



МИР

цифровой обработки

С.В. Умняшкин

Основы теории
цифровой обработки
сигналов

Издание шестое,
исправленное

ТЕХНОСФЕРА

Москва
2021

УДК 519.6:621.391

ББК 32.811

У54

Рецензент:

Член-корреспондент РАН, доктор технических наук А.В. Деоржович

У54 Умняшкин С.В.

Основы теории цифровой обработки сигналов: Учебное пособие.

Издание шестое, исправленное

Москва: ТЕХНОСФЕРА, 2021. – 550 с. ISBN 978-5-94836-617-3

Учебное пособие предназначено для студентов вузов, обучающихся по направлениям подготовки бакалавров и магистров «Прикладная математика», «Информатика и вычислительная техника», и направлено на изучение теоретических основ цифровой обработки сигналов (ЦОС). Помимо базовых разделов ЦОС, связанных с цифровым представлением сигналов, анализом и синтезом цифровых фильтров, в пособии рассматриваются вопросы многократной обработки сигналов, адаптивной фильтрации, дискретного спектрального анализа. Значительное внимание уделено таким специальным разделам ЦОС, как эффективное представление сигналов (сжатие данных), теория и приложения дискретных вейвлет-преобразований.

Пособие также может быть рекомендовано в качестве дополнительного учебного материала для других инженерных направлений подготовки, связанных с ЦОС (радиотехника, электроника, телекоммуникации и связь, автоматическое управление и др.). При этом базовая часть материала первых глав представляет собой основу для вводного курса (уровень бакалавриата), а в полном объеме учебное пособие ориентировано на более углубленное изучение теории ЦОС в магистратуре.

УДК 519.6:621.391

ББК 32.811

© Умняшкин С.В., 2021

© АО «РИЦ «ТЕХНОСФЕРА», оригинал-макет, оформление, 2021

ISBN 978-5-94836-617-3

Содержание

| | |
|--|------------|
| Предисловие..... | 6 |
| Список сокращений и обозначений..... | 10 |
| Глава 1. Элементы функционального анализа и спектрального представления функций..... | 12 |
| 1.1. Линейные нормированные пространства..... | 12 |
| 1.2. Пространства со скалярным произведением..... | 19 |
| 1.3. Примеры ортогональных систем в пространстве L_2 | 27 |
| 1.4. Тригонометрические ряды Фурье. Явление Гиббса..... | 33 |
| 1.5. Интеграл Фурье..... | 38 |
| 1.6. Принцип неопределенности время-частотного представления сигналов..... | 44 |
| 1.7. Обобщенное преобразование Фурье..... | 50 |
| 1.8. Энергетический спектр. Спектр мощности..... | 54 |
| Глава 2. Дискретизация и квантование сигналов. Дискретные ортогональные преобразования..... | 59 |
| 2.1. Преобразование непрерывных сигналов в дискретные..... | 59 |
| 2.2. Дискретизация по критерию наибольшего отклонения..... | 61 |
| 2.3. Частотный критерий выбора шага дискретизации..... | 62 |
| 2.4. Спектр дискретного сигнала..... | 67 |
| 2.5. Дискретизация узкополосных сигналов..... | 72 |
| 2.6. Дискретное преобразование Фурье..... | 79 |
| 2.7. Быстрое преобразование Фурье. Алгоритм БПФ с прореживанием по времени..... | 87 |
| 2.8. Алгоритм БПФ с прореживанием по частоте..... | 100 |
| 2.9. Дискретное преобразование Уолша..... | 105 |
| 2.10. Дискретное преобразование Хаара..... | 111 |
| 2.11. Некоторые применения дискретных ортогональных преобразований..... | 115 |
| 2.12. Квантование дискретных сигналов..... | 118 |
| Глава 3. Описание и анализ линейных дискретных систем..... | 129 |
| 3.1. Нормирование временной и частотной осей..... | 129 |
| 3.2. Z -преобразование..... | 130 |
| 3.3. Линейные дискретные фильтры..... | 136 |
| 3.4. Соединения и структурные схемы фильтров..... | 140 |
| 3.5. Устойчивость ЛДФ..... | 149 |
| 3.6. Частотная характеристика ЛДФ..... | 151 |
| 3.7. Минимально-фазовые системы. Фазовые звенья..... | 159 |
| 3.8. Нахождение отклика фильтра с использованием БПФ..... | 165 |



| | |
|---|------------|
| 3.9. Многоскоростная обработка сигналов..... | 167 |
| 3.10. Изменение частоты дискретизации сигналов с помощью полифазных фильтров..... | 173 |
| 3.11. Эффекты квантования в цифровых системах..... | 180 |
| 3.12. Согласованный дискретный фильтр..... | 190 |
| 3.13. Линейная дискретная система как генератор случайных сигналов. Полусная модель сигналов..... | 198 |
| 3.14. Фильтр Калмана..... | 203 |
| Глава 4. Введение в методы синтеза цифровых фильтров..... | 211 |
| 4.1. Этапы разработки цифровых фильтров..... | 211 |
| 4.2. КИХ-фильтры с линейной фазой. Синтез КИХ-фильтров методом частотной выборки..... | 217 |
| 4.3. Оконный метод синтеза КИХ-фильтров..... | 226 |
| 4.4. Синтез оптимальных КИХ-фильтров..... | 237 |
| 4.5. Специальные КИХ-фильтры: преобразователь Гильберта и цифровой дифференциатор..... | 246 |
| 4.6. Основные характеристики аналоговых линейных систем и их связь с характеристиками ЛДС..... | 255 |
| 4.7. Синтез БИХ-фильтров по аналоговым прототипам..... | 261 |
| 4.8. Синтез БИХ-фильтров методом инвариантности импульсной характеристики..... | 273 |
| 4.9. Синтез БИХ-фильтров методом билинейного Z-преобразования..... | 278 |
| 4.10. Выбор структуры для реализации фильтра..... | 286 |
| 4.11. Адаптивная фильтрация. Фильтр Винера..... | 291 |
| 4.12. Построение фильтра Винера в частотной области..... | 304 |
| Глава 5. Основы прикладной теории информации..... | 310 |
| 5.1. Мера количества информации для дискретного источника сообщений без памяти..... | 310 |
| 5.2. Основные теоремы о кодировании источника без памяти..... | 316 |
| 5.3. Эффективное кодирование дискретного источника без памяти по методам Шеннона — Фано и Хаффмана..... | 326 |
| 5.4. Кодирование длин серий..... | 332 |
| 5.5. Арифметическое кодирование..... | 335 |
| 5.6. Условная энтропия..... | 344 |
| 5.7. Кодирование дискретного источника с памятью..... | 348 |
| 5.8. Статистическое моделирование источника..... | 355 |
| 5.9. Непрерывный источник сообщений. Дифференциальная энтропия..... | 356 |
| 5.10. Передача дискретного сообщения по каналу с помехами..... | 364 |
| 5.11. Словарные методы кодирования..... | 371 |

| | |
|--|------------|
| Глава 6. Применение дискретных ортогональных преобразований для компрессии и спектрального анализа сигналов | 378 |
| 6.1. Корреляция как мера статистической зависимости данных. Преобразование Карунена—Лоэва..... | 378 |
| 6.2. Эффективность использования дискретных ортогональных преобразований для кодирования коррелированных данных..... | 384 |
| 6.3. ДПФ в вещественной форме. Дискретное преобразование Хартли..... | 391 |
| 6.4. Дискретное косинусное преобразование..... | 394 |
| 6.5. Компрессия изображений на основе двумерного ДКП..... | 403 |
| 6.6. Дискретное псевдокосинусное преобразование..... | 409 |
| 6.7. Оптимизация алгоритмов сжатия данных с потерями..... | 418 |
| 6.8. Аппроксимационный подход к выбору преобразований для представления дискретных сигналов. Частотная трактовка..... | 425 |
| 6.9. Время-частотный анализ. Оконное преобразование Фурье..... | 432 |
| 6.10. Использование ДПФ для спектрального анализа..... | 439 |
| Глава 7. Вейвлет-преобразования и их приложения для обработки дискретных сигналов | 449 |
| 7.1. Кратно-масштабный анализ..... | 449 |
| 7.2. Проектирование функций на подпространства КМА..... | 455 |
| 7.3. Вычисление дискретных вейвлет-преобразований..... | 461 |
| 7.4. Квадратурно-зеркальные фильтры..... | 465 |
| 7.5. Свойства КЗФ..... | 471 |
| 7.6. Построение масштабирующих функций и вейвлетов по масштабирующим уравнениям..... | 478 |
| 7.7. Вейвлеты Добеши..... | 482 |
| 7.8. Биортогональные вейвлет-преобразования..... | 488 |
| 7.9. Применение дискретных вейвлет-преобразований для сжатия сигналов..... | 492 |
| 7.10. Подавление шумов фильтрацией в базисе дискретных вейвлет-преобразований..... | 496 |
| 7.11. Двумерные дискретные вейвлет-преобразования..... | 499 |
| 7.12. Метод сжатия цифровых изображений JPEG 2000..... | 507 |
| 7.13. Вейвлет-пакеты..... | 514 |
| 7.14. Вычисление ДВП по схеме лифтинга..... | 526 |
| Заключение | 544 |
| Литература | 545 |

Предисловие

В последние два десятилетия методы цифровой обработки сигналов (ЦОС) в радиотехнике, электронике, системах связи, контроля и управления стали преобладающими, активно вытесняя методы аналоговой обработки. Этому способствовала стремительно увеличивавшаяся производительность вычислительной техники, которая уже проникла практически во все области человеческой деятельности. Сегодня, например, говоря о записи или обработке аудио- и видеоинформации, мы не уточняем, что речь идет о цифровых форматах, подразумевая это само собой разумеющимся. Важность изучения методов ЦОС трудно переоценить: вопросы, связанные с цифровой обработкой и представлением сигналов, давно перестали быть узкоспециальными. Основы знаний в данной области требуются большинству инженеров, а для специалистов в области электроники, радиотехники и телекоммуникаций, информатики и вычислительной техники необходимо более глубокое понимание основных методов ЦОС и математической теории, лежащей в их основе. Предлагаемое вниманию читателя учебное пособие посвящено изучению данных вопросов.

Изначально ЦОС развивалась как ветвь, растущая из теории обработки аналоговых сигналов, и рассматривалась как раздел электроники и радиотехники. Однако внедрение в системы обработки сигналов цифровых программируемых процессоров и контроллеров, расширение спектра и увеличение сложности алгоритмов ЦОС, реализация которых стала возможна в том числе в реальном масштабе времени, превратили ЦОС в политехническую дисциплину. Сегодня для математиков-программистов, специалистов в области информатики и вычислительной техники эта область знаний представляет собой постоянно расширяющееся поле для приложения усилий. Данную тенденцию необходимо учитывать при подготовке инженерных кадров. В предлагаемом вниманию читателя учебном пособии рассматриваются основы теории ЦОС, изложение которой по формату и содержанию ориентировано в первую очередь именно

на студентов, обучающихся по инженерным направлениям «Прикладная математика» и «Информатика и вычислительная техника» (бакалавриат и магистратура). Однако при написании пособия автор старался опираться лишь на курс высшей математики, общий для всех инженерных направлений подготовки, поэтому оно может быть рекомендовано также для студентов, обучающихся по профилям подготовки в области радиотехники, связи и телекоммуникаций.

Первая глава носит вводный характер и содержит изложение тех основных положений функционального анализа и теории преобразования Фурье, которые потребуются далее в последующих главах. Помимо этого, в первой главе вводятся популярные в ЦОС функциональные системы Уолша и Хаара.

Вторая глава посвящена вопросам дискретизации непрерывных сигналов и преобразований, прежде всего преобразования Фурье. Значительное внимание уделено частотным аспектам дискретизации (как широкополосных, так и узкополосных сигналов) и построению быстрых алгоритмов вычислений дискретных преобразований Фурье, Уолша, Хаара. Кроме того, рассматривается скалярное и векторное квантование сигналов.

В третьей главе вводятся основные понятия теории линейных дискретных систем (фильтров). Основное внимание уделено анализу систем. Рассматриваются вопросы передискретизации сигналов и влияния эффектов квантования в цифровых системах. Также изучаются некоторые вопросы согласованной фильтрации и скалярных фильтров Калмана.

В четвертой главе изложены основные классические методы синтеза КИХ- и БИХ-фильтров по заданной частотной характеристике. Завершает главу рассмотрение основ теории адаптивных фильтров.

Пятая глава представляет собой введение в теорию информации, а также содержит описание ряда используемых на практике методов эффективного статистического (энтропийного) кодирования.

Первые пять глав включают в себя рассмотрение общих теоретических вопросов ЦОС. В сокращенном варианте этот материал может быть использован для базовой подготовки бакалавров по различным инженерным направлениям обучения. В полном объеме материал этих глав может составить основу учебного курса для подготовки магистров, имеющих профиль обучения со специализацией в ЦОС, прежде всего, по направлениям «Информатика и вычислительная техника» и «Прикладная математика».

Две последние главы (шестая и седьмая) представляют собой более углубленное теоретическое рассмотрение таких специальных разделов ЦОС, как эффективное представление и компрессия сигналов, цифровой спектральный анализ. Седьмая глава целиком посвящена применению дискретных вейвлет-преобразований в ЦОС и может составить основу отдельного спецкурса.

Таким образом, в зависимости от потребностей читателя учебное пособие можно использовать при изучении ЦОС для различных уровней «глубины погружения» в предметную область.

1. Для первоначального освоения основных понятий ЦОС следует изучить главу 1 и параграфы 2.1–2.4, 2.6–2.8, 2.11–2.12, 3.1–3.6, 3.8. Полезно также разобрать материал параграфов 5.1–5.4. Этот уровень изучения можно считать начальным, соответствующим бакалавриату различных инженерных направлений обучения.
2. Для получения тех знаний теории ЦОС, которые, на взгляд автора, необходимы для магистерского уровня инженерной подготовки профильных направлений обучения (включая связь и телекоммуникации, радиотехнику и электронику), следует полностью изучить главы 1–4 и параграфы 5.1–5.4, 6.9, 6.10. Рекомендуется также ознакомиться с материалом, изложенным в параграфах 5.5–5.11.
3. В полном объеме содержание учебного пособия ориентировано на студентов магистратуры, обучающихся по направлению «Прикладная математика» и избравших область ЦОС своей специализацией.

В книге используется двойная нумерация для рисунков, формул, примеров и теорем: первая цифра обозначает главу, вторая — порядковый номер формулы (примера, теоремы) в главе. При нумерации (обозначении) аксиом и свойств используется значок \circ , например 1° . Начало и окончание доказательств теорем, решений примеров обозначается соответственно символами \blacktriangleleft и \blacktriangleright . Упражнения (задачи) для самостоятельного решения приводятся непосредственно в тех местах, где их появление логически наиболее связано с излагаемым материалом, а не в конце глав или разделов, как это чаще всего практикуется в учебной литературе.

Основу данного пособия составляют учебные курсы, читаемые автором на протяжении многих лет в Национальном исследовательском университете «Московский институт электронной техники» (МИЭТ) для бакалавров и магистров, обучающихся по направлению «Прикладная математика». Большое значение в работе над пособием имело обсуждение его содержания с коллегами. Автор выражает глубокую признательность доценту кафедры высшей математики №1 МИЭТ В.В. Лесину и рецензенту, внимательно ознакомившимся с текстом рукописи и высказавшим ряд ценных замечаний, которые были учтены при подготовке издания.

Список сокращений и обозначений

АЦП — аналого-цифровое преобразование
АЧХ — амплитудно-частотная характеристика
БИХ — бесконечная импульсная характеристика
БПУ — быстрое преобразование Уолша
БПФ — быстрое преобразование Фурье
БПХ — быстрое преобразование Хаара
ВДПФ — вещественное дискретное преобразование Фурье
ВЧ — верхние частоты
ГВЗ — групповое время задержки
ДВП — дискретное вейвлет-преобразование
ДКП — дискретное косинусное преобразование
ДПГ — дискретный преобразователь Гильберта
ДПКП — дискретное псевдокосинусное преобразование
ДПЛ — дискретное преобразование Лапласа
ДПУ — дискретное преобразование Уолша
ДПФ — дискретное преобразование Фурье
ДПХ — дискретное преобразование Хаара
ИХ — импульсная характеристика
КДС — кодирование длин серий
КИХ — конечная импульсная характеристика
КЗФ — квадратурно-зеркальные фильтры
КМА — кратно-масштабный анализ
КФ — ковариационная функция
ЛДС — линейная дискретная система
ЛДФ — линейный дискретный фильтр
ЛИВС — линейная инвариантная во времени система
ЛНП — линейное нормированное пространство
НОД — наибольший общий делитель
НЧ — нижние частоты
ОДКП — обратное ДКП
ОДПФ — обратное ДПФ
ОПФ — оконное преобразование Фурье
ПФ — передаточная функция

- ЧХ — частотная характеристика
 ФВЧ — фильтр верхних частот
 ФНЧ — фильтр нижних частот
 ФЧХ — фазочастотная характеристика
 ЦД — цифровой дифференциатор
 ЦОС — цифровая обработка сигналов
 ЦФ — цифровой фильтр
 ЭНП — элемент наилучшего приближения
- \mathbb{C} — множество комплексных чисел
 \mathbb{N} — множество натуральных чисел
 \mathbb{R} — множество действительных чисел
 \mathbb{Z} — множество целых чисел
 $x \gg y$ — число x много больше числа y
 $[x]$ — целая часть числа $x \geq 0$
 $\text{round}(x) = \text{sign}(x)[|x| + 0,5]$ — округление x до ближайшего целого
 $x \bmod p$ — остаток от деления целого числа $x \geq 0$ на число $p \in \mathbb{N}$
 $a \oplus b = (a + b) \bmod 2$ — сложение по модулю 2, где $a \in \{0, 1\}$ и $b \in \{0, 1\}$
 $A \oplus B$ — ортогональная сумма подпространств A и B
 $\|x\|$ — норма элемента x
 $\langle x, y \rangle$ — скалярное произведение элементов x, y
 \bar{A} — замыкание множества A
 $\bar{z} = \text{Re } z - i \text{Im } z$ — комплексное сопряжение числа $z \in \mathbb{C}$
 $\text{Res}[f(z), z_0]$ — вычет функции $f(z)$ в точке z_0
 $\mathcal{F}\{f(t)\}$ — преобразование Фурье функции $f(t)$
 $\mathcal{F}^{-1}\{F(v)\}$ — обратное преобразование Фурье функции $F(v)$
 $\mathcal{L}\{f(t)\}$ — преобразование Лапласа функции $f(t)$
 $\mathcal{L}^{-1}\{G(p)\}$ — обратное преобразование Лапласа функции $G(p)$
 $Z\{x(n)\}$ — Z -преобразование последовательности $x(n)$
 $Z^{-1}\{X(z)\}$ — обратное Z -преобразование функции $X(z)$
 $M(X)$ — математическое ожидание случайной величины X
 $D(X)$ — дисперсия случайной величины X

ГЛАВА I

ЭЛЕМЕНТЫ ФУНКЦИОНАЛЬНОГО АНАЛИЗА И СПЕКТРАЛЬНОГО ПРЕДСТАВЛЕНИЯ ФУНКЦИЙ

I.1. Линейные нормированные пространства

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

Определение. Множество E элементов произвольной природы называется *линейным пространством*, если в нем однозначно 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°. $\forall x, y \in E: x+y = y+x$.
- 2°. $\forall x, y, z \in E: (x+y)+z = x+(y+z)$.
- 3°. $\exists \theta \in E, \forall x \in E: x+\theta = x$ (существование нулевого элемента θ).
- 4°. $\forall x \in E, \forall \lambda, \mu$ (λ, μ — 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°. $\forall x, y \in E, \forall \lambda: \lambda(x+y) = \lambda x + \lambda y$.

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

В курсе линейной алгебры изучались линейные пространства \mathbb{R}^n арифметических векторов размерности n . Приведенное выше аксиоматическое определение обобщает понятие линейного пространства \mathbb{R}^n на множества произвольной природы. По аналогии, элементы любого линейного пространства также 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 $(n+1)$ ненулевых 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 θ 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 сколь 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 характер (см., например, [17]).

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. $\forall x \in E: \|x\| \geq 0, \|x\| = 0 \Leftrightarrow x = \theta$.
- 2°. Однородность нормы. $\forall \lambda \in \mathbb{R}, \forall x \in E: \|\lambda x\| = |\lambda| \cdot \|x\|$.
- 3°. Неравенство треугольника. $\forall x, y \in E: \|x + y\| \leq \|x\| + \|y\|$.

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

Пример 1.2. Рассмотрим векторное пространство $C[a; b]$ из примера 1.1. Покажите самостоятельно, что приводимые ниже способы вычисления нормы удовлетворяют аксиомам 1°–3°:

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

Указание: в пункте б) для доказательства аксиомы треугольника воспользуйтесь неравенством Минковского (1.1).

