin{pmatrix} 0,2990 & 0,5870 & 0,1140 \\ 0,5000 & -0,4187 & -0,0813 \\ -0,1687 & -0,3313 & 0,5000 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix},$$

где компонента Y называется *яркостной*, а компоненты Cr , Cb — *цветоразностными* составляющими. По сравнению с яркостной составляющей, чувствительность человеческого глаза к компонентам Cr , Cb значительно ниже, что учитывается при построении алгоритмов компрессии цветных изображений. Для цветоразностных компонент в методе JPEG применяется дополнительное прореживание массивов пикселей и затем используется более грубая обработка по той же общей схеме, которая была рассмотрена выше для яркостной компоненты. При восстановлении изображений используется обратное преобразование цветового пространства:

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1,4020 & 0 \\ 1 & -0,7141 & -0,3441 \\ 1 & 0 & 1,772 \end{pmatrix} \begin{pmatrix} Y \\ Cr - 128 \\ Cb - 128 \end{pmatrix}.$$

■ 6.6. Аппроксимационный подход к выбору преобразований для кодирования дискретных сигналов. Частотная трактовка

Ранее отмечалось, что целью использования преобразований является перевод коррелированного набора данных \mathbf{X} в менее коррелированный (в идеале — независимый) набор данных — вектор \mathbf{Y} . Именно с этой точки зрения и рассматривался вопрос эффективности применения преобразований в схемах компрессии дискретных сигналов, попутно была отмечена (см. заключительное замечание в разделе 6.2) важность неравномерного распределения энергии по компонентам преобразованного вектора \mathbf{Y} , когда основная часть энергии дискретного сигнала концентрируется в малом количестве компонент-трансформант. На этом принципе будем основывать дальнейшие рассуждения.

Рассмотренные ранее дискретные ортогональные преобразования (Фурье, Уолша, Хаара, ДКП) были получены как результат дискретизации разложений в ряд по некоторым ортогональным системам функций $\{\phi_k(t)\}_{k=0}^{\infty}$ (для упрощения изложения будем считать, что система ортонормированна, $\|\phi_k(t)\| = \sqrt{\langle \phi_k(t), \phi_k(t) \rangle} = 1$), когда

исходный аналоговый сигнал представляется в виде

$$x(t) = \sum_{k=0}^{\infty} c_k \phi_k(t), \text{ а дискретный — в виде}$$

$$x_j = x(j\Delta t) = \sum_{k=0}^{N-1} y_k \phi_k(j\Delta t),$$

где $c_k = \langle x(t), \phi_k(t) \rangle \approx \frac{y_k}{\sqrt{N}}$ (см. раздел 6.4 и замечания к формулам (3.12), (3.24), (3.25)). При этом для энергии сигналов в силу ортонормированности систем и ортогональности преобразований имеем

в непрерывном случае $\langle x(t), x(t) \rangle = \sum_{k=0}^{\infty} c_k^2$ (см. теорему 1.8), а в

дискретном случае $\sum_{k=0}^{N-1} |x_k|^2 = \sum_{k=0}^{N-1} |y_k|^2$ (см. лемму 3.1).

Для концентрации основной энергии сигнала в малом числе коэффициентов обобщенного спектра необходимо построить такой полный ортогональный базис $\{\phi_k(t)\}_{k=0}^{\infty}$, который позволяет достаточно точно представить непрерывный сигнал $x(t) = \sum_{k=0}^{\infty} c_k \phi_k(t)$ в виде конечной суммы с малым числом слагаемых:

$x(t) \approx x_n(t) = \sum_{k=0}^n c_k \phi_k(t)$. При этом квадрат ошибки, связанный с отбрасыванием членов ряда, равен $\|x(t) - x_n(t)\|_E^2 = \sum_{k=n+1}^{\infty} c_k^2$,

т. е. искомый базис $\{\phi_k(t)\}_{k=0}^{\infty}$ в идеале должен обеспечивать минимальную ошибку $\|x(t) - x_n(t)\|_E$, для заданного класса сигналов $x(t)$ и фиксированного числа n . Сформулируем требования к базису $\{\phi_k(t)\}_{k=0}^{\infty}$ в терминах частотных характеристик.

Частотный подход к исследованию реальных сигналов основан на том, что для сигналов $f(t)$, представимых в виде интеграла Фурье (2.5), $f(t) = \int_{-\infty}^{+\infty} S(v) e^{2\pi i v t} dv$, необходимым условием существования данного несобственного интеграла является

$\lim_{|v| \rightarrow \infty} S(v)e^{2\pi i v t} = 0$, т. е. «затухание» спектральной плотности с

увеличением частоты¹: $\lim_{v \rightarrow \infty} S(v) = 0$. Основная энергия сигнала во

многих случаях локализована в области низких частот, так как высокочастотные составляющие спектра вносят ощутимый вклад в общую энергию лишь для быстро изменяющихся во времени сигналов. Поэтому на практике частотный спектр сигнала можно с той или иной точностью считать ограниченным некоторой полосой $v \in (-F; F)$.

Из приведенных рассуждений следует, что обычно, выполнив ДПФ дискретного сигнала, можно без особого ущерба для точности последующего восстановления отбросить (положить равными нулю) высокочастотные коэффициенты дискретного спектра (3.14) $\{y_k\}_{k=0}^{N-1}$, т. е. те из них, номера индексов которых близки к $N/2$ (см. пример 3.11).

Помимо тригонометрической системы, для представления сигналов в виде ряда $x(t) = \sum_{k=0}^{\infty} c_k \phi_k(t)$ можно использовать и дру-

гие функциональные базисы $\{\phi_k(t)\}_{k=0}^{\infty}$. Основной вопрос при этом заключается в следующем: какую базисную систему функций использовать предпочтительнее? Приведенные выше рассуждения о природе спектров реальных сигналов позволяют предположить, что частотные спектры используемых базисных функций $\{\phi_k(t)\}_{k=0}^{\infty}$ должны различаться, образуя выраженные низкочастотные и высокочастотные области локализации энергии в спектральной области. Тогда заранее можно предположить, что коэффициенты Фурье, соответствующие «высокочастотным» базисным функци-

ям, будут вносить малый вклад в разложение $x(t) = \sum_{k=0}^{\infty} c_k \phi_k(t)$.

Рассмотренные нами в главе 1 системы Уолша, Хаара обладают такими общими свойствами: с увеличением номера базисной

¹ При разложении функций в тригонометрический ряд (2.1) или (2.3) соответствующее свойство коэффициентов Фурье $\{c_k\}$ принимает вид: $\lim_{k \rightarrow \infty} c_k = 0$.

функции локализация ее энергии в частотной области смещается в сторону высоких частот.

Исторически для представления сигнала на конечном интервале наблюдения длительности T все начиналось с тригонометрической системы (или базиса комплексных экспонент) (см. (2.3)):

$$\left\{ \phi_k(t) = e^{i \frac{2\pi}{T} k t} \right\}_{k=-\infty}^{\infty}. \quad (6.17)$$

Тригонометрический базис является наилучшим выбором для представления информации (исходного аналогового сигнала или последовательности отсчетов дискретного сигнала) во многих практических задачах. Этот базис оказывается в определенном смысле оптимальным при работе со стационарными случайными процессами: известно, что коэффициенты Фурье в этом случае представляют собой некоррелированные величины, а это как раз то свойство обрабатываемых данных, к которому мы стремимся при использовании корреляционной модели дискретного сигнала.

Однако, как отмечалось в разделе 6.1, в силу возможного неизменства статистических свойств дискретного сигнала его описание через ковариационную матрицу не всегда адекватно: реальные, интересующие нас сигналы во многих случаях являются нестационарными случайными процессами, имеют «особенности» во временной области («срывы» стационарности). Например, для цифровых изображений (двумерных массивов яркостей точек-пикселей) такие особенности проявляются вблизи контуров объектов, а стационарный характер данных наблюдается в фоновых областях и областях текстур. Для представления этих особенностей тригонометрический базис неприемлем и вызывает хорошо известный эффект Гиббса, связанный с распространением влияния ошибки при отбрасывании части спектральных коэффициентов на всю область тригонометрического разложения сигнала. Приведем иллюстрацию данного эффекта для дискретного случая.

Пример 6.1 (дискретный аналог эффекта Гиббса). Пусть дискретный сигнал $\mathbf{X} = (1, 0, 0, \dots, 0)$ представляет собой N -компонентный вектор. Найти ошибку представления сигнала при отбрасывании какой-либо спектральной составляющей.

◀ Данный дискретный сигнал является наиболее ярким примером «особенности» — дискретным аналогом δ -функции. Для его спектра, который находится с помощью ДПФ (3.14), получаем $\forall k = 0, \dots, N - 1$:

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N} kj} = \frac{1}{\sqrt{N}}.$$

Отбрасывание (обнуление) спектральной составляющей y_m вносит такую ошибку $\mathbf{X} - \hat{\mathbf{X}}$, что

$$\varepsilon_j = x_j - \hat{x}_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} y_k e^{\frac{2\pi i}{N} kj} - \frac{1}{\sqrt{N}} \sum_{\substack{k=0, \\ k \neq m}}^{N-1} y_k e^{\frac{2\pi i}{N} kj} = \frac{1}{N} e^{\frac{2\pi i}{N} mj},$$

$j = 0, \dots, N - 1$. Следовательно, вектор ошибки при отбрасывании произвольной спектральной компоненты y_m представляет собой гармоническое колебание, т. е. влияние ошибки распространяется на весь вектор дискретного сигнала. ►

Совершенно аналогичные проблемы с распространением ошибки возникают, если вместо тригонометрического использовать другие базисы, имеющие неограниченный носитель во временной области. Действительно, представив сигнал в виде ряда

Фурье: $x(t) = \sum_{k=0}^{\infty} c_k \varphi_k(t)$ и отбросив коэффициент ряда с номером

m , получаем ошибку $\varepsilon(t) = c_m \varphi_m(t)$, локализация которой во временной области полностью определяется базисной функцией $\varphi_m(t)$. В итоге, для эффективного представления сигналов к функциональному базису $\{\varphi_k(t)\}_{k=0}^{\infty}$ должны предъявляться достаточно противоречивые требования, а именно:

- локализация (в идеале — ограниченность носителя) базисных функций в частотной области (поскольку базисные функции, имеющие высокочастотную локализацию, вносят малый вклад в разложение большинства реальных сигналов, и соответствующие им коэффициенты Фурье перед кодированием можно квантовать более грубо или же вообще отбросить, заменив при восстановлении нулевыми значениями);

- локализация (в идеале — ограниченность носителя) базисных функций во временной области для возможности эффективного представления локальных особенностей сигнала, когда высокочастотные составляющие разложения значимы и не должны быть опущены.

Как отмечалось в разделе 2.3, ни для какой функции нельзя добиться высокой локализации энергии одновременно и в частотной, и во временной областях. Поэтому, как следует из принципа неопределенности представления сигнала на плоскости времени×частота, приведенные требования к искомому базису $\{\varphi_k(t)\}_{k=0}^{\infty}$ находятся в противоречии.

■ 6.7. Время-частотный анализ. Оконное преобразование Фурье

Для преодоления трудностей, связанных с необходимостью использования функциональных базисов, имеющих хорошую локализацию как во временной, так и в частотной областях, в 40-х годах прошлого века для анализа сигналов Габором было предложено *оконное преобразование Фурье* (ОПФ). Предполагается, что исследуемый сигнал является *квазистационарным*, т. е. в пределах временного интервала некоторой длительности T его можно считать стационарным, и для каждого такого интервала в отдельности спектр находится по формуле

$$S(v, n) = \int_{-\infty}^{+\infty} p(t - n\tau) f(t) e^{-2\pi i vt} dt, \quad (6.18)$$

где n — номер интервала наблюдения, $p(t)$ — некоторая вещественная и четная функция «окна», осуществляющая выбор интервала для частотного анализа сигнала, τ — определенным образом выбранный параметр. Выражение (6.18) называют *оконным преобразованием Фурье*.

В простейшем случае $p(t) = \Pi_{\tau/2}(t) = \begin{cases} 1 & \text{при } t \in [-\tau/2; \tau/2]; \\ 0 & \text{при } t \notin [-\tau/2; \tau/2] \end{cases}$

$\tau = T$ («окно» — прямоугольное, интервалы наблюдения не перекрываются), тогда формула (6.18) примет вид:

$$S(v, n) = \int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} f(t) e^{-2\pi i vt} dt.$$

Взяв отсчеты спектральной плотности в точках $v_k = k/T$ и обозначив $c_{k,n} = S(v_k, n)/T$, функцию $f(t)$ тогда можно представить (см. формулы (2.3), (2.4)) в виде:

$$f(t) = \sum_{n=-\infty}^{\infty} \Pi_{T/2}(t - nT) \sum_{k=-\infty}^{\infty} c_{k,n} e^{\frac{i2\pi}{T} kt}.$$

В общем случае интервалы наблюдения сигнала в ОПФ могут перекрываться, а весовая функция $p(t)$ может иметь неограниченный носитель, например, в случае гауссова импульса $p(t) = C e^{-\beta t^2}$. Общей идеей ОПФ остается использование вместо нелокализованного во временной и абсолютно локализованного в частотной области базиса (6.17) нового базиса

$$\left\{ \varphi_{k,n}(t) = p(t - n\tau) e^{\frac{i2\pi}{T} kt} \right\}_{n,k=-\infty}^{\infty} \quad (6.19)$$

для представления сигнала в виде:

$$f(t) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{k,n} \varphi_{k,n}(t) = \sum_{n=-\infty}^{\infty} p(t - n\tau) \sum_{k=-\infty}^{\infty} c_{k,n} e^{\frac{i2\pi}{T} kt}. \quad (6.20)$$

При надлежащем выборе весовой функции $p(t)$ и величин T, τ можно построить ортонормированный базис, тогда

$$c_{k,n} = \langle f, \varphi_{k,n} \rangle = \int_{-\infty}^{\infty} f(t) p(t - n\tau) e^{-\frac{i2\pi}{T} kt} dt.$$

В случае вещественного характера функции-окна $p(t)$ последнее соотношение устанавливает следующую связь коэффициентов ряда Фурье (6.20) по системе (6.19) и ОПФ (6.18):

$$c_{k,n} = S\left(\frac{k}{T}, n\right) = \int_{-\infty}^{+\infty} p(t - n\tau) f(t) e^{-\frac{i2\pi}{T} kt} dt.$$

Пример 6.2. Найти выражение для спектральной плотности функции $\varphi_{k,n}(t) = p(t - n\tau) e^{\frac{i2\pi}{T} kt}$ по заданной спектральной плотности $S_p(v) = \Phi\{p(t)\}$.

◀ На основании свойства 4° преобразования Фурье (см. раздел 2.1) для спектра сдвинутой оконной функции справедливо равенство

$$S_{p_n}(v) = \Phi\{\varphi_{0,n}(t) = p(t - n\tau)\} = S_p(v) e^{-2\pi i n \tau v}.$$

Используя свойство 5° преобразования Фурье (см. раздел 2.1), окончательно получим:

$$S_{\varphi_{k,n}}(v) = \Phi\{\varphi_{k,n}(t)\} = S_{p_n}\left(v - \frac{k}{T}\right) = S_p\left(v - \frac{k}{T}\right) e^{-2\pi i n \tau \left(v - \frac{k}{T}\right)}. \quad ▶$$

Как следует из свойства 5° преобразования Фурье, умножение оконной функции $p(t - n\tau)$ на комплексное гармоническое колебание $e^{\frac{i2\pi}{T} kt}$ в частотной области соответствует сдвигу функции спектральной плотности, при котором нулевое значение аргумента (частоты) переходит в точку $v = k/T$. Форма графика амплитудного спектра сохраняется, поскольку $|S_{\varphi_{k,n}}(v)| = |S_p\left(v - \frac{k}{T}\right)|$. Поэтому свойства локализованности базисных функций (6.19) полностью определяются выбором весовой функции $p(t)$ не только для временной, но и для частотной областей. Обозначив $\Delta t, \Delta v$ — длины интервалов, характеризующих локализацию весовой функции $p(t)$ во временной и в частотной областях соответственно¹, можно схематично изобразить следующее покрытие плоскости время \times частота базисными функциями $\{\varphi_{k,n}(t)\}_{k,n=-\infty}^{\infty}$ (реально соседние ячейки могут частично перекрываться) (рис. 6.8).

¹ То, каким образом могут выбираться данные величины, обсуждалось в разделе 2.3.

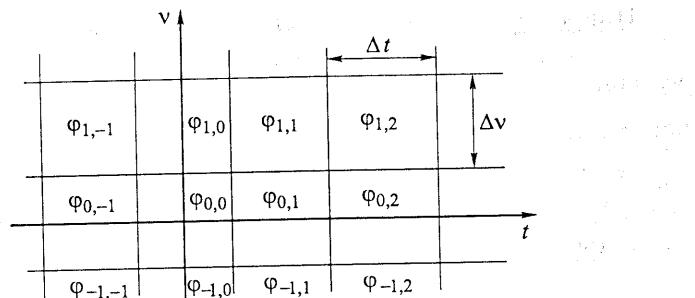


Рис. 6.8. Локализация на плоскости время×частота базисных функций, используемых для разложения (6.20)

Чем меньше значения величин Δt , Δv , тем выше *разрешение* базисных функций во временной и частотной областях соответственно. Разрешение базисных функций в представлении сигнала (6.20) ограничено принципом неопределенности (см. теорему 2.4). Если понимать величины Δt и Δv в смысле эффективных значений, то наибольшим разрешением будет обладать базис, для которого весовая функция представляет собой гауссов импульс (см. пример 2.3). В частности, выбрав весовую функцию

$$p_\alpha(t) = \frac{1}{2\sqrt{\pi}\alpha} e^{-\frac{t^2}{4\alpha}},$$

получим в (6.18) частный случай ОПФ —

преобразование Габора. Изменяя параметр α , можно менять соотношение между параметрами Δt , Δv .

Отметим, что рассмотренный в разделе 6.5 метод сжатия цифровых изображений JPEG на основе ДКП также использует идеи ОПФ, адаптированные для двумерного дискретного случая: окном (весовой функцией) является квадрат 8×8 пикселей, а вместо преобразования Фурье используется его модификация — косинусное преобразование.

Несмотря на использование базиса, имеющего как временную, так и частотную локализацию базисных функций, ОПФ обладает недостатком, вытекающим из общих свойств широкого класса сигналов, представляющих практический интерес. Природа многих реальных сигналов такова, что временной интервал стационарности (квазипостоянства характеристик амплитуды и фазы) для низкочастотных составляющих больше, а для высокочастотных — меньше. То есть для эффективного представления сигнала высокочастотные базисные функции, характерно проявляющиеся при спектральном разложении в областях резких изменений сигнала (в выраженных «особенностях»), должны обладать более высокой временной локализацией по сравнению с низкочастотными базисными функциями. Поскольку повышение локализации базисных функций во временной области возможно лишь за счет снижения локализации в частотной области, то разбиение плоскости время \times частота для конструируемого базиса должно соответствовать общим идеям, схематично отображенными на рис. 6.9. Такими свойствами переменного времена-частотного разрешения обладают функциональные базисы *вейвлетов*, к рассмотрению которых мы перейдем в следующей главе.