Определение. *Расстоянием* между элементами x, y нормированного векторного пространства E назовем 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 имеем:

$$\rho(x, y) = \|x - y\| = \|(x - z) + (z - y)\| \leq \|x - z\| + \|z - y\| = \rho(x, z) + \rho(y, z). \blacktriangleright$$

Расстояние между элементами называют *метрикой* пространства. Пространство (не обязательно нормированное), каждой паре x, y элементов которого поставлено в соответствие вещественное число $\rho(x, y)$ (расстояние), обладающее свойствами 1°–3°, называется *метрическим*¹.

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

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

Понятия нормы, расстояния, окрестности являются исходными для построения анализа 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 $x \in M, x \neq a$. То есть $\forall r > 0: (S_r(a) \setminus a) \cap M \neq \emptyset$.

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 положительных элементов, например $\varepsilon_k = 1/k, k = 1, 2, \dots$. Так как a — предельная точка M , то, по определению, $\forall \varepsilon_k > 0 \exists x_k \in M, x_k \neq a: x_k \in S_{\varepsilon_k}(a)$. Поскольку $\rho(x_k, a) = \|x_k - a\| < \varepsilon_k = 1/k$ и $\lim_{k \rightarrow \infty} \|x_k - a\| = 0$, то 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 , $L \subset E$, назовем *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 $u \in L$, где L — подпространство из ЛНП E , называется *элементом наилучшего приближения* (ЭНП) для заданного элемента $x \in E$, если $\rho(x, L) = \|x - u\|$.

Элемент наилучшего приближения существует не всегда, а также 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 следующим образом: $\|x\| = |\xi_1| + |\xi_2|$ (убедитесь самостоятельно, что аксиомы нормы выполняются). Рассмотрим подмножество $L \subset \mathbb{R}^2, L = \{(\xi_1, \xi_2) | \xi_1 = \xi_2\} = \{(\alpha, \alpha) | \alpha \in \mathbb{R}\}$. Тогда:

- 1) L — подпространство в E ;
- 2) для элемента $x = (-1, 1)$ 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

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

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

Очевидно, $\inf_{t \in \mathbb{R}} f(t) = \min_{t \in \mathbb{R}} f(t) = 2$. 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 $L^* = \{(\alpha, \alpha) \mid -1 \leq \alpha \leq 1\}$, $L^* \subset L$. ▶

Определение. Пусть X — ЛНП. Последовательность $\{x_n\} \subset X$ называется *фундаментальной*, если $\forall \varepsilon > 0 \exists N = N(\varepsilon): \forall n > N, \forall p \in \mathbb{N} \|x_{n+p} - x_n\| < \varepsilon$. (\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 $\exists \lim_{n \rightarrow \infty} x_n = x \in X$, и $\forall \varepsilon > 0 \exists N = N(\varepsilon), \forall n > N: \|x_n - x\| < \varepsilon/2$. Тогда также $\forall p \in \mathbb{N}: \|x_{n+p} - x\| < \varepsilon/2$, 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 сколь угодно малым. Следовательно, последовательность $\{x_n\} \subset X$ является фундаментальной. ▶

Упражнение. Покажите самостоятельно, что если последовательность $\{x_n\}$ — фундаментальна, то последовательность $\{\lambda x_n\}$ также является фундаментальной.

Возникает вопрос: а 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 $\|x\| = \sqrt{\int_0^T |x(t)|^2 dt}$ не является банаховым.

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

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

(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 $\{s_n\}$ — фундаментальна 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 составленная 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 .

1.2. 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
- $$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle = \|x\|^2 - 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2.$$

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

Докажем теперь 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 лемму 1.3, получаем:

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2, \quad \text{или} \quad \|x + y\| \leq \|x\| + \|y\|.$$

Определения. Пусть E — линейное пространство с введенным скалярным произведением. *Ортогональными элементами*

пространства E называются такие элементы $x, y \in E$, что $\langle x, y \rangle = 0$. Ортогональность элементов будем обозначать $x \perp y$. (Очевидно, нулевой элемент ортогонален всем элементам пространства.) *Ортогональной системой* в E назовем 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 $\forall j = 1, \dots, m: 0 = \langle x_j, \theta \rangle = \langle x_j, \sum_{k=1}^m \lambda_k x_k \rangle = \sum_{k=1}^m \lambda_k \langle x_j, x_k \rangle = \lambda_j \langle x_j, x_j \rangle$.

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 $\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle x, y \rangle$.

◀ С учетом леммы 1.3 имеем:

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

Выражение в правой части последнего 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:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

$$\blacktriangleleft \quad \|x + y\|^2 + \|x - y\|^2 = \langle x + y, x + y \rangle + \langle x - y, x - y \rangle = \dots = 2\|x\|^2 + 2\|y\|^2.$$

Проделайте опущенные выкладки самостоятельно. \blacktriangleright

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 $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ как $\langle x, y \rangle = \sum_{k=1}^n x_k y_k$, — полное, т. е. гильбертово.

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

\blacktriangleleft Для функций $x(t)$ и $y(t)$, заданных на отрезке $t \in [a, b]$, определим скалярное произведение как

$$\langle x, y \rangle = \int_a^b x(t)y(t)dt$$

(убедитесь, что аксиомы скалярного произведения выполняются).

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 $t \in [0; T]$ 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 $\|x\| = \sqrt{\int_a^b |x(t)|^2 dt}$, 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 нашего рассмотрения (подробнее см., напр., [45]), 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 неотъемлемо 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. ►

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.

$$\|x - y\| = \inf_{u \in L} \|x - u\|. \tag{1.3}$$

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 $\forall \epsilon > 0 \exists u_\epsilon \in L$: $d \leq \|x - u_\epsilon\| < d + \epsilon$. Тогда, взяв числовую последовательность $\epsilon_k = 1/k$, $k = 1, 2, \dots$, 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

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

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

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

Таким образом, величину $\|u_m - u_n\|$ можно сделать сколь угодно малой за счет выбора достаточно большого числа N , т. е. последовательность $\{u_k\}$ — фундаментальная, и вследствие полноты H

$\exists \lim_{k \rightarrow \infty} u_k = y \in H$. 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) получаем:

$$4d^2 = 2\|x - y\|^2 + 2\|x - \tilde{y}\|^2 = \|y - \tilde{y}\|^2 + 4\left\|x - \frac{y + \tilde{y}}{2}\right\|^2 \geq \|y - \tilde{y}\|^2 + 4d^2,$$

откуда $\|y - \tilde{y}\|^2 = 0$ 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 : $\tilde{y} \in L$, так как $y \in L, u \in L$. Имеем:

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

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 $y - a + z - b = \theta$, и

$$\langle y - a + z - b, y - a \rangle = 0 = \langle y - a, y - a \rangle + \langle z, y - a \rangle - \langle b, y - a \rangle = \|y - a\|^2,$$

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

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).

Теорема 1.4 определяет способ нахождения ЭНП для элемента $x \in H$ 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

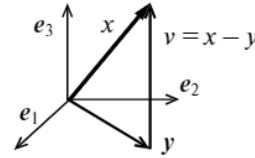


Рис. 1.1

$\forall k: g_k \in L, \langle x - y, g_k \rangle = 0$, 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

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

Определитель системы линейных уравнений (1.4) есть определитель матрицы Грама $G = \{\langle g_j, g_k \rangle\}_{k,j=1}^n$, причем $\det G \neq 0$ в силу линейной независимости элементов $\{g_1, g_2, \dots, g_n\}$. (Напомним, что $\det G = 0$ 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):

$$\lambda_j = \frac{\langle x, g_j \rangle}{\langle g_j, g_j \rangle}. \tag{1.5}$$

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 .

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

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

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

То есть 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. $\forall u \in L_i, \forall v \in L_j: \langle u, v \rangle = 0$ при $i \neq j$.
- 2) $\forall x \in H$ существует разложение $x = \sum_{i=1}^n x_i$, 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 , теорем 1.3–1.5 и их следствий.

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 $\langle x - y, y \rangle = 0$ (см. теорему 1.4), 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 $\|y\|^2 = \left\langle \sum_{j=1}^n \lambda_j g_j, \sum_{m=1}^n \lambda_m g_m \right\rangle = \sum_{j=1}^n \lambda_j \sum_{m=1}^n \lambda_m \langle g_j, g_m \rangle = \sum_{j=1}^n \lambda_j^2 \|g_j\|^2$. ▶

Пусть теперь в H задана бесконечная последовательность ненулевых ортогональных векторов $\{\varphi_k\}_{k=1}^\infty \subset H$. Это означает, что H — бесконечномерное пространство, так как ортогональные элементы линейно независимы. Рассматривая первые элементы $\{\varphi_k\}_{k=1}^n$ как базис, получаем некоторое линейное многообразие L_n , «натянутое» на $\{\varphi_k\}_{k=1}^n$. Можно показать, что L_n — замкнуто, т.е. является подпространством. Так как L_n — конечномерно, то с учетом теоремы 1.6 для

заданного элемента $x \in H$ и его ЭНП $y_n = \sum_{j=1}^n \lambda_j \phi_j$, $y_n \in L_n$, имеем:

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

Числовая последовательность $s_n = \|y_n\|^2 = \sum_{j=1}^n \lambda_j^2 \|\phi_j\|^2$ 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 $s = \sum_{j=1}^{\infty} \lambda_j^2 \|\phi_j\|^2$, причем (см. (1.6))

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

Соотношение (1.7) называется *неравенством Бесселя*.

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

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

т. е. $\lim_{n \rightarrow \infty} \|x - \sum_{k=1}^n \lambda_k \phi_k\| = 0$. Ряд (1.8) называется *рядом Фурье* (по ортогональной системе $\{\phi_k\}$), 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 непрерывности скалярного произведения и ортогональности системы $\{g_k\}_{k=1}^{\infty}$ имеем:

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

откуда следует формула (1.5). ▶

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

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

которое называется *равенством Парсеваля — Стеклова*.

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

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

что эквивалентно равенству $\|x\|^2 - \sum_{j=1}^{\infty} \lambda_j^2 \|\varphi_j\|^2 = 0$, которое получается предельным переходом $n \rightarrow \infty$ в соотношении (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.

1.3. 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 алгебраических 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: $r_k(x) = r_0(2^k x)$, $k = 1, 2, \dots$ (см. рис. 1.2).

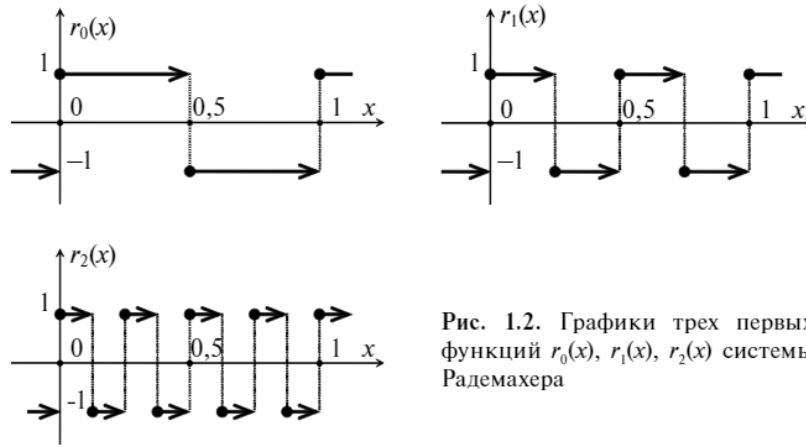


Рис. 1.2. Графики трех первых функций $r_0(x)$, $r_1(x)$, $r_2(x)$ системы Радемахера

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. $\forall x \in \Delta_m^{k+1}: r_k(x) = \text{const} = (-1)^m$.

На более мелких подынтервалах, естественно, функции также постоянны:

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

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

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

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

3°. Система функций $\{r_k(x)\}_{k=0}^{\infty}$ — ортонормирована 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:

$$\langle r_k, r_m \rangle = \int_{\Delta_0^k} r_k(x)r_m(x)dx = \sum_{j=0}^{2^k-1} \int_{\Delta_j^k} r_k(x)r_m(x)dx = \sum_{j=0}^{2^k-1} \underbrace{c(m,k,j)}_{\substack{c(m,k,j) \\ \text{см. (1.9)}}} \underbrace{\int_{\Delta_j^k} r_k(x)dx}_{\substack{\pm 1 \\ \text{0, см. (1.10)}}} = 0. \blacktriangleright$$

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.

Упражнение. Покажите самостоятельно 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 при помощи 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, $n \leq 2^{l+1} - 1$. Таким образом, функция Уолша $w_n(x)$ определяется как произведение функций Радемахера с номерами, которые соответствуют 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

| 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 | $w_4(x) = r_2(x)$ |

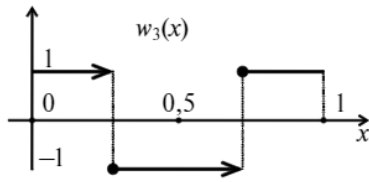


Рис. 1.3. Функция Уолша $w_3(x)$

$\{w_n(x)\}_{n=0}^{\infty}$ построением ее первых функций, см. таблицу. График функции $w_3(x) = r_0(x)r_1(x)$ приведен на рис. 1.3.

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

Упражнение. Постройте самостоятельно по определению (1.11) графики функций $w_4(x), \dots, w_7(x)$.

Теорема 1.9. Система функций Уолша (1.11) — ортонормирована на интервале $x \in [0, 1)$.

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

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

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

$$\begin{aligned} \langle w_k, w_n \rangle &= \int_0^1 r_{\tilde{j}}(x) \prod_{\substack{j:n_j \neq k_j \\ j < \tilde{j}}} r_j(x) dx = \sum_{m=0}^{2^{\tilde{j}}-1} \int_{\Delta_m^{\tilde{j}}} r_{\tilde{j}}(x) \overbrace{\prod_{\substack{j:n_j \neq k_j \\ j < \tilde{j}}} r_j(x)}^{\text{константа } c(k,n,m), \text{ см. (1.9)}} dx = \\ &= \sum_{m=0}^{2^{\tilde{j}}-1} \underbrace{c(k,n,m)}_{\pm 1} \underbrace{\int_{\Delta_m^{\tilde{j}}} r_{\tilde{j}}(x) dx}_{0, \text{ см. (1.10)}} = 0. \blacktriangleright \end{aligned}$$

Теорема 1.10. Система Уолша (1.11) является полной 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 в виде: $n = 2^k + m$,

где целые числа $k \geq 0$, $0 \leq m \leq 2^k - 1$ однозначно определяются по номеру $n > 0$. Тогда

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

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

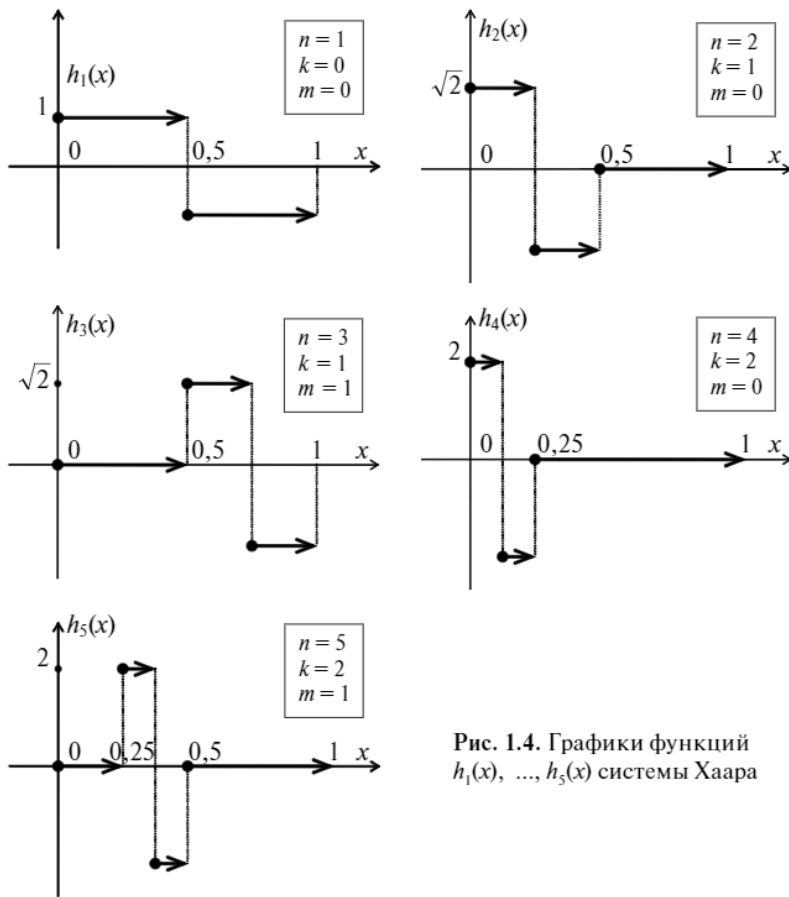


Рис. 1.4. Графики функций $h_1(x), \dots, h_5(x)$ системы Хаара

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

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 (1.12) — ортонормирована на интервале $x \in [0; 1)$.

◀ В соответствии с определением (1.12)

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

Рассмотрим теперь 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

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

как следует из приведенных выше свойств системы Хаара.

Случай 2. Пусть $\gamma = k$, но $m \neq \mu$. Так как (см. (1.12)) $h_n(x) = 0$ при $x \notin \Delta_m^k, h_v(x) = 0$ при $x \notin \Delta_\mu^k$, 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.)

Упражнение. Разложите функцию

$$f(x) = \begin{cases} 1, & x \in [1/4; 1/2) \\ 2, & x \in [1/2; 3/4) \\ 0, & x \notin [1/4; 3/4) \end{cases}$$

в ряд Фурье по системам Хаара и Уолша. Проверьте выполнение равенства Парсеваля.

1.4. Тригонометрические ряды Фурье. Явление Гиббса

Напомним следующую теорему.

Теорема 1.13. Если функция $f(t)$ имеет период T и является кусочно-гладкой, 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.

$$\frac{f(t+0) + f(t-0)}{2} = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos \frac{2\pi kt}{T} + b_k \sin \frac{2\pi kt}{T} \right), \quad (1.13)$$

где коэффициенты Фурье находятся по формулам:

$$\begin{aligned} a_k &= \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos \left(\frac{2\pi k}{T} t \right) dt, \quad k = 0, 1, \dots \\ b_k &= \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin \left(\frac{2\pi k}{T} t \right) dt, \quad k = 1, 2, \dots \end{aligned} \quad (1.14)$$

Упражнение. Убедитесь, что формулы (1.14) являются частным случаем (1.5) для $f \in L_2[-T/2; T/2]$, см. также пример 1.8.

Теорема 1.13 определяет условия *поточечной сходимости* ряда Фурье, т. е. те условия, при выполнении которых периодическая функция $f(t)$ может быть *точно* представлена рядом (1.13) в каждой точке числовой оси $t \in \mathbb{R}$. Так как система ортогональных функций $\{1, \cos(2\pi kt/T), \sin(2\pi kt/T)\}_{k=1}^{\infty}$ является полной в гильбертовом пространстве L_2 на любом отрезке длины T (см. пример 1.8), то последовательность частичных сумм

$$f_N(t) = \frac{a_0}{2} + \sum_{k=1}^N (a_k \cos(2\pi kt/T) + b_k \sin(2\pi kt/T))$$

сходится в норме (1.2), т. е. $\lim_{N \rightarrow \infty} \|f(t) - f_N(t)\| = 0$, и для $f(t)$ содержит в смысле этой нормы элементы наилучшего приближения из конечномерных подпространств с базисами $\{1, \cos(2\pi kt/T), \sin(2\pi kt/T)\}_{k=1}^N$.

¹ Если не говорится, какая система функций рассматривается в виде базиса для построения ряда Фурье (см. раздел 1.2), то традиционно подразумевается тригонометрическая система.

В целом ряде практических приложений ЦОС помимо нормы (1.2) приближение T -периодических функций частичными суммами ряда Фурье (1.13) рассматривается в смысле нормы $\|x\| = \max|x(t)|$ (максимального уклонения). Мы встретимся с такими задачами далее в главе 4. В случае если аппроксимируемая функция $f(t)$ является разрывной, поведение частичных сумм $f_N(t)$ ряда Фурье (1.13) характеризуется «всплесками», дающими максимальное уклонение $\delta = |f(t) - f_N(t)|$ именно вблизи точек разрыва. Причем величина этого максимального уклонения практически не зависит от количества слагаемых в частичной сумме. Рассмотрим это явление, известное как *эффект Гиббса*, на примере.

Пример 1.13. Для следующей функции периода $T = 2\pi$:

$$f(x) = \begin{cases} 1/2, & -\pi/2 < x < \pi/2, \\ -1/2, & -\pi \leq x < -\pi/2 \text{ или } \pi/2 < x \leq \pi, \\ 0, & x = -\pi/2 \text{ или } x = \pi/2, \end{cases} \quad f(x) = f(x + 2\pi),$$

оценить величину $A_K = \max_{x \in [-\pi; \pi]} |f_K(x)|$, где $f_K(x)$ — K -я частичная сумма ряда Фурье (1.13), для больших значений $K \gg 1$.