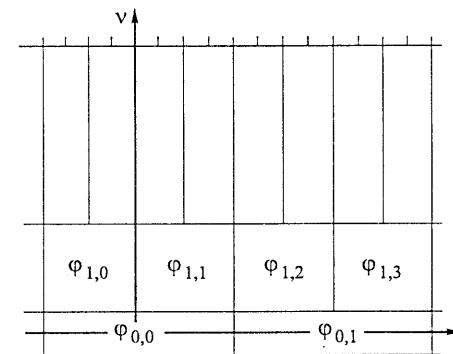


Рис. 6.9. Локализация базисных функций с переменным времена-частотным разрешением

Глава 7

Вейвлет-преобразования и их приложения для обработки дискретных сигналов

В данной главе будет рассматриваться гильбертово пространство $L_2(\mathbb{R})$ (см. пример 1.7) с нормой (1.2), индуцированной скалярным произведением

$$\langle f, g \rangle = \int_{-\infty}^{+\infty} f(t)g(t)dt.$$

■ 7.1. Кратно-масштабный анализ

Определение. Последовательность подпространств

$$\{V_m\} \subset L_2(\mathbb{R}), m \in \mathbb{Z}$$

образует *кратно-масштабный анализ* (КМА), если она обладает следующими свойствами:

1°. Подпространства вложены, $\forall m \in \mathbb{Z}: V_m \subset V_{m+1}$.

2°. Если функция $f(x) \in V_m$, то $f(2x) \in V_{m+1}$, и наоборот, т.е.

$$\forall m \in \mathbb{Z}: f(x) \in V_m \Leftrightarrow f(2x) \in V_{m+1}.$$

3°. Найдется некоторая функция $\phi(x) \in V_0$, целочисленные сдвиги которой $\{\phi(x-n)\}_{n \in \mathbb{Z}} \subset V_0$ образуют ортонормированный базис подпространства V_0 . Функция $\phi(x)$ называется *масштабирующей*.

4°. Единственный общий для всех подпространств элемент — нулевой: $\bigcap_{m \in \mathbb{Z}} V_m = \{0\}$.

5°. Замыкание множества подпространств является пространством $L_2(\mathbb{R})$: $\overline{\bigcup_{m \in \mathbb{Z}} V_m} = L_2(\mathbb{R})$.

Замечание. С учетом свойства 1° кажется возможным записать свойство 5° в следующем виде: $\lim_{m \rightarrow +\infty} V_m = L_2(\mathbb{R})$, используя некоторое понятие предела последовательности подпространств. Однако из-за отсутствия определения такого предела мы используем ту формулировку свойства 5°, которая приведена выше.

Лемма 7.1. Пусть для некоторого КМА выполнены условия 1°—3°. Тогда система функций $\{\phi_{m,n}(x) = 2^{m/2}\phi(2^m x - n)\}_{n \in \mathbb{Z}}$, где масштабирующая функция $\phi(x) \in V_0$, образует ортонормированный базис в подпространстве V_m .

◀ Действительно, $\forall f(x) \in V_m$ в силу свойства 2° имеем $f(2^{-m}x) \in V_0$. Поэтому на основании свойства 3° существует разложение: $f(2^{-m}x) = \sum_{n=-\infty}^{\infty} c_n \phi(x-n)$. т. е. $f(x) = \sum_{n=-\infty}^{\infty} c_n \phi(2^m x - n)$ и система $\{\phi(2^m x - n)\}_{n \in \mathbb{Z}}$ образует базис в подпространстве V_m . В силу ортонормированности системы $\{\phi(x-n)\}_{n \in \mathbb{Z}} \subset V_0$ можно записать:

$$\begin{aligned} & \int_{-\infty}^{\infty} \phi(2^m x - n) \phi(2^m x - k) dx = \\ & = \frac{1}{2^m} \int_{-\infty}^{\infty} \phi(x-n) \phi(x-k) dx = \begin{cases} 1/2^m & \text{при } k=n; \\ 0 & \text{при } k \neq n. \end{cases} \end{aligned}$$

Тогда $\{\phi_{m,n}(x) = 2^{m/2}\phi(2^m x - n)\}_{n \in \mathbb{Z}}$ — ортонормированный базис в V_m . ►

Поскольку $V_0 \subset V_1$, масштабирующую функцию

$$\phi(x) = \phi_{0,0}(x) \in V_0$$

можно представить в виде разложения по базису $\{\phi_{1,n}(x) = \sqrt{2}\phi(2x - n)\}_{n \in \mathbb{Z}} \subset V_1$:

$$\phi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} h_n \phi(2x - n), \quad (7.1)$$

где $h_n = \langle \phi, \phi_{1,n} \rangle$. Уравнение (7.1) называется *масштабирующим* и может быть записано также в более общем виде следующим образом:

$$\phi_{m,0}(x) = \sum_{n \in \mathbb{Z}} h_n \phi_{m+1,n}(x),$$

причем коэффициенты Фурье остаются теми же:

$$\begin{aligned} h_n &= \langle \phi_{m,0}, \phi_{m+1,n} \rangle = 2^{m+\frac{1}{2}} \int_{-\infty}^{\infty} \phi(2^m x) \phi(2^{m+1} x - n) dx = \\ &= \sqrt{2} \int_{-\infty}^{\infty} \phi(x) \phi(2x - n) dx = \langle \phi, \phi_{1,n} \rangle. \end{aligned}$$

Свойство 5° КМА означает, что множество $\Omega = \bigcup_{m \in \mathbb{Z}} V_m$ *плотно* в $L_2(\mathbb{R})$ и $\forall f \in L_2(\mathbb{R}), \forall \varepsilon > 0 \exists f_m \in V_m: \|f - f_m\| < \varepsilon$, т. е. произвольную функцию $f(x) \in L_2(\mathbb{R})$ можно с любой точностью описать элементом $f_m(x)$ из некоторого подпространства V_m . Для рассматриваемого нами гильбертова пространства $L_2(\mathbb{R})$ с этой целью естественно в качестве $f_m(x)$ выбирать в подпространстве V_m элемент наилучшего приближения для $f(x)$. Тогда из свойства 1° следует, что $\|f - f_m\| \geq \|f - f_{m+1}\| \geq \dots$, т. е. с ростом индекса m подпространства V_m , из которого выбирается элемент наилучшего приближения $f_m(x)$, точность представления функции $f(x)$ повышается, а с уменьшением m — понижается (рис. 7.1).

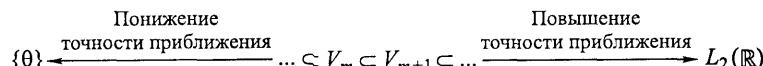


Рис. 7.1

Обозначив через A_m оператор проектирования элемента $f \in L_2(\mathbb{R})$ на подпространство V_m , $f_m = A_m(f)$, в базисе подпространства V_m можем записать:

$$f_m(x) = A_m(f) = \sum_{n \in \mathbb{Z}} a_{m,n} \phi_{m,n}(x),$$

где коэффициенты Фурье $a_{m,n} = a_{m,n}(f) = \langle f, \phi_{m,n} \rangle$.

Ортогональное дополнение подпространства V_m до V_{m+1} обозначим W_m :

$$V_m \oplus W_m = V_{m+1}, \quad V_m \perp W_m.$$

Используя данное рекуррентное соотношение далее, получаем $\forall N > m$

$$V_N = \underbrace{V_m \oplus W_m}_{V_{m+1}} \oplus W_{m+1} \oplus \dots \oplus W_{N-2} \oplus W_{N-1}.$$

Далее по индукции для $m \rightarrow -\infty$ получим, что $V_N = \bigoplus_{m=-\infty}^{N-1} W_m$, откуда на основании свойства 5° КМА следует:

$$L_2(\mathbb{R}) = \overline{\bigoplus_{m \in \mathbb{Z}} W_m}. \quad (7.2)$$

Одно из основополагающих утверждений КМА, которое мы примем без доказательства, состоит в том, что для масштабирующей функции (7.1) найдется такая функция $\psi(x) \in W_0$, называемая *материнским вейвлетом*, что множество функций

$$\{\psi_{m,n}(x) = 2^{m/2} \psi(2^m x - n)\}_{n \in \mathbb{Z}} \subset W_m$$

образует ортонормированный базис в подпространстве W_m . При этом подпространства вейвлетов $\{W_j\}_{j \in \mathbb{Z}}$ наследуют масштабирующее свойство, аналогичное свойству 2°, а именно:

$$\forall m \in \mathbb{Z}: f(x) \in W_m \Leftrightarrow f(2x) \in W_{m+1}.$$

В силу (7.2) произвольную функцию $f(x) \in L_2(\mathbb{R})$ можно представить в виде разложения по ортогональному базису вейвлетов:

$$f(x) = \sum_{m \in \mathbb{Z}} \sum_{\substack{n \in \mathbb{Z} \\ y_m(x) \in W_m}} c_{m,n} \psi_{m,n}(x),$$

где $c_{m,n}(f) = \langle f, \psi_{m,n} \rangle$. Так как $W_0 \subset V_1$, то материнский вейвлет также можно представить через базис

$$\{\phi_{1,n}(x) = \sqrt{2}\phi(2x-n)\}_{n \in \mathbb{Z}} \subset V_1:$$

$$\psi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} g_n \phi(2x-n). \quad (7.3)$$

Уравнение (7.3) называется *масштабирующим для вейвлетов* и по аналогии с (7.1) может быть записано в более общем виде следующим образом:

$$\psi_{m,0}(x) = \sum_{n \in \mathbb{Z}} g_n \phi_{m+1,n}(x),$$

$$\text{где } g_n = \langle \psi_{m,0}, \phi_{m+1,n} \rangle = \langle \psi, \phi_{1,n} \rangle.$$

Позднее (в разделе 7.5) покажем, что коэффициенты Фурье в масштабирующих уравнениях (7.1) и (7.3) удовлетворяют одному из соотношений, $\forall n \in \mathbb{Z}: g_n = (-1)^{1-n} h_{1-n-2k}$, либо $g_n = (-1)^{-n} h_{1-n-2k}$, где k — целое число, которое можно выбрать произвольно. Поэтому для того чтобы задать КМА, достаточно знать один из наборов коэффициентов Фурье, $\{h_n\}$ или $\{g_n\}$. Функция $\phi(x)$, как мы увидим позднее, может быть построена по коэффициентам $\{h_n\}$.

Особый интерес для практики цифровой обработки сигналов представляют случаи, когда КМА определяется масштабирующей функцией $\phi(x)$, которая имеет ограниченный носитель. Тогда в представлении масштабирующей функции через ее сжатые сдвиги сумма (7.1) будет конечной, и, следовательно, конечная сумма из того же числа слагаемых будет получена в уравнении (7.3).

Пример 7.1 (вейвлет Хаара). Масштабирующей функцией в этом случае является $\phi(x) = \begin{cases} 1 & \text{при } x \in [0, 1]; \\ 0 & \text{при } x \notin [0, 1]. \end{cases}$ Очевидно, что масштабирующее уравнение (7.1) принимает вид: $\phi(x) = \phi(2x) + \phi(2x-1)$,

откуда $h_0 = h_1 = 1/\sqrt{2}; \forall k \notin \{0, 1\} h_k = 0$. Тогда коэффициенты для уравнения (7.3) можно выбрать так: $g_0 = h_1 = 1/\sqrt{2}$, $g_1 = -h_0 = -1/\sqrt{2}$, $g_k = 0 \quad \forall k \notin \{0, 1\}$, и масштабирующее уравнение для вейвлетов принимает вид: $\psi(x) = \phi(2x) - \phi(2x-1)$. Постройте самостоятельно графики масштабирующей функции $\phi(x) = \phi_{0,0}(x)$, материнского вейвлета $\psi(x) = \psi_{0,0}(x)$, вейвлетов $\psi_{1,0}(x), \psi_{1,1}(x)$ и сравните их с графиками функций системы Хаара из примера 1.12. Что можно сказать о свойствах функций $f(x) \in V_m$?

◀ Базисом подпространства V_m являются (см. лемму 7.1) функции $\{\phi_{m,n}(x)\}_{n \in \mathbb{Z}}$, которые в данном случае кусочно постоянны:

$$\phi_{m,n}(x) = \begin{cases} 2^{m/2} & \text{при } x \in \Delta_n^m; \\ 0 & \text{при } x \notin \Delta_n^m, \end{cases} \quad \text{где } \Delta_n^m = \left[\frac{n}{2^m}, \frac{n+1}{2^m} \right].$$

Поскольку любая функция $f(x)$ из подпространства V_m должна быть представима в виде линейной комбинации базисных функций $\{\phi_{m,n}(x)\}_{n \in \mathbb{Z}} \subset V_m$, то все функции $f(x) \in V_m$ также должны быть кусочно постоянными: $f(x) = \text{const}$ при $x \in \Delta_n^m$. ►

Укажем на время-частотные свойства функций из подпространств $V_j, W_j, j \in \mathbb{Z}$. Если Δv_0 — понимаемая в каком-либо смысле ширина частотной полосы спектров масштабирующих функций $S_{0,n}(v) = \Phi\{\phi_{0,n}(x)\}$ (см. раздел 2.3), то для масштабирующих функций $\{\phi_{j,n}(x)\}_{n \in \mathbb{Z}}$ — базиса подпространства V_j — частотная полоса Δv_j будет шире в 2^j раза. Например, для эффективного значения (2.10) ширины частотной полосы в силу свойств $2^\circ, 3^\circ$ интеграла Фурье (см. раздел 2.1) имеем:

$$\begin{aligned} \Delta v_j &= \sqrt{\int_{-\infty}^{\infty} v^2 |S_{j,0}(v)|^2 dv} = \sqrt{\frac{1}{2^j} \int_{-\infty}^{\infty} v^2 |S_{0,0}\left(\frac{v}{2^j}\right)|^2 dv} = \\ &= \sqrt{2^{2j} \int_{-\infty}^{\infty} v^2 |S_{0,0}(v)|^2 dv} = 2^j \Delta v_0, \end{aligned}$$

поскольку $S_{j,0}(v) = \Phi\{2^{j/2}\phi_{0,0}(2^j x)\} = \frac{1}{2^{j/2}}S_{0,0}\left(\frac{v}{2^j}\right)$. При этом (убедитесь!) эффективное значение (2.11) длительности для базисных функций $\{\phi_{j,n}(x) = 2^{j/2}\phi_{0,0}(2^j x - n)\}$ уменьшается и $\Delta t_j = \Delta t_0 / 2^j$.

С частотной точки зрения понижение точности проектирования при уменьшении номера m подпространства V_m можно объяснить сужением частотной полосы базисных функций, и как следствие, своего рода низкочастотной «сглаживающей» фильтрацией исходного сигнала, при которой отбрасываются верхние частоты исходного спектра $S(v) = \Phi\{f(x)\}$. Переход от $f_m \in V_m$ к более точной проекции $f_{m+1} \in V_{m+1}$, напротив, означает двукратное расширение частотной полосы базисных функций и, соответственно, добавление дополнительных высокочастотных составляющих в представление проекции $f_{m+1} \in V_{m+1}$.

Поскольку $V_m \oplus W_m = V_{m+1}$, то любую функцию $f_{m+1} \in V_{m+1}$ можно представить в виде ее «сглаженной» проекции $f_m \in V_m$, частотная полоса которой вдвое уже по сравнению с $f_{m+1} \in V_{m+1}$, и некоторой функции $y_m \in W_m$ из пространства вейвлетов: $f_{m+1}(x) = f_m(x) + y_m(x)$, причем $f_m \perp y_m$. Разность между функцией f_{m+1} и ее сглаженной копией f_m , т. е. $y_m = f_{m+1} - f_m$ — это как раз недостающие высокочастотные составляющие функции $f_{m+1} \in V_{m+1}$, те «детали», которые уточняют проекцию f_m до f_{m+1} . Можно сказать, что в разложении $f_{m+1} = f_m + y_m$ первое слагаемое соответствует нижней, а второе — верхней половине частотной полосы спектра функции $f_{m+1} \in V_{m+1}$. Поэтому базисные функции пространства V_m должны иметь отличное от нуля среднее, т.е. содержать самую низкочастотную (постоянную) составляющую. Более того, можно показать, что масштабирующая функция $\phi(x)$ порождает КМА, только если выполняется условие¹

$$\int_{-\infty}^{+\infty} \phi(x) dx = 1. \quad (7.4)$$

Напротив, для вейвлета, как мы увидим позднее, необходимо выполнение требования

$$\int_{-\infty}^{+\infty} \psi(x) dx = 0.$$

■ 7.2. Проектирование функций на подпространства КМА

Как будет показано далее, в качестве исходного объекта для вейвлет-анализа вместо функции $f(x) \in L_2(\mathbb{R})$ оказывается удобнее рассматривать элемент наилучшего приближения $f_M = A_M(f) \in V_M$, где A_M — оператор проектирования функции на подпространство $V_M \subset L_2(\mathbb{R})$:

$$f_M(x) = \sum_{n=-\infty}^{\infty} a_{M,n} \phi_{M,n}(x). \quad (7.5)$$

В разделе 7.1 отмечалось, что за счет надлежащего выбора подпространства V_M можно получить как угодно высокую точность приближения исходной функции $f(x)$ ее проекцией $f_M(x)$. Поэтому для любой практической задачи найдется такое подпространство V_M , которое позволит адекватно описать сигнал $f(x)$ элементом наилучшего приближения $f_M(x) \in V_M$.

Пример 7.2. Для КМА, построенного в примере 7.1, оценить ошибку приближения $\varepsilon = \|f - f_N\|$ для дифференцируемой функции $f(x)$, заданной на отрезке $x \in [0; 1]$.

◀ В пространстве V_N проекция $f_N = A_N(f)$ функции $f(x)$, заданной на отрезке $x \in [0; 1]$, для КМА из примера 7.1 может быть представлена в виде:

$$f_N(x) = \sum_{n=0}^{2^N-1} a_{N,n} \phi_{N,n}(x),$$

причем $\phi_{N,n}(x) = \begin{cases} 2^{N/2} & \text{при } x \in \Delta_n^N; \\ 0 & \text{при } x \notin \Delta_n^N, \end{cases} \quad \Delta_n^N = \left[\frac{n}{2^N}, \frac{n+1}{2^N}\right],$ и поэтому

¹ См. утверждение 5.8 в монографии: Блаттер К. Вейвлет-анализ. Основы теории. М., 2004.

$$a_{N,n} = \langle f, \phi_{N,n} \rangle = \int_{\Delta_n^N} f(x) \phi_{N,n}(x) dx = 2^{N/2} \int_{n/2^N}^{(n+1)/2^N} f(x) dx = 2^{-N/2} \bar{f}_n,$$

где $\bar{f}_n = 2^N \int_{n/2^N}^{(n+1)/2^N} f(x) dx$ — среднее значение функции $f(x)$ на интервале $x \in \Delta_n^N$. Заметим, что в силу дифференцируемости $f(x)$ — непрерывна, поэтому $\forall n=0, \dots, 2^N-1 \exists x_n \in \Delta_n^N : \bar{f}_n = f(x_n)$; с использованием данного обозначения можно записать:

$$f_N(x) = \frac{1}{2^{N/2}} \sum_{n=0}^{2^N-1} f(x_n) \phi_{N,n}(x),$$

где некоторые точки $x_n \in \Delta_n^N$, $n = 0, \dots, 2^N - 1$. Таким образом $f_N(x) = f(x_n)$ для $x \in \Delta_n^N$ (в силу кусочного постоянства функции $f_N(x)$).

Обозначим $h = |\Delta_n^N| = 1/2^N$. Для квадрата ошибки проектирования

$$\varepsilon^2 = \|f - f_N\|^2 = \int_0^1 (f(x) - f_N(x))^2 dx = \sum_{n=0}^{2^N-1} \int_{\Delta_n^N} (f(x) - f_N(x))^2 dx$$

найдем оценку для слагаемого

$$\varepsilon_n^2 = \int_{\Delta_n^N} (f(x) - f_N(x))^2 dx = \int_{\Delta_n^N} (f(x) - f(x_n))^2 dx.$$

По теореме о среднем для интегралов в силу непрерывности по-дышральной функции $\exists \xi_n \in \Delta_n^N$:

$$\varepsilon_n^2 = \int_{\Delta_n^N} (f(x) - f(x_n))^2 dx = (f(\xi_n) - f(x_n))^2 h.$$

Для дифференцируемой функции $f(x)$ с учетом формулы конечных приращений имеем: $f(\xi_n) - f(x_n) = f'(\gamma_n)(\xi_n - x_n)$, где некоторая точка $\gamma_n \in \Delta_n^N$ расположена между точками $x_n \in \Delta_n^N$ и $\xi_n \in \Delta_n^N$.