◀ Так как заданная функция $f(x)$ является четной, то $b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(2\pi kx/T) dx = 0$. Очевидно также, что $a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx = 0$. Поэтому ряд (1.13) также является четной функцией, принимая вид $f(x) = \sum_{k=1}^{\infty} a_k \cos(kx)$, где

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx = \frac{2}{\pi} \int_0^{\pi} f(x) \cos(kx) dx = \\ &= \frac{2}{\pi} \left(\frac{1}{2} \int_0^{\pi/2} \cos(kx) dx - \frac{1}{2} \int_{\pi/2}^{\pi} \cos(kx) dx \right) = \frac{2}{\pi k} \sin \frac{\pi k}{2}. \end{aligned}$$

Учитывая, что для четных значений $k = 2n$ коэффициенты ряда $a_{2n} = 0$, а для нечетных индексов $a_{2n-1} = \frac{2(-1)^{n-1}}{\pi(2n-1)}$, получаем следующее представление для ряда Фурье заданной функции:

$$f(x) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} \cos((2n-1)x).$$

Запишем его K -ю частичную сумму, полагая число $K = 2N$ (четным):

$$f_{2N}(x) = \frac{2}{\pi} \sum_{n=1}^{2N} \frac{(-1)^{n-1}}{2n-1} \cos((2n-1)x)$$

и построим графики данной функции (см. рис. 1.5) для значений $N = 1, 3, 10$. С увеличением числа N происходит приближение пика отклонения значения частичной суммы к точке разрыва аппроксимируемой функции $f(x)$, но видимого изменения абсолютной величины $A_{2N} = \max_{x \in [-\pi, \pi]} |f_{2N}(x)|$ с увеличением N на графиках не наблюдается.

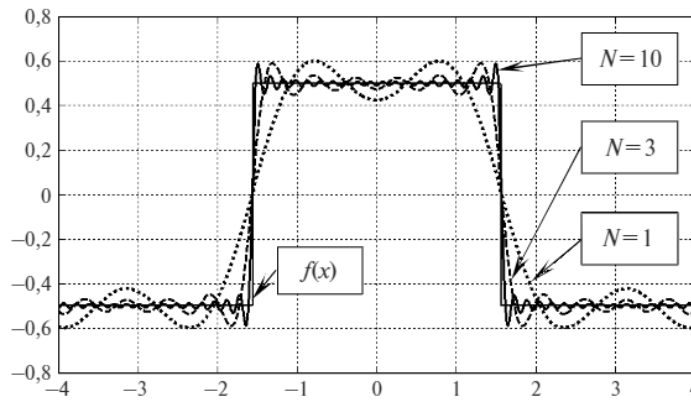


Рис. 1.5. Графики значений частичной суммы ряда Фурье $f_{2N}(x)$ для функции $f(x)$ (тонкая сплошная линия) из примера 1.13

В силу четности функций $f(x)$ и $f_{2N}(x)$ достаточно рассмотреть их на половине периода, для значений аргумента $x \in [0; \pi]$. Для определения точки максимального отклонения $A_{2N} = \max_{x \in [-\pi, \pi]} |f_{2N}(x)|$ частичной суммы $f_{2N}(x)$ найдем ее локальные экстремумы, ближайшие к разрыву $f(x)$ в точке $x = \pi/2$. Эти локальные экстремумы, как видно из графиков, соответствуют глобальным экстремумам частичной суммы $f_{2N}(x)$.

Исследуем производную частичной суммы:

$$\begin{aligned} f'_{2N}(x) &= \frac{2}{\pi} \sum_{n=1}^{2N} (-1)^n \sin((2n-1)x) = \\ &= \frac{2}{\pi} \sum_{n=1}^N (-1)^{2n-1} \sin((2(2n-1)-1)x) + (-1)^{2n} \sin((2(2n)-1)x) = \\ &= \frac{2}{\pi} \sum_{n=1}^N \sin((4n-1)x) - \sin((4n-3)x) = \frac{4 \sin x}{\pi} \sum_{n=1}^N \cos((2n-1)2x), \end{aligned}$$

так как $\sin \alpha - \sin \beta = 2 \cos \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}$. Используя формулу для суммы геометрической прогрессии, находим, что

$$\begin{aligned} \sum_{n=1}^N \cos((2n-1)2x) &= \operatorname{Re} \left(\sum_{n=1}^N e^{i(2n-1)2x} \right) = \operatorname{Re} \left(e^{-i2x} \sum_{n=1}^N e^{i4nx} \right) = \\ &= \operatorname{Re} \left(e^{i2x} \frac{e^{i4Nx} - 1}{e^{i4x} - 1} \right) = \operatorname{Re} \left(e^{i2x} \frac{e^{i2Nx} (e^{i2Nx} - e^{-i2Nx})}{e^{i2x} (e^{i2x} - e^{-i2x})} \right) = \\ &= \operatorname{Re} \left(\frac{e^{i2Nx} \sin(2Nx)}{\sin(2x)} \right) = \frac{\cos(2Nx) \sin(2Nx)}{\sin(2x)} = \frac{\sin(4Nx)}{2 \sin(2x)}. \end{aligned}$$

Поэтому окончательно

$$f'_{2N}(x) = \frac{2 \sin(x) \sin(4Nx)}{\pi \sin(2x)}.$$

Экстремумы частичной суммы $f_{2N}(x)$ удовлетворяют условию $f'_{2N}(x) = 0$. Ближайшие к точке разрыва $x = \pi/2$ экстремумы $f_{2N}(x)$ являются глобальными, выбираем их из общего набора решений уравнения $\sin(4Nx) = 0$: $x_k = \frac{\pi k}{4N}$, $k \in \mathbb{Z}$. Очевидно, что это две точки $x_{2N \mp 1} = \frac{\pi(2N \mp 1)}{4N} = \frac{\pi}{2} \mp \frac{\pi}{4N}$, лежащие слева (максимум $f_{2N}(x)$) и справа (минимум $f_{2N}(x)$) от разрыва $f(x)$.

Найдем значения частичной суммы в точках глобальных экстремумов $x_{\max} = \frac{\pi}{2} - \frac{\pi}{4N}$ и $x_{\min} = \frac{\pi}{2} + \frac{\pi}{4N}$:

$$f_{2N} \left(\frac{\pi}{2} \mp \frac{\pi}{4N} \right) = \frac{2}{\pi} \sum_{n=1}^{2N} \frac{(-1)^{n-1}}{2n-1} \cos \left(\frac{\pi(2n-1)}{2} \mp \frac{\pi(2n-1)}{4N} \right).$$

Воспользуемся соотношением $\cos(\alpha \mp \beta) = \cos \alpha \cos \beta \pm \sin \alpha \sin \beta$, тогда

$$\begin{aligned} &\cos \left(\frac{\pi(2n-1)}{2} \mp \frac{\pi(2n-1)}{4N} \right) = \\ &= \underbrace{\cos \frac{\pi(2n-1)}{2}}_0 \cos \frac{\pi(2n-1)}{4N} \pm \underbrace{\sin \frac{\pi(2n-1)}{2}}_{(-1)^{n-1}} \sin \frac{\pi(2n-1)}{4N} = \pm (-1)^{n-1} \sin \frac{\pi(2n-1)}{4N}, \\ f_{2N} \left(\frac{\pi}{2} \mp \frac{\pi}{4N} \right) &= \pm \frac{2}{\pi} \sum_{n=1}^{2N} \frac{1}{2n-1} \sin \frac{\pi(2n-1)}{4N} = \pm \sum_{n=1}^{2N} \left(\sin \frac{\pi(2n-1)}{4N} / \frac{\pi(2n-1)}{4N} \right) \frac{1}{2N}. \end{aligned}$$

Полученное выражение представляет собой интегральную сумму, в данном случае совпадающую с квадратурной формулой прямоугольников, для интеграла $\pm \int_0^1 \frac{\sin \pi x}{\pi x} dx = \pm \frac{1}{\pi} \int_0^\pi \frac{\sin x}{x} dx$. Поэтому при

достаточно больших значениях N для частичных сумм ряда Фурье получаем наибольшие отклонения от оси абсцисс

$$f_{2N}\left(\frac{\pi}{2} \mp \frac{\pi}{4N}\right) \approx \pm \lim_{N \rightarrow \infty} f_{2N}\left(\frac{\pi}{2} \mp \frac{\pi}{4N}\right) = \pm \frac{1}{\pi} \int_0^{\pi} \frac{\sin x}{x} dx \approx \pm 0,589490.$$

Соответствующий интеграл (так называемый интегральный синус) находится численно, например, как сумма быстро сходящегося знакопеременного ряда:

$$\int_0^{\pi} \frac{\sin x}{x} dx = \int_0^{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1} x^{2n-2}}{(2n-1)!} dx = -\frac{1}{\pi} \sum_{n=1}^{\infty} \frac{(-\pi^2)^n}{(2n-1)!(2n-1)} \approx 1,851937.$$

Как видим, с увеличением количества $2N$ слагаемых в частичной сумме $f_{2N}(x)$ ряда Фурье ее «пики» не уменьшаются, но приближаются к точке разрыва:

$$x_{\max} = \frac{\pi}{2} - \frac{\pi}{4N} \rightarrow \frac{\pi}{2} - 0, \quad x_{\min} = \frac{\pi}{2} + \frac{\pi}{4N} \rightarrow \frac{\pi}{2} + 0.$$

Амплитуда пиков $A_{2N} = \max_{x \in [-\pi, \pi]} |f_{2N}(x)| \approx 0,589490$, а отклонение от $f(x)$:

$$\delta_{2N} = |f(x_{\max}) - f_{2N}(x_{\max})| = |f(x_{\min}) - f_{2N}(x_{\min})| = A_{2N} - 0,5 \approx 0,08949.$$

Размах пульсации частичной суммы вблизи точки $x = \pi/2$ составляет величину $2A_{2N} = f_{2N}(x_{\max}) - f_{2N}(x_{\min}) \approx 1,17898$, т.е. примерно на 18% больше «скачка»

$$\left| f\left(\frac{\pi}{2} + 0\right) - f\left(\frac{\pi}{2} - 0\right) \right| = 1$$

функции $f(x)$ в точке разрыва $x = \pi/2$. ►

Рассмотренный в примере 1.13 частный случай явления Гиббса может быть обобщен [52] на случай произвольной (разрывной) функции $f(t)$, представимой в виде ряда Фурье (1.13). Вблизи точек разрыва $\{t_k\}$ величина максимальных пульсаций частичных сумм $f_k(t)$ ряда (1.13) практически не изменяется с увеличением количества слагаемых K в частичной сумме, и размах пульсации составляет величину около 118% от величины «скачка» $|f(t_k + 0) - f(t_k - 0)|$. С увеличением количества слагаемых K в частичной сумме ряда Фурье пики уклонения $f_k(t)$ от аппроксимируемой функции $f(t)$ приближаются к точкам разрыва.

Упражнение. Покажите, что для точек $\{x_{\text{extr}}\}$ локальных экстремумов частичной суммы ряда Фурье $f_k(x) = f_{2N}(x)$ из примера 1.13, которые

являются ближайшими на оси абсцисс к найденным в примере 1.13 точкам глобальных экстремумов, при $N \gg 1$ верна оценка:

$$f_{2N}(x_{\text{extr}}) \approx \pm \lim_{N \rightarrow \infty} f_{2N}(x_{\text{extr}}) = \pm \frac{1}{\pi} \int_0^{2\pi} \frac{\sin x}{x} dx \approx \pm 0,451412.$$

(Это означает, что вторые по величине пики уклонений частичной суммы $|f(x_{\text{extr}}) - f_{2N}(x_{\text{extr}})| \approx 0,5 - 0,451412 = 0,048588$ примерно вдвое меньше максимальных отклонений $\delta_{2N} \approx 0,08949$.) Как изменяется положение на оси абсцисс точек $\{x_{\text{extr}}\}$ с увеличением количества слагаемых $K = 2N$ в частичной сумме $f_K(x) = f_{2N}(x)$?

Вместо (1.13) часто удобнее использовать комплексную форму ряда Фурье:

$$\frac{f(t+0) + f(t-0)}{2} = \sum_{k=-\infty}^{\infty} c_k e^{i \frac{2\pi k}{T} t}, \quad (1.15)$$

где

$$c_k = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i \frac{2\pi k}{T} t} dt. \quad (1.16)$$

Несложно убедиться, что для вещественной функции $f(t)$ комплексные коэффициенты (1.16) ряда (1.15) обладают свойством сопряженной симметрии: $c_{-k} = (a_k + ib_k)/2$, $c_k = (a_k - ib_k)/2 = \bar{c}_{-k}$, где вещественные коэффициенты $\{a_k\}_{k=0}^{\infty}$, $\{b_k\}_{k=1}^{\infty}$ находятся по формулам (1.14).

Упражнение. Для функции единичного периода $f(t) = f(t+1)$, где $f(t) = t$ при $t \in [-1/2; 1/2)$, найдите разложение в ряд Фурье в форме (1.13) и (1.15).

1.5. Интеграл Фурье

Реальные сигналы чаще всего представляют собой аperiodические функции, искусственная периодизация которых, необходимая для корректного использования разложений (1.13) или (1.15), представляет собой неоднозначную процедуру, приводящую к искажению 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 (1.15) и (1.16) получаем:

$$\sum_{k=-\infty}^{\infty} c_k e^{i\frac{2\pi k}{T}t} = \sum_{k=-\infty}^{\infty} \left(\int_{-T/2}^{T/2} f(u) \exp(-i2\pi v_k u) du \right) e^{i2\pi v_k t} \Delta v_k.$$

Далее непериодический сигнал представим как периодический с бесконечно большим периодом, см. рис. 1.6.

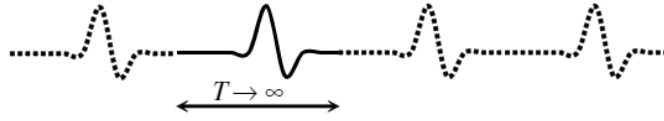


Рис. 1.6. Переход от периодического сигнала к непериодическому

Предположим, что существует интеграл (см. (1.16))

$$S(v_k) = \lim_{T \rightarrow \infty} c_k T = \lim_{T \rightarrow \infty} \int_{-T/2}^{+T/2} f(u) e^{-i\frac{2\pi k}{T}u} du = \int_{-\infty}^{+\infty} f(u) e^{-i2\pi v_k u} du.$$

При формальном переходе к пределу при $T \rightarrow \infty$ из ряда (1.15) получим:

$$\lim_{T \rightarrow \infty} \sum_{k=-\infty}^{\infty} c_k e^{i\frac{2\pi k}{T}t} = \lim_{\Delta v \rightarrow 0} \sum_{k=-\infty}^{\infty} S(v_k) e^{i2\pi v_k t} \Delta v = \int_{-\infty}^{+\infty} S(v) \exp(2\pi i v t) dv.$$

В случае существования последнего интеграла он понимается в смысле главного значения по Коши:

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

Данный интеграл носит название *интеграла Фурье*. Условия, которые гарантируют возможность представления функции в виде интеграла Фурье, определяет следующая теорема.

Теорема 1.14. Пусть функция $f(t)$ абсолютно интегрируема на всей числовой оси, т.е. $\int_{-\infty}^{+\infty} |f(t)| dt < \infty$, является кусочно-гладкой на любом конечном отрезке $t \in [a, b] \subset (-\infty; \infty)$ и в точках разрыва $f(t) = (f(t+0) + f(t-0))/2$. Тогда она представима в виде интеграла Фурье:

$$f(t) = \lim_{A \rightarrow \infty} \int_{-A}^{+A} S(v) \exp(2\pi i v t) dv = \int_{-\infty}^{+\infty} S(v) e^{2\pi i v t} dv, \quad (1.17)$$

где

$$S(v) = \int_{-\infty}^{+\infty} f(t) e^{-2\pi i v t} dt. \quad (1.18)$$

При этом $S(v)$ является непрерывной функцией.

Функция $S(v)$ из (1.18) носит название *частотного спектра*, или *спектральной плотности*, или *спектральной характеристики*

функции (сигнала) $f(t)$. Представления (1.18) и (1.17) называют соответственно прямым и обратным преобразованиями Фурье сигнала $f(t)$. Их записывают также с использованием в качестве аргумента спектральной плотности циклической частоты $\omega = 2\pi\nu$:

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

Таким образом, при выполнении условий теоремы 1.14 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: $f(t) \leftrightarrow S(\nu)$ (или $f(t) \leftrightarrow \hat{S}(\omega)$).

Отметим ряд важных свойств интегрального преобразования Фурье.

1°. Сопряженная симметрия. Для любой вещественной функции $f(t)$: $S(\nu) = S(-\nu)$. (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. $\forall f(t) \leftrightarrow S(\nu), \forall \alpha > 0$:

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

$$\begin{aligned} \blacktriangleleft S_\alpha(\nu) &= \int_{-\infty}^{+\infty} f(\alpha t)e^{-2\pi i \nu t} dt = \frac{1}{\alpha} \int_{-\infty}^{+\infty} f(\alpha t)e^{-2\pi i \frac{\nu}{\alpha}(\alpha t)} d(\alpha t) = \\ &= \frac{1}{\alpha} \int_{-\infty}^{+\infty} f(u)e^{-2\pi i (\nu/\alpha)u} du = \frac{1}{\alpha} S(\nu / \alpha). \quad \blacktriangleright \end{aligned}$$

4°. Задержка сигнала. $\forall f(t) \leftrightarrow S(\nu), \forall t_0: f(t - t_0) \leftrightarrow e^{-2\pi i \nu t_0} S(\nu)$. (Докажите самостоятельно.)

5°. Сдвиг спектра. $\forall f(t) \leftrightarrow S(\nu), \forall \alpha: f(t)e^{2\pi i \alpha t} \leftrightarrow S(\nu - \alpha)$. (Докажите самостоятельно.)

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

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

$$\begin{aligned} \blacktriangleleft S(\nu) &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} u(t)w(x-t)dt \right) e^{-2\pi i \nu x} dx = \int_{-\infty}^{+\infty} u(t) \underbrace{\int_{-\infty}^{+\infty} w(x-t)e^{-2\pi i \nu x} dx}_{\text{см. свойство 4}^\circ} dt = \\ &= \int_{-\infty}^{+\infty} u(t)S_w(\nu)e^{-2\pi i \nu t} dt = S_u(\nu)S_w(\nu). \quad \blacktriangleright \end{aligned}$$

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

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

(Докажите по аналогии с доказательством свойства 6°.)

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

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

(величину E называют *энергией сигнала*).

$$\begin{aligned} \blacktriangleleft E &= \int_{-\infty}^{\infty} \overline{f(t)}f(t)dt = \int_{-\infty}^{\infty} \overline{f(t)} \int_{-\infty}^{\infty} S(v)e^{2\pi ivt} dv dt = \int_{-\infty}^{\infty} S(v) \int_{-\infty}^{\infty} \overline{f(t)}e^{2\pi ivt} dt dv = \\ &= \int_{-\infty}^{\infty} S(v) \int_{-\infty}^{\infty} f(t)e^{-2\pi ivt} dt dv = \int_{-\infty}^{\infty} S(v)\overline{S(v)}dv = \int_{-\infty}^{\infty} |S(v)|^2 dv. \quad \blacktriangleright \end{aligned}$$

9°. Дифференцирование во временной области. Если $f(t) \leftrightarrow S(v)$ и функция $f(t)$ дифференцируема, причем $\lim_{t \rightarrow \pm\infty} f(t) = 0$, то

$$f'(t) \leftrightarrow 2\pi ivS(v).$$

◀ Интегрируя по частям, имеем:

$$\int_{-\infty}^{\infty} f'(t)e^{-2\pi ivt} dt = \underbrace{f(t)e^{-2\pi ivt}}_0 \Big|_{t=-\infty}^{t=+\infty} - \int_{-\infty}^{\infty} (-2\pi iv)f(t)e^{-2\pi ivt} dt = 2\pi ivS(v). \quad \blacktriangleright$$

10°. Дифференцирование в частотной области. Если $f(t) \leftrightarrow S(v)$ и функция $S(v)$ дифференцируема, причем $\lim_{v \rightarrow \pm\infty} S(v) = 0$, то

$$-2\pi itf(t) \leftrightarrow S'(v).$$

(Докажите самостоятельно, аналогично доказательству свойства 9°.)

Упражнение. Покажите, что для функции и ее спектра $f(t) \leftrightarrow S(v)$

$$\forall m \in \mathbb{N} \quad \text{условие} \quad \left. \frac{d^m}{dv^m} S(v) \right|_{v=0} = 0 \quad \text{эквивалентно} \quad \text{условию} \quad \int_{-\infty}^{\infty} t^m f(t) dt = 0.$$

Упражнение. Сформулируйте свойства 1°–10° пары преобразований Фурье $f(t) \leftrightarrow \hat{S}(\omega)$, записанных для циклической частоты ω .

Определение. Амплитудным спектром сигнала $f(t)$ называется модуль $|S(v)|$ спектральной плотности $S(v)$, а фазовым спектром — главное значение ее аргумента $\varphi(v) = \arg S(v) \in (-\pi; \pi]$.

Амплитудный и фазовый спектры позволяют записать спектральную плотность в показательной форме: $S(\nu) = |S(\nu)|e^{i\varphi(\nu)}$. Если $S(\nu) = 0$, то значение $\varphi(\nu)$ не определено.

Замечания. Из свойства 4° следует, что сдвиги сигнала во временной области влияют в частотной области лишь на фазовый спектр, но не изменяют амплитудный спектр сигнала. Из свойства 1° следует, что для вещественных сигналов амплитудный спектр является четной, а фазовый спектр — нечетной функцией.

Пример 1.14. Найти амплитудный и фазовый спектры сигнала, представляющего собой прямоугольный импульс длительности T :

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

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

$$S_g(\nu) = \int_{-\infty}^{+\infty} g(t)e^{-2\pi i\nu t} dt = \int_{-T/2}^{+T/2} e^{-2\pi i\nu t} dt = \frac{e^{\pi i\nu T} - e^{-\pi i\nu T}}{2\pi i\nu} = \frac{\sin(\pi\nu T)}{\pi\nu},$$

где полагаем $S_g(0) = T$. Так как $f(t) = g(t - 0,5T)$, то на основании свойства 4° получаем

$$S_f(\nu) = e^{-\pi i\nu T} \frac{\sin(\pi\nu T)}{\pi\nu},$$

откуда амплитудный спектр

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

Для фазового спектра $\varphi(\nu)$ рассмотрим сначала частоты $\nu > 0$.

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

При $\nu \in (2k/T; (2k+1)/T)$, $k = 0, 1, \dots$, получаем: $|\sin(\pi\nu T)| = \sin(\pi\nu T)$,

$$S_f(\nu) = \frac{\sin(\pi\nu T)}{\pi\nu} e^{-\pi i\nu T} = \left| \frac{\sin(\pi\nu T)}{\pi\nu} \right| e^{-\pi i\nu T} = |S_f(\nu)| e^{i\varphi(\nu)}, \quad \text{т.е.}$$

$$\varphi(\nu) = \arg(e^{-i\pi\nu T}), \quad \text{причем} \quad \varphi(\nu) = \varphi\left(\nu + \frac{2}{T}\right).$$

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

$$|\sin(\pi\nu T)| = -\sin(\pi\nu T),$$

$$S_f(\nu) = \frac{\sin(\pi\nu T)}{\pi\nu} e^{-\pi\nu T} = - \left| \frac{\sin(\pi\nu T)}{\pi\nu} \right| e^{-\pi\nu T} = |S_f(\nu)| e^{i\varphi(\nu)}, \quad \text{т.е.}$$

$$e^{i\varphi(\nu)} = -e^{-\pi\nu T} = e^{-i\pi\nu T + i\pi},$$

$$\varphi(\nu) = \arg(e^{-i\pi(\nu T - 1)}), \quad \text{и вновь} \quad \varphi(\nu) = \varphi\left(\nu + \frac{2}{T}\right).$$

В силу полученной периодичности фазового спектра для положительных частот $\nu > 0$ достаточно привести один период $\varphi(\nu)$. Для $\nu \in (0; 2/T)$ имеем:

$$\varphi(\nu) = \begin{cases} -\pi\nu T, & \text{при } \nu \in (0; 1/T) \\ -\pi\nu T + \pi, & \text{при } \nu \in (1/T; 2/T) \end{cases}.$$

Так как $S_f(0) = T$, то $\varphi(0) = 0$. Вид функции $\varphi(\nu)$ для $\nu < 0$ находим на основании того, что фазовый спектр является нечетной функцией.

Графики амплитудного и фазового спектров приведены на рис. 1.7. ▶

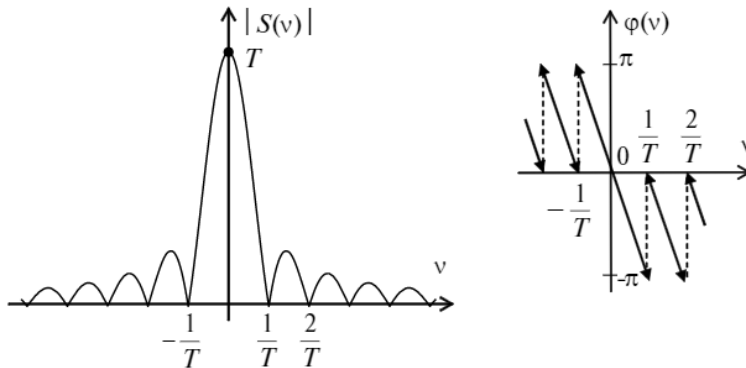


Рис. 1.7. Графики амплитудного (слева) и фазового (справа) спектров сигнала из примера 1.14

Упражнение. Используя свойства 2°–4° преобразования Фурье и результаты решения примера 1.14, найдите амплитудный и фазовый спектры функции

$$f(x) = \begin{cases} 1, & x \in [-1; 0) \\ -1, & x \in [0; 1) \\ 0, & x \notin [-1; 1) \end{cases}.$$

1.6. Принцип неопределенности время–частотного представления сигналов

Определение. *Носителем* Ω функции $f(x)$ назовем замыкание множества аргументов x , при которых $f(x)$ принимает ненулевые значения, т. е. $\Omega = \overline{\{x \mid f(x) \neq 0\}}$. Обозначаем: $\Omega = \text{supp } f(x)$. Будем говорить, что функция имеет *ограниченный* (или *компактный*) носитель, если существует конечный отрезок $[a, b]$, полностью содержащий этот 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 1.15. Пусть выполнены условия теоремы 1.14 и ненулевые функции $f(t)$ и $S(v)$ связаны соотношениями (1.17) и (1.18). Тогда пара функций $f(t) \leftrightarrow S(v)$ не может одновременно иметь ограниченные носители.

◀ Допустим противное, т. е. ненулевые функции $f(t) \leftrightarrow S(v)$ имеют ограниченные носители одновременно. Тогда существуют конечные отрезки $[-a; a] \supset \text{supp } f(t)$, $[-b; b] \supset \text{supp } S(v)$ и интегралы (1.17), (1.18) можно записать в виде:

$$f(t) = \int_{-b}^b S(v) e^{2\pi i vt} dv, \quad S(v) = \int_{-a}^a f(t) e^{-2\pi i vt} dt.$$

На основании теоремы 1.14 спектральная плотность $S(v)$ является непрерывной функцией, поэтому согласно теореме о дифференцировании интеграла, зависящего от параметра, существуют интегралы:

$$f^{(n)}(t) = \int_{-b}^b \frac{\partial^n}{\partial t^n} (S(v) e^{2\pi i vt}) dv = (2\pi i)^n \int_{-b}^b v^n S(v) e^{2\pi i vt} dv,$$

$n = 0, 1, \dots$, так как подынтегральные функции непрерывны по переменной v и по параметру t на множестве $\{(t, v) \mid v \in [-b; b], t \in \mathbb{R}\}$.

Таким образом, функция $f(t)$ в каждой точке $t_0 \in (-\infty; \infty)$ имеет производные любого порядка и, следовательно, может быть представлена рядом Тейлора:

$$f(t) = \sum_{n=0}^{\infty} \frac{f^{(n)}(t_0)}{n!} (t - t_0)^n.$$

Так как $f(t)$ имеет ограниченный носитель, то найдется некоторый отрезок $t \in [c; d]$, для которого $f(t) \equiv 0$. Для любой внутренней точки этого отрезка, например его середины $t_0 = (c+d)/2$, имеем $f^{(n)}(t_0) = 0$, $n = 0, 1, \dots$, и разложение в ряд Тейлора дает $\forall t \in \mathbb{R}$

$$f(t) = \sum_{n=0}^{\infty} \frac{(t-t_0)^n}{n!} \underbrace{f^{(n)}(t_0)}_0 \equiv 0.$$

Мы получили противоречие условию $f(t) \neq 0$, $t \in \mathbb{R}$, следовательно, функция $f(t)$ и ее спектр $S(\nu)$ не могут одновременно иметь ограниченные носители. ►

Из теоремы 1.15 следует, что сигналы конечной длительности имеют неограниченную *частотную полосу* (т.е. носитель спектральной плотности). Так, в примере 1.14 мы рассматривали функцию, имеющую конечный носитель во временной области, $\text{supp } f(t) = [0; T]$, и видели, что в частотной области носитель спектра $S(\nu) = e^{-\pi\nu T} \frac{\sin \pi\nu T}{\pi\nu\sqrt{T}}$ совпадает со всей числовой осью, $\text{supp } S(\nu) = (-\infty; \infty)$. Однако на практике почти всегда необходимо задаваться требованиями ограниченной (конечной) частотной полосы. То есть для произвольного сигнала (функции) $f(t)$ необходимо каким-то образом определить его частотную полосу $\nu \in [\nu_1; \nu_2]$, вполне характеризующую сигнал в частотной области. Под *шириной полосы* спектра тогда понимается величина $\Delta\nu = \nu_2 - \nu_1$.

Единого строгого подхода для определения частотной полосы сигнала, реально имеющего бесконечную ширину спектра, нет. На практике обычно выбирают на оси частот такой отрезок $\nu \in [\nu_1; \nu_2]$, который содержит основную часть $E_{\Delta\nu}$ энергии сигнала E , т.е.

$$E = \int_{-\infty}^{+\infty} |S(\nu)|^2 d\nu \approx E_{\Delta\nu} = \int_{\nu_1}^{\nu_2} |S(\nu)|^2 d\nu.$$

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

$$\tilde{S}(\nu) = \begin{cases} S(\nu), & \text{при } \nu \in [\nu_1; \nu_2] \\ 0, & \text{при } \nu \notin [\nu_1; \nu_2] \end{cases}$$

и соответствующий ему искаженный сигнал $\tilde{f}(t)$, то, очевидно, в силу свойства 2° преобразования Фурье имеем $S(v) - \tilde{S}(v) \leftrightarrow f(t) - \tilde{f}(t)$, а на основании свойства 8° энергия ошибки $\epsilon(t) = f(t) - \tilde{f}(t)$:

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

Так, для сигнала из примера 1.14 в качестве полосы сигнала можно было бы взять отрезок $v \in [-1/T; 1/T]$ (тогда $E_{\Delta v}/E \approx 0,90$) или $v \in [-2/T; 2/T]$ (тогда $E_{\Delta v}/E \approx 0,95$).

Для пояснения принципа неопределенности время-частотного представления сигналов более удобен иной подход к определению ширины полосы в частотной области и длительности сигнала во временной области. Положим, что энергия вещественного 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° преобразования Фурье. Поэтому для первого начального момента (среднего значения распределения энергии в частотной области) всегда получим

$$m_v = \int_{-\infty}^{+\infty} v |S(v)|^2 dv = 0$$

в силу того, что подынтегральная функция нечетна. Локализацию энергии в частотной области можно характеризовать по величине второго центрального момента

$$D_v = \int_{-\infty}^{+\infty} (v - m_v)^2 |S(v)|^2 dv = \int_{-\infty}^{+\infty} v^2 |S(v)|^2 dv, \quad (1.19)$$

который представляет собой меру «разброса» энергии в частотной области относительно m_v . Величина $\sqrt{D_v}$, определяемая из выражения (1.19), представляет собой среднеквадратичное значение распределения энергии в частотной области, поэтому естественно назвать соответствующую частотную полосу $v \in [-\sqrt{D_v}, \sqrt{D_v}]$ среднеквадратичной частотной полосой, ширина которой $\Delta v = 2\sqrt{D_v}$.

Во временной области функцией плотности распределения энергии сигнала является $\gamma(t) = |f(t)|^2$. Аналогично, в качестве меры длительности сигнала возьмем удвоенную величину среднеквадратичной длительности $\Delta t = 2\sqrt{D_t}$, определяемую по значению второго центрального момента

$$D_t = \int_{-\infty}^{+\infty} (t - m_t)^2 |f(t)|^2 dt = \int_{-\infty}^{+\infty} t^2 |f(t)|^2 dt - m_t^2, \quad (1.20)$$

где $m_t = \int_{-\infty}^{+\infty} t |f(t)|^2 dt$ — среднее значение для распределения энергии сигнала во временной области. Будем также называть величины $\Delta\nu = 2\sqrt{D_\nu}$ и $\Delta t = 2\sqrt{D_t}$, определенные при помощи формул (1.19) и (1.20), *эффективными* значениями ширины полосы $\Delta\nu$ и длительности Δt сигнала соответственно.

Величины (1.19) и (1.20) характеризуют локализацию энергии сигнала: чем меньше среднеквадратичная полоса (длительность), тем выше концентрация энергии в частотной (временной) области. *Принцип неопределенности* гласит, что добиться высокой локализации энергии одновременно и во временной, и в частотной областях невозможно. Так, верна следующая теорема.

Теорема 1.16. Для дифференцируемых вещественных сигналов $f(t)$ единичной энергии таких, что сходится интеграл $\int_{-\infty}^{+\infty} t |f(t)|^2 dt$ и $\lim_{t \rightarrow \pm\infty} t (f(t))^2 = 0$, произведение ширины полосы $\Delta\nu = 2\sqrt{D_\nu}$ и длительности $\Delta t = 2\sqrt{D_t}$ ограничено снизу:

$$\Delta t \Delta \nu \geq 1/\pi. \quad (1.21)$$

◀ 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, $f(t) \rightarrow f(t+m)$, не изменяющий его амплитудный спектр, см. свойство 4° преобразования Фурье).

Поскольку $f'(t) \leftrightarrow 2\pi i \nu S(\nu)$ (см. свойство 9°), на основании равенства Парсеваля (свойство 8°) имеем: $\int_{-\infty}^{+\infty} (f'(t))^2 dt = (2\pi)^2 \int_{-\infty}^{+\infty} \nu^2 |S(\nu)|^2 d\nu$.

Тогда

$$D_\nu D_t = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} (f'(t))^2 dt \int_{-\infty}^{+\infty} t^2 (f(t))^2 dt = \frac{1}{(2\pi)^2} \|f'(t)\|^2 \|tf(t)\|^2,$$

где норма вещественной функции индуцирована скалярным произведением $\langle f(t), g(t) \rangle = \int_{-\infty}^{+\infty} f(t)g(t)dt$ (см. раздел 1.2). Так как на основании неравенства Коши — Буняковского (лемма 1.3)

то

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

$$\Delta v \Delta t = 4\sqrt{D_v D_t} \geq \frac{2}{\pi} \left| \int_{-\infty}^{\infty} f'(t) tf(t) dt \right| = \frac{1}{\pi} \left| \underbrace{t(f(t))^2}_{0-0} \Big|_{t=-\infty}^{t=+\infty} - \int_{-\infty}^{\infty} (f(t))^2 dt \right| = \frac{1}{\pi} \blacktriangleright$$

Упражнение. Покажите, что для сигнала с энергией $E = \int_{-\infty}^{\infty} |f(t)|^2 dt$ оценка (1.21) принимает следующий общий вид: $\Delta t \Delta v \geq E/\pi$.

Пример 1.15. Показать, что равенство в оценке (1.21) достигается для гауссова импульса, т.е. для функции вида $f(t) = Ce^{-kt^2}$, где C, k — некоторые константы ($k > 0$).

◀ Достаточно рассмотреть случай единичной энергии сигнала, положив $C = \sqrt[4]{2k/\pi}$ — убедитесь, что в этом случае энергия сигнала $E = \int_{-\infty}^{\infty} |f(t)|^2 dt = 1$, учитывая значение интеграла Пуассона $\int_{-\infty}^{\infty} \exp(-t^2/2) dt = \sqrt{2\pi}$. 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 $\frac{1}{\sqrt[4]{2\pi}} \int_{-\infty}^{\infty} t^2 \exp(-t^2/2) dt = 1$ (как выражение для дисперсии стандартного нормального закона), то и для эффективной длительности сигнала во временной области также получаем $D_t = \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt = 1$, т.е. $\Delta t = 2\sqrt{D_t} = 2$.

Для производной спектральной плотности имеем:

$$S'(v) = \frac{1}{\sqrt[4]{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/4} e^{-2\pi i vt} (-2\pi i t) dt = \frac{4\pi i}{\sqrt[4]{2\pi}} \int_{-\infty}^{\infty} e^{-2\pi i vt} d(e^{-t^2/4}) =$$

$$= \frac{4\pi i}{\sqrt[4]{2\pi}} \left(\underbrace{e^{-2\pi i vt - (t^2/4)}}_{0-0} \Big|_{t=-\infty}^{t=+\infty} - \int_{-\infty}^{\infty} e^{-t^2/4} (-2\pi i v) e^{-2\pi i vt} dt \right) = -8\pi^2 v S(v).$$

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

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

Окончательно, получаем спектральную плотность $S(\nu) = \sqrt[4]{8\pi} e^{-4\pi^2\nu^2}$. Тогда для эффективной полосы имеем:

$$D_\nu = \sqrt{8\pi} \int_{-\infty}^{+\infty} \nu^2 e^{-8\pi^2\nu^2} d\nu = \frac{\sqrt{8\pi}}{4\pi(4\pi)^2} \int_{-\infty}^{+\infty} (4\pi\nu)^2 \underbrace{e^{-(4\pi\nu)^2/2}}_{\sqrt{2\pi}} d(4\pi\nu) = \frac{1}{(4\pi)^2}.$$

Отсюда $\Delta\nu = 2\sqrt{D_\nu} = \frac{1}{2\pi}$ и $\Delta\nu\Delta t = 1/\pi$. ▶

Упражнение. Используя свойства 2° и 3° интеграла Фурье и результаты примера 1.15, убедитесь, что для произвольного числа $k > 0$ гауссову импульсу $f(t) = \sqrt[4]{2k/\pi} e^{-kt^2}$ соответствует спектр $S(\nu) = \sqrt[4]{2\pi/k} e^{-\pi^2\nu^2/k}$.

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. Построив прямоугольную систему координат, осями которой являются время и частота, каждому сигналу на полученной плоскости «Время×Частота» можно поставить в соответствие некоторую прямоугольную область Ω локализации сигнала с длинами сторон Δt и $\Delta\nu$. При этом, если понимать Δt и $\Delta\nu$ в смысле эффективных значений, то для сигнала единичной энергии площадь данной области в соответствии с теоремой 1.16 не может быть меньше величины $1/\pi$.

Пример 1.16. Изобразить на плоскости «Время×Частота» область $\Omega = \Delta t \times \Delta\nu$ локализации сигнала из примера 1.14 для эффективных значений полосы и длительности.

◀ 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:

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

В частотной области

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

т.е. эффективная полоса сигнала в данном случае неограничена: $v \in (-\infty; \infty)$. Область $\Omega = \Delta t \times \Delta v$ отражена на рис. 1.8. ►

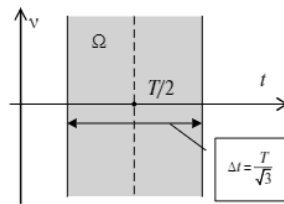


Рис. 1.8. Время-частотная локализация функции из примера 1.14

1.7. Обобщенное преобразование Фурье

Введем сначала важное для многих теоретических вопросов цифровой обработки сигналов понятие δ -функции Дирака. Положим для $\tau > 0$:

$$u_\tau(t) = \begin{cases} 1/\tau, & \text{при } t \in [-\tau/2; \tau/2] \\ 0, & \text{при } t \notin [-\tau/2; \tau/2] \end{cases}$$

Тогда для любой непрерывной функции $f(t)$ на основании интегральной теоремы о среднем найдется такая точка $\xi \in [-\tau/2; \tau/2]$, что

$$\int_{-\infty}^{+\infty} u_\tau(t) f(t) dt = \frac{1}{\tau} \int_{-\tau/2}^{+\tau/2} f(t) dt = f(\xi).$$

Поэтому если $f(t)$ — непрерывная в окрестности точки $t = 0$ функция, то

$$\lim_{\tau \rightarrow 0} \int_{-\infty}^{+\infty} u_\tau(t) f(t) dt = f(0).$$

Обозначим формально $\delta(t) = \lim_{\tau \rightarrow 0} u_\tau(t) = \begin{cases} +\infty, & \text{при } t = 0 \\ 0, & \text{при } t \neq 0 \end{cases}$ и будем называть $\delta(t)$ *дельта-функцией Дирака* (или δ -функцией). Основное свойство δ -функции описывается равенством:

$$\int_{-\infty}^{+\infty} \delta(t) f(t) dt = \lim_{\tau \rightarrow 0} \int_{-\infty}^{+\infty} u_{\tau}(t) f(t) dt = f(0),$$

где $f(t)$ — непрерывная в точке $t = 0$ функция.