Поэтому

$$\varepsilon_n^2 = (f(\xi_n) - f(x_n))^2 h \leq (f'(\gamma_n))^2 h^3 \leq h^3 \max_{x \in \Delta_n^N} (f'(x))^2.$$

Так как $2^N = 1/h$, то для ошибки проектирования имеем:

$$\begin{aligned} \varepsilon^2 &= \sum_{n=0}^{2^N-1} \varepsilon_n^2 \leq h^3 \sum_{n=0}^{2^N-1} \max_{x \in \Delta_n^N} (f'(x))^2 \leq h^2 \max_{x \in [0,1]} (f'(x))^2 = h^2 \left(\max_{x \in [0,1]} |f'(x)| \right)^2, \\ \varepsilon &= \|f - f_N\| \leq h \max_{x \in [0,1]} |f'(x)|. \end{aligned} \quad (7.6)$$

Видим, что $\lim_{N \rightarrow \infty} \|f - f_N\| = \lim_{h \rightarrow 0} \|f - f_N\| = 0$ и порядок малости ошибки $\varepsilon = O(h)$. ►

Задать функцию $f_M(x) \in V_M$, как видно из (7.5) и приведенного примера 7.2, можно набором коэффициентов

$$a_{M,n} = \langle f, \phi_{M,n} \rangle = \int_{-\infty}^{\infty} f(x) \phi_{M,n}(x) dx, \quad (7.7)$$

$n \in \mathbb{Z}$. Однако вычисление интегралов (7.7) при аппаратной обработке сигналов представляет собой сложную для реализации процедуру. Оказывается, набор коэффициентов (7.7) можно достаточно точно найти в результате обычной дискретизации непрерывного сигнала $f(x)$. Поясним данное утверждение.

Масштабирующие функции, используемые для построения КМА, имеют, как правило, выраженную локализацию во временной области и часто обладают ограниченным носителем (см. пример 7.1). В силу этого интеграл (7.7) можно приближенно представить в виде:

$$a_{M,n} = \langle f, \phi_{M,n} \rangle \approx \int_{\Delta_{M,n}} f(x) \phi_{M,n}(x) dx, \quad (7.8)$$

где конечный отрезок $\Delta_{M,n}$ — интервал локализации функции $\phi_{M,n}(x)$. В частности, если носитель функции $\phi_{M,n}(x)$ ограничен, то $\Delta_{M,n}$ представляет собой отрезок, содержащий этот носитель, и тогда в (7.8) следует записать строгое равенство. На практике вейвлеты и порождающие их масштабирующие функции, которые обладают ограниченным носителем во временной области,

наиболее часто используются в цифровой обработке сигналов и потому представляют особый интерес.

Если ограниченный носитель для масштабирующей функции $\phi(x) = \phi_{0,0}(x)$ представляет собой отрезок $\Delta_{0,0} = \text{supp } \phi(x)$, то носители $\{\Delta_{M,n}\}_{n \in \mathbb{Z}}$ функций $\{\phi_{M,n}(x) = 2^{M/2} \phi(2^M x - n)\}_{n \in \mathbb{Z}}$, очевидно, будут отрезками длины $|\Delta_{M,n}| = |\Delta_{0,0}|/2^M$, причем отрезок $\Delta_{M,n+1}$ получается из отрезка $\Delta_{M,n}$ сдвигом вправо по числовой оси на величину 2^{-M} . Отметим, что отрезки $\{\Delta_{M,n}\}_{n \in \mathbb{Z}}$ в общем случае частично перекрывают друг друга. При использовании для исходного описания сигнала его проекции $f_M(x) \in V_M$ с достаточно большим значением индекса M можно добиться как угодно малой длины отрезка $\Delta_{M,n} = \text{supp } \phi_{M,n}(x)$ и для непрерывной функции $f(x)$ обеспечить $\forall \varepsilon > 0, \forall x \in \Delta_{M,n} |f(x) - f(x_{M,n})| < \varepsilon$, где $x_{M,n}$ — произвольная точка из отрезка $\Delta_{M,n}$, например его середина. Для дифференцируемой функции $f(x)$ и $\forall x \in \Delta_{M,n}$ имеем тогда $|f(x) - f(x_{M,n})| = O(|\Delta_{M,n}|) = O(2^{-M})$, т. е. $f(x) = f(x_{M,n}) + O(2^{-M})$. Выбрав достаточно большое число M , с точностью порядка $\varepsilon = O(2^{-M})$ можно считать $\forall x \in \Delta_{M,n}: f(x) \approx f(x_{M,n})$. Тогда, вынося константу из-под знака интеграла в (7.8), имеем:

$$\begin{aligned} a_{M,n} &= \int_{\Delta_{M,n}} f(x) \phi_{M,n}(x) dx \approx \tilde{a}_{M,n} = f(x_{M,n}) \int_{\Delta_{M,n}} \phi_{M,n}(x) dx = \\ &= f(x_{M,n}) 2^{M/2} \int_{\Delta_{M,n}} \phi(2^M x - n) dx = \frac{f(x_{M,n})}{2^{M/2}} \int_{\Delta_{M,n}} \phi(2^M x - n) d(2^M x - n) = \\ &= \frac{f(x_{M,n})}{2^{M/2}} \int_{\Delta_{0,0}} \phi(x) dx = f(x_{M,n}) 2^{-M/2}, \end{aligned}$$

где последний переход осуществлен на основании (7.4). Таким образом, с любой достаточной для практики точностью коэффициенты разложения (7.5) можно получить в результате равномерной дискретизации сигнала $f(x)$ с шагом $h = x_{M,n+1} - x_{M,n} = 2^{-M}$, если последний выбран из условия $f(x) \approx f(x+h)$:

$$a_{M,n} \approx \tilde{a}_{M,n} = f(x_{M,n}) 2^{-M/2}. \quad (7.9)$$

Как оценить эту точность?

Лемма 7.2. При использовании формулы (7.9) для приближенного вычисления коэффициентов (7.7) дифференцируемой функции $f(x)$, для КМА с ограниченным носителем масштабирующей функции $\Delta_{0,0} = \text{supp } \phi(x)$ верна оценка:

$$|a_{M,n} - \tilde{a}_{M,n}| \leq \frac{H^{3/2}}{2} \max_{x \in \Delta_n^M} |f'(x)|, \quad (7.10)$$

где $\tilde{a}_{M,n} = f(x_{M,n}) 2^{-M/2}$ (приближенное значение соответствующего коэффициента $a_{M,n}$); $x_{M,n}$ — середина отрезка $\Delta_{M,n} = \text{supp } \phi_{M,n}(x)$, а $H = |\Delta_{M,n}| = |\Delta_{0,0}|/2^M$ — его длина.

◀ Для погрешности $\delta_n = |a_{M,n} - \tilde{a}_{M,n}|$ на основании леммы 1.3 получим:

$$\begin{aligned} \delta_n^2 &= \left| \int_{\Delta_{M,n}} (f(x) - f(x_{M,n})) \phi_{M,n}(x) dx \right|^2 = \left| \langle f(x) - f(x_{M,n}), \phi_{M,n}(x) \rangle \right|^2 \leq \\ &\leq \underbrace{\int_{\Delta_{M,n}} |f(x) - f(x_{M,n})|^2 dx}_{\|f(x) - f(x_{M,n})\|^2} \underbrace{\int_{\Delta_{M,n}} |\phi_{M,n}(x)|^2 dx}_{\|\phi_{M,n}(x)\|^2=1} = H |f(\xi_n) - f(x_{M,n})|^2, \end{aligned}$$

где некоторая точка $\xi_n \in \Delta_{M,n}$. Поскольку функция $f(x)$ — дифференцируема, то $\exists \gamma_n \in \Delta_{M,n}: f(\xi_n) - f(x_{M,n}) = f'(\gamma_n)(\xi_n - x_{M,n})$. Так как точка $x_{M,n}$ является серединой отрезка $\Delta_{M,n}$, то $|f(\xi_n) - f(x_{M,n})| \leq |f'(\gamma_n)|H/2$ и

$$\delta_n \leq \sqrt{H} |f(\xi_n) - f(x_{M,n})| \leq \frac{H^{3/2}}{2} \max_{x \in \Delta_n^M} |f'(x)|. ▶$$

Следствие. Если носитель $\Delta_{0,0}$ масштабирующей функции $\phi(x)$ конечен, то в пространстве $L_2[0; 1]$ для дифференцируемой

функции $f(x)$ погрешность, связанная с использованием формулы (7.9), оценивается как

$$\|f_M - \tilde{f}_M\| = O(h), \quad (7.11)$$

где $h = x_{M,n+1} - x_{M,n} = 2^{-M}$ — шаг дискретизации.

◀ Не умаляя общности рассуждений, положим, что $\Delta_{0,0} = \text{supp } \phi_{0,0}(x) \subset [0; K]$, где K — некоторое натуральное число. Тогда $\Delta_{M,n} = \text{supp } \phi_{M,n}(x) \subset [n/2^M; (K+n)/2^M]$, причем $\Delta_{M,n} \cap [0; 1] = \emptyset$ для $n \notin \{-K, -K+1, \dots, 2^M\}$. Поэтому в пространстве $L_2[0; 1]$ можно представить функцию $f_M(x) \in V_M$ в виде конечной суммы:

$$f_M(x) = \sum_{n=-K}^{2^M} a_{M,n} \phi_{M,n}(x).$$

Так как

$$\begin{aligned} f_M - \tilde{f}_M &= \sum_{n=-K}^{2^M} a_{M,n} \phi_{M,n}(x) - \sum_{n=-K}^{2^M} \tilde{a}_{M,n} \phi_{M,n}(x) = \\ &= \sum_{n=-K}^{2^M} (a_{M,n} - \tilde{a}_{M,n}) \phi_{M,n}(x), \end{aligned}$$

то, воспользовавшись равенством Парсеваля, получаем:

$$\|f_M - \tilde{f}_M\|^2 = \sum_{n=-K}^{2^M} (a_{M,n} - \tilde{a}_{M,n})^2 = O(h^2),$$

поскольку каждое слагаемое в данной сумме имеет порядок (см. (7.10)) $O(H^3) = O(h^3) = O(2^{-3M})$, а число слагаемых имеет порядок $O(2^M) = O(h^{-1})$. Отсюда следует оценка (7.11). ►

Таким образом, мы пришли к важному выводу: при достаточно малом шаге дискретизации $h = x_{M,n+1} - x_{M,n} = 2^{-M}$ коэффициенты разложения (7.5) можно трактовать как масштабированные отсчеты дискретного сигнала: $a_{M,n} \approx \sqrt{h} f(x_{M,n})$, так как

$$a_{M,n} = \sqrt{h} f(x_{M,n}) + O(h^{3/2}) \quad (7.12)$$

(см. (7.9) и (7.10), где нужно положить $H = |\Delta_{0,0}| h$). При этом (см. (7.11)) в норме L_2 погрешность $\delta = \|f_M - \tilde{f}_M\| = O(h)$.

Отметим, что общие оценки (7.10) и (7.11) в ряде случаев можно улучшить, увеличив порядок малости ошибки в формуле (7.12).

Пример 7.3. Для дважды дифференцируемой функции $f(x)$, заданной на отрезке $x \in [0, 1]$, и КМА, построенного в примере 7.1, оценить ошибку $\delta = \|f_N - \tilde{f}_N\|$, где $\tilde{f}_N(x) = \sum_{n=0}^{2^N-1} \tilde{a}_{N,n} \phi_{N,n}(x)$ — приближенное проектирование, при котором коэффициенты разложения получены по формуле (7.9) в результате масштабирования и дискретизации функции $f(x)$ с шагом $h = 1/2^N$:

$$\tilde{a}_{N,n} = \sqrt{h} f((n+0,5)h), \quad n = 0, 1, \dots, 2^N - 1.$$

◀ В силу ортонормированности базиса $\{\phi_{N,n}(x)\}_{n \in \mathbb{Z}}$ для функции $f_N(x) - \tilde{f}_N(x)$ равенство Парсеваля примет вид:

$$\delta^2 = \|f_N - \tilde{f}_N\|^2 = \sum_{n=0}^{2^N-1} (a_{N,n} - \tilde{a}_{N,n})^2.$$

Оценим величину $\delta_n = |a_{N,n} - \tilde{a}_{N,n}|$. Имеем (см. пример 7.2):

$$\begin{aligned} \delta_n &= \left| \underbrace{\frac{1}{\sqrt{h}} \int_{nh}^{(n+1)h} f(x) dx}_{a_{N,n}} - \underbrace{\frac{f((n+0,5)h)}{\sqrt{h}} \int_{nh}^{(n+1)h} dx}_{\tilde{a}_{N,n}} \right| = \\ &= \frac{1}{\sqrt{h}} \left| \int_{nh}^{(n+1)h} (f(x) - f(x_n)) dx \right|, \end{aligned}$$

где $x_n = (n+0,5)h$ — середина интервала

$$\Delta_{N,n} = \Delta_n = [nh, (n+1)h].$$

С использованием формулы Тейлора

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(x_n)}{2}(x - x_n)^2 + O((x - x_n)^3)$$

получим:

$$\begin{aligned}\delta_n &= \frac{1}{\sqrt{h}} \left| \int_{nh}^{(n+1)h} \left(f'(x_n)(x - x_n) + \frac{f''(x_n)}{2}(x - x_n)^2 + O(h^3) \right) dx \right| = \\ &= \frac{1}{\sqrt{h}} \left(\frac{h^3 |f''(x_n)|}{24} + O(h^4) \right),\end{aligned}$$

поскольку $\int_{nh}^{(n+1)h} (x - x_n) dx = 0$, $\int_{nh}^{(n+1)h} (x - x_n)^2 dx = \frac{h^3}{12}$.

Видим, что $\delta_n = O(h^{5/2})$. Поэтому $\delta^2 = \sum_{n=0}^{2^N-1} \delta_n^2 = O(h^4)$ (так

как $2^N = 1/h$), и окончательно:

$$\delta = \|f_N - \tilde{f}_N\| = O(h^2).$$

Как видим, замена точного проектирования приближенным вносит дополнительную погрешность $\delta = O(h^2)$ более высокого порядка малости, чем порядок малости погрешности точного проектирования (7.6): $\epsilon = O(h)$. То есть на практике использование точного проектирования с применением формулы (7.7) для КМА из примера 7.1 при достаточно малых значениях h нецелесообразно, так как это лишь увеличивает вычислительные затраты. ►

7.3. Вычисление вейвлет-преобразований

Рассмотрим вопросы практического вычисления вейвлет-преобразования произвольной функции $f(x) \in L_2(\mathbb{R})$, т. е. нахождения коэффициентов Фурье $\{c_{m,n}\}$ ряда

$$f(x) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \psi_{m,n}(x). \quad (7.13)$$

Оказывается, можно не вычислять интегралы

$$c_{m,n} = \langle f, \psi_{m,n} \rangle = \int_{-\infty}^{\infty} f(x) \psi_{m,n}(x) dx, \quad (7.14)$$

если в качестве исходного объекта для вейвлет-анализа вместо функции $f(x) \in L_2(\mathbb{R})$ задан некоторый элемент наилучшего приближения (7.5), представленный набором коэффициентов (7.7).

Поскольку $V_M = \bigoplus_{j=-\infty}^{M-1} W_j$, то вейвлет-разложение (7.13) для проекции $f_M(x) \in V_M$ примет вид:

$$f_M(x) = \sum_{m=-\infty}^{M-1} \sum_{n=-\infty}^{\infty} c_{m,n} \psi_{m,n}(x) = \sum_{m=-\infty}^{M-1} y_m(x), \quad (7.15)$$

где $y_m(x) = \sum_{n=-\infty}^{\infty} c_{m,n} \psi_{m,n}(x)$, $y_m(x) \in W_m$. Задача, которая будет рассматриваться в данном разделе, состоит в следующем: как найти коэффициенты $\{c_{m,n}\}$ разложения (7.15) по заданному набору (7.7) и как восстановить значения (7.7) по коэффициентам выражения (7.15).

Представим некоторый элемент $f_m(x) \in V_m = V_{m-1} \oplus W_{m-1}$ в виде суммы $f_m(x) = f_{m-1}(x) + y_{m-1}(x)$, где $f_{m-1}(x) \in V_{m-1}$, $y_{m-1}(x) \in W_{m-1}$. Тогда раскладывая функции $f_m(x)$, $f(x)_{m-1}$, $y(x)_{m-1}$ по базисам соответствующих подпространств V_m , V_{m-1} , W_{m-1} , можем записать:

$$\underbrace{\sum_{n=-\infty}^{\infty} a_{m,n} \phi_{m,n}(x)}_{f_m \in V_m} = \underbrace{\sum_{n=-\infty}^{\infty} a_{m-1,n} \phi_{m-1,n}(x)}_{f_{m-1} \in V_{m-1}} + \underbrace{\sum_{n=-\infty}^{\infty} c_{m-1,n} \psi_{m-1,n}(x)}_{y_{m-1} \in W_{m-1}}. \quad (7.16)$$

Далее заметим, что из масштабирующего уравнения (7.1) следует соотношение

$$\phi_{m-1,k}(x) = \sum_{j \in \mathbb{Z}} h_{j-2k} \phi_{m,j}(x). \quad (7.17)$$

Действительно (см. (7.1)),

$$\begin{aligned} \phi_{m,k}(x) &= 2^{m/2} \phi(2^m x - k) = 2^{m/2} \sqrt{2} \sum_{j \in \mathbb{Z}} h_j \phi\left(2(2^m x - k) - j\right) = \\ &= \sum_{j \in \mathbb{Z}} h_j 2^{\frac{m+1}{2}} \phi\left(2^{m+1} x - j - 2k\right) = \sum_{j \in \mathbb{Z}} h_j \phi_{m+1,j+2k}(x) = \sum_{j \in \mathbb{Z}} h_{j-2k} \phi_{m+1,j}(x). \end{aligned}$$

Совершенно аналогично из (7.3) следует, что

$$\psi_{m-1,k}(x) = \sum_{j \in \mathbb{Z}} g_{j-2k} \phi_{m,j}(x). \quad (7.18)$$

Умножив равенство (7.16) скалярно на элемент $\phi_{m-1,k}(x)$,

$$\sum_{n \in \mathbb{Z}} a_{m,n} \langle \phi_{m,n}(x), \phi_{m-1,k}(x) \rangle = a_{m-1,k},$$

и подставив в полученное уравнение представление (7.17) для функции $\phi_{m-1,k}(x)$, имеем

$$a_{m-1,k} = \sum_{n \in \mathbb{Z}} a_{m,n} \left\langle \phi_{m,n}(x), \sum_{j \in \mathbb{Z}} h_{j-2k} \phi_{m,j}(x) \right\rangle = \sum_{n \in \mathbb{Z}} a_{m,n} \sum_{j \in \mathbb{Z}} h_{j-2k} \langle \phi_{m,n}(x), \phi_{m,j}(x) \rangle.$$

В силу ортонормированности системы $\{\phi_{m,k}(x)\}_{k \in \mathbb{Z}} \subset V_m$ получим окончательно $\forall k$:

$$a_{m-1,k} = \sum_{n \in \mathbb{Z}} a_{m,n} h_{n-2k}. \quad (7.19)$$

Аналогично, умножив уравнение (7.16) скалярно на функцию $\psi_{m-1,k}(x)$ и затем представив последнюю при помощи выражения (7.18), получаем $\forall k$:

$$c_{m-1,k} = \sum_{n \in \mathbb{Z}} a_{m,n} g_{n-2k}. \quad (7.20)$$

Формулы (7.19) и (7.20) задают вычислительную процедуру, рекуррентное применение которой к последовательностям $\{a_{m,n}\}_{n \in \mathbb{Z}}$ для $m = M, M-1, M-2, \dots$ позволяет найти коэффици-

енты разложения (7.15) по известному набору (7.7), т. е. выполнить вейвлет-преобразование функции $f(x) \in L_2(\mathbb{R})$, заданной своей проекцией $f_M(x) \in V_M$. Отметим, что вычисления по формулам (7.19), (7.20) носят дискретный характер, сводятся к обработке числовых последовательностей. Причем для КМА, образованного масштабирующей функцией с конечным носителем, суммы в (7.19) и в (7.20) также конечны. Алгоритм соответствующих вычислений можно изобразить в виде схемы, приведенной на рис. 7.2.