При выполнении условий теоремы 1.14 между сигналом и его спектром существует взаимно однозначное соответствие:

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

Рассмотрим функцию $g(t) = 1$. Условия теоремы 1.14, очевидно, для нее не выполнены, и спектр $S_g(\nu)$, т. е. понимаемый в традиционном смысле интеграл $\int_{-\infty}^{+\infty} e^{-2\pi i \nu t} dt$, не существует. Однако если положить, что $S_g(\nu) = \delta(\nu)$, то с учетом рассмотренного выше свойства δ -функции запись обратного преобразования Фурье не вызывает затруднений и дает точное восстановление функции $g(t)$:

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

Для того чтобы расширить класс функций, для которых применимо интегральное преобразование Фурье (1.18), положим, по определению, что

$$\int_{-\infty}^{+\infty} e^{-2\pi i \nu t} dt = \delta(\nu). \quad (1.22)$$

Замечание. Эквивалентными соотношению (1.22) являются следующие определения. В силу вещественности δ -функции:

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

В силу симметрии выражения (1.22) относительно переменных ν и t :

$$\int_{-\infty}^{+\infty} e^{2\pi i \nu t} d\nu = \int_{-\infty}^{+\infty} e^{-2\pi i \nu t} d\nu = \delta(t).$$

Покажем, что интеграл $\int_{-\infty}^{+\infty} e^{-2\pi i \nu t} dt$ действительно проявляет свойства дельта-функции. Пусть $\varphi(t)$ — некоторая непрерывная в окрестности точки $t = 0$ функция, отвечающая условиям теоремы 1.14. Тогда

$$\int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} e^{-2\pi i \nu t} d\nu \right) \varphi(t) dt = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \varphi(t) e^{-2\pi i \nu t} dt \right) d\nu = \int_{-\infty}^{+\infty} S_{\varphi}(\nu) d\nu = \varphi(0),$$

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

Вводя в рассмотрение представление (1.22), мы сразу же расширили область применимости интегральных преобразований (1.17) и (1.18), превратив ряды Фурье для периодических функций ((1.15), (1.16)) в частный случай интегральных преобразований. Действительно, исходя из (1.22), функции $\varphi_k(t) = e^{i\frac{2\pi}{T}kt}$ соответствует обобщенный спектр

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

Поэтому для произвольной функции периода T , представимой в виде ряда (1.15), $f(t) = \sum_{k=-\infty}^{+\infty} c_k \varphi_k(t)$, обобщенное интегральное преобразование дает спектр

$$S(\nu) = \int_{-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} c_k \varphi_k(t) \right) e^{-2\pi i\nu t} dt = \sum_{k=-\infty}^{\infty} c_k \int_{-\infty}^{\infty} \varphi_k(t) e^{-2\pi i\nu t} dt = \sum_{k=-\infty}^{\infty} c_k \delta\left(\nu - \frac{k}{T}\right),$$

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

$$\begin{aligned} f(t) &= \int_{-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} c_k \delta\left(\nu - \frac{k}{T}\right) \right) e^{2\pi i\nu t} d\nu = \sum_{k=-\infty}^{\infty} c_k \left(\int_{-\infty}^{\infty} \delta\left(\nu - \frac{k}{T}\right) e^{2\pi i\nu t} d\nu \right) = \\ &= \sum_{k=-\infty}^{\infty} c_k e^{i\frac{2\pi}{T}kt} = \sum_{k=-\infty}^{\infty} c_k \varphi_k(t). \end{aligned}$$

Таким образом, интегралы в преобразовании Фурье (1.17) и (1.18) мы будем далее понимать обобщенно. Для обозначения преобразования Фурье будем использовать следующие сокращенные записи:

$$S(\nu) = \mathcal{F}\{f(t)\} \quad \text{— прямое преобразование (1.18),}$$

$$f(t) = \mathcal{F}^{-1}\{S(\nu)\} \quad \text{— обратное преобразование (1.17).}$$

Определение. Решетчатой будем называть функцию вида $g(x) = \sum_{m=-\infty}^{\infty} g_m \delta(x - m\Delta x)$, где $\{g_m\}$ — вещественные или комплексные числа, а константа (шаг аргумента) $\Delta x > 0$.

Из рассмотренного выше следует важное наблюдение: если функция (сигнал) $f(t)$ является T -периодической, то ее спектр $S(\nu) = \sum_{k=-\infty}^{\infty} c_k \delta\left(\nu - \frac{k}{T}\right)$ является решетчатой функцией и принимает ненулевые значения лишь для определенных равноотстоящих значений аргумента ν , а именно $\nu_k = k/T$, $k \in \mathbb{Z}$. Таким образом, спектр периодических функций полностью характеризуется набором коэффициентов (1.16) ряда (1.15):

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{i \frac{2\pi}{T} kt} \leftrightarrow S(\nu) = \sum_{k=-\infty}^{\infty} c_k \delta\left(\nu - \frac{k}{T}\right). \quad (1.23)$$

По этой причине под амплитудным спектром для периодических функций понимают набор модулей коэффициентов Фурье $\{|c_k|\}$, а под фазовым спектром — набор их аргументов $\{\arg c_k\}$.

Упражнение. Найдите обобщенный спектр функций $\cos t$, $\sin t$.

Преобразования (1.17) и (1.18) имеют сходную природу, differing только знаком при мнимом показателе подынтегральной экспоненты. Вследствие этого преобразования (1.17), (1.18) 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 T -периодическому сигналу соответствует решетчатый спектр с дискретным шагом частоты $\Delta\nu = 1/T$, мы можем ожидать, 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 $S(\nu)$ имеет период, равный $1/\Delta t$. То есть периодическая функция в одной области соответствует решетчатой функции в другой области, и наоборот.

Упражнение. Используя свойства 2° и 4° преобразования Фурье, покажите, что спектр решетчатой функции $f(t) = \sum_{k=-\infty}^{\infty} f_k \delta(t - k\Delta t)$ имеет вид $S(\nu) = \sum_{k=-\infty}^{\infty} f_k e^{-2\pi i \nu k \Delta t} = S\left(\nu + \frac{1}{\Delta t}\right)$.

Установленное выше влияние периодического или решетчатого характера функции на свойства ее преобразования Фурье делает практически очевидным вывод, что периодическая решетчатая функция имеет спектр, который также представляет собой решетчатую периодическую функцию. Убедимся в этом.

Пример 1.17. Найти спектр функции $f(t) = \Delta t \sum_{k=-\infty}^{\infty} f_k \delta(t - k\Delta t)$, где $f_k = f_{k+N}$.

◀ Очевидно, что функция $f(t) = f(t + T)$, где период $T = N\Delta t$:

$$f(t + N\Delta t) = \Delta t \sum_{k=-\infty}^{\infty} f_k \delta(t - (k - N)\Delta t) = \Delta t \sum_{k=-\infty}^{\infty} \underbrace{f_{k+N}}_{=f_k} \delta(t - k\Delta t) = f(t).$$

На основании (1.23) ее спектр представляет собой решетчатую функцию $S(\nu) = \sum_{k=-\infty}^{\infty} c_k \delta(\nu - k\Delta\nu)$, где $\Delta\nu = 1/T$, причем

$$\begin{aligned} c_k &= \frac{1}{T} \int_{-\Delta t/2}^{T-\Delta t/2} f(t) e^{-i\frac{2\pi k}{T}t} dt = \frac{\Delta t}{T} \int_{-\Delta t/2}^{\left(N-\frac{1}{2}\right)\Delta t} \underbrace{\left(\sum_{m=-\infty}^{\infty} f_m \delta(t - m\Delta t) \right)}_{\text{Учитываем } m: m\Delta t \in \left[\frac{-\Delta t}{2}; \left(N-\frac{1}{2}\right)\Delta t\right]} e^{-i\frac{2\pi k}{T}t} dt = \\ &= \frac{1}{N} \sum_{m=0}^{N-1} f_m \int_{-\Delta t/2}^{\left(N-\frac{1}{2}\right)\Delta t} \delta(t - m\Delta t) e^{-i\frac{2\pi k}{N\Delta t}t} dt = \frac{1}{N} \sum_{m=0}^{N-1} f_m e^{-i\frac{2\pi}{N}km} = c_{k+N}. \end{aligned}$$

Поэтому спектральная плотность также является периодической функцией, $S(\nu) = S(\nu + N\Delta\nu) = S\left(\nu + \frac{1}{\Delta t}\right)$. ▶

Упражнение. Покажите, что функции $G(t) = \Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t)$ в частотной области соответствует спектр $S_G(\nu) = \sum_{k=-\infty}^{\infty} \delta\left(\nu - \frac{k}{\Delta t}\right)$, причем

$$\Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t) = \sum_{k=-\infty}^{\infty} \exp\left(i\frac{2\pi k}{\Delta t}t\right). \quad (1.24)$$

Указание: воспользуйтесь результатами решения примера 1.17, положив $N = 1$ и $\forall k: f_k = 1$, после чего запишите оба соотношения из (1.23).

1.8. Энергетический спектр. Спектр мощности

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

$$B(\tau) = \int_{-\infty}^{+\infty} f(t)f(t-\tau)dt. \quad (1.25)$$

Функция $B(\tau)$ характеризует «степень сродства» сигнала и его сдвига по аргументу: чем больше значение интеграла (1.25), тем это сродство выше. Так, максимальное сродство наблюдается при $\tau = 0$, когда значение корреляционной функции (1.25) максимально и равно энергии сигнала E , вообще же $\forall \tau: |B(\tau)| \leq B(0) = E$. Действительно,

поскольку скалярным произведением функций $f(t), g(t)$ (элементов пространства $L_2(\mathbb{R})$) является величина $\langle f, g \rangle = \int_{-\infty}^{+\infty} f(t)g(t)dt$, то на основании неравенства Коши — Буняковского (лемма 1.3) получаем для функции (1.25) $\forall \tau$:

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

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 $B(\tau)$ (1.25) — четная функция.

Найдем преобразование Фурье для корреляционной функции (1.25):

$$\begin{aligned} \mathcal{F}(B(\tau)) &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(t)f(t-\tau)dt \right) e^{-2\pi i\nu\tau} d\tau = \\ &= \int_{-\infty}^{+\infty} f(t) \left(\int_{-\infty}^{+\infty} f(t+\tau)e^{2\pi i\nu\tau} d\tau \right) dt = \int_{-\infty}^{+\infty} f(t) \left(\int_{-\infty}^{+\infty} f(t+\tau)e^{2\pi i\nu(\tau+t)} d\tau \right) e^{-2\pi i\nu t} dt = \\ &= \int_{-\infty}^{+\infty} f(t)\overline{S_f(\nu)}e^{-2\pi i\nu t} dt = \overline{S_f(\nu)}S_f(\nu) = |S_f(\nu)|^2, \end{aligned} \quad (1.26)$$

где $S_f(\nu) = \mathcal{F}\{f(t)\}$. Таким образом, для $f(t) \in L_2(\mathbb{R})$ преобразование Фурье функции (1.25) представляет собой квадрат амплитудного спектра сигнала $f(t)$. Функция $\mathcal{F}\{B(\tau)\} = |S_f(\nu)|^2$ называется *энергетическим спектром* сигнала.

Для многих важных моделей детерминированных сигналов энергия $E = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \infty$, т. е. интеграл (1.25) для $\tau = 0$ не существует (например, для периодических сигналов). В этом случае вместо $B(\tau)$ (1.25) корреляционной будем называть функцию

$$R(\tau) = \lim_{T \rightarrow \infty} R_T(\tau), \quad (1.27)$$

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

если предел в (1.27) 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 $R(\tau)$ (1.27) — также четная функция и $|R(\tau)| \leq R(0)$.

Спектром мощности $s_f(\nu)$ детерминированного сигнала $f(t)$ назовем преобразование Фурье функции $R(\tau)$ (1.27), т. е.

$$s_f(\nu) = \mathcal{F}\{R(\tau)\}. \quad (1.28)$$

Пример 1.18. Найти спектр мощности сигнала $f(t) = \cos(2\pi\nu_0 t + \gamma)$, $\nu_0 \neq 0$.

◀ Данный сигнал представляет собой гармоническое колебание с частотой ν_0 и начальной фазой γ . Используя определение (1.27), получаем:

$$\begin{aligned} R(\tau) &= \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-T/2}^{T/2} \cos(2\pi\nu_0 t + \gamma) \cos(2\pi\nu_0(t - \tau) + \gamma) dt = \\ &= \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T/2}^{T/2} \cos(2\pi\nu_0(2t - \tau) + 2\gamma) + \cos(2\pi\nu_0 \tau) dt = \\ &= \lim_{T \rightarrow +\infty} \frac{1}{2T} \left(\underbrace{\int_{-T/2}^{T/2} \cos(2\pi\nu_0(2t - \tau) + 2\gamma) dt}_{\text{значение интеграла ограничено}} + T \cos(2\pi\nu_0 \tau) \right) = \frac{1}{2} \cos(2\pi\nu_0 \tau). \end{aligned}$$

Тогда в соответствии с определением (1.28) спектр мощности

$$s_f(\nu) = \mathcal{F}\left\{\frac{1}{2} \cos(2\pi\nu_0 \tau)\right\} = \frac{1}{4} \mathcal{F}\{e^{i2\pi\nu_0 \tau} + e^{-i2\pi\nu_0 \tau}\} = \frac{\delta(\nu - \nu_0) + \delta(\nu + \nu_0)}{4}.$$

Видим, что спектр мощности не зависит от начальной фазы колебания γ . ▶

Как и в рассмотренном примере 1.18, в общем случае энергетический спектр (1.26) и спектр мощности (1.28) также являются вещественными функциями, не зависящими от сдвига сигнала во временной области.

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

$$R_X(\tau) = M(X(t)X(t - \tau)) = \text{cov}(X(t), X(t - \tau)) + m_X^2, \quad (1.29)$$

где $m_X = M(X(t))$ — математическое ожидание процесса, а ковариационной функцией (КФ) называется

$$K_X(\tau) = \text{cov}(X(t), X(t - \tau)). \quad (1.30)$$

Функции (1.29) и (1.30) 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. Для центрированного процесса ($m_X = 0$), таким образом, его дисперсия характеризует мощность сигнала.

Второй начальный момент (1.29) для эргодического относительно ковариационной функции процесса $X(t)$ можно найти по формуле (1.27), полагая $f(t) = x(t)$, где $x(t)$ — любая реализация процесса $X(t)$. В силу такой общности формул (1.27) и (1.29) логично было бы определить спектр мощности случайного процесса $X(t)$ как преобразование Фурье второго начального момента $R_X(\tau)$ (1.29) по аналогии с определением (1.28). Тогда получили бы

$$\mathcal{F}\{R_X(\tau)\} = \mathcal{F}\{K_X(\tau) + m_X^2\} = \underbrace{\mathcal{F}\{K_X(\tau)\}}_{\text{Обозначим } \tilde{s}_X(\nu)} + \mathcal{F}\{m_X^2\} = \tilde{s}_X(\nu) + m_X^2\delta(\nu),$$

т. е. при $m_X \neq 0$ для частоты $\nu = 0$ имеем бесконечное значение спектральной плотности, которую можно понимать только в смысле обобщенного преобразования Фурье. Поэтому, по определению, *спектром мощности* $\tilde{s}_X(\nu)$ стационарного случайного процесса $X(t)$ называют преобразование Фурье КФ (1.30) (а не второго начального момента (1.29)), т. е.

$$\tilde{s}_X(\nu) = \mathcal{F}\{K_X(\tau)\}, \quad K_X(\tau) = \mathcal{F}^{-1}\{\tilde{s}_X(\nu)\}.$$

Пара данных соотношений, связывающих КФ сигнала и его спектр мощности посредством преобразования Фурье, носит название *теоремы Винера — Хинчина*.

Для центрированного ($m_X = 0$) случайного процесса $K_X(\tau) = R_X(\tau)$ и $\tilde{s}_X(\nu) = \mathcal{F}\{K_X(\tau)\} = \mathcal{F}\{R_X(\tau)\}$.

Упражнение. Найдите спектр мощности стационарного случайного процесса, имеющего ковариационную функцию $K_X(\tau) = e^{-|\tau|}$.

Если спектр мощности стационарного случайного сигнала постоянен, $\forall \nu \in \mathbb{R}: \tilde{s}_X(\nu) = C > 0$, то такой сигнал называют *белым шумом*. Ковариационная функция (1.30) белого шума с точностью до коэффициента C представляет собой дельта-функцию:

$K_X(\tau) = \mathcal{F}^{-1}\{\tilde{s}_X(\nu) = C\} = C\delta(\tau)$. Таким образом, даже как угодно близкие сечения белого шума некоррелированы, так как

$$\text{cov}(X(t), X(t - \tau)) = C\delta(\tau). \quad (1.31)$$

Белый шум — это физически нереализуемый процесс, поскольку его дисперсия (мощность) бесконечна: $\sigma_X^2 = C\delta(0)$, однако белый шум является полезной математической моделью, которая часто используется на практике для описания случайных сигналов, спектр мощности которых можно считать постоянным в некотором рассматриваемом диапазоне частот.

Дискретным белым шумом называют стационарный случайный процесс $X(n)$ с дискретным временем $n \in \mathbb{Z}$, который представляет собой некоррелированную последовательность, т. е. $\forall k, n \in \mathbb{Z}$:

$$\text{cov}(X(n), X(n - k)) = \begin{cases} \sigma_X^2, & \text{при } k = 0 \\ 0, & \text{при } k \neq 0 \end{cases}. \quad (1.32)$$

(Отметим, что иногда вместо некоррелированности (1.32) предполагается выполнение более жесткого условия независимости элементов случайной последовательности $X(n)$.) Дискретный белый шум является физически реализуемым случайным процессом.

ГЛАВА 2

ДИСКРЕТИЗАЦИЯ И КВАНТОВАНИЕ СИГНАЛОВ. ДИСКРЕТНЫЕ ОРТОГОНАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ

2.1. Преобразование непрерывных сигналов в дискретные

Под термином *сигнал* мы традиционно понимаем 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 — это функция дискретного аргумента $y = y(nT)$ с областью определения $D = \{nT | n \in \mathbb{Z}\}$. Значения $y(nT)$ называются *отсчетами* 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 $y(n) = \{0, 0,5, -0,7, 1\}$. Во многих случаях, в частности при спектральном анализе, для дискретного сигнала удобно использовать представление в форме решетчатой функции $\tilde{y}(t) = T \sum_{k=-\infty}^{\infty} y(kT) \delta(t - kT)$, с чем мы уже сталкивались ранее в разделе 1.7.

Под *цифровым* сигналом понимаем дискретный сигнал, который может принимать значения из конечного множества чисел, $\forall n y(n) \in \{d_1, \dots, d_M\}$. Возможные значения сигнала d_j ($j = 1, \dots, M$) называем *уровнями* 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 сигнала $f(t)$ производится равномерная выборка с периодом (шагом) дискретизации $T = \Delta t$: $f_n = f(n\Delta t)$, $n = 0, 1, \dots, N$. Для восстановления аналогового сигнала по отсчетам $\{f_0, f_1, \dots, f_N\}$ чаще всего используются интерполяционные методы, например, восстановленный сигнал можно представить в виде алгебраического многочлена $\tilde{f}(t) = \sum_{k=0}^N a_k t^k$, или в более общем виде полинома $\tilde{f}(t) = \sum_{k=0}^N c_k \varphi_k(t)$, 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 нужно 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. период дискретизации Δt нужно сделать по возможности бóльшим. Однако, увеличивая Δt , мы можем неадекватно описать исходный аналоговый сигнал через дискретное представление, т. е. получим неприемлемую ошибку $\|f(t) - \tilde{f}(t)\|$ при восстановлении сигнала $\tilde{f}(t) \approx f(t)$.

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

2.2. 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

$$\|f(t) - \tilde{f}(t)\| \leq \varepsilon_0. \quad (2.1)$$

В качестве нормы в (2.1) обычно берется наибольшее отклонение, $\|\varepsilon(t)\| = \max_{t \in [a; b]} |\varepsilon(t)|$, а для восстановления непрерывного сигнала $\tilde{f}(t)$ по дискретному $\{f_k\}$ используются алгебраические интерполяционные многочлены. Известно, что по $n+1$ точкам $\{(t_k, f_k)\}_{k=0}^{n}$ 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 [7, 15]:

$$|\varepsilon(t)| = |f(t) - \tilde{f}(t)| \leq \frac{M_{n+1}}{(n+1)!} \left| \prod_{k=0}^n (t - t_k) \right|, \quad (2.2)$$

$$\|\varepsilon(t)\| = \max_{t \in [a; b]} |\varepsilon(t)| \leq \frac{M_{n+1}}{(n+1)!} \max_{t \in [a; b]} \left| \prod_{k=0}^n (t - t_k) \right|, \quad (2.3)$$

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 аппаратные требования.

Итак, положим, что восстановление аналогового сигнала происходит с использованием степенного интерполяционного полинома. Тогда на основании формул (2.1) и (2.3) получаем следующий критерий выбора шага дискретизации:

$$\max_{t \in [a; b]} \left| \prod_{k=0}^n (t - t_k) \right| \leq \frac{(n+1)!}{M_{n+1}} \varepsilon_0. \quad (2.4)$$

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

Пример 2.1. 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 период 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 (2.4), т.е. $\max_{t \in [t_k; t_{k+1}]} |(t - t_k)(t - t_{k+1})| \leq 2\varepsilon_0 / M_2$. Так как

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

2.3. Частотный критерий выбора шага дискретизации

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

Пример 2.2. «Эффект наложения частот». Показать, что гармонические колебания $x_1(t) = \cos(2\pi\nu_1 t - \gamma)$ и $x_2(t) = \cos(2\pi\nu_2 t + \gamma)$ при выборе частоты дискретизации $f_{\text{д}} = 1/\Delta t = \nu_1 + \nu_2$ неотличимы, т.е. $\forall k \in \mathbb{Z}: x_1(k\Delta t) = x_2(k\Delta t)$.

◀ Обозначим $\alpha = \frac{\nu_2 - \nu_1}{2}$. Тогда $\nu_1 = \frac{f_{\text{д}}}{2} - \alpha$, $\nu_2 = \frac{f_{\text{д}}}{2} + \alpha$. Так как $\cos x - \cos y = 2 \sin \frac{x+y}{2} \sin \frac{y-x}{2}$, $f_{\text{д}} = 1/\Delta t$, то

$$x_1(k\Delta t) - x_2(k\Delta t) = \underbrace{2 \sin(\pi f_{\text{д}} k \Delta t)}_{\sin(\pi k) = 0} \sin(2\pi \alpha k \Delta t + \gamma) = 0.$$

Иллюстрация эффекта наложения частот для случая $\nu_1 = 2$, $\nu_2 = 6$, $\gamma = 0$ приведена на рис. 2.1. ▶

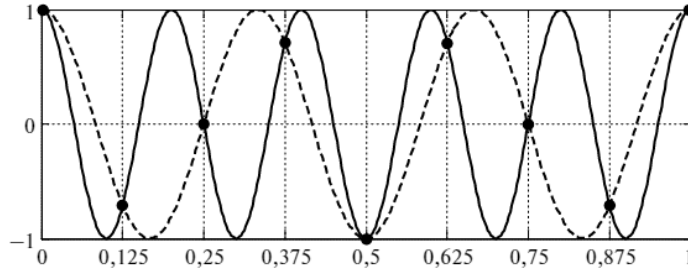


Рис. 2.1. Эффект наложения частот: функции $\varphi_1(t) = \cos(10\pi t)$ (сплошная линия) и $\varphi_2(t) = \cos(6\pi t)$ (пунктирная линия) совпадают в точках $t_k = k/8, k \in \mathbb{Z}$

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

Теорема 2.1. (Теорема отсчетов.) Если функция $f(t)$ имеет спектр с ограниченной частотной полосой, т.е. $S(\nu) = 0$ при $|\nu| \geq F$, то $f(t)$ может быть точно восстановлена по своим значениям в точках $t_k = k\Delta t$ ($k \in \mathbb{Z}$), где $\Delta t = 1/(2F)$, при помощи ряда Котельникова:

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\Delta t) \frac{\sin(2\pi F(t - k\Delta t))}{2\pi F(t - k\Delta t)}. \quad (2.5)$$

◀ Выразим функцию через ее спектр, $f(t) = \int_{-F}^F S(\nu)e^{2\pi i\nu t} d\nu$, тогда с использованием обозначения $\Delta t = 1/2F$ получим $\forall k \in \mathbb{Z}$:

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

Спектральную плотность $S(\nu)$, принимающую ненулевые значения только на интервале $\nu \in (-F; F)$, продолжим с периодом $2F$ на всю ось частот и полученную в результате функцию $\tilde{S}(\nu) = \sum_{m=-\infty}^{\infty} S(\nu - 2Fm)$ представим в виде ряда Фурье (1.15):

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

Сравнивая полученную запись для c_k с выражением для $f(k\Delta t)$, видим, что $c_k = \frac{1}{2F} f(-k\Delta t)$. Тогда

$$\begin{aligned} \tilde{S}(v) &= \frac{1}{2F} \sum_{k=-\infty}^{\infty} f(-k\Delta t) e^{\frac{2\pi j}{2F} kv} = \frac{1}{2F} \sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-2\pi jvk\Delta t}, \\ f(t) &= \int_{-F}^F \tilde{S}(v) e^{2\pi jvt} dv = \frac{1}{2F} \int_{-F}^F \left(\sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-2\pi jvk\Delta t} \right) e^{2\pi jvt} dv. \end{aligned}$$

Меняя в последнем выражении порядок интегрирования и суммирования, имеем:

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

Отсюда, учитывая, что $\int_{-F}^F e^{2\pi jv(t-k\Delta t)} dv = \frac{\sin(2\pi F(t-k\Delta t))}{\pi(t-k\Delta t)}$, получаем утверждение теоремы. ▶

Таким образом, при помощи формулы (2.5) 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 $t_k = k\Delta t$, которые полностью определяют функцию. Формула (2.5) задает функциональный ряд $f(t) = \sum_{k=-\infty}^{\infty} \lambda_k \varphi_k(t)$ по системе функций

$$\left\{ \varphi_k(t) = \frac{\sin(2\pi F(t-k\Delta t))}{2\pi F(t-k\Delta t)} \right\}_{k \in \mathbb{Z}}, \quad (2.6)$$

представляющих собой сдвиги по оси абсцисс на величины $k\Delta t = k/(2F)$ функции $\varphi_0(t) = \varphi(t) = \frac{\sin 2\pi Ft}{2\pi Ft}$. Базисные функции $\{\varphi_k(t)\}$ 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 $\varphi_k(k\Delta t) = 1$, см. рис. 2.2. То есть в представлении (2.5) для $f(t)$ получаем, что в точках $t = t_m = m\Delta t$ в сумме (2.5) имеется единственное ненулевое слагаемое для $k = m$.

Удобным свойством системы функций (2.6) является ее ортогональность на числовой оси $t \in (-\infty; \infty)$.

Теорема 2.2. Для системы функций (2.6) при $\Delta t = 1/(2F)$ верно:

$$\langle \varphi_k(t), \varphi_m(t) \rangle = \int_{-\infty}^{\infty} \varphi_k(t) \varphi_m(t) dt = \begin{cases} 1/(2F) & \text{при } k = m \\ 0 & \text{при } k \neq m \end{cases}.$$

◀ Заметим, что спектр функции $\varphi_0(t) = \varphi(t) = \frac{\sin 2\pi Ft}{2\pi Ft}$ имеет вид:

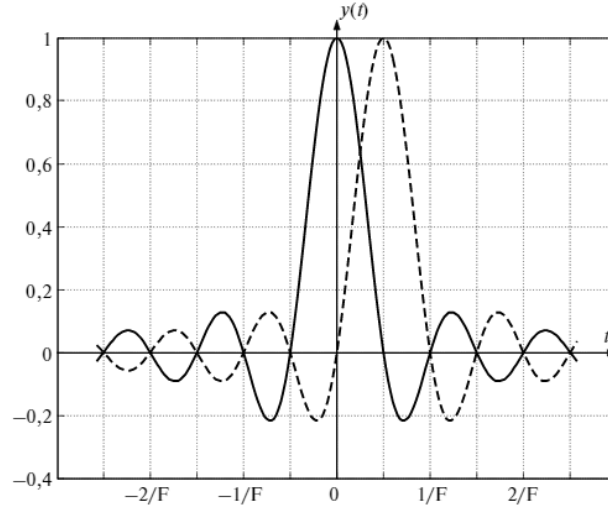


Рис. 2.2. Графики функций $\varphi_0(t)$ (сплошная линия) и $\varphi_1(t)$ (пунктирная линия) системы (2.6)

$$S_\varphi(\nu) = \begin{cases} 1/(2F), & \text{при } \nu \in [-F; F] \\ 0, & \text{при } \nu \notin [-F; F] \end{cases}$$

Действительно, выполняя обратное преобразование Фурье, получаем: $\int_{-\infty}^{\infty} S_\varphi(\nu) e^{2\pi i \nu t} d\nu = \frac{1}{2F} \int_{-F}^F e^{2\pi i \nu t} d\nu = \frac{\sin 2\pi Ft}{2\pi Ft}$.

Теперь рассмотрим скалярное произведение

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

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

$S_g(\nu) = S_\varphi(\nu) S_\varphi(\nu) = \frac{1}{2F} S_\varphi(\nu)$, и в силу линейности преобразования Фурье (свойство 2°) $g(u) = \frac{1}{2F} \varphi(u)$. Поэтому (см. рис. 2.2)

$$\langle \varphi_k(t), \varphi_m(t) \rangle = g((m - k)\Delta t) = \frac{1}{2F} \varphi((m - k)\Delta t) = \begin{cases} 1/(2F), & \text{при } k = m \\ 0, & \text{при } k \neq m \end{cases} \blacktriangleright$$

Таким образом, сдвиги функции $\varphi(t)$ на $k\Delta t$, $k \in \mathbb{Z}$, образуют функциональную систему, которая является ортогональным базисом для представления функций, спектр которых ограничен частотной полосой $\nu \in \left(-\frac{1}{2\Delta t}; \frac{1}{2\Delta t}\right)$. Коэффициенты Фурье ряда Котельникова (2.5), т.е. коэффициенты разложения по базису (2.6), находятся очень просто, равномерной выборкой отсчетов анализируемой функции: $\lambda_k = f(k\Delta t)$, $k \in \mathbb{Z}$. По этому набору отсчетов мы можем восстановить значение функции в любой точке числовой оси $t \in (-\infty; \infty)$ — соответствующее правило восстановления определяет ряд (2.5). Соотношение (2.5) можно рассматривать как интерполяционную формулу, имеющую при правильном выборе периода дискретизации нулевую ошибку интерполяции для функций с ограниченным носителем спектра.

Итак, при шаге дискретизации $\Delta t = 1/(2F)$ можно точно восстановить аналоговый сигнал по его отсчетам, если спектр сигнала ограничен: $\max_{\nu: S(\nu) \neq 0} |\nu| = F_{\max} < F$. Отсюда, окончательно, критерий для выбора шага дискретизации аналогового сигнала при известной максимальной («верхней») частоте F_{\max} спектрального представления принимает следующий вид:

$$\Delta t < \frac{1}{2F_{\max}}. \quad (2.7)$$

Величину $f_{\text{д}} = 1/\Delta t$ называют *частотой дискретизации*. В этих обозначениях условие (2.7) можно переписать в эквивалентном виде:

$$f_{\text{д}} > 2F_{\max}.$$

Неравенство (2.7) требует ряда комментариев. Прежде всего, нужно 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 $\Omega_{\nu} = (-\infty; \infty)$, как это следует из теоремы 1.15.

Чаше всего для определения ширины полосы спектра верхнюю частоту F_{\max} выбирают так, чтобы в энергетическом смысле (см. раздел 1.6) ограничение на ширину полосы не вносило больших искажений: $\int_{-F_{\max}}^{F_{\max}} |S(\nu)|^2 d\nu = \beta \int_{-\infty}^{\infty} |S(\nu)|^2 d\nu$, 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 частотой полного подавления не выше F_{\max} . В противном случае при восстановлении сигнала могут возникнуть недопустимые искажения.

Пример 2.3. Для представления речи достаточно полосы звуковых колебаний с верхней частотой $F_{\max} = 3500$ Гц. В этом случае используется стандартная частота дискретизации $f_{\text{д}} = 1/\Delta t = 8000$ Гц, что удовлетворяет условию (2.7). Однако перед дискретизацией звуковой сигнал обязательно нужно подвергнуть НЧ-фильтрации с частотой полного подавления $f_s \leq 1/(2\Delta t) = f_{\text{д}}/2 = 4000$ Гц. В противном случае может возникнуть эффект наложения (смещения) 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 (2.5)) может появиться тон с частотой $f = 2000$ Гц, которого, возможно, в исходном звуковом сигнале вообще не было. Это непосредственно следует из решения примера 2.2, где нужно положить $\nu_1 = 2000$, $\nu_2 = 6000$. ▶

Таким образом, в подвергаемом дискретизации сигнале верхняя частота должна отвечать следующему из (2.7) условию: $F_{\max} < f_{\text{д}}/2$. Граничное значение $F = f_{\text{д}}/2$ носит название *частоты Найквиста*.

Упражнение. Последовательность $x(n) = \cos(\pi n/4)$ получена в результате дискретизации аналогового сигнала $u(t) = \cos(2\pi f_0 t)$ с частотой дискретизации $f_{\text{д}} = 1000$ Гц. Найдите два минимальных значения частоты f_0 , при которых это возможно.

2.4. Спектр дискретного сигнала

Выбрав определенным образом шаг дискретизации Δt , можно получить из аналогового сигнала дискретный: $f(t) \rightarrow \{f_k = f(k\Delta t)\}$.

Каким образом можно 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 (2.5), $\{f_k\} \rightarrow \tilde{f}(t)$.

Если исходный непрерывный сигнал $f(t)$ имел ограниченную частотную полосу и период дискретизации был выбран в соответствии с условием (2.7), то использование формулы (2.5) дает точное восстановление аналогового сигнала, т.е. $f(t) = \tilde{f}(t)$, поэтому в силу свойств 2° и 4° преобразования Фурье (см. раздел 1.5) спектр $S(v) = \mathcal{F}\{f(t)\} = \mathcal{F}\{\tilde{f}(t)\}$ можно записать в следующем виде:

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

Так как (см. доказательство теоремы 2.2)

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

то с использованием обозначения $\Pi_F(v) = \begin{cases} 1, & v \in [-F; F] \\ 0, & v \notin [-F; F] \end{cases}$ имеем:

$$S(v) = \Pi_F(v) \sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-2\pi i v k \Delta t} \Delta t = \Pi_F(v) \tilde{S}(v), \quad (2.8)$$

$$\tilde{S}(v) = \sum_{k=-\infty}^{\infty} f(k\Delta t) e^{-2\pi i v k \Delta t} \Delta t. \quad (2.9)$$

Выражение (2.9) уже встречалось нам при доказательстве теоремы 2.1 и представляло собой периодическое продолжение ограниченного спектра $S(v)$ на всю частотную ось. Получаем, что для функции со спектром, ограниченным частотной полосой $|v| < F$, выражение (2.9) для $v \in (-F; F)$ совпадает (см. (2.8)) с интегралом Фурье (1.18). Заметим также, что выражение (2.9) представляет собой квадратурную формулу прямоугольников для интеграла Фурье (1.18), в которой узлы взяты в отсчетах дискретного сигнала $t_k = k\Delta t$.

Функцию $\tilde{S}(v)$ (2.9) будем называть *спектром дискретного сигнала*¹ $\{f(k\Delta t)\}$. Действительно, если преобразовать дискретный сигнал $\{f_k = f(k\Delta t)\}$ 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: $\mathcal{F}\{\tilde{f}\} = \tilde{S}(v)$. Отличие (2.8) от (2.9) состоит в том, что спектр дискретного сигнала (2.9) 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 2.1. Представив решетчатую функцию $\tilde{f}(t) = f(t)\Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t)$ на основании соотношения (1.24) в виде $\tilde{f}(t) = f(t) \sum_{k=-\infty}^{\infty} \exp(i2\pi kt/\Delta t)$, ее спектр (1.18), с учетом свойств 2° и 5° преобразования Фурье, можем тогда записать как

$$\tilde{S}(v) = \mathcal{F}\{\tilde{f}(t)\} = \sum_{k=-\infty}^{\infty} \mathcal{F}\left\{f(t)\exp\left(i\frac{2\pi k}{\Delta t}t\right)\right\} = \sum_{k=-\infty}^{\infty} S\left(v - \frac{k}{\Delta t}\right). \quad (2.10)$$

То есть спектр дискретного сигнала представляет собой сумму сдвинутых копий спектра непрерывного сигнала. Как легко видеть из (2.9) и (2.10), $\tilde{S}(v) = \tilde{S}\left(v + \frac{1}{\Delta t}\right)$, $\Delta t = \frac{1}{2F}$, см. рис. 2.3.

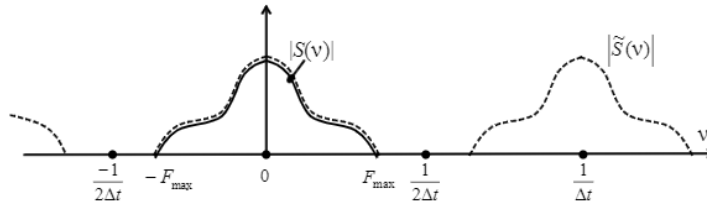


Рис. 2.3. Правильный выбор частоты дискретизации:

$$f_{\text{д}} = \frac{1}{\Delta t} > 2F_{\text{max}}$$

Если частота дискретизации была выбрана с нарушением условия (2.7), то период функции (2.9), (2.10) окажется меньше ширины

¹ Определение (2.9) чаще используется в математической литературе (см. напр., [17, 27]), а в технической литературе обычно используется определение спектра дискретного сигнала, отличающееся нормировкой: $\tilde{S}(v) = \sum_{k=-\infty}^{\infty} f(k\Delta t)e^{-2\pi vk\Delta t}$.

полосы $(-F; F)$ исходного сигнала, и копии спектра непрерывного сигнала перекрываются. При этом $S(v) \neq \Pi_F(v)\tilde{S}(v)$ и возникают необратимые искажения сигнала, связанные с эффектом наложения частот, см. рис. 2.4. Вследствие этого при восстановлении непрерывного сигнала по отсчетам $\{f_k = f(k\Delta t)\}$ при помощи формулы (2.5) произойдет «подмена» гармоник спектра с частотами $v_+ = \frac{f_d + \alpha}{2}$, $\alpha > 0$, на гармоники с частотами $v_- = \frac{f_d - \alpha}{2}$, см. примеры 2.2 и 2.3.

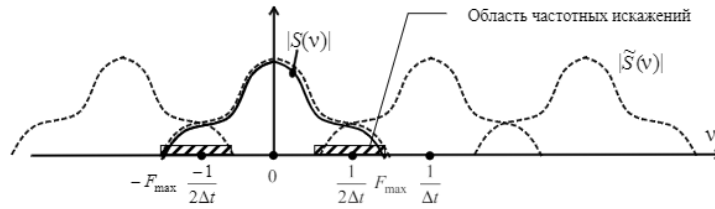


Рис. 2.4. Неправильный выбор частоты дискретизации:

$$f_d = \frac{1}{\Delta t} < 2F_{\max}$$

Упражнение. Убедитесь, что $\forall K \in \mathbb{Z}$ сигналы $u_k(t) = \cos(2\pi(v_0 + Kf_d)t)$ после дискретизации с периодом $\Delta t = 1/f_d$ совпадают; при этом спектр (2.9) полученного дискретного сигнала $x(n) = u_k(n\Delta t)$ имеет вид:

$$\tilde{S}(v) = \frac{1}{2} \sum_{k=-\infty}^{\infty} (\delta(v - v_0 + kf_d) + \delta(v + v_0 + kf_d)).$$

Спектр дискретного сигнала, определяемый по формулам (2.9) и (2.10), соответствует обобщенному спектру решетчатой функции $\hat{f}(t) = f(t)\Delta t \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t)$. На практике при дискретизации сигнала приходится иметь дело со стробирующими (дискретизирующими) импульсами конечной длительности, отличными от дельта-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 τ которого 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 вновь 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:

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

где $\tilde{S}(v)$ — исходный спектр дискретного сигнала (2.9), а $S_{\sigma}(v) = \mathcal{F}\{\sigma(t)\}$ — спектр стробирующего импульса. Видим, что отличие формы дискретизирующего импульса от дельта-функции (когда $S_{\sigma}(v) \neq 1$) влечет мультипликативные искажения в спектре дискретного сигнала. Данными искажениями можно пренебречь, если для основной полосы частот $v \in [0; 1/2\Delta t]$ (которая характеризует непрерывный сигнал, см. также рис. 2.3) имеем: $S_{\sigma}(v) \approx 1$. Это достигается, если длительность стробирующего импульса τ выбрана значительно меньше периода дискретизации Δt .

Пример 2.4. Оценить искажения спектра дискретного сигнала, которые обусловлены прямоугольным стробирующим импульсом:

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

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

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

Для $\tau = \Delta t/3$ получаем $s_{\min} \approx 0,955$; для $\tau = \Delta t/10$: $s_{\min} \approx 0,996$. ▶

Упражнение. Повторите оценки, аналогичные приведенным в примере 2.4, и найдите значения спектральной плотности $s_{\min} = S_{\sigma}\left(\frac{1}{2\Delta t}\right)$ для треугольного стробирующего импульса

$$\sigma(t) = \begin{cases} \frac{2}{\tau} + \frac{4}{\tau^2}t, & -\frac{\tau}{2} \leq t \leq 0 \\ \frac{2}{\tau} - \frac{4}{\tau^2}t, & 0 \leq t \leq \frac{\tau}{2} \\ 0, & t \notin \left[-\frac{\tau}{2}; \frac{\tau}{2}\right] \end{cases}$$

длительности $\tau = \Delta t/3$, $\tau = \Delta t/10$. *Указание:* для нахождения спектра $S_{\sigma}(v)$ используйте свойство 6° преобразования Фурье, учитывая, что треугольный импульс $\sigma(t)$ можно представить как интегральную свертку двух одинаковых прямоугольных импульсов длительности $\tau/2$.

2.5. Дискретизация узкополосных сигналов

Рассмотренная в разделе 2.3 теорема Котельникова, определяющая частоту дискретизации $f_{\text{д}}$ по максимальной частоте F_{max} спектра обрабатываемого сигнала как $f_{\text{д}} > 2F_{\text{max}}$, в ряде случаев налагает невыполнимые на практике условия, предъявляемые к быстродействию аппаратуры дискретизации.