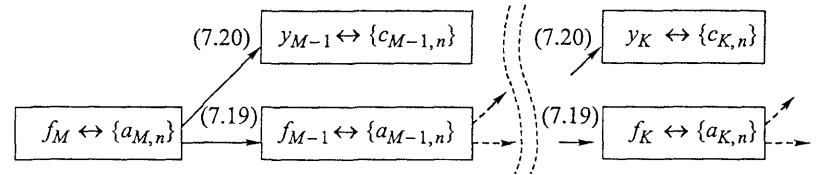


Рис. 7.2. Схема вычисления дискретного вейвлет-преобразования

На практике число шагов алгоритма вейвлет-преобразования (число повторений процедур (7.19), (7.20)) должно быть конечно. Для некоторого числа $m = K$ вычисления прекращаются, т. е. вместо разложения (7.15) мы ограничиваемся представлением:

$$\begin{aligned} f_M(x) &= \underbrace{\sum_{n=-\infty}^{\infty} a_{K,n} \phi_{K,n}(x)}_{f_K(x) \in V_K} + \sum_{m=K}^{M-1} \underbrace{\sum_{n=-\infty}^{\infty} c_{m,n} \psi_{m,n}(x)}_{y_m(x) \in W_m} = \\ &= f_K(x) + y_K(x) + y_{K+1}(x) + \dots + y_{M-1}(x). \end{aligned} \quad (7.21)$$

Поскольку из (7.15) следует сходимость последовательности частичных сумм $S_K(x) = \sum_{m=K}^{M-1} y_m(x)$ (для $K \rightarrow -\infty$), то для любой заданной точности $\varepsilon > 0$ можно выбрать число K так, чтобы $\|f_K(x)\| = \|f_M(x) - S_K(x)\| < \varepsilon$, тогда слагаемое $f_K(x)$ в сумме (7.21) можно отбросить.

Замечание. Для приложений, связанных с компрессией (сжатием) сигналов, число шагов вычисления дискретного вейвлет-преобразования выбирают небольшим и коэффициенты $\{a_{K,n}\}$ не отбрасывают. Число K обычно выбирается так, чтобы коэффици-

коэффициенты $a_{K,n} = \int_{\Delta_{K,n}} f(x)\phi_{K,n}(x)dx$ (здесь $\Delta_{K,n} = \text{supp } \phi_{K,n}(x)$ — конечный отрезок) «грубой» проекции $f_K(x) \in V_K$ можно было счи-

тать некоррелированными. Так, если L — длина интервала корреляции для сигнала $f(x)$ (т. е. $\forall x_0, \forall \tau > L$ значения функции в точках $f(x_0)$ и $f(x_0 + \tau)$ считаются некоррелированными), то выби-

рать K следует так, чтобы длина носителя $\Delta_{K,n}$ была много

больше интервала корреляции: $L \ll |\Delta_{K,n}| = |\Delta_{0,0}|/2^K$.

Формулы (7.19), (7.20) представляют собой соотношения композиции, используемые для получения вейвлет-разложения функций. Рассмотрим теперь обратную задачу (композиции) и получим выражения, которые обращают выражения (7.19) и (7.20), т. е. по заданным своими коэффициентами Фурье разложением

$$f_{m-1}(x) = \sum_{n \in \mathbb{Z}} a_{m-1,n} \phi_{m-1,n}(x),$$

$$y_{m-1}(x) = \sum_{n \in \mathbb{Z}} c_{m-1,n} \psi_{m-1,n}(x)$$

позволяют найти коэффициенты $\{a_{m,n}\}_{n \in \mathbb{Z}}$ представления функции $f_m(x) = \sum_{n \in \mathbb{Z}} a_{m,n} \phi_{m,n}(x)$. Для этого в правую часть уравнения (7.16) подставим выражения для $\phi_{m-1,n}(x)$ (7.17) и $\psi_{m-1,n}(x)$ (7.18), а затем полученное равенство скалярно умножим на функцию $\phi_{m,k}(x)$:

$$\begin{aligned} & \sum_{n \in \mathbb{Z}} a_{m,n} \langle \phi_{m,n}(x), \phi_{m,k}(x) \rangle = \\ & = \sum_{n \in \mathbb{Z}} a_{m-1,n} \left(\underbrace{\sum_{j \in \mathbb{Z}} h_{j-2n} \phi_{m,j}(x), \phi_{m,k}(x)}_{\phi_{m-1,n}(x)} \right) + \sum_{n \in \mathbb{Z}} c_{m-1,n} \left(\underbrace{\sum_{j \in \mathbb{Z}} g_{j-2n} \phi_{m,j}(x), \phi_{m,k}(x)}_{\psi_{m-1,n}(x)} \right). \end{aligned}$$

Отсюда вследствие ортонормированности системы

$$\{\phi_{m,k}(x)\}_{k \in \mathbb{Z}} \subset V_m$$

получаем $\forall k$:

$$a_{m,k} = \sum_{n \in \mathbb{Z}} a_{m-1,n} h_{k-2n} + \sum_{n \in \mathbb{Z}} c_{m-1,n} g_{k-2n}. \quad (7.22)$$

Последовательно применяя формулу (7.22) для значений $m = K+1, \dots, M$, можно восстановить коэффициенты $\{a_{M,n}\}_{n \in \mathbb{Z}}$, задающие проекцию $f_M(x) \in V_M$. Соответствующую последовательность вычислений можно представить в виде следующей схемы (рис. 7.3).

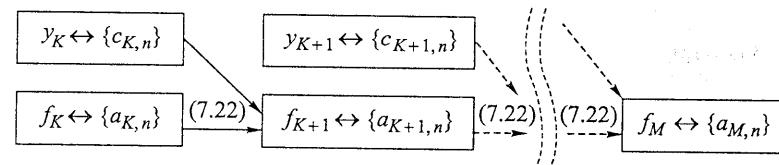


Рис. 7.3. Схема вычисления обратного дискретного вейвлет-преобразования

■ 7.4. Квадратурно-зеркальные фильтры (КЗФ)

Вернемся к рассмотрению формулы (7.19), которая задает преобразование некоторой входной последовательности $x(n) = a_{m,n}$ в выходную $y(k) = a_{m-1,k}$. Обозначим $\tilde{h}(m) = \tilde{h}_m = h_{-m}$. Пусть последовательность $u(j)$ ($j \in \mathbb{Z}$) такова, что для четных номеров $u(2k) = a_{m-1,k}$, $k \in \mathbb{Z}$. Тогда можно переписать (7.19) в виде:

$$y(k) = u(2k) = \sum_{n=-\infty}^{\infty} x(n) \tilde{h}(2k-n). \quad (7.23)$$

Сравнивая выражение (7.23) с формулой (4.9), приходим к выводу, что преобразование последовательности $x(n)$ в $y(k)$ можно трактовать как обработку входной последовательности $x(n)$ с помощью дискретного фильтра с ИХ $\{\tilde{h}(n) = \tilde{h}_n\}$ и последующее прореживание отклика фильтра, при котором сохраняются только отсчеты с четными номерами, а отсчеты с нечетными номерами отбрасываются. Используя для операции прореживания, которая отбрасывает нечетные номера последовательности, обозначение $\downarrow 2$, вычисления по формуле (7.19) (или (7.23)) можно представить в виде схемы, приведенной на рис. 7.4.

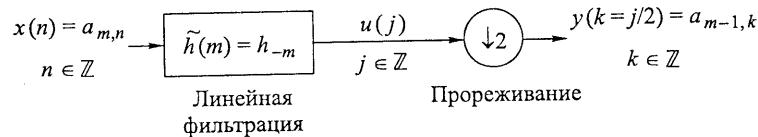


Рис. 7.4. Схема реализации процедуры (7.19) при помощи дискретного фильтра и последующего прореживания последовательности

Замечание. При рассмотрении операции (7.19) с точки зрения теории линейных дискретных систем возможна ситуация, когда ИХ $\tilde{h}(n) = \tilde{h}_n \neq 0$ при $n < 0$, что соответствует физически нереализуемой системе вследствие нарушения причинно-следственной связи: отклик на выходе появляется раньше сигнала на входе. Однако рассмотрение подобных систем корректно и противоречия в этом нет, поскольку номера элементов обрабатываемых последовательностей не следует обязательно отождествлять с моментами времени, а линейное преобразование последовательностей по формуле (7.19) — с физическим устройством, обрабатывающим последовательно поступающие отсчеты дискретного сигнала в реальном масштабе времени. Во многих случаях в обрабатываемой последовательности $x(n)$ необходимые «опережающие» отсчеты известны заранее (например, это могут быть заданные массивы в памяти компьютера).

Мы будем рассматривать далее только КИХ-фильтры, для которых ИХ $h(n) = \{h_N, \dots, h_M\}$ удовлетворяет условиям: $h(n) = 0$ при $n < N$ или $n > M$, причем допустимы ситуации, когда $N < 0$, $M < 0$. Для таких фильтров под передаточной функцией и частотной характеристикой мы будем понимать соответственно следующие конечные суммы:

$$H(z) = \sum_{k=N}^M h(k)z^{-k} = \sum_{k=-\infty}^{\infty} h(k)z^{-k},$$

$$K_H(\omega) = H(e^{i\omega}) = \sum_{k=N}^M h(k)e^{-i\omega k} = \sum_{k=-\infty}^{\infty} h(k)e^{-i\omega k},$$

где бесконечные границы суммирования указаны только для удобства и упрощения записи.

Все рассуждения, касающиеся трактовки формулы (7.19), можно повторить и для формулы (7.20), тогда один шаг алгоритма вычисления коэффициентов вейвлет-разложения (см. рис. 7.2) можно представить как параллельную обработку фильтрами с импульсными характеристиками $\{\tilde{h}(n) = h_{-n}\}$, $\{\tilde{g}(n) = g_{-n}\}$ и последующее прореживание выходных последовательностей (рис. 7.5).

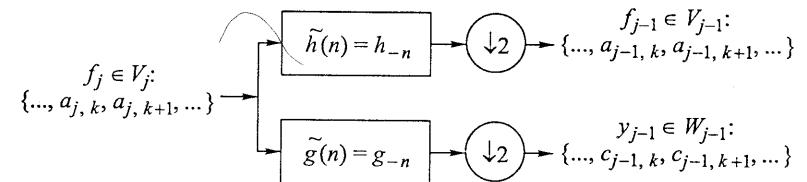


Рис. 7.5. Реализация одного шага вейвлет-преобразования (7.19), (7.20)

Пример 7.4. Для КМА из примера 7.1 найти выходные последовательности в схеме обработки рис. 7.5 при входном воздействии $\{x(0), x(1), \dots, x(2L-1)\}$.

◀ Воспользовавшись результатами примера 7.1, получим следующие ИХ:

$$\begin{aligned} \tilde{h}(n) &= \left\{ \dots, 0, 0, \underbrace{1/\sqrt{2}}_{\tilde{h}(-1)}, \underbrace{1/\sqrt{2}}_{\tilde{h}(0)}, 0, 0, \dots \right\}, \\ \tilde{g}(n) &= \left\{ \dots, 0, 0, \underbrace{-1/\sqrt{2}}_{\tilde{g}(-1)}, \underbrace{1/\sqrt{2}}_{\tilde{g}(0)}, 0, 0, \dots \right\}. \end{aligned}$$

Отклики фильтров находим по формулам свертки, которые примут вид:

$$\begin{aligned} u(k) &= \sum_{n=-1}^0 \tilde{h}(n)x(k-n) = \frac{x(k) + x(k+1)}{\sqrt{2}}, \\ v(k) &= \sum_{n=-1}^0 \tilde{g}(n)x(k-n) = \frac{x(k) - x(k+1)}{\sqrt{2}}. \end{aligned}$$

После прореживания, сохраняющего лишь четные номера выходных последовательностей, $\{u(2k)\}$ и $\{v(2k)\}$, окончательно полу-

чим из одной входной последовательности $\{x(0), x(1), \dots, x(2L-1)\}$ две выходные: $\left\{\frac{x(2k) + x(2k+1)}{\sqrt{2}}\right\}_{k=0}^{L-1}, \left\{\frac{x(2k) - x(2k+1)}{\sqrt{2}}\right\}_{k=0}^{L-1}$.

Повторив последовательно преобразования по схеме, приведенной на рис. 7.5, необходимое число раз, можно выполнить вейвлет-разложение на любую «глубину», т.е. до любого уровня разрешения проекции $f_k(x) \in V_k$, заданной набором коэффициентов $\{a_{k,k}\}$ (рис. 7.6).

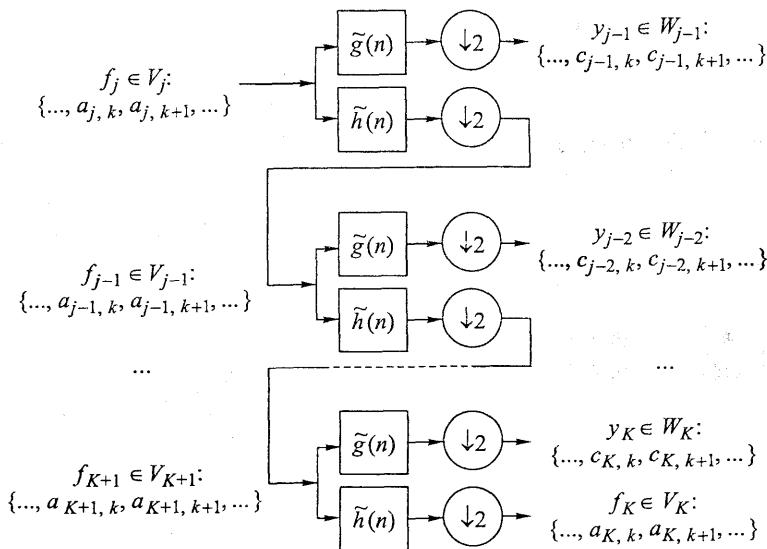


Рис. 7.6. Вычисление вейвлет-преобразования при помощи блоков фильтров

Рассмотрим теперь формулу (7.22), которая обращает процедуру, определяемую формулами (7.19), (7.20) и схемой рис. 7.5. Обозначим $j = 2n$, $b(2n) = a_{m-1,n}$, $b(2n+1) = 0$, $h_m = h(m)$, тогда для первой суммы в формуле (7.22) получим:

$$S_1(k) = \sum_{n=-\infty}^{\infty} a_{m-1,n} h_{k-2n} = \sum_{n=-\infty}^{\infty} b(2n) h(k-2n) = \sum_{j=-\infty}^{\infty} b(j) h(k-j), \quad (7.24)$$

что можно представить как отклик линейного фильтра с ИХ $h(n)$, на вход которого подана последовательность $\{b(j)\}$, полученная

добавлением одного нуля после каждого элемента последовательности $\{a_{m-1,n}\}_{n \in \mathbb{Z}}$. Аналогично, вторую сумму в формуле (7.22) можем записать:

$$S_2(k) = \sum_{n=-\infty}^{\infty} c_{m-1,n} g_{k-2n} = \sum_{j=-\infty}^{\infty} d(j) g(k-j),$$

где $g_m = g(m)$, $d(j) = \begin{cases} c_{m-1,j/2}, & \text{если } j \text{ — четное число;} \\ 0, & \text{если } j \text{ — нечетное число.} \end{cases}$

Операцию вставки нулевого элемента на каждую вторую позицию в последовательности обозначим $\uparrow 2$. Тогда вычислениям по формуле (7.22) соответствует схема обработки, которая реализуется с помощью двух фильтров с импульсными характеристиками $h(n)$ и $g(n)$ (рис. 7.7).

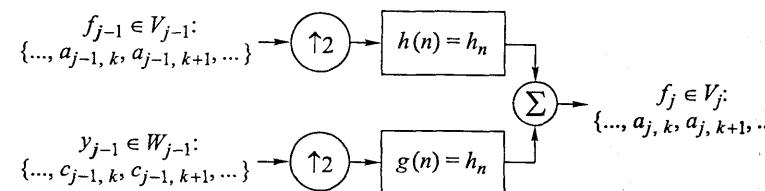


Рис. 7.7. Реализация процедуры (7.22) при помощи дискретных фильтров

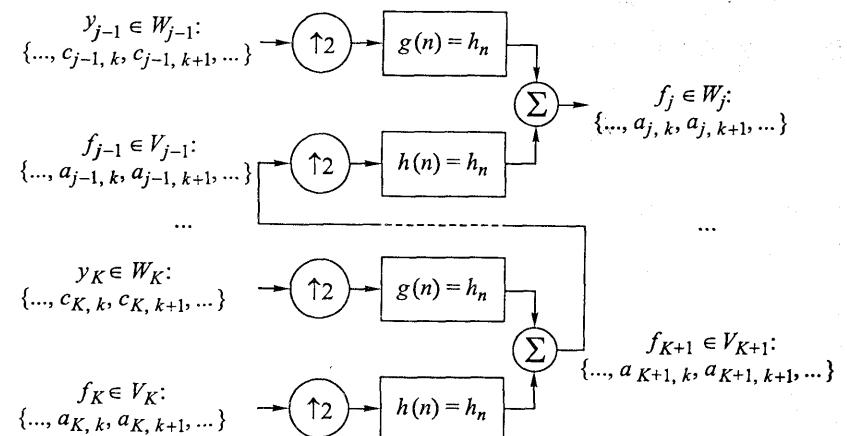


Рис. 7.8. Вычисление обратного вейвлет-преобразования при помощи блоков зеркальных фильтров

Схема обработки, представленная на рис. 7.7, реализует обратную процедуру по сравнению со схемой обработки последовательности, приведенной ранее на рис. 7.5. Поэтому обращение многоуровневого преобразования, схема которого отражена на рис. 7.6, также достигается последовательным соединением соответствующих структурных блоков фильтров (рис. 7.8).

Для импульсных характеристик фильтров на рис. 7.5 и 7.7 верно: $\tilde{g}(n) = g(-n)$, $\tilde{h}(n) = h(-n)$, т. е. соответствующие ИХ являются зеркальными. Поэтому четыре рассмотренных фильтра с ИХ $\{\tilde{h}(n)\}$, $\{h(n)\}$, $\{\tilde{g}(n)\}$, $\{g(n)\}$ называют *квадратурно-зеркальными фильтрами* (КЗФ).

7.5. Свойства КЗФ

Как отмечалось в разделе 7.1, представление функции $f_{m+1} \in V_{m+1}$ в виде $f_{m+1}(x) = f_m(x) + y_m(x)$, где $f_m \in V_m$, $y_m \in W_m$, есть разделение сигнала на низкочастотную $f_m(x)$ и высокочастотную $y_m(x)$ составляющие. Поэтому можно предположить, что в паре фильтров, осуществляющих такое разложение сигнала, фильтр с ИХ $\{\tilde{h}(n) = h_{-n}\}$ должен быть низкочастотным, а фильтр с ИХ $\{\tilde{g}(n) = g_{-n}\}$ — высокочастотным. Описать данные фильтры можно не только при помощи ИХ, но и задав передаточные функции $\tilde{H}(z)$, $\tilde{G}(z)$, $H(z)$, $G(z)$, или частотные характеристики $K_{\tilde{H}}(\omega)$, $K_{\tilde{G}}(\omega)$, $K_H(\omega)$, $K_G(\omega)$. Установим, как связаны между собой частотные характеристики и передаточные функции КЗФ, которые определяют схему обработки сигналов, представленную рис. 7.9.