Оказывается, для многих важных и часто встречающихся приложений требование к частоте дискретизации можно ослабить, допустив нарушение условия $f_{\text{д}} > 2F_{\text{max}}$. При этом эффекта наложения частот можно избежать, если подвергаемый дискретизации сигнал является узкополосным.

Определение. Пусть вещественный сигнал $f(t)$ имеет спектр $S(v) = \mathcal{F}\{f(t)\}$, для которого найдутся такие частоты F_{min} и F_{max} , $0 < F_{\text{min}} < F_{\text{max}}$, что можно считать $\forall v \notin (-F_{\text{max}}; -F_{\text{min}}) \cup (F_{\text{min}}; F_{\text{max}}): S(v) = 0$. Такой сигнал с ограниченной частотной полосой¹ $v \in (-F_{\text{max}}; -F_{\text{min}}) \cup (F_{\text{min}}; F_{\text{max}})$ называется *узкополосным*, если ширина частотной полосы $\Delta F = F_{\text{max}} - F_{\text{min}}$ много меньше средней частоты $F_{\text{cp}} = (F_{\text{max}} + F_{\text{min}})/2$, т. е. $\Delta F/F_{\text{cp}} \ll 1$.

Узкополосными обычно являются *радиосигналы* — так называют сигналы, описываемые функцией следующего вида:

$$x(t) = u(t) \cos(2\pi v_0 t + \gamma(t)), \quad (2.11)$$

где v_0 — так называемая *несущая частота* (часто это середина частотной полосы F_{cp} радиосигнала $x(t)$), а функции $u(t)$ и/или $\gamma(t)$

¹ Для реальных сигналов условие ограниченности частотной полосы обычно выполняется приближенно, о способах определения частотной полосы см. раздел 1.6.

содержат передаваемую сигналом информацию. Радиосигнал (2.11) можно представить в следующем виде:

$$x(t) = \operatorname{Re}\{U(t)\exp(i2\pi\nu_0 t)\} = \frac{1}{2}(U(t)e^{i2\pi\nu_0 t} + \overline{U(t)}e^{-i2\pi\nu_0 t}), \quad (2.12)$$

где $U(t) = u(t)e^{i\gamma(t)}$ — так называемая *комплексная огибающая*, или *комплексная амплитуда*, содержащая передаваемую радиосигналом информацию. Функцию $u(t)$ называют *вещественной огибающей* радиосигнала, которую можно считать всегда неотрицательной (смену знака $U(t)$ при необходимости можно осуществить, изменив значение функции $\gamma(t)$ на величину $\pm\pi$). Полагая $u(t) \geq 0$, имеем $u(t) = |U(t)|$.

Определим, как связаны спектры радиосигнала (2.12) $S_x(\nu) = \mathcal{F}\{x(t)\}$ и комплексной огибающей $S_U(\nu) = \mathcal{F}\{U(t)\}$. На основании свойств 2° и 5° преобразования Фурье (см. раздел 1.5), применяя его к выражению (2.12), получаем:

$$S_x(\nu) = (S_U(\nu - \nu_0) + S_{\overline{U}}(\nu + \nu_0))/2, \text{ где } S_U(\nu) = \mathcal{F}\{U(t)\} = \int_{-\infty}^{\infty} U(t)e^{-2\pi i\nu t} dt,$$

$$S_{\overline{U}}(\nu) = \mathcal{F}\{\overline{U(t)}\} = \int_{-\infty}^{\infty} \overline{U(t)}e^{-2\pi i\nu t} dt = \overline{\int_{-\infty}^{\infty} U(t)e^{2\pi i\nu t} dt} = \overline{S_U(-\nu)}.$$

Поэтому

$$S_x(\nu) = \frac{1}{2}(S_U(\nu - \nu_0) + \overline{S_U(-\nu - \nu_0)}). \quad (2.13)$$

Как несложно видеть, $S_x(-\nu) = (S_U(-\nu - \nu_0) + \overline{S_U(\nu - \nu_0)})/2 = \overline{S_x(\nu)}$ — свойство сопряженной симметрии имеет место в силу вещественного характера сигнала (2.11).

Обозначая $S_A(\nu) = \overline{S_U(-\nu - \nu_0)}/2$, $S_B(\nu) = S_U(\nu - \nu_0)/2$, спектр (2.13) радиосигнала (2.11) можно представить в виде $S_x(\nu) = S_A(\nu) + S_B(\nu)$, а спектр комплексной огибающей выразить как

$$S_U(\nu) = 2S_B(\nu + \nu_0) = 2\overline{S_A(-\nu - \nu_0)}. \quad (2.14)$$

Очевидно, что $S_A(-\nu) = \overline{S_B(\nu)}$, $|S_A(-\nu)| = |S_B(\nu)|$. Если ширина спектра $S_U(\nu) = \mathcal{F}\{U(t)\}$ комплексной огибающей $U(t)$ много меньше несущей частоты ν_0 , то сигнал (2.11) является узкополосным и его амплитудный спектр схематически можно изобразить в виде, приведенном на рис. 2.5а.

Рассмотренная нами ранее теорема 2.1 гарантирует отсутствие наложения частот, если частота дискретизации выбрана

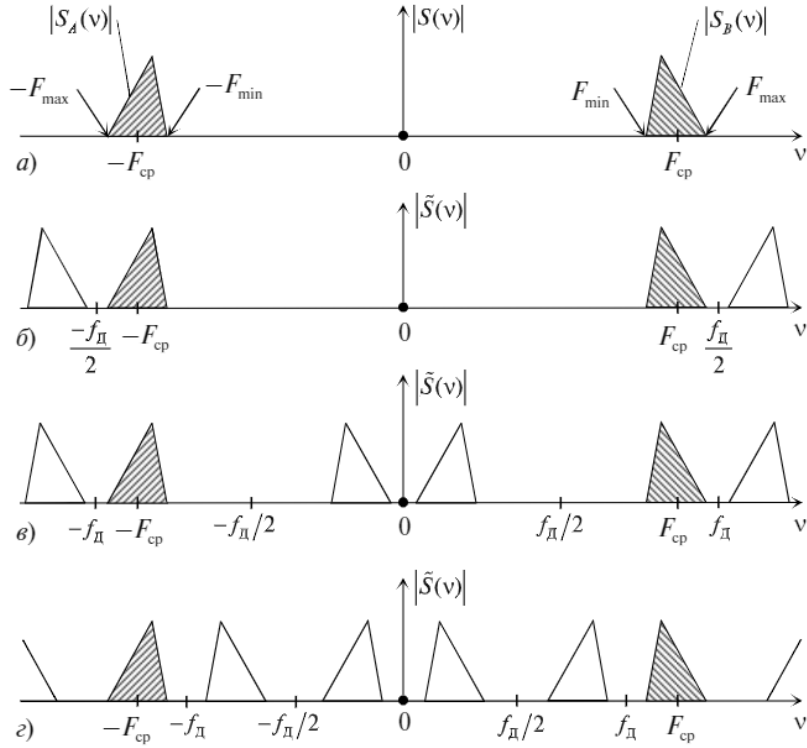


Рис. 2.5. Амплитудные спектры сигналов, полученных в результате дискретизации узкополосного сигнала с различными частотами дискретизации f_d , не приводящими к наложению частот: а) спектр исходного аналогового сигнала, б) $2F_{\max} < f_d$, в) $F_{\max} < f_d < 2F_{\min}$, г) $2F_{\max}/3 < f_d < F_{\min}$

из условия $f_d > 2F_{\max}$. Однако для узкополосных сигналов данное требование может быть нарушено без необратимых последствий. Например, для некоторого узкополосного сигнала $y(t)$, имеющего амплитудный спектр приведенного на рис. 2.5а вида, в результате дискретизации получим сигналы, амплитудные спектры которых для ряда частот дискретизации отображены на рис. 2.5б–г. При этом в двух последних случаях (рис. 2.5в, г) условие $f_d > 2F_{\max}$ не выполняется, однако наложения частот в спектре дискретного сигнала (2.10) $\tilde{S}_y(v) = \sum_{k=-\infty}^{\infty} S_y(v - kf_d)$ не происходит. Это оказалось

возможным потому, что спектр $S_y(v)$ исходного (узкополосного) сигнала $y(t)$ не содержит составляющих в полосе частот $v \in [-F_{\min}; F_{\min}]$.

Сформулируем условия, которые позволяют понизить частоту дискретизации при обработке узкополосных сигналов.

Теорема 2.3. (Обобщенная теорема Котельникова.) Пусть вещественный аналоговый сигнал $y(t)$ имеет спектр $S_y(v) = \mathcal{F}\{y(t)\}$, отличный от нуля только в области частот $v \in (-F_{\max}; -F_{\min}) \cup (F_{\min}; F_{\max})$, $0 < F_{\min} < F_{\max}$. Тогда если для какого-то неотрицательного целого числа q найдется такая частота дискретизации $f_{\text{д}} = 1/\Delta t$, что

$$\frac{2F_{\max}}{q+1} < f_{\text{д}} < \frac{2F_{\min}}{q}, \quad (2.15)$$

то в спектре $\tilde{S}_y(v) = \mathcal{F}\{\tilde{y}(t)\}$ соответствующего дискретного сигнала $\tilde{y}(t) = \Delta t \sum_{k=-\infty}^{\infty} y(k\Delta t)\delta(t - k\Delta t)$ эффект наложения частот отсутствует и, в частности, $\forall v \in (-F_{\max}; -F_{\min}) \cup (F_{\min}; F_{\max})$: $\tilde{S}_y(v) = S_y(v)$.

◀ Спектр $S_y(v)$ произвольного вещественного узкополосного аналогового сигнала $y(t)$ можно представить в виде суммы двух слагаемых (см. также рис. 2.5а): $S_y(v) = S_A(v) + S_B(v)$, где $S_A(v) = 0$ при $v \in (-F_{\max}; -F_{\min})$, а $S_B(v) = 0$ при $v \in (F_{\min}; F_{\max})$.

После дискретизации аналогового сигнала с некоторой частотой $f_{\text{д}}$ спектр дискретного сигнала можно записать с использованием соотношения (2.10):

$$\tilde{S}_y(v) = \sum_{k=-\infty}^{\infty} S_y(v - kf_{\text{д}}) = \sum_{k=-\infty}^{\infty} \underbrace{S_A(v - kf_{\text{д}})}_{\text{обозначим } S_{A_k}(v)} + \sum_{k=-\infty}^{\infty} \underbrace{S_B(v - kf_{\text{д}})}_{\text{обозначим } S_{B_k}(v)}.$$

Функции $\tilde{S}_A(v) = \sum_{k=-\infty}^{\infty} S_{A_k}(v)$ и $\tilde{S}_B(v) = \sum_{k=-\infty}^{\infty} S_{B_k}(v)$ периода $f_{\text{д}}$ представляют собой «гребешки», «зубья» которых — функции $S_{A_k}(v)$ и $S_{B_k}(v)$ — равномерно расположены на оси частот; $\forall k \in \mathbb{Z}$: $S_{A_k}(v) = 0$ при $v \in (-F_{\max} + kf_{\text{д}}; -F_{\min} + kf_{\text{д}})$, $S_{B_k}(v) = 0$ при $v \in (F_{\min} + kf_{\text{д}}; F_{\max} + kf_{\text{д}})$.

Для того чтобы избежать наложения частот в спектре $\tilde{S}_y(v)$, функции $\{S_{A_k}(v)\}_{k \in \mathbb{Z}}$ и $\{S_{B_k}(v)\}_{k \in \mathbb{Z}}$ не должны взаимно перекрываться, см. рис. 2.5б–г. Так как функции $\tilde{S}_A(v)$ и $\tilde{S}_B(v)$ имеют периодическую структуру, то достаточно рассмотреть одну функцию $S_{B_0}(v) = S_B(v)$ из $\tilde{S}_B(v) = \sum_{k=-\infty}^{\infty} S_{B_k}(v)$ и определить условия, при которых окружающие ее соседние сдвиги $S_{A_q}(v)$ и $S_{A_{q+1}}(v)$ с некоторыми номерами $k = q$ и $k = q+1$ из другого «гребешка» $\tilde{S}_A(v) = \sum_{k=-\infty}^{\infty} S_{A_k}(v)$ не перекрывают $S_{B_0}(v) = S_B(v)$.

Поскольку функции $S_{A_k}(v) = S_A(v - kf_{\text{Д}})$, $k \in \mathbb{Z}$, могут принимать ненулевое значение только для частот $v \in (-F_{\text{max}} + kf_{\text{Д}}; -F_{\text{min}} + kf_{\text{Д}})$, а функция $S_{B_0}(v) = S_B(v)$ отлична от нуля только для $v \in (F_{\text{min}}; F_{\text{max}})$, то условие отсутствия наложения частот при дискретизации узкополосного сигнала выполняется (см. также рис. 2.5), если

$$-F_{\text{min}} + qf_{\text{Д}} < F_{\text{min}} \quad \text{и} \quad F_{\text{max}} < -F_{\text{max}} + (q+1)f_{\text{Д}}.$$

Отсюда получаем условие (2.15). ►

Значение числа q в (2.15) определяет количество сдвинутых копий функции $S_A(v)$ (или $S_B(v)$), которые попали в интервал $v \in [-F_{\text{min}}; F_{\text{min}}]$. В приведенных на рис. 2.5 примерах для рис. 2.5б имеем $q = 0$; для рис. 2.5в, г имеем соответственно $q = 1$ и $q = 2$. Заметим, что значение $q = 0$ в (2.15) соответствует ранее установленному условию (2.7) теоремы 2.1 — классической теоремы Котельникова.

Упражнение. Покажите, что возможные значения целочисленного параметра q в (2.15) могут быть выбраны из следующего набора:

$$q = 0, 1, \dots, \left[\frac{F_{\text{max}}}{F_{\text{max}} - F_{\text{min}}} \right] - 1, \quad (2.16)$$

где квадратные скобки обозначают целую часть стоящего в них числа.

Отметим, что для практики более важным результатом теоремы 2.3 является не столько сохранение без наложений компонент оригинального спектра $S_y(v) = \mathcal{F}\{y(t)\}$ в полосе $v \in (-F_{\text{max}}; -F_{\text{min}}) \cup (F_{\text{min}}; F_{\text{max}})$ при дискретизации сигнала в соответствии с (2.15), сколько перенос неискаженных копий спектра $S_y(v)$ в полосу частот $v \in [-F_{\text{min}}; F_{\text{min}}]$ при $q > 0$.

Действительно, симметричные (относительно нулевой частоты) сдвиги составляющих $S_A(v) = \overline{S_U(-v - v_0)}/2$ и $S_B(v) = S_U(v - v_0)/2$ спектров $S_y(v) = S_A(v) + S_B(v)$ для узкополосных радиосигналов вида (2.11)–(2.12), как это видно из формул (2.13) и (2.14), означают лишь изменение несущей частоты v_0 радиосигнала и не изменяют его комплексную огибающую $U(t)$, которая и переносит заключенную в сигнале (2.11) информацию. Поэтому при выполнении условия (2.15) использование пониженной частоты дискретизации при-

водит к дополнительному эффекту уменьшения несущей частоты радиосигнала.

Пример 2.5. Пусть аналоговый радиосигнал $x(t) = u(t)\cos(2\pi\nu_0 t + \gamma(t))$ имеет амплитудный спектр, приведенный на рис. 2.5а. Выберем некоторый параметр α из условия $2F_{\max}/3 < \alpha < F_{\min}$ и рассмотрим второй радиосигнал $y(t) = u(t)\cos(2\pi(\nu_0 - \alpha)t + \gamma(t))$, имеющий более низкую несущую частоту, равную $\nu_0 - \alpha$. Покажем, что после дискретизации с частотой $f_{\Delta} = \alpha$ дискретные сигналы $\tilde{x}(t) = \Delta t \sum_{k=-\infty}^{\infty} x(k\Delta t)\delta(t - k\Delta t)$ и $\tilde{y}(t) = \Delta t \sum_{k=-\infty}^{\infty} y(k\Delta t)\delta(t - k\Delta t)$, где $\Delta t = 1/f_{\Delta}$, совпадают, а их амплитудный спектр имеет вид, изображенный на рис. 2.5г.

◀ Для отсчетов сигналов $x(t)$ и $y(t)$ в точках $k\Delta t = k/f_{\Delta} = k/\alpha$ имеем:

$$x(k\Delta t) = u\left(\frac{k}{\alpha}\right) \cos\left(\frac{2\pi\nu_0 k}{\alpha} + \gamma\left(\frac{k}{\alpha}\right)\right),$$

$$y(k\Delta t) = u\left(\frac{k}{\alpha}\right) \cos\left(\frac{2\pi(\nu_0 - \alpha)k}{\alpha} + \gamma\left(\frac{k}{\alpha}\right)\right) = u\left(\frac{k}{\alpha}\right) \cos\left(\frac{2\pi\nu_0 k}{\alpha} - 2\pi k + \gamma\left(\frac{k}{\alpha}\right)\right),$$

т. е. $x(k\Delta t) = y(k\Delta t)$ и после дискретизации сигналы совпадают. Спектры полученных дискретных сигналов одинаковы: $\tilde{S}_x(\nu) = \tilde{S}_y(\nu)$.

Условие выбора частоты дискретизации для сигнала $x(t)$ из условия $2F_{\max}/3 < f_{\Delta} < F_{\min}$ означает, что в (2.15) было выбрано значение $q = 2$. Это означает, в полосу частот $\nu \in [-F_{\min}; F_{\min}]$ спектра $\tilde{S}_x(\nu)$ дискретного сигнала $\tilde{x}(t)$ попадут две копии спектра $S_x(\nu)$ аналогового сигнала $x(t)$. Поэтому, если амплитудный спектр аналогового сигнала $x(t)$ имел вид, представленный на рис. 2.5а, то дискретный сигнал $\tilde{x}(t)$ будет иметь такой спектр, как на рис. 2.5г. ▶

В результате появления после дискретизации узкополосного аналогового сигнала $x(t)$ сдвигов (см. (2.10)) оригинального спектра $S_x(\nu) = S_A(\nu) + S_B(\nu)$ возможно возникновение ситуации, когда оказавшиеся самыми низкочастотными, ближайшими к нулевой частоте, копии $S_{A_k}(\nu) = S_A(\nu - Kf_{\Delta})$ и $S_{B_k}(\nu) = S_B(\nu + Kf_{\Delta})$, $K \in \mathbb{N}$, «меняются местами». То есть составляющая $S_{A_k}(\nu)$ оказывается локализованной в области положительных, а $S_{B_k}(\nu)$ — отрицательных частот, в полосах $\nu \in [0; f_{\Delta}/2]$ и $\nu \in [-f_{\Delta}/2; 0]$ соответственно. Пример подобной ситуации иллюстрирует рис. 2.5в — такой эффект приводит к искажению комплексной огибающей и носит название *инверсии*

спектра. Для ее устранения спектр полученного дискретного сигнала $\tilde{S}_x(v)$ нужно сместить по оси частот на величину $\pm f_d/2$. Во временной области это соответствует (см. свойство 5° преобразования Фурье в разделе 1.5) умножению сигнала $\tilde{x}(t)$ на комплексную функцию $\exp(\mp 2\pi i(f_d/2)t) = \exp(\mp i\pi t/\Delta t)$:

$$\tilde{S}_y(v) = \tilde{S}_x\left(v \pm \frac{f_d}{2}\right) \leftrightarrow \tilde{y}(t) = e^{\mp i\pi t/\Delta t} \tilde{x}(t).$$

Поэтому для устранения инверсии спектра дискретный сигнал $\tilde{x}(t) = \Delta t \sum_{k=-\infty}^{\infty} x(k\Delta t)\delta(t - k\Delta t)$ нужно преобразовать (убедитесь!) в сигнал $\tilde{y}(t) = \Delta t \sum_{k=-\infty}^{\infty} y(k\Delta t)\delta(t - k\Delta t) = \Delta t \sum_{k=-\infty}^{\infty} (-1)^k x(k\Delta t)\delta(t - k\Delta t)$.

Упражнение. Определите общее правило: при каких допустимых значениях q из набора (2.16) возникает эффект инверсии спектра?

Пример 2.6. Пусть спектр радиосигнала имеет полосу $[F_{\min}; F_{\max}]$, где $F_{\min} = 63$ МГц, $F_{\max} = 77$ МГц. Найти допустимые диапазоны для частоты дискретизации данного сигнала.

◀ Ширина полосы для данного сигнала $\Delta F = F_{\max} - F_{\min} = 14$ МГц, центральная частота $F_{\text{cp}} = (F_{\max} + F_{\min})/2 = 70$ МГц. Сигнал является узкополосным, так как $\Delta F/F_{\text{cp}} = 0,2 \ll 1$. Найдем максимальное для (2.15) значение q из правила (2.16): $\left\lceil \frac{F_{\max}}{F_{\max} - F_{\min}} \right\rceil - 1 = \left\lceil \frac{77}{14} \right\rceil - 1 = 4$.