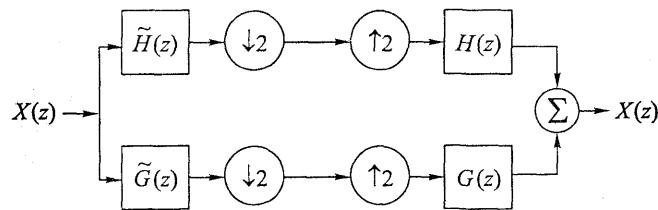


Рис. 7.9. Общая схема системы обработки сигналов при помощи КЗФ

Прежде всего заметим, что если $K_H(\omega) = \sum_{k=-\infty}^{\infty} h(k)e^{-i\omega k}$ — ЧХ фильтра (см. раздел 4.5) с вещественной ИХ $h(n) = h_n$, то для зеркального фильтра, очевидно, ЧХ комплексно сопряжена:

$$K_{\tilde{H}}(\omega) = \sum_{k=-\infty}^{\infty} h(-k)e^{-i\omega k} = \sum_{k=-\infty}^{\infty} h(k)e^{i\omega k} = \overline{K_H(\omega)} = K_H(-\omega).$$

$$\text{Аналогично, } K_G(\omega) = \sum_{k=-\infty}^{\infty} g(k)e^{-i\omega k}, \quad K_{\tilde{G}}(\omega) = \overline{K_G(\omega)} = K_G(-\omega).$$

Таким образом, АЧХ зеркальных фильтров одинаковы, а ФЧХ отличаются знаком. Так как $K_{\tilde{H}}(\omega) = \tilde{H}(e^{i\omega}) = K_H(-\omega) = H(e^{-i\omega})$, то для передаточных функций $\tilde{H}(z) = H(z^{-1})$ и, аналогично, $\tilde{G}(z) = G(z^{-1})$.

С учетом свойств 3° и 4° преобразования Фурье (см. раздел 2.1) его применение для масштабирующего уравнения (7.1) дает:

$$S_\phi(\omega) = \frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} h_n e^{-i\omega n/2} S_\phi(\omega/2) = \frac{1}{\sqrt{2}} K_H(\omega/2) S_\phi(\omega/2), \quad (7.25)$$

$$\text{где } S_\phi(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(t) e^{-i\omega t} dt.$$

Рассмотрим интеграл $I(u) = \int_{-\infty}^{\infty} \phi(t) \phi(t-u) dt$. В силу ортонормированности целочисленных сдвигов масштабирующей функции $\forall n \in \mathbb{Z}$:

$$I(n) = \begin{cases} 1, & n = 0; \\ 0, & n \neq 0. \end{cases}$$

Найдем преобразование Фурье для функции $I(u)$.

$$\begin{aligned} S_I(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(t) \phi(t-u) dt e^{-i\omega u} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(t) \int_{-\infty}^{\infty} \phi(t-u) e^{-i\omega u} du dt = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(t) e^{-i\omega t} \int_{-\infty}^{\infty} \phi(t-u) e^{-i\omega(u-t)} du dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(t) e^{-i\omega t} \int_{-\infty}^{\infty} \phi(x) e^{i\omega x} dx dt = \end{aligned}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(t) e^{-i\omega t} dt \int_{-\infty}^{\infty} \phi(x) e^{i\omega x} dx = \sqrt{2\pi} S_{\phi}(\omega) \overline{S_{\phi}(\omega)} = \sqrt{2\pi} |S_{\phi}(\omega)|^2.$$

Выполнив затем обратное преобразование Фурье, получим $\forall n \in \mathbb{Z}$:

$$\begin{aligned} I(n) &= \int_{-\infty}^{\infty} |S_{\phi}(\omega)|^2 e^{i\omega n} d\omega = \sum_{j=-\infty}^{\infty} \int_{2\pi j}^{2\pi(j+1)} |S_{\phi}(\omega)|^2 e^{i\omega n} d\omega = \\ &= \sum_{j=-\infty}^{\infty} \int_0^{2\pi} |S_{\phi}(\omega + 2\pi j)|^2 e^{i(\omega + 2\pi j)n} d\omega = \int_0^{2\pi} \sum_{j=-\infty}^{\infty} |S_{\phi}(\omega + 2\pi j)|^2 e^{i\omega n} d\omega. \end{aligned}$$

В правой части полученного равенства функция

$$F(\omega) = \sum_{j=-\infty}^{\infty} |S_{\phi}(\omega + 2\pi j)|^2$$

имеет период 2π , причем ее коэффициенты Фурье (2.4) равны:

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} F(\omega) e^{-i\omega k} d\omega = \frac{I(-k)}{2\pi} = \begin{cases} \frac{1}{2\pi}, & k = 0; \\ 0, & k \neq 0. \end{cases}$$

Следовательно,

$$F(\omega) = \sum_{j \in \mathbb{Z}} |S_{\phi}(\omega + 2\pi j)|^2 = \frac{1}{2\pi} \quad (7.26)$$

(возможно, за исключением отдельных точек $\omega \in \mathbb{R}$, т. е. *почти всюду*).

Запишем формулу (7.26) для функции $F(2\omega)$ и воспользуемся соотношением (7.25):

$$\begin{aligned} \frac{1}{2\pi} &= \sum_{j \in \mathbb{Z}} |S_{\phi}(2\omega + 2\pi j)|^2 = \frac{1}{2} \sum_{j \in \mathbb{Z}} |K_H(\omega + \pi j) S_{\phi}(\omega + \pi j)|^2 = \\ &= \frac{1}{2} \left(\sum_{j \in \mathbb{Z}} |K_H(\omega + 2\pi j) S_{\phi}(\omega + 2\pi j)|^2 + \sum_{j \in \mathbb{Z}} |K_H(\omega + (2j+1)\pi) S_{\phi}(\omega + (2j+1)\pi)|^2 \right). \end{aligned}$$

С учетом 2π -периодичности ЧХ $K_H(\omega)$ получим:

$$\frac{1}{\pi} = \underbrace{|K_H(\omega)|^2 \sum_{j \in \mathbb{Z}} |S_{\phi}(\omega + 2\pi j)|^2}_{F(\omega) = \frac{1}{2\pi}} + \underbrace{|K_H(\omega + \pi)|^2 \sum_{j \in \mathbb{Z}} |S_{\phi}(\omega + \pi + 2\pi j)|^2}_{F(\omega + \pi) = \frac{1}{2\pi}},$$

откуда

$$|K_H(\omega)|^2 + |K_H(\omega + \pi)|^2 = 2 \quad (7.27)$$

почти всюду. Очевидно, что соотношение (7.27) верно и для зеркального фильтра \tilde{H} .

Проинтегрируем уравнение (7.1):

$$\begin{aligned} \int_{-\infty}^{\infty} \phi(x) dx &= \sqrt{2} \sum_{n \in \mathbb{Z}} h_n \int_{-\infty}^{\infty} \phi(2x - n) dx = \\ &= \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} h_n \int_{-\infty}^{\infty} \phi(x - n) dx = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} \phi(x) dx \sum_{n \in \mathbb{Z}} h_n. \end{aligned}$$

Отсюда, в силу (7.4), получим $\forall j \in \mathbb{Z}$:

$$K_H(2\pi j) = \sum_{k=-\infty}^{\infty} h(k) e^{-i2\pi j k} = \sum_{k=-\infty}^{\infty} h(-k) e^{-i2\pi j k} = K_{\tilde{H}}(2\pi j) = \sum_{k \in \mathbb{Z}} h_k = \sqrt{2}. \quad (7.28)$$

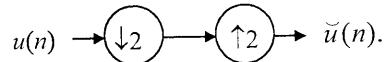
Тогда на основании (7.27), (7.28) $\forall j \in \mathbb{Z}$:

$$K_H((2j+1)\pi) = \sum_{k=-\infty}^{\infty} h(k) e^{-i\pi k(2j+1)} = K_{\tilde{H}}((2j+1)\pi) = \sum_{k \in \mathbb{Z}} (-1)^k h_k = 0. \quad (7.29)$$

Напомним, что нормированные ЧХ дискретных фильтров имеют период 2π , причем частоты $\omega_j = 2\pi j$ ($j \in \mathbb{Z}$) соответствуют постоянной составляющей, а частоты $\omega_j = (2j+1)\pi$ ($j \in \mathbb{Z}$) — верхней частоте обрабатываемого сигнала (см. главу 4). Из выражений (7.27)–(7.29) следует, что АЧХ $|K_{\tilde{H}}(\omega)| = |K_H(\omega)|$ принимает максимальное значение для постоянной составляющей, а верхние частоты дискретного сигнала полностью подавляются,

т. е. соответствующие фильтры должны быть фильтрами нижних частот.

Запишем условия, выполнение которых необходимо для точного восстановления сигнала, обработанного системой КЗФ (рис. 7.9). Предварительно установим выражение для Z -образа сигнала, у которого отсчеты с нечетными номерами принудительно обнулены, т.е. сигнал подвергнут обработке по следующей схеме:



Очевидно, что

$$\begin{aligned}\check{U}(z) &= \sum_{n \in \mathbb{Z}} \check{u}(n) z^{-n} = \sum_{n \in \mathbb{Z}} u(2n) z^{-2n} = \\ &= \sum_{n \in \mathbb{Z}} u(n) \frac{z^{-n} + (-z)^{-n}}{2} = \frac{U(z) + U(-z)}{2}.\end{aligned}$$

Поэтому для точного восстановления сигнала при обработке по схеме, представленной на рис. 7.9, необходимо выполнение следующего условия:

$$\frac{X(z)\tilde{H}(z) + X(-z)\tilde{H}(-z)}{2} H(z) + \frac{X(z)\tilde{G}(z) + X(-z)\tilde{G}(-z)}{2} G(z) = X(z).$$

Отсюда, учитывая, что $\tilde{H}(z) = H(z^{-1})$, $\tilde{G}(z) = G(z^{-1})$, получим:

$$\begin{aligned}X(z)(H(z^{-1})H(z) + G(z^{-1})G(z)) + \\ + X(-z)(H(-z^{-1})H(z) + G(-z^{-1})G(z)) = 2X(z).\end{aligned}$$

Поскольку последнее соотношение должно выполняться для любых сигналов $x(n)$ и соответствующих им Z -образов $X(z)$, окончательно условия точного восстановления сигнала примут вид:

$$\begin{cases} H(z^{-1})H(z) + G(z^{-1})G(z) = 2; \\ H(-z^{-1})H(z) + G(-z^{-1})G(z) = 0. \end{cases} \quad (7.30)$$

Рассмотрим полученные соотношения (7.30) для $z = e^{i\omega}$.

Учитывая, что

$$\begin{aligned}H(e^{i\omega}) &= K_H(\omega), \\ H(e^{-i\omega}) &= \sum_{n \in \mathbb{Z}} h(n) e^{i\omega n} = \overline{K_H(\omega)}, \\ H(-e^{-i\omega}) &= \sum_{n \in \mathbb{Z}} h(n) (-e^{i\omega})^n = \sum_{n \in \mathbb{Z}} h(n) e^{i(\omega+\pi)n} = \overline{K_H(\omega+\pi)}\end{aligned}$$

и проводя аналогичные замены для G -фильтра, получим из (7.30) условия точного восстановления сигнала в терминах частотных характеристик:

$$\begin{cases} |K_H(\omega)|^2 + |K_G(\omega)|^2 = 2; \\ \overline{K_H(\omega+\pi)}K_H(\omega) + \overline{K_G(\omega+\pi)}K_G(\omega) = 0. \end{cases} \quad (7.31)$$

Эти равенства должны выполняться для любых значений частоты $\omega \in \mathbb{R}$. Из (7.27) и первого уравнения системы (7.31) следует, что

$$|K_G(\omega)| = |K_H(\omega + \pi)|.$$

Отсюда с учетом (7.28) и 2π -периодичности ЧХ получим $\forall j \in \mathbb{Z}$:

$$K_{\tilde{G}}(2\pi j) = K_G(2\pi j) = \sum_{n \in \mathbb{Z}} g_n = 0, \quad (7.32)$$

а на основании (7.29) имеем $\forall j \in \mathbb{Z}$:

$$|K_{\tilde{G}}((2j+1)\pi)| = |K_G((2j+1)\pi)| = \left| \sum_{k \in \mathbb{Z}} (-1)^k g_k \right| = \sqrt{2}, \quad (7.33)$$

т. е. в фильтрах G и \tilde{G} постоянная составляющая сигнала подавляется, а для верхней частоты АЧХ принимает максимальное возможное значение. Поэтому фильтры G и \tilde{G} являются фильтрами верхних частот.

Итак, пусть задан набор коэффициентов $\{h_n\}$ в масштабирующем уравнении (7.1) и тем самым определены низкочастотные фильтры H и \tilde{H} . Тогда фильтры G и \tilde{G} , ИХ которых определяют материнский вейвлет по уравнению (7.3), являются высокочастотными, причем их ЧХ должны удовлетворять условиям (7.32) и (7.33).

Выразим ЧХ G -фильтра через ЧХ H -фильтра. Будем искать $K_G(\omega)$ в виде

$$K_G(\omega) = e^{ib(\omega)} \overline{K_H(\omega + \pi)},$$

где $b(\omega)$ — некоторая вещественная функция, такая что $e^{ib(\omega)} = e^{ib(\omega+2\pi)}$. Подставив выражение для $K_G(\omega)$ во второе уравнение системы (7.31), получим с учетом 2π -периодичности функции $K_H(\omega)$:

$$\overline{K_H(\omega + \pi)} K_H(\omega) + e^{-ib(\omega+\pi)} K_H(\omega) e^{ib(\omega)} \overline{K_H(\omega + \pi)} = 0.$$

Поскольку ЧХ $K_H(\omega)$ отлична от тождественного нуля, после сокращений получим из последнего равенства соотношение $e^{-i(b(\omega+\pi)-b(\omega))} = -1$, откуда $\forall \omega$:

$$b(\omega + \pi) - b(\omega) = (2k-1)\pi, \quad k \in \mathbb{Z}.$$

Полученное уравнение задает линейную функцию $b(\omega) = (2k-1)\omega + C$, где k — любое целое число, а C — произвольная вещественная константа, поэтому

$$K_G(\omega) = e^{i((2k-1)\omega+C)} \overline{K_H(\omega + \pi)}.$$

Отсюда

$$\begin{aligned} K_G(\omega) &= e^{i((2k-1)\omega+C)} \sum_{n \in \mathbb{Z}} h_n e^{in(\omega+\pi)} = e^{iC} \sum_{n \in \mathbb{Z}} h_n (-1)^n e^{i(2k+n-1)\omega} = \\ &= e^{iC} \sum_{n \in \mathbb{Z}} h_{n-2k+1} (-1)^{n-2k+1} e^{i\omega n} = e^{iC} \sum_{n \in \mathbb{Z}} h_{n-2k+1} (-1)^{n+1} e^{i\omega n} = \\ &= e^{iC} \sum_{n \in \mathbb{Z}} h_{1-n-2k} (-1)^{1-n} e^{-i\omega n} = \sum_{n \in \mathbb{Z}} g_n e^{-i\omega n}. \end{aligned}$$

Последнее равенство всегда будет выполняться, если ИХ G -фильтра задана по ИХ H -фильтра следующим образом: $g_n = e^{iC} h_{1-n-2k} (-1)^{1-n}$. Поскольку рассматриваются фильтры с вещественными ИХ, константу C нужно выбирать так, чтобы множитель $\alpha = e^{iC}$ также был вещественным числом, т. е. $C = \pi m$, $m \in \mathbb{Z}$, и $\alpha \in \{-1, 1\}$.

Окончательно получим, что условие точного восстановления сигнала после фильтрации по схеме рис. 7.9 устанавливает сле-

дующую связь ИХ КЗФ (и, соответственно, коэффициентов в масштабирующих уравнениях (7.1) и (7.3)):

$$g_n = \alpha h_{1-n-2k} (-1)^{1-n}, \quad (7.34)$$

где число $k \in \mathbb{Z}$ и множитель $\alpha \in \{-1, 1\}$ выбираются одними и теми же для всех значений индекса $n \in \mathbb{Z}$. При этом для ЧХ имеем

$$K_G(\omega) = \alpha e^{i(2k-1)\omega} \overline{K_H(\omega + \pi)}, \quad (7.35)$$

а для передаточных функций —

$$G(z) = \alpha z^{2k-1} H(-z^{-1}). \quad (7.36)$$

Соотношение (7.34) уже использовалось ранее при построении вейвлета Хаара в примере 7.1.

Пример 7.5. Построить графики АЧХ $|K_H(\omega)|$ и $|K_G(\omega)|$ для КЗФ, определяемых вейвлетом Хаара (см. пример 7.1).

◀ Записывая по определению ЧХ H -фильтра и учитывая (7.35), получаем:

$$|K_H(\omega)| = |K_H(\omega)| = \left| \frac{1}{\sqrt{2}} + \frac{e^{-i\omega}}{\sqrt{2}} \right| = \sqrt{2} \left| e^{-i\omega/2} \left| \frac{e^{i\omega/2} + e^{-i\omega/2}}{2} \right| \right| = \sqrt{2} \left| \cos \frac{\omega}{2} \right| = \sqrt{1 + \cos \omega},$$

$$|K_G(\omega)| = |K_G(\omega)| = |K_H(\omega + \pi)| = \sqrt{1 + \cos(\omega + \pi)} = \sqrt{1 - \cos \omega} = \sqrt{2} \left| \sin \frac{\omega}{2} \right|.$$

Графики АЧХ приведены на рис. 7.10. ►

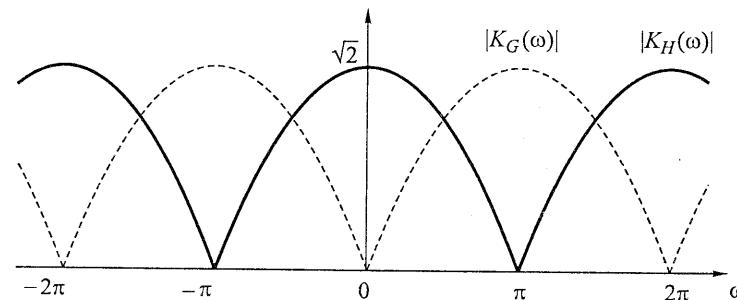


Рис. 7.10. Графики АЧХ для КЗФ, порождаемых вейвлетом Хаара

Данный пример наглядно иллюстрирует отмеченное ранее общее свойство КЗФ: \tilde{H} и H фильтры являются низкочастотными, а \tilde{G} и G — высокочастотными. Процедуру фильтрации дискретного сигнала парой фильтров (рис. 7.5) с частотной точки зрения можно пояснить как «расщепление» частотного спектра дискретного сигнала — последовательности $\{\dots, a_{j,k}, a_{j,k+1}, \dots\}$ — на два частотных диапазона равной ширины, низкочастотного (отклик \tilde{H} -фильтра) и высокочастотного (отклик \tilde{G} -фильтра). При этом (в идеале!), если для входной последовательности максимально возможная верхняя частота сигнала равна $F_B = f_d/2$ (половина частоты дискретизации), то частотный спектр отклика \tilde{H} -фильтра будет иметь вдвое меньшую верхнюю частоту.

■ 7.6. Построение масштабирующих функций и вейвлетов по масштабирующему уравнениям

Как упоминалось в разделе 7.1, набор коэффициентов $\{h_n\}$ уравнения (7.1) позволяет задать масштабирующую функцию $\phi(x)$. Приведем два рекуррентных численных метода, позволяющих найти функцию $\phi(x)$ по заданным коэффициентам уравнений (7.1) и (7.3). Поскольку коэффициенты масштабирующего уравнения для вейвлетов (7.3) выражаются через $\{h_n\}$ с помощью соотношения (7.34), то задания набора $\{h_n\}$ достаточно также и для построения функции вейвлета.

Ранее отмечалось, что особый интерес для практики представляет КМА, порождаемый масштабирующей функцией с ограниченным носителем. Пусть носитель масштабирующей функции $\Delta_{0,0} = \text{supp } \phi(t) = [0; M-1]$, где целое число $M \geq 2$. Тогда при $k \notin \{0, \dots, M-1\}$ носители

$$\Delta_{1,k} = \text{supp } \phi(2t-k) = [k/2; (M-1+k)/2] \subsetneq \Delta_{0,0}.$$

Поэтому уравнение (7.1) можно записать в виде конечной суммы:

$$\phi(x) = \sqrt{2} \sum_{n=0}^{M-1} h_n \phi(2x-n). \quad (7.37)$$

Возьмем четное число M , тогда, обозначив $2k = 2 - M$ на основании (7.34), для уравнения (7.3) получаем:

$$\psi(x) = \sqrt{2} \sum_{n=0}^{M-1} g_n \phi(2x-n) = \sqrt{2} \sum_{n=0}^{M-1} (-1)^n h_{M-1-n} \phi(2x-n). \quad (7.38)$$

Заметим, что запись (7.38) означает также, что $\text{supp } \psi(t) \subset [0; M-1]$.

Перейдем к описанию методов нахождения функции $\phi(x)$.

Первый метод. Из формул (7.9) и (7.12) следует, что

$$f(x_n) 2^{-N/2} = a_{N,n} + O(2^{-3N/2}),$$

где $a_{N,n} = \langle f, \phi_{N,n} \rangle$, x_n — некоторая точка отрезка $\Delta_{N,n} = \text{supp } \phi_{N,n}(x)$ (например, середина), причем $\forall n$:

$$h = x_{n+1} - x_n = 2^{-N}.$$

Поэтому для отсчетов функции $f(x_n)$, взятых с шагом h , можно записать:

$$f(x_n) = 2^{N/2} a_{N,n} + O(2^{-N}). \quad (7.39)$$

Отсюда видно, что, имея набор коэффициентов $\{a_{N,n}\}_{n \in \mathbb{Z}}$, который определяет проекцию $f_N = A_N(f) \in V_N$, можно с погрешностью порядка $O(h)$ найти значения функции $f(x)$ в точках x_n , взятых с шагом $h = 1/2^N$. То есть, увеличивая N , можно как угодно точно описать функцию $f(x)$: погрешность определения значения функции в отсчетах $f(x_n)$ по формуле (7.39) будет уменьшаться, а сами отсчеты будут располагаться на числовой оси все более часто.

С учетом (7.39) процедура нахождения масштабирующей функции $\phi(x)$ при известном наборе коэффициентов $\{h_n\}$ состоит в следующем.

Сначала заметим, что $\phi(x) \in V_0 \subset V_1 \subset \dots$ и $\phi(x) \perp W_0$, $\phi(x) \perp W_1, \dots$. Так как $\phi(x) = \phi_{0,0}(x)$ — базисная функция подпространства V_0 , то для разложения (7.5) функции $\phi(x)$ в V_0 получим $a_{0,0} = 1$ и $\forall k \neq 0$: $a_{0,k} = 0$. При этом (т. к. $\phi(x) \perp W_0$) $\forall k$: $c_{0,k} = \langle \phi(x), \psi_{0,k}(x) \rangle = 0$. Поэтому с использованием формулы

(7.22) можно найти коэффициенты $\{a_{l,k}\}_{k \in \mathbb{Z}}$ разложения масштабирующей функции $\phi(x) \in V_0$ по базису пространства V_1 , получим:
 $a_{l,k} = h_k, k \in \mathbb{Z}$.

Рекуррентно применяя полученную из (7.22) формулу

$$a_{m,k} = \sum_{n \in \mathbb{Z}} a_{m-1,n} h_{k-2n}$$

далее для $m = 2, 3, \dots, N$, можно получить как угодно точное описание масштабирующей функции $\phi(x)$ через ее отсчеты в точках $x_n \in \Delta_{N,n}$. Для этого необходимо воспользоваться выражением (7.39), положив, что $\phi(x) = f(x)$.

Второй метод. Запишем масштабирующее уравнение (7.37) для целых значений аргумента $j = 0, \dots, M-1$:

$$\phi(j) = \sqrt{2} \sum_{n=0}^{M-1} h_n \phi(2j-n).$$

Или в матричном виде:

$$\begin{pmatrix} \phi(0) \\ \phi(1) \\ \dots \\ \phi(M-1) \end{pmatrix} = \mathbf{B} \begin{pmatrix} \phi(0) \\ \phi(1) \\ \dots \\ \phi(M-1) \end{pmatrix}, \quad (7.40)$$

где

$$\mathbf{B} = \sqrt{2} \begin{pmatrix} h_0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ h_1 & h_0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ h_2 & h_3 & h_2 & h_1 & h_0 & \dots & 0 & 0 & 0 & 0 \\ \dots & \dots \\ h_{M-4} & h_{M-5} & h_{M-6} & h_{M-7} & h_{M-8} & \dots & h_1 & h_0 & 0 & 0 \\ h_{M-2} & h_{M-3} & h_{M-4} & h_{M-5} & h_{M-6} & \dots & h_3 & h_2 & h_1 & h_0 \\ 0 & h_{M-1} & h_{M-2} & h_{M-3} & h_{M-4} & \dots & h_5 & h_4 & h_3 & h_2 \\ 0 & 0 & 0 & h_{M-1} & h_{M-2} & \dots & h_7 & h_6 & h_5 & h_4 \\ \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & h_{M-1} & h_{M-2} & h_{M-3} & h_{M-4} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & h_{M-1} & h_{M-2} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & h_{M-1} \end{pmatrix}.$$

Равенство (7.40) означает, что матрица \mathbf{B} имеет собственное число $\lambda = 1$, которому соответствует собственный вектор

$$\Phi_0 = (\phi(0), \phi(1), \dots, \phi(M-1))^T.$$

Уравнение (7.40) определяет следующий способ рекуррентного построения масштабирующей функции $\phi(x)$ по набору коэффициентов $\{h_n\}_{n=0}^{M-1}$.

Сначала находим собственный вектор Φ_0 матрицы \mathbf{B} , который соответствует собственному числу $\lambda = 1$. Так как собственный вектор Φ_0 матрицы \mathbf{B} определяется с точностью до масштабного множителя c (вектор $c\Phi_0$ также будет собственным для матрицы \mathbf{B}), то для выбора $\Phi_0 = (\phi(0), \phi(1), \dots, \phi(M-1))^T$ из собственных векторов матрицы \mathbf{B} , соответствующих собственному числу $\lambda = 1$, введем дополнительное требование, которое мы сформулируем в виде следующей леммы.

Лемма 7.3. Масштабирующая функция $\phi(x)$ должна удовлетворять условию

$$\sum_{n \in \mathbb{Z}} \phi(x-n) = 1. \quad (7.41)$$

◀ Рассмотрим функцию $u(x) = \sum_{n \in \mathbb{Z}} \phi(x-n)$. Учитывая, что

данная функция имеет единичный период, представим ее в виде ряда Фурье:

$$u(x) = \sum_{m \in \mathbb{Z}} c_m e^{i2\pi mx}.$$

Здесь

$$\begin{aligned} c_m &= \int_0^1 u(x) e^{-i2\pi mx} dx = \sum_{n \in \mathbb{Z}} \int_0^1 \phi(x-n) e^{-i2\pi mx} dx = \sum_{n \in \mathbb{Z}} \int_{-n}^{1-n} \phi(t) e^{-i2\pi m(t+n)} dt = \\ &= \sum_{n \in \mathbb{Z}} e^{-i2\pi mn} \int_{-n}^{1-n} \phi(t) e^{-i2\pi mt} dt = \int_{-\infty}^{+\infty} \phi(t) e^{-i2\pi mt} dt = \sqrt{2\pi} S_\phi(2\pi m), \end{aligned}$$

где

$$S_\phi(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(t) e^{-i\omega t} dt.$$

Так как $S_\phi(0) = \frac{1}{\sqrt{2\pi}}$ (см. (7.4)), то из соотношения (7.26) для $\omega = 0$ следует, что $\forall m \in \mathbb{Z}$:

$$S_\phi(2\pi m) = \begin{cases} 1/\sqrt{2\pi} & \text{при } m = 0; \\ 0 & \text{при } m \neq 0. \end{cases}$$

Поэтому $c_m = \begin{cases} 1 & \text{при } m = 0; \\ 0 & \text{при } m \neq 0 \end{cases}$ и $u(x) = 1$. ►

Положив в (7.41) $x = 0$, находим, что искомый собственный вектор Φ_0 матрицы \mathbf{B} из уравнения (7.40) должен удовлетворять условию $\sum_{n=0}^{M-1} \phi(n) = 1$. После определения собственного вектора Φ_0 , отвечающего данному требованию, далее последовательно находим значения масштабирующей функции в точках $t_j = j2^{-m}$, $j \in \mathbb{Z}$, рекуррентно используя уравнение (7.37) для $m = 1, 2, \dots$:

$$\phi(j2^{-m}) = \sqrt{2} \sum_{n=0}^{M-1} h_n \phi(j2^{-m+1} - n). \quad (7.42)$$

В результате получаем значения масштабирующей функции на множестве двоично-рациональных значений аргумента. Реально вычисления по формуле (7.42) прекращаются при достижении достаточно большого значения $m = K$, когда для масштабирующей функции известен вектор отсчетов $\Phi_K = (\phi(0), \phi(2^{-K}), \dots, \phi(M-1))^T$, заданных на сетке значений аргумента t с шагом $\Delta t = 2^{-K}$: $t_j = j\Delta t$, $j = 0, 1, \dots, 2^K(M-1)$.

Задание. По аналогии с описанными процедурами итерационного построения масштабирующей функции опишите процедуры нахождения материнского вейвлета $\psi(x)$ по заданному набору коэффициентов $\{h_n\}$ масштабирующих уравнений (7.37) и (7.38).

■ 7.7. Примеры синтеза вейвлетов

Один из способов построения вейвлетов состоит в определении условий, которым должны удовлетворять коэффициенты $\{h_n\}$ из уравнения (7.1), и поиске таких наборов $\{h_n\}$, которые этим условиям отвечают. Ранее для коэффициентов $\{h_n\}$ были установлены два необходимых требования, которые определяются выражениями (7.28) и (7.29). Дополнительное ограничение на коэффициенты масштабирующего уравнения (7.1) вытекает из ортонормированности системы функций $\{\phi(x-n)\}_{n \in \mathbb{Z}}$. Обозначая

$$\delta_{m,n} = \begin{cases} 1 & m = n; \\ 0 & m \neq n, \end{cases} \quad \text{получим } \forall m \in \mathbb{Z}:$$

$$\begin{aligned} \delta_{0,m} &= \langle \phi(x), \phi(x-m) \rangle = 2 \int_{-\infty}^{\infty} \sum_{j \in \mathbb{Z}} h_j \phi(2x-j) \sum_{k \in \mathbb{Z}} h_k \phi(2(x-m)-k) dx = \\ &= \sum_{j \in \mathbb{Z}} h_j \sum_{k \in \mathbb{Z}} h_k \int_{-\infty}^{\infty} \phi(2x-j) \phi(2x-(2m+k)) d(2x) = \\ &= \sum_{j \in \mathbb{Z}} h_j \sum_{k \in \mathbb{Z}} h_k \langle \phi(x-j), \phi(x-(2m+k)) \rangle = \sum_{j \in \mathbb{Z}} h_j \sum_{k \in \mathbb{Z}} h_k \delta_{j,2m+k} = \sum_{j \in \mathbb{Z}} h_j h_{j-2m}. \end{aligned}$$

Таким образом, $\forall m \in \mathbb{Z}$:

$$\sum_{j \in \mathbb{Z}} h_j h_{j-2m} = \delta_{0,m}. \quad (7.43)$$

Далее мы убедимся, что уравнения (7.28), (7.29) и (7.43) оставляют достаточный простор для выбора коэффициентов $\{h_n\}$. Дополнительное условие, которое налагается на $\{h_n\}$, может вытекать, например, из требования точной аппроксимации произвольного степенного полинома (по возможности более высокого порядка N) при разложении по базису в пространстве V_0 . Чем выше степень N полинома, для которого возможно точное представление

$$P_N(t) = \sum_{m=0}^N b_m t^m = \sum_{n \in \mathbb{Z}} a_{0,n} \phi(t-n), \quad (7.44)$$

тем выше порядок формулы Тейлора, которую можно использовать для описания функций в подпространстве V_0 и, следовательно, тем точнее можно задать произвольную N -кратно дифференцируемую функцию $f(t) \in L_2(\mathbb{R})$ ее проекцией $f_0(t) \in V_0$.

Из возможности представления (7.44) для произвольного набора коэффициентов $\{b_m\}_{m=0}^{m=N}$ и ортогональности соответствующего полинома $P_N(t) \in V_0$ материнскому вейвлету $\psi(t) \in W_0$ следует требование

$$\int_{-\infty}^{+\infty} t^m \psi(t) dt = 0, \quad m = 0, \dots, N. \quad (7.45)$$

При поиске вейвлетов, удовлетворяющих условию (7.45), будем исходить из требований ограниченности носителя масштабирующей функции отрезком $[0; M-1]$ ($M = 2L$, $L \in \mathbb{N}$), когда масштабирующие уравнения (7.1) и (7.3) записываются в виде (7.37), (7.38).

Лемма 7.4. Для того чтобы выполнялось условие (7.45), необходимо и достаточно выполнения требования¹

$$\sum_{n=0}^{M-1} n^m g_n = 0, \quad m = 0, \dots, N. \quad (7.46)$$

◀ Подставим выражение для материнского вейвлета (7.38) в (7.45):

$$\int_{-\infty}^{+\infty} t^m \sum_{n=0}^{M-1} g_n \phi(2t-n) dt = 0.$$

Выполняя замену переменной $x = 2t - n$, получим из (7.45) эквивалентные выражения:

$$\sum_{n=0}^{M-1} g_n \int_0^{M-1} (x+n)^m \phi(x) dx = 0, \quad m = 0, \dots, N$$

(т. к. $\phi(x) = 0$ при $x \notin [0; M-1]$). Рассмотрим полученные равенства последовательно для $m = 0, \dots, N$.

При $m = 0$ имеем:

$$\left(\sum_{n=0}^{M-1} g_n \right) \int_0^{M-1} \phi(x) dx = 0,$$

что с учетом (7.4) эквивалентно условию (7.46) для $m = 0$:

$$\sum_{n=0}^{M-1} g_n = 0.$$

При $m = 1$:

$$\begin{aligned} & \sum_{n=0}^{M-1} g_n \int_0^{M-1} (x+n) \phi(x) dx = \\ & = \underbrace{\left(\sum_{n=0}^{M-1} g_n \right) \int_0^{M-1} x \phi(x) dx}_{0} + \underbrace{\left(\sum_{n=0}^{M-1} g_n n \right) \int_0^{M-1} \phi(x) dx}_{1} = 0, \end{aligned}$$

что равносильно требованию (7.46) для $m = 1$:

$$\sum_{n=0}^{M-1} g_n n = 0.$$

Продолжая рассуждения по индукции, в предположении эквивалентности утверждений (7.45) и (7.46) для $m = 0, \dots, N-1$, получим для $m = N$:

$$\begin{aligned} & \sum_{n=0}^{M-1} g_n \int_0^{M-1} (x+n)^N \phi(x) dx = \sum_{n=0}^{M-1} g_n \int_0^{M-1} \left(\sum_{k=0}^N C_N^k x^{N-k} n^k \right) \phi(x) dx = \\ & = \sum_{k=0}^N C_N^k \left(\int_0^{M-1} x^{N-k} \phi(x) dx \underbrace{\sum_{n=0}^{M-1} n^k g_n}_{=0 \text{ при } k=0, \dots, N-1} \right) = \underbrace{\int_0^{M-1} \phi(x) dx}_{1} \sum_{n=0}^{M-1} n^N g_n = 0, \end{aligned}$$

что равносильно выполнению условия (7.46) при $m = N$.

Справедливость утверждения леммы доказана. ►

¹ Считаем, что при $n = m = 0$: $n^m = 1$.

Теорема 7.1. Условие (7.45) эквивалентно требованию

$$\sum_{n=0}^{M-1} n^m (-1)^n h_n = 0, \quad m = 0, \dots, N. \quad (7.47)$$

◀ Учитывая связь $g_n = (-1)^n h_{M-1-n}$, $n = 0, \dots, M-1$, $M = 2L$ (см. (7.38)), в силу четности числа M можно записать: $h_n = (-1)^{1-n} g_{M-1-n}$. Тогда $\forall m = 0, \dots, N$ имеем:

$$\begin{aligned} \sum_{n=0}^{M-1} n^m (-1)^n h_n &= \sum_{n=0}^{M-1} n^m (-1)^n g_{M-1-n} (-1)^{1-n} = - \sum_{n=0}^{M-1} n^m g_{M-1-n} = \\ &= - \sum_{n=0}^{M-1} (M-1-n)^m g_n = (-1)^{m+1} \sum_{n=0}^{M-1} (n+1-M)^m g_n = \\ &= (-1)^{m+1} \sum_{n=0}^{M-1} g_n \sum_{j=0}^m C_m^j n^j (1-M)^{m-j} = (-1)^{m+1} \sum_{j=0}^m C_m^j (1-M)^{m-j} \sum_{n=0}^{M-1} g_n n^j. \end{aligned}$$

Рассмотрим полученное равенство последовательно для $m = 0, \dots, N$. Для $m = 0$ условия (7.46) и (7.47) эквивалентны, поскольку имеем:

$$\sum_{n=0}^{M-1} (-1)^n h_n = - \sum_{n=0}^{M-1} g_n = 0.$$

Для $m = 1$ вновь получаем равносильность требований (7.46) и (7.47), так как

$$\sum_{n=0}^{M-1} n(-1)^n h_n = C_1^0 (1-M) \underbrace{\sum_{n=0}^{M-1} g_n}_0 + C_1^1 (1-M)^0 \sum_{n=0}^{M-1} g_n n,$$

т. е.

$$\sum_{n=0}^{M-1} n(-1)^n h_n = \sum_{n=0}^{M-1} g_n n = 0.$$

Последовательно доказав эквивалентность (7.46) и (7.47) для значений $m = 2, 3, \dots, N-1$, получим и для $m = N$:

$$\sum_{n=0}^{M-1} n^N (-1)^n h_n = (-1)^{N+1} C_N^N (1-M)^0 \sum_{n=0}^{M-1} g_n n^N = (-1)^{N+1} \sum_{n=0}^{M-1} g_n n^N = 0.$$

Равносильность условий (7.46) и (7.47) доказана. ►

Таким образом, мы определили необходимые (обязательные) требования, которым должны отвечать коэффициенты масштабирующего уравнения — равенства (7.28), (7.29), (7.43), и дополнительные — условия (7.47). Заметим также, что условие (7.29) эквивалентно записи (7.47) для $m = 0$. Поэтому будем искать коэффициенты масштабирующих уравнений (7.37), (7.38), исходя из условий (7.28), (7.43) и (7.47).

Пример 7.6 (вейвлет Хаара — «Добеши-1»). В примере 7.1, рассмотренном ранее, было «угадано» возможное решение уравнения (7.1). Найти данное решение можно по рассмотренной выше методике.