Для условия (2.15) получаем пять возможных диапазонов частоты дискретизации (в МГц):

$$154 < f_d < \infty \quad (q = 0); \quad 77 < f_d < 126 \quad (q = 1);$$

$$51,33(3) < f_d < 63 \quad (q = 2); \quad 38,5 < f_d < 42 \quad (q = 3); \quad 30,8 < f_d < 31,5 \quad (q = 4).$$

Чаще всего выбирается наименьшая частота дискретизации из возможного набора значений (например, в нашем случае это может быть 31 МГц), поскольку меньшая частота дискретизации, как правило, влечет менее жесткие требования к аппаратуре. Однако при меньшей частоте дискретизации, когда копии оригинального спектра более плотно заполняют полосу $v \in [-F_{\min}; F_{\min}]$, могут оказаться более высокими требования к аналоговым фильтрам, используемым перед дискретизацией узкополосного сигнала для предотвращения наложения частот. Кроме того, частоту дискретизации

иногда требуется выбрать только из некоторого возможного набора значений. На выбор частоты дискретизации могут оказывать влияние и другие факторы, определяемые техническими требованиями конкретных приложений. ►

2.6. Дискретное преобразование Фурье

Как было установлено в разделе 2.4, дискретный сигнал $\{f_k = f(k\Delta t)\}_{k \in \mathbb{Z}}$, представленный в виде решетчатой функции $\tilde{f}(t) = \Delta t \sum_{k=-\infty}^{\infty} f_k \delta(t - k\Delta t)$, обладает периодическим спектром $\tilde{S}(v) = \mathcal{F}\{\tilde{f}(t)\} = \tilde{S}(v + \frac{1}{\Delta t})$, который определяется выражением (2.9). Если, кроме того, во временной области дискретный сигнал является периодическим, т.е. $f_k = f_{k+N}$ и $\tilde{f}(t) = \tilde{f}(t + T)$, где $T = N\Delta t$, то его спектр (2.9) представляет собой решетчатую функцию:

$$\tilde{S}(v) = \Delta v \sum_{k=-\infty}^{\infty} \tilde{s}_k \delta(v - k\Delta v), \quad (2.17)$$

где $\Delta v = 1/T = 1/(N\Delta t)$. При этом $\tilde{s}_k = \tilde{s}_{k+N}$ и

$$\tilde{s}_k = \Delta t \sum_{m=0}^{N-1} f_m e^{-i\frac{2\pi}{N}km}. \quad (2.18)$$

Упражнение. Убедитесь в справедливости соотношений (2.17) и (2.18), воспользовавшись результатом решения примера 1.17.

Таким образом, дискретный периодический сигнал $\{f_k = f(k\Delta t)\}_{k \in \mathbb{Z}}$, $f_k = f_{k+N}$, имеет периодический дискретный спектр (2.17), компоненты которого $\{\tilde{s}_k = \tilde{s}_{k+N}\}_{k \in \mathbb{Z}}$ могут быть найдены по формуле (2.18). Рассматривая один период последовательности $\{\tilde{s}_k\}$ (2.18) для $k = 0, 1, \dots, N-1$, назовем вектор спектральных компонент $(\tilde{s}_0, \dots, \tilde{s}_{N-1})$ *дискретным преобразованием Фурье (ДПФ)* вектора (f_0, \dots, f_{N-1}) , который представляет собой один период дискретного сигнала во временной области.

Отметим, что верхним (максимальным) частотам дискретного сигнала соответствуют компоненты дискретного спектра, находящиеся в середине вектора $(\tilde{s}_0, \dots, \tilde{s}_{N-1})$, номера индексов которых близки к $N/2$. В частности, при четном значении N компонен-

та ДПФ $\tilde{s}_{N/2}$ соответствует частоте Найквиста (половине частоты дискретизации).

Пример 2.7. Пусть дискретный сигнал (не обязательно периодический) представлен в виде решетчатой функции $\tilde{x}(t) = \Delta t \sum_{k \in \mathbb{Z}} x_k \delta(t - k\Delta t)$, а $\tilde{S}_x(v) = \mathcal{F}\{\tilde{x}(t)\}$ – соответствующий спектр (2.9). Задавшись некоторым числом N частотных отсчетов, равномерно располагаемых на периоде спектральной плотности $\tilde{S}_x(v) = \tilde{S}_x(v + \frac{1}{\Delta t})$, выполним дискретизацию:

$$\tilde{S}_x(v) \rightarrow \tilde{S}_f(v) = \tilde{S}_x(v) \left(\Delta v \sum_{k \in \mathbb{Z}} \delta(v - k\Delta v) \right) = \Delta v \sum_{k \in \mathbb{Z}} \tilde{S}_x(k\Delta v) \delta(v - k\Delta v),$$

где шаг дискретизации в частотной области $\Delta v = (1/\Delta t)/N = 1/(N\Delta t)$. Какие искажения во временной области вызовет такая дискретизация спектра и как связан сигнал $\tilde{f}(t) = \mathcal{F}^{-1}\{\tilde{S}_f(v)\}$ с исходным $\tilde{x}(t)$?

◀ В соответствии с (1.24) дискретизацию спектра сигнала в частотной области представим как

$$\tilde{S}_f(v) = \tilde{S}_x(v) \left(\Delta v \sum_{m \in \mathbb{Z}} \delta(v - m\Delta v) \right) = \sum_{m \in \mathbb{Z}} \tilde{S}_x(v) \exp\left(i \frac{2\pi m}{\Delta v} v \right).$$

Воспользовавшись далее свойствами 2° и 4° преобразования Фурье (см. раздел 1.5), получаем

$$\begin{aligned} \tilde{f}(t) &= \mathcal{F}^{-1}\{\tilde{S}_f(v)\} = \sum_{m \in \mathbb{Z}} \tilde{x}\left(t + \frac{m}{\Delta v}\right) = \sum_{m \in \mathbb{Z}} \tilde{x}(t + mN\Delta t) = \\ &= \Delta t \sum_{m \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} x_k \delta(t + mN\Delta t - k\Delta t) = \Delta t \sum_{m \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} x_{k+mN} \delta(t - k\Delta t) = \\ &= \Delta t \sum_{k \in \mathbb{Z}} \delta(t - k\Delta t) \sum_{m \in \mathbb{Z}} x_{k+mN} = \Delta t \sum_{k \in \mathbb{Z}} f_k \delta(t - k\Delta t), \end{aligned}$$

где $f_k = \sum_{m \in \mathbb{Z}} x_{k+mN} = f_{k+N}$. ▶

Как следует из решения примера 2.7, преобразование непрерывного спектра в решетчатую функцию $\tilde{S}_x(v) \rightarrow \tilde{S}_f(v)$ с шагом дискретизации $\Delta v = 1/(N\Delta t)$ означает во временной области размножение с наложением сдвинутых копий сигнала $\tilde{x}(t)$, что приводит к периодическому сигналу: $\tilde{x}(t) \rightarrow \tilde{f}(t) = \sum_{m \in \mathbb{Z}} \tilde{x}(t + mT) = \tilde{f}(t + T)$, где период $T = 1/\Delta v = N\Delta t$.

Упражнение. Аналоговый сигнал $x(t)$ наблюдается на конечном интервале $t \in [0; T)$, и после дискретизации с шагом $\Delta t = T/N$ получен конечный набор отсчетов $\{x_k = x(k\Delta t)\}_{k=0}^{N-1}$; считаем $\forall k \neq 0, 1, \dots, N-1$:

$x_k = 0$. Представив дискретный сигнал в виде решетчатой функции $\tilde{x}(t) = \Delta t \sum_{k=0}^{N-1} x_k \delta(t - k\Delta t)$, обозначим его спектр $\tilde{S}_x(v) = \mathcal{F}\{\tilde{x}(t)\}$. Убедитесь, что полученному в результате дискретизации с шагом $\Delta v = 1/(N\Delta t)$ спектру

$$\tilde{S}_f(v) = \tilde{S}_x(v) \left(\Delta v \sum_{k=-\infty}^{\infty} \delta(v - k\Delta v) \right) = \Delta v \sum_{k=-\infty}^{\infty} \tilde{S}_x(k\Delta v) \delta(v - k\Delta v)$$

соответствует дискретный сигнал $\tilde{f}(t) = \Delta t \sum_{k \in \mathbb{Z}} f_k \delta(t - k\Delta t)$, который получается в результате периодического продолжения конечной последовательности $\{x_k\}_{k=0}^{N-1}$, а именно: $f_{m+jN} = x_m$, $m = 0, 1, \dots, N-1$; $j \in \mathbb{Z}$.

Заметим, что в упражнении выше величина $T = N\Delta t$ представляет собой длину интервала наблюдения сигнала $x(t)$ во временной области. Тогда формулу для ДПФ (2.18) можно переписать в следующем виде:

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

Видим, что с точностью до множителя $1/T$ ДПФ (2.18) представляет собой квадратурную формулу прямоугольников для интеграла (1.16) — коэффициента c_j ряда Фурье (1.15) функции периода T :

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

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 2.4. (Обратное ДПФ.) Вектор (f_0, \dots, f_{N-1}) можно восстановить по дискретному спектру (2.18) при помощи обратного ДПФ (ОДПФ), которое определяется формулой:

$$f_k = \Delta v \sum_{j=0}^{N-1} \tilde{s}_j e^{\frac{2\pi i}{N} kj}, \quad (2.19)$$

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

где $\Delta v = 1/(N\Delta t)$.

◀ Рассмотрим правую часть равенства (2.19): $a(k) = \Delta v \sum_{j=0}^{N-1} \tilde{s}_j e^{\frac{2\pi i}{N} kj}$. Имеем, см. (2.18):

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

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 $a(k) = \frac{1}{N} \sum_{m=0}^{N-1} f_m v_{k,m} = f_k$. ▶

Итак, (2.18) и (2.19) — это 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 (временные и частотные отсчеты), то величины Δt в (2.18) и Δv в (2.19) есть просто некоторые масштабные множители, такие, что $\Delta t \Delta v = 1/N$ (заметим, данное равенство также несет в себе смысл время-частотной неопределенности дискретного представления сигнала). Для симметрии положим $\Delta t = \Delta v = 1/\sqrt{N}$, и парой ДПФ—ОДПФ далее будем называть преобразования векторов $(x_0, \dots, x_{N-1}) \leftrightarrow (y_0, \dots, y_{N-1})$, где ДПФ и ОДПФ определяются соответственно:

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N}kj}, \quad (2.20)$$

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

$$x_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} y_k e^{\frac{2\pi i}{N}kj}, \quad (2.21)$$

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

Как уже отмечалось, в радиотехнике и электронике частотное описание аналоговых сигналов во многих случаях более предпочтительно. Мы показали, что спектр дискретного сигнала определяется выражением (2.9), причем для конечных интервалов наблюдения сигнала последний можно выразить через ОДПФ (2.21) частотных

отсчетов спектральной функции (2.9), которые находятся с помощью ДПФ (2.20). Таким образом, определяемый через ДПФ (2.20) дискретный спектр является аналогом интегрального преобразования Фурье (1.18). Точно так же термин *амплитудный спектр* употребляется для модулей $\{|y_k|\}_{k=0}^{N-1}$ спектральных коэффициентов (2.20), а *фазовый спектр* — для аргументов $\{\arg y_k\}_{k=0}^{N-1}$.

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 (2.20), $\mathbf{X} \leftrightarrow \mathbf{Y}$. Отметим ряд важных свойств ДПФ.

1°. Сопряженная симметрия. Для любого вещественного вектора $\mathbf{X} = \bar{\mathbf{X}}$ в векторе-спектре $y_0 = \bar{y}_0$ (компонента 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

$$\mathbf{Y} = (y_0, y_1, \dots, y_{N-1}) \rightarrow \mathbf{Y}' = (y'_0, y'_1, \dots, y'_{N-1}), \text{ что } \forall k: y'_k = e^{-i\frac{2\pi k}{N}} y_k.$$

$$\begin{aligned} \leftarrow y'_k &= \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x'_j e^{-\frac{2\pi i}{N}kj} = \frac{1}{\sqrt{N}} \left(x_{N-1} + \sum_{j=0}^{N-2} x_j e^{-\frac{2\pi i}{N}k(j+1)} \right) = \\ &= \frac{e^{-\frac{2\pi i}{N}k}}{\sqrt{N}} \left(x_{N-1} e^{-\frac{2\pi i}{N}(N-1)k} + \sum_{j=0}^{N-2} x_j e^{-\frac{2\pi i}{N}kj} \right) = \frac{e^{-\frac{2\pi i}{N}k}}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N}kj} = e^{-\frac{2\pi i}{N}k} y_k. \rightarrow \end{aligned}$$

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

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

$$Y = (y_0, y_1, \dots, y_{N-1}) \rightarrow Y' = (y'_0, y'_1, \dots, y'_{N-1}), \text{ что } \forall k: y'_k = e^{i\frac{2\pi k}{N}m} y_k.$$

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

$$c_k = \sum_{j=0}^{N-1} a_j b_{k-j} = \sum_{j=0}^{N-1} a_{k-j} b_j \quad (2.22)$$

($k = 0, 1, \dots, 2N-1$), 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 $\hat{A} = (a_0, \dots, a_{N-1}, \underbrace{0, \dots, 0}_N)$, $\hat{B} = (b_0, \dots, b_{N-1}, \underbrace{0, \dots, 0}_N)$,

$$\hat{A} = (\hat{a}_0, \hat{a}_1, \dots, \hat{a}_{2N-1}) \text{ — ДПФ вектора } \hat{A}, \hat{a}_k = \frac{1}{\sqrt{2N}} \sum_{j=0}^{2N-1} \hat{a}_j e^{-\frac{2\pi i}{2N}kj},$$

$$\hat{B} = (\hat{b}_0, \hat{b}_1, \dots, \hat{b}_{2N-1}) \text{ — ДПФ вектора } \hat{B}, \hat{b}_k = \frac{1}{\sqrt{2N}} \sum_{j=0}^{2N-1} \hat{b}_j e^{-\frac{2\pi i}{2N}kj}.$$

Тогда для вектора C — свертки (2.22) векторов A и B — ДПФ $\hat{C} = (\hat{c}_0, \hat{c}_1, \dots, \hat{c}_{2N-1})$ можно найти перемножением компонент векторов \hat{A} и \hat{B} , $\forall k = 0, 1, \dots, 2N-1: \hat{c}_k = \sqrt{2N} \hat{a}_k \hat{b}_k$.

◀ 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

$$\sum_{l=0}^{N-1} b_l \omega^{kl} = \sum_{l=0}^{2N-1} \hat{b}_l \omega^{kl} = \sqrt{2N} \hat{b}_k \text{ и, аналогично, } \sum_{m=0}^{N-1} a_m \omega^{km} = \sqrt{2N} \hat{a}_k. \blacktriangleright$$

6°. Совмещенное ДПФ вещественных векторов.

Пусть вектор $\mathbf{B} = (b_0, \dots, b_{N-1})$ представляет собой ДПФ (2.20) вектора $\mathbf{A} = (a_0, \dots, a_{N-1})$, а вектор $\mathbf{V} = (v_0, \dots, v_{N-1})$ — ДПФ (2.20) вектора $\mathbf{U} = (u_0, \dots, u_{N-1})$, причем компоненты векторов \mathbf{A} и \mathbf{U} — вещественные. Тогда, вычислив ДПФ (2.20) вектора $\mathbf{X} = \mathbf{A} + i\mathbf{U}$, получим вектор-спектр \mathbf{Y} , из которого можно восстановить векторы \mathbf{B} и \mathbf{V} с использованием следующих соотношений:

$$\begin{aligned} b_0 &= \operatorname{Re} y_0, \quad v_0 = \operatorname{Im} y_0; \quad \forall k = 1, \dots, N-1: \\ b_k &= \frac{\operatorname{Re} y_k + \operatorname{Re} y_{N-k}}{2} + i \frac{\operatorname{Im} y_k - \operatorname{Im} y_{N-k}}{2}, \\ v_k &= \frac{\operatorname{Im} y_k + \operatorname{Im} y_{N-k}}{2} + i \frac{\operatorname{Re} y_{N-k} - \operatorname{Re} y_k}{2}. \end{aligned}$$

◀ 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

$$\begin{aligned} b_0 &= \operatorname{Re} y_0, \quad v_0 = \operatorname{Im} y_0; \quad \forall k = 1, \dots, N-1: \\ y_k &= \underbrace{(\operatorname{Re} b_k - \operatorname{Im} v_k)}_{\operatorname{Re} y_k} + i \underbrace{(\operatorname{Re} v_k + \operatorname{Im} b_k)}_{\operatorname{Im} y_k}, \\ y_{N-k} &= \underbrace{(\operatorname{Re} b_k + \operatorname{Im} v_k)}_{\operatorname{Re} y_{N-k}} + i \underbrace{(\operatorname{Re} v_k - \operatorname{Im} b_k)}_{\operatorname{Im} y_{N-k}}, \end{aligned}$$

откуда $\forall k = 1, \dots, N-1$:

$$\begin{aligned} \operatorname{Re} b_k &= \frac{1}{2}(\operatorname{Re} y_k + \operatorname{Re} y_{N-k}), \quad \operatorname{Im} b_k = \frac{1}{2}(\operatorname{Im} y_k - \operatorname{Im} y_{N-k}), \\ \operatorname{Re} v_k &= \frac{1}{2}(\operatorname{Im} y_k + \operatorname{Im} y_{N-k}), \quad \operatorname{Im} v_k = \frac{1}{2}(\operatorname{Re} y_{N-k} - \operatorname{Re} y_k). \quad \blacktriangleright \end{aligned}$$

Упражнение. Найдите ДПФ (2.20) каждой из приведенных ниже конечных последовательностей длины N , считая, что число N — четное.

$$\begin{aligned} \text{а) } x(n) &= \tilde{\delta}(n) = \begin{cases} 1, & n=0 \\ 0, & n \neq 0 \end{cases}; & \text{б) } x(n) &= \tilde{\delta}(n-n_0), \quad 0 \leq n_0 \leq N-1; \\ \text{в) } x(n) &= \begin{cases} 1, & n - \text{четное} \\ 0, & n - \text{нечетное} \end{cases}; & \text{г) } x(n) &= \begin{cases} 1, & 0 \leq n \leq N/2-1 \\ 0, & N/2 \leq n \leq N-1 \end{cases}. \end{aligned}$$

ДПФ (2.20) и ОДПФ (2.21) 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: $\mathbf{Y} = \mathbf{W}\mathbf{X}$, $\mathbf{X} = \overline{\mathbf{W}}\mathbf{Y}$, где матрица ДПФ

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

а матрица ОДПФ (убедитесь!):

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

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

Очевидно, что 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 2.1. Матрица \mathbf{W} унитарна тогда и только тогда, когда для любого вектора \mathbf{A} его евклидова норма равна норме вектора $\mathbf{B} = \mathbf{W}\mathbf{A}$: $\|\mathbf{A}\|_E = \|\mathbf{B}\|_E$.

◀ Действительно, если матрица \mathbf{W} унитарна, то для любого вектора \mathbf{A} : $\|\mathbf{B}\|_E^2 = \sum_{k=0}^{N-1} |b_k|^2 = \sum_{k=0}^{N-1} \overline{b_k} b_k = \overline{\mathbf{B}}^T \mathbf{B} = (\overline{\mathbf{W}\mathbf{A}})^T (\mathbf{W}\mathbf{A}) = \overline{\mathbf{A}}^T \underbrace{\overline{\mathbf{W}}^T \mathbf{W}}_{\mathbf{E}} \mathbf{A} = \overline{\mathbf{A}}^T \mathbf{A} = \|\mathbf{A}\|_E^2$,

и наоборот, выполнение данного равенства для произвольного вектора \mathbf{A} означает, что $\overline{\mathbf{W}}^T \mathbf{W} = \mathbf{E}$, 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.

2.7. Быстрое преобразование Фурье. Алгоритм БПФ с прореживанием по времени

Обозначим $x(k) = x_k$, $y(k) = y_k$ и запишем ДПФ (2.20) в виде:

$$y(k) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x(j) e^{-\frac{2\pi i}{N} kj}, \quad k = 0, \dots, N-1. \quad (2.23)$$

Хотя преобразование (2.23) 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, $\tilde{y}(k) = \sqrt{N} y(k)$:

$$\tilde{y}(k) = \sum_{j=0}^{2^n-1} x(j) \omega_n^{kj}, \quad k = 0, 1, \dots, 2^n-1. \quad (2.24)$$

Вычисление каждого коэффициента $\tilde{y}(k)$, как легко видеть, требует около 2^n комплексных сложений с умножениями. Поэтому для реализации всего преобразования (2.24) необходимо около $2^n \times 2^n = 2^{2n}$ комплексных сложений с умножениями.

Пусть $\mathbf{X}_0 = (x(0), x(2), \dots, x(2^n-2)) = (x_0(0), x_0(1), \dots, x_0(2^{n-1}-1))$ и $\mathbf{X}_1 = (x(1), x(3), \dots, x(2^n-1)) = (x_1(0), x