◀ Если положить $M = 2$ (в масштабирующем уравнении — два ненулевых коэффициента: h_0, h_1), то условия (7.43), (7.28) примут вид системы уравнений

$$\begin{cases} h_0^2 + h_1^2 = 1; \\ h_0 + h_1 = \sqrt{2}, \end{cases}$$

которая имеет единственное решение: $h_0 = h_1 = 1/\sqrt{2}$. Заметим, что полученное решение удовлетворяет также и соотношению (7.47) для $m = 0$. ►

Пример 7.7 (вейвлет «Добеши-2»).

◀ Положим $M = 4$. Необходимые условия (7.43), (7.28) дают три уравнения, а четвертое уравнение запишем, положив в (7.47) $m = 0$. Получим систему уравнений:

$$\begin{cases} h_0^2 + h_1^2 + h_2^2 + h_3^2 = 1; \\ h_0 h_2 + h_1 h_3 = 0; \\ h_0 + h_1 + h_2 + h_3 = \sqrt{2}; \\ h_0 - h_1 + h_2 - h_3 = 0. \end{cases}$$

Заменим третье и четвертое уравнения системы на их сумму и разность:

$$\begin{cases} h_0^2 + h_1^2 + h_2^2 + h_3^2 = 1; \\ h_0 h_2 + h_1 h_3 = 0; \\ h_0 + h_2 = 1/\sqrt{2}; \\ h_1 + h_3 = 1/\sqrt{2}. \end{cases}$$

Тогда из второго, третьего и четвертого уравнений следует первое:

$$1 = (h_0 + h_2)^2 + (h_1 + h_3)^2 = h_0^2 + h_1^2 + h_2^2 + h_3^2 + 2(h_0 h_2 + h_1 h_3).$$

Следовательно, первое уравнение в системе — избыточное, его можно исключить. Оставшиеся три уравнения системы дополним четвертым, которое представляет собой условие (7.47) для $m = 1$. Полученная система уравнений

$$\begin{cases} -h_1 + 2h_2 - 3h_3 = 0; \\ h_0 h_2 + h_1 h_3 = 0; \\ h_0 + h_2 = 1/\sqrt{2}; \\ h_1 + h_3 = 1/\sqrt{2} \end{cases}$$

имеет две пары «зеркальных» решений (убедитесь!):

$$\begin{cases} h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}} \end{cases},$$

$$\begin{cases} h_0 = \frac{1-\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3-\sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1+\sqrt{3}}{4\sqrt{2}} \end{cases}.$$

Обычно вейвлетом «Добеши-2» называют функцию, порождающую первым набором решений. Для построения графиков масштабирующей функции и вейвлета (рис. 7.11) можно воспользоваться одним из двух методов, описанных в предыдущем разделе. ►

Увеличивая число $M = 2L$, можно построить другие вейвлеты и масштабирующие функции семейства «Добеши- L » (Ингрид Добеши — математик, которой принадлежат многие основополагающие результаты теории вейвлет-преобразований и КМА). На

рис. 7.12 приведен пример масштабирующей функции и вейвлета, которые получены для $L = 10$.

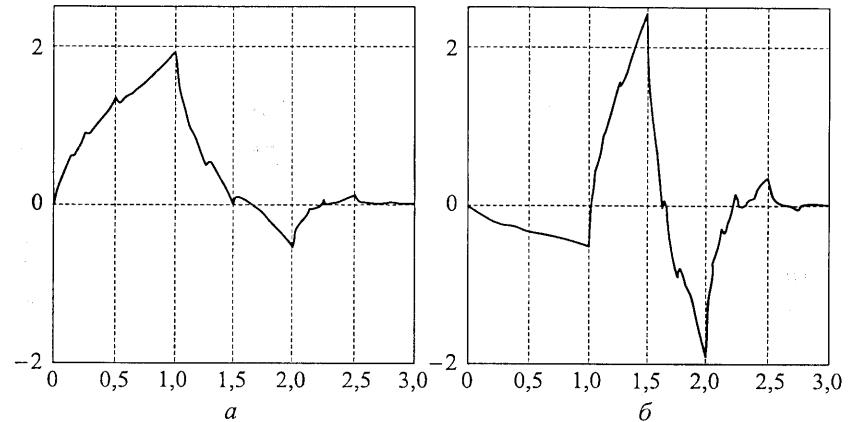


Рис. 7.11. Графики масштабирующей функции (а) и материального вейвлета (б), соответствующие КЗФ «Добеши-2»

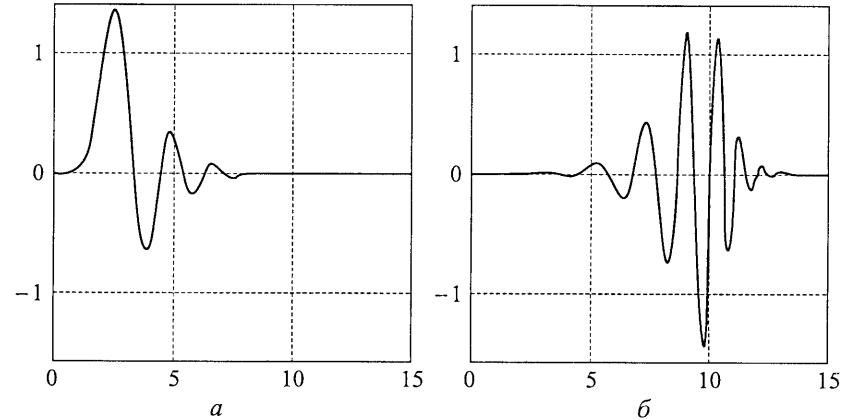


Рис. 7.12. Графики масштабирующей функции (а) и материального вейвлета (б), соответствующие КЗФ «Добеши-10»

Рассмотренный подход, основанный на решении системы уравнений (7.28), (7.43), (7.47) — не единственный. Так, например, построение вейвлетов и масштабирующих функций (соответственно, синтез КЗФ) можно проводить, накладывая различные условия не на импульсные, а на частотные характеристики.

■ 7.8. Биортогональные вейвлет-преобразования

Требование конечности сумм в формулах (7.19), (7.20) (и соответственно, (7.1), (7.3)), необходимое для численной реализации дискретного вейвлет-преобразования, существенно снижает возможности выбора вейвлет-базисов, ограничивая его базисами с конечным носителем. Поэтому многие приведенные ранее рассуждения основывались на предположении об ограниченности носителей вейвлета и масштабирующей функции.

Ортогональные базисы вейвлетов с ограниченным носителем не всегда оказываются самыми удобными для представления сигналов. Более широкое распространение на практике получили вейвлет-базисы с «ослабленными» требованиями к ортогональности, так называемые *биортогональные* базисы. В этом случае функциональные базисы масштабирующих функций $\{\phi_{m,n}\} \subset V_m$ и вейвлетов $\{\psi_{m,n}\} \subset W_m$ уже не являются ортогональными, но выбираются так, чтобы в тех же пространствах для них существовала дуальная пара базисов (также неортогональных) $\{\tilde{\phi}_{m,n}\} \subset V_m$, $\{\tilde{\psi}_{m,n}\} \subset W_m$ таких, что $\langle \tilde{\phi}_{m,n}, \phi_{m,k} \rangle = \delta_{n,k}$, $\langle \tilde{\psi}_{m,n}, \psi_{m,k} \rangle = \delta_{n,k}$, где $\delta_{n,k} = \begin{cases} 1 & \text{при } n=k; \\ 0 & \text{при } n \neq k. \end{cases}$

Биортогональные базисы обладают взаимной симметрией. Поиск коэффициентов разложения по одному из базисов дуальной пары состоит в вычислении скалярных произведений $f(x)$ и функций другого дуального базиса. Так, если

$$f(x) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} \tilde{\psi}_{m,n}(x), \quad (7.48)$$

то $c_{m,n} = \langle f, \tilde{\psi}_{m,n} \rangle$ (убедитесь!).

При использовании биортогональных базисов коэффициенты в формулах декомпозиции (7.19), (7.20) и реконструкции (7.22) могут отличаться, хотя схема вычислений остается совершенно той же. Необходимые вычислительные процедуры вновь можно трактовать как линейную обработку сигнала по схеме рис. 7.9. Декомпозиция осуществляется с использованием формул (7.19) и (7.20), а формула восстановления (7.22) примет вид:

$$a_{m,k} = \sum_{n \in \mathbb{Z}} a_{m-1,n} \tilde{h}_{k-2n} + \sum_{n \in \mathbb{Z}} c_{m-1,n} \tilde{g}_{k-2n},$$

причем для биортогональных преобразований в общем случае $\tilde{h}_j \neq h_j$, $\tilde{g}_j \neq g_j$ и, более того, число ненулевых коэффициентов в импульсных характеристиках $\{\tilde{h}(n) = h_{-n}\}$, $\{\tilde{g}(n) = g_{-n}\}$, $\{h(n) = \tilde{h}_n\}$, $\{g(n) = \tilde{g}_n\}$ может быть различным.

Можно показать, что между коэффициентами биортогональных фильтров существует следующая связь: $g_n = (-1)^{1-n} \tilde{h}_{1-n-2m}$, $\tilde{g}_n = (-1)^{1-n} h_{1-n-2m}$, где m — произвольное целое число (одинаковое для всех значений индекса $n \in \mathbb{Z}$). Кроме того, $\sum_n h_n \tilde{h}_{n-2k} = \delta_{k,0}$ и $\sum_n g_n \tilde{g}_{n-2k} = \delta_{k,0}$ (оба последних условия устанавливаются аналогично тому, как это было сделано при выводе формулы (7.43)). Поэтому для того чтобы определить биортогональное вейвлет-преобразование, необходимо задать уже два набора коэффициентов: $\{h_n\}$ и $\{\tilde{h}_n\}$ или $\{g_n\}$ и $\{\tilde{g}_n\}$. Этим наборам на основании (7.1) и (7.3) соответствуют пары функций $\phi(x) \leftrightarrow \{h_n\}$, $\psi(x) \leftrightarrow \{g_n\}$, используемых для анализа, и две функции $\tilde{\phi}(x) \leftrightarrow \{\tilde{h}_n\}$, $\tilde{\psi}(x) \leftrightarrow \{\tilde{g}_n\}$, которые используются для синтеза сигнала.

Таким образом, требования к коэффициентам фильтров декомпозиции (см. рис. 7.5) и реконструкции (см. рис. 7.7) ослабеваются, соответствующие им базисы принадлежат более широкому классу функций и возможности выбора преобразований для разложения сигналов существенно расширяются. Недостатком биортогональных базисов является потеря возможности анализа ошибки представления сигнала в области преобразований при отбрасывании (или квантовании) коэффициентов вейвлет-разложения, так как равенство Парсеваля в этом случае не выполняется, и для представления (7.48) в общем случае $\langle f, f \rangle = \|f\|^2 \neq \sum_{m,n} c_{m,n}^2$. Обычно

стремятся строить такие биортогональные базисы, для которых равенство Парсеваля все же выполняется с некоторой приемлемой для практики точностью.

На сегодняшний день предложено большое количество биортогональных базисов и соответствующих им КЗФ, которые реализованы в том числе в различных программных средах. В качестве примера приведем набор коэффициентов биортогональных КЗФ, который получил широкое распространение в приложениях цифровой обработки сигналов.

Пример 7.8 (биортогональные сплайн-вейвлеты 9/7).

Импульсная характеристика $\{\tilde{h}(n) = h_{-n}\}$ НЧ-фильтра, используемого при декомпозиции сигнала, является симметричной и имеет девять ненулевых коэффициентов:

$$h_0 = 0,852699, \quad h_{\pm 1} = 0,377402, \quad h_{\pm 2} = -0,110624, \\ h_{\pm 3} = -0,023849, \quad h_{\pm 4} = 0,037828.$$

Импульсная характеристика $\{h(n) = \check{h}_n\}$ НЧ-фильтра синтеза также симметрична и имеет семь ненулевых коэффициентов:

$$\check{h}_0 = 0,788486, \quad \check{h}_{\pm 1} = 0,418092, \quad \check{h}_{\pm 2} = -0,040689, \quad \check{h}_{\pm 3} = -0,064539.$$

Коэффициенты импульсных характеристик ВЧ-фильтров анализа (декомпозиции) и синтеза (реконструкции) находятся соответственно по формулам $g_n = (-1)^{1-n} \check{h}_{1-n}$ и $\check{g}_n = (-1)^{1-n} h_{1-n}$. Графики функций анализа и синтеза, которые соответствуют приведенным значениям коэффициентов фильтров, изображены на рис. 7.13 и 7.14.

Рассмотренные в примере 7.8 фильтры применяются, в частности, в схеме компрессии данных, которая описывается стандартом сжатия фотографических изображений JPEG-2000.

■ 7.9. Применение вейвлет-преобразований для сжатия сигналов

Сжатие (компрессия) сигналов с потерей информации является наиболее важной областью применения вейвлет-преобразований. Основная идея большинства алгоритмов вейвлет-компрессии состоит в следующем.

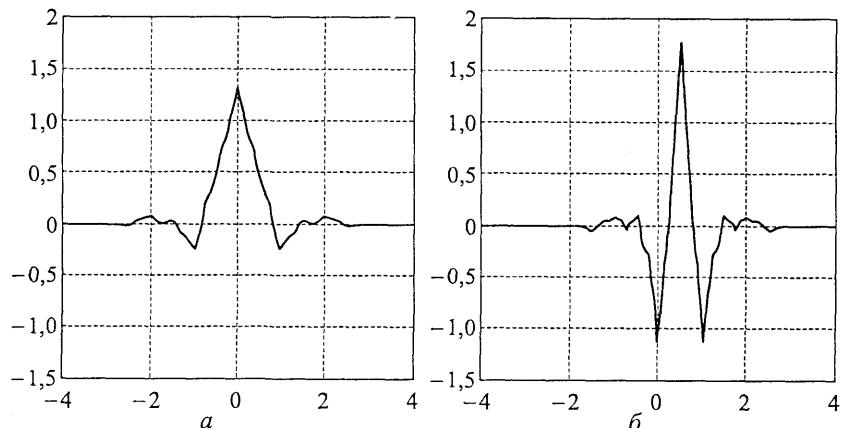


Рис. 7.13. Графики масштабирующей функции $\phi(x)$ (а) и материнского вейвлета $\psi(x)$ (б), соответствующие фильтрам анализа (декомпозиции) биортогонального преобразования 9/7

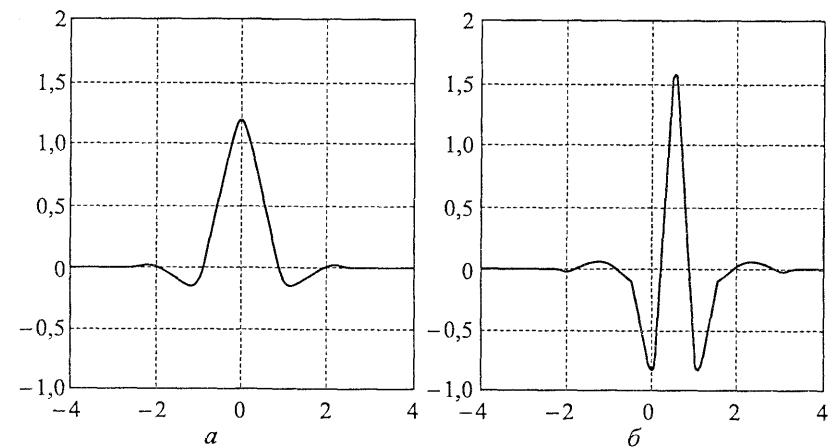


Рис. 7.14. Графики масштабирующей функции $\check{\phi}(x)$ (а) и материнского вейвлета $\check{\psi}(x)$ (б), соответствующие фильтрам синтеза (композиции) биортогонального преобразования 9/7

Для исходного дискретного сигнала f_M , заданного конечным набором (т. е. вектором) коэффициентов $\{a_{M,n}\}$ (7.5), сначала в результате выполнения дискретного вейвлет-преобразования по схеме, приведенной на рис. 7.2, получают коэффициенты разложения для представления сигнала в виде (7.21). Число коэффициентов для сглаженной проекции $f_K \leftrightarrow \{a_{K,n}\}$ будет существенно меньшим, чем в исходном векторе $\{a_{M,n}\}$, поскольку каждый шаг, соответствующий снижению уровня разрешения проекции f_j , $j = M-1, M-2, \dots, K$, соответствует двукратному сокращению числа компонент вектора $\{a_{j,n}\}$. Коэффициенты $\{a_{K,n}\}$ далее квантуются и статистически кодируются, обычно независимо друг от друга, поскольку представляют собой практически не коррелированные данные.

Вейвлет-коэффициенты $\{c_{K,n}\}, \{c_{K+1,n}\}, \dots, \{c_{M-2,n}\}, \{c_{M-1,n}\}$ обрабатываются по схеме, которая основана на специфических свойствах локализации функциональных вейвлет-базисов. Базисы вейвлет-функций $\{\psi_{m,n}(t)\}$ и соответствующие им коэффициенты $\{c_{m,n}(t)\}$ разложения обрабатываемого сигнала удобно упорядочить в виде древовидной структуры (рис. 7.15).

При таком упорядочивании каждый узел $\psi_{j,m}$ дерева (за исключением листвьев) имеет по два непосредственных (ближайших) «потомка» $\{\psi_{j+1,2m}, \psi_{j+1,2m+1}\}$, причем носители $\Delta_{j+1,2m} = \text{supp } \psi_{j+1,2m}(t)$ и $\Delta_{j+1,2m+1} = \text{supp } \psi_{j+1,2m+1}(t)$ соответствующих им базисных функций представляют собой области, лежащие внутри области носителя $\Delta_{j,m} = \text{supp } \psi_{j,m}(t)$ базисной функции-родителя: $\Delta_{j+1,2m} \subset \Delta_{j,m}$, $\Delta_{j+1,2m+1} \subset \Delta_{j,m}$. Действительно, пусть $\Delta_{0,0} = \text{supp } \psi(t) = [0; A]$. Заметим, что $A \geq 1$, так как в противном случае целочисленные сдвиги $\{\psi(t-n)\}_{n \in \mathbb{Z}}$ не покроют полностью числовую ось $t \in \mathbb{R}$. Тогда

$$\Delta_{j,m} = \left[\frac{m}{2^j}; \frac{m}{2^j} + \frac{A}{2^j} \right];$$

$$\Delta_{j+1,2m} = \left[\frac{2m}{2^{j+1}}; \frac{2m+A}{2^{j+1}} \right] = \left[\frac{m}{2^j}; \frac{m}{2^j} + \frac{A}{2^{j+1}} \right] \subset \Delta_{j,m};$$

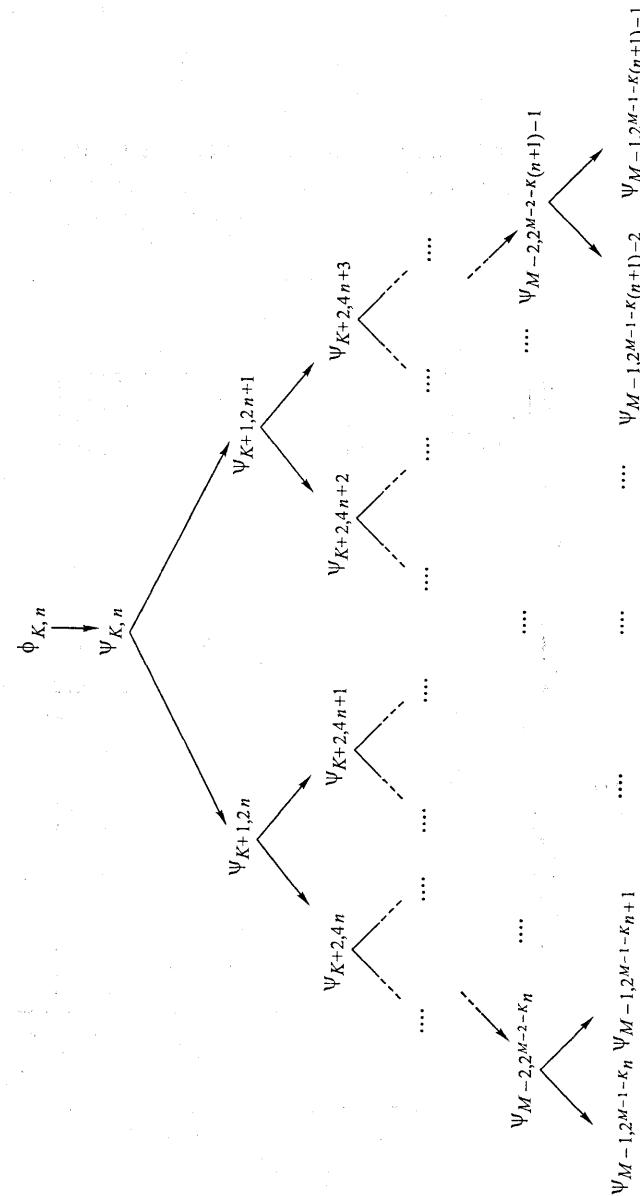


Рис. 7.15. Схема упорядочивания базисных функций в виде древовидной структуры

$$\Delta_{j+1,2m+1} = \left[\frac{2m+1}{2^{j+1}}, \frac{2m+1+A}{2^{j+1}} \right] = \left[\frac{m}{2^j} + \frac{1}{2^{j+1}}, \frac{m}{2^j} + \frac{A+1}{2^{j+1}} \right] \subset \Delta_{j,m},$$

так как для $A \geq 1$ справедливо: $\frac{A+1}{2^{j+1}} \leq \frac{A}{2^j}$.

Схематически локализация базисных функций на плоскости времени×частота отражена на рис. 7.16. Вейвлет-базис обладает «переменным» времея-частотным разрешением базисных функций (см. также раздел 6.7 и рис. 6.9).

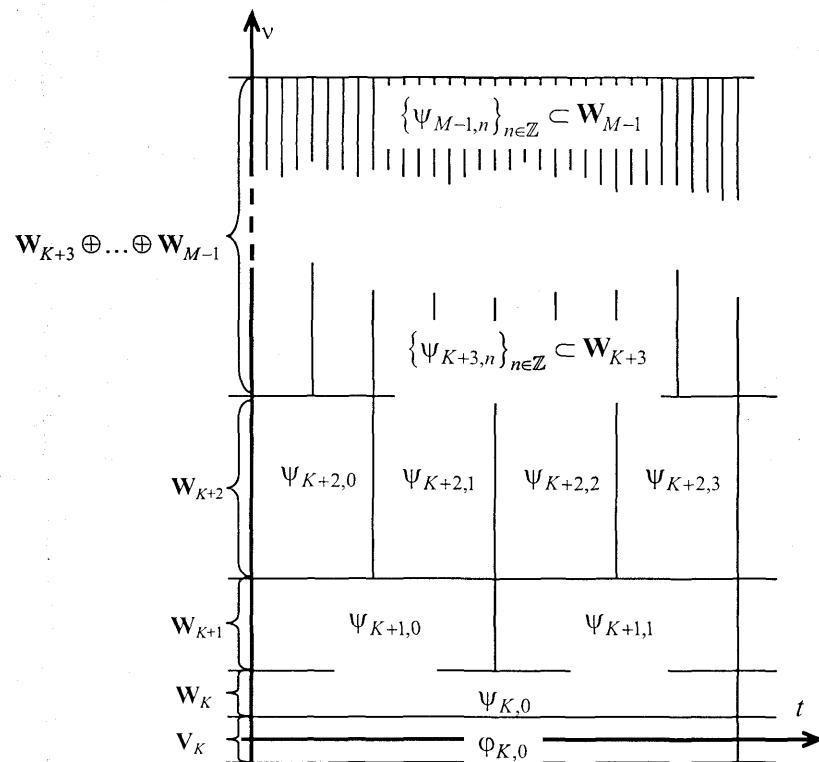


Рис. 7.16. Время-частотная локализация базисных функций, используемых для представления сигнала в виде (7.21). (Сравните с рис. 6.9)

Основное предположение, лежащее в основе многих алгоритмов вейвлет-кодирования, состоит в том, что если некоторый

коэффициент $c_{j,m} = \langle f(t), \psi_{j,m}(t) \rangle$ мал по абсолютной величине, то в области аргумента $t \in \Delta_{j,m} = \text{supp } \psi_{j,m}(t)$ сигнал $f(t)$, скорее всего, обладает малой энергией (и несет малую информацию) и все вейвлет-коэффициенты, которые являются прямыми (не только непосредственными) потомками узла $\psi_{j,m}$, также малы по абсолютной величине. Тогда «ветвь дерева» $T_{j,m}$, выходящую из узла $\psi_{j,m}$ и включающую в себя всех его прямых потомков (см. рис. 7.15), можно «подрезать», не кодируя входящие в нее вейвлет-коэффициенты, а при декодировании заменить отброшенные коэффициенты нулями.

Если используемое вейвлет-преобразование является ортогональным, то евклидова норма ошибки, вносимая в дискретный сигнал $f_M \leftrightarrow \{a_{M,n}\}$ при подрезании ветви $T_{j,m}$, может быть найдена в области трансформант:

$$\|f_M - \tilde{f}_M\|^2 = \sum_n |a_{M,n} - \tilde{a}_{M,n}|^2 = \sum_{y_{k,l} \in T_{j,m}} c_{k,l}^2,$$

где $\tilde{f}_M \leftrightarrow \{\tilde{a}_{M,n}\}$ — восстановленный после декодирования и обратного вейвлет-преобразования дискретный сигнал. Различные алгоритмы вейвлет-компрессии различаются, в основном, правилами, по которым принимается решение о подрезании ветвей, и способами кодирования структуры подрезанного дерева, а также процедурами квантования и статистического кодирования тех вейвлет-коэффициентов, которые остались после подрезания ветвей.

■ 7.10. Двумерные дискретные вейвлет-преобразования

Для обработки двумерных дискретных сигналов, которыми описываются цифровые изображения, используется следующая схема, обобщающая одномерный случай. С помощью одномерной масштабирующей функции $\phi(x)$ и соответствующего ей материнского вейвлета $\psi(x)$ определяются двумерная масштабирующая функция

$$\phi(x, y) = \phi(x)\phi(y)$$

и три двумерных вейвлета:

$$\psi^H(x, y) = \phi(x)\psi(y), \quad \psi^V(x, y) = \psi(x)\phi(y), \quad \psi^D(x, y) = \psi(y)\psi(x).$$

Соответственно вместо двух фильтров декомпозиции (см. рис. 7.5) получим четыре. На практике это означает, что сначала двумерная матрица дискретного изображения построчно подвергается одномерной фильтрации по схеме рис. 7.5 (после прореживаний отклик \tilde{h} -фильтра записывается в левую половину вектора-строки, отклик \tilde{g} -фильтра — в правую), а затем одномерная фильтрация проводится вдоль столбцов (после прореживаний отклик \tilde{h} -фильтра записывается в верхнюю половину вектора-столбца, отклик \tilde{g} -фильтра — в нижнюю). В результате получаем из матрицы изображения (рис. 7.17, *a*) четыре группы коэффициентов разложения (рис. 7.17, *б*). Восстановить исходное изображение после такой обработки можно, последовательно обработав столбцы и строки с помощью пары фильтров реконструкции (см. рис. 7.7).

Процедура декомпозиции может быть продолжена над левой верхней четвертью полученных коэффициентов (рис. 7.17, *б*), которые соответствуют двумерным масштабирующим функциям. Коэффициенты, лежащие в оставшихся трех наборах, относятся к функциям двумерных вейвлетов и далее не обрабатываются. Соответствующие блоки коэффициентов называем V-, H-, D- (vertical, horizontal, diagonal) или HL-, LH-, HH- (high-low, low-high, high-high) поддиапазонами¹. Три вейвлет-поддиапазона, которые получаются на следующем шаге (рис. 7.17, *в*), представляют более «грубый» уровень пространственного (т. е. в области изображения) разрешения.

Аналогично одномерному случаю (см. раздел 7.9), базисные функции двумерного вейвлет-спектра также обладают важным свойством, широко используемым в современных алгоритмах компрессии изображений. А именно, базисные функции более высокого уровня разрешения получаются из базисных функций более низкого уровня разрешения путем сжатия и сдвига, причем

«потомки» имеют носитель, не выходящий за область носителя «родителя», но у каждого родителя имеется уже по четыре потомка. Так, если $\psi(x, y)$ — родительский вейвлет, то порождаются следующие четыре потомка:

$$\psi(x, y) \rightarrow \begin{cases} 2\psi(2x, 2y), & 2\psi(2x-1, 2y), \\ 2\psi(2x, 2y-1), & 2\psi(2x-1, 2y-1) \end{cases}. \quad (7.49)$$

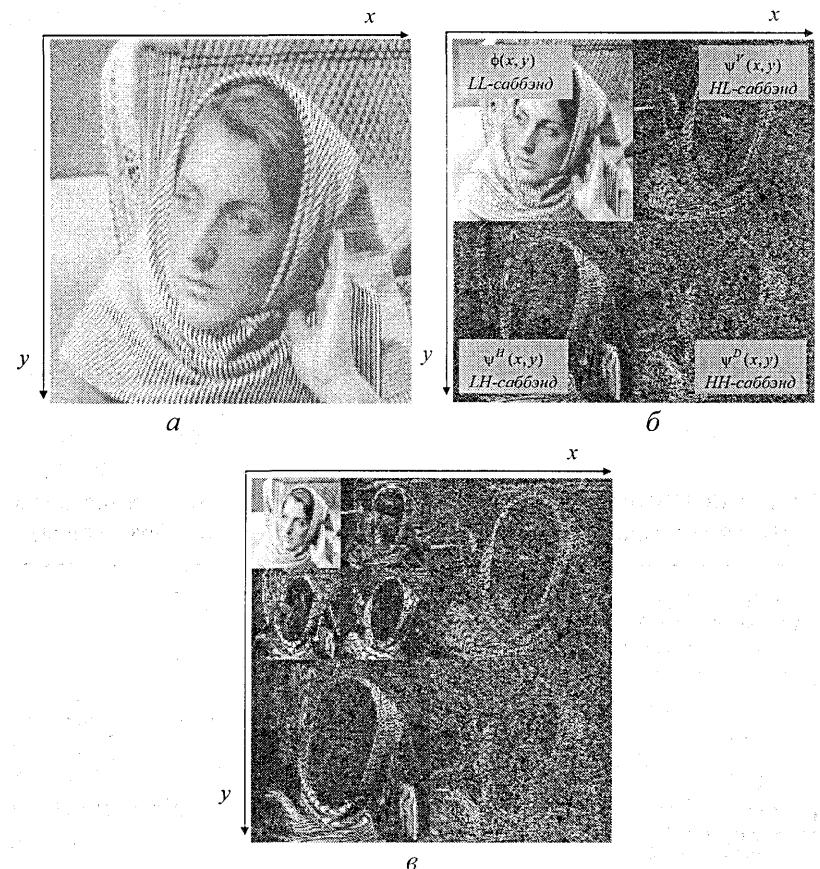


Рис. 7.17. Двумерное дискретное вейвлет-преобразование:
 а — исходное изображение; б — результат выполнения первого шага двумерной декомпозиции; в — результат выполнения двух шагов

¹ Часто поддиапазон называют «саббэндом», от оригинального слова *subband*.

Схема связей «родитель—потомки» отражена на рис. 7.18. Видим, что коэффициенты спектра могут быть упорядочены в древовидную структуру, как это показано на следующем рис. 7.19. Родителем трех вейвлетов самого грубого разрешения по определению считается единственная масштабирующая функция (начало дерева), дальнейшие связи устанавливаются в соответствии с (7.49). Таким образом, в двумерном случае коэффициенты вейвлет-преобразования также могут быть упорядочены в древовидную структуру, которая наследует свойства одномерной структуры, описанной в разделе 7.9. Соответственно, и основные идеи алгоритмов компрессии двумерных сигналов остаются теми же, что и в одномерном случае.

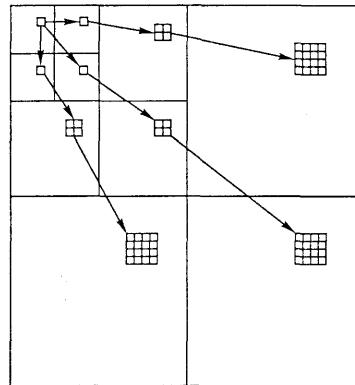


Рис. 7.18. Структура связей «родитель—потомки» между базисными функциями вейвлет-спектра (выполнено три шага преобразования)

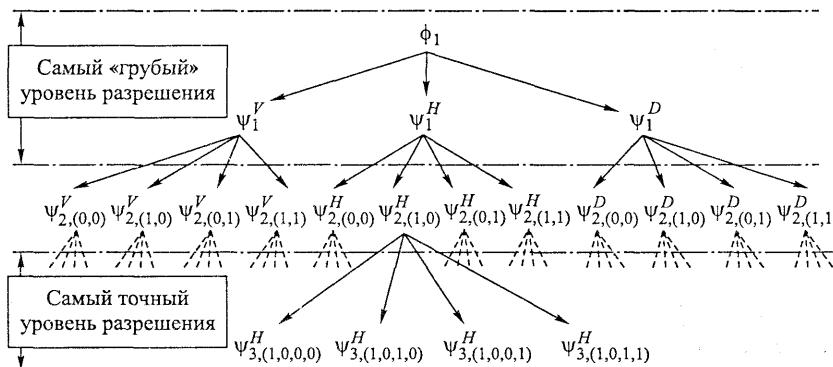


Рис. 7.19. Способ упорядочивания базисных функций дискретного вейвлет-преобразования в древовидную структуру

Список литературы

1. Ахмед Н., Рао К. Ортогональные преобразования при обработке цифровых сигналов: Пер. с англ. М.: Связь, 1980.
2. Блаттер К. Вейвлет-анализ. Основы теории М.: Техносфера, 2004.
3. Быстрые алгоритмы в цифровой обработке изображений. Преобразования и медианные фильтры /Под ред Т.С. Хуанга. М.: Радио и связь, 1984.
4. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: ДИАЛОГ-МИФИ, 2002.
5. Вентцель Е.С. Теория вероятностей: Учеб. для вузов. М.: Академия, 2003.
6. Волков Е.А. Численные методы: Учеб. пособие для вузов. СПб.: Лань, 2004.
7. Гольденберг Л.М., Матюшкин Б.Д., Поляк М.Н. Цифровая обработка сигналов: Учеб. пособие. М.: Радио и связь, 1990.
8. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005.
9. Дмитриев В.И. Прикладная теория информации: Учеб. для вузов. М.: Высшая школа, 1989.
10. Добеши И. Десять лекций по вейвлетам. Ижевск: Регулярная и хаотическая динамика, 2004.
11. Дьяконов В.П. Вейвлеты. От теории к практике. М.: СОЛОН-Р, 2004.
12. Ефимов А.В., Золотарев Ю.Г., Тертигорева В.М. Математический анализ (специальные разделы). М.: Высш. шк., 1980.
13. Иванов М.Т., Сергиенко А.Б., Ушаков В.Н. Теоретические основы радиотехники: Учеб. пособие / Под. ред. В.Н.Ушакова. М.: Высш. шк., 2002.
14. Купер Дж., Макгиллем К. Вероятностные методы анализа сигналов и систем. М.: Мир, 1989.
15. Лесин В.В. Теория случайных процессов: Учеб. пособие. М.: МИЭТ, 2005.
16. Макс Ж. Методы и техника обработки сигналов при физических измерениях: в 2 т. М.: Мир, 1983.
17. Малла С. Вейвлеты в обработке сигналов. М.: Мир, 2005.

18. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера. 2-е изд., испр. М.: Физматлит, 2003.
19. Нуссбаумер Г. Быстрое преобразование Фурье и алгоритмы вычисления сверток: Пер. с англ. М.: Радио и связь, 1985.
20. Солонина А.И., Улахович Д.А., Арбузов С.М., Соловьев Е.Б. Основы цифровой обработки сигналов: Курс лекций. СПб.: БХВ-Петербург, 2005.
21. Оттенгейм А., Шафер М. Цифровая обработка сигналов. М.: Физматлит, 2006.
22. Петухов А.П. Введение в теорию базисов всплесков. СПб.: Изд-во СПбГТУ, 1999.
23. Потапов В.Н. Теория информации. Кодирование дискретных вероятностных источников. Новосибирск: Изд-во НГУ, 1999.
24. Рабинер Л., Голд Б. Теория и применение цифровой обработки сигналов. М: Мир, 1978.
25. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. М.: Техносфера, 2005.
26. Сиберт У.М. Цепи, сигналы, системы: в 2 ч. М.: Мир, 1988.
27. Смирнов А.В. Основы цифрового телевидения. М.: Горячая линия-телеком, 2001.
28. Сэломон Д. Сжатие данных, изображений и звука. М.: Техносфера, 2004.
29. Треногин В.А. Функциональный анализ. М.: Физматлит, 2003.
30. Уэлстид С. Фракталы и вейвлеты для сжатия изображений в действии. Учеб. пособ. М.: Триумф, 2003.
31. Быков Р.Е., Фрайер Р., Иванов К.В., Манцевцов А.А. Цифровое преобразование изображений / Под ред. проф. Р.Е. Быкова. М.: Горячая линия-телеком, 2003.

Сергей Владимирович Умняшкин

Теоретические основы цифровой обработки и представления сигналов

Учебное пособие

Редактор Н.В. Скугаревская

Корректор Н.Н. Морозова

Компьютерная верстка С.Ч. Соколовского

Оформление серии К.В. Пономарева

Сдано в набор 16.04.2007. Подписано в печать 25.11.2007.

Формат 60 × 90^{1/16}. Бумага офсетная. Усл. печ. л. 19.

Уч.-изд. л. 19,5. Тираж 2000 экз.

Заказ № 8554.

ЛР № 071629 от 20.04.98

Издательский Дом «ФОРУМ»

101000, Москва-Центр, Колпачный пер., 9а

Тел./факс: (495) 625-39-27

E-mail: forum-books@mail.ru

ЛР № 070824 от 21.01.93

Издательский Дом «ИНФРА-М»

127282, Москва, Полярная ул., 31в

Тел.: (495) 380-05-40; факс: (495) 363-92-12

E-mail: books@infra-m.ru

<http://www.infra-m.ru>

По вопросам приобретения книг обращайтесь:

Отдел продаж «ИНФРА-М»

127282, Москва, ул. Полярная, д. 31в

Тел.: (495) 363-42-60

Факс: (495) 363-92-12

E-mail: books@infra-m.ru

Центр комплектования библиотек

119019, Москва, ул. Моховая, д.16

(Российская государственная библиотека, кор. К)

Тел.: (495) 202-93-15

Магазин «Библиосфера» (розничная продажа)

109147, Москва, ул. Марксистская, д. 9

Тел.: (495) 670-52-18, (495) 670-52-19

Отпечатано с готовых диапозитивов в ОАО ордена «Знак Почета»

«Смоленская областная типография им. В. И. Смирнова».

214000, г. Смоленск, проспект им. Ю. Гагарина, 2.



**УМНЯШКИН
СЕРГЕЙ
ВЛАДИМИРОВИЧ**

Окончил Московский государственный институт электронной техники (технический университет) – МИЭТ в 1994 году, где продолжает работать по настоящее время. Профессор, доктор физико-математических наук. Ученый в области цифровой обработки сигналов и изображений.

ISBN 978-5-8199-0357-5

A standard linear barcode representing the ISBN 978-5-8199-0357-5.

9 785819 903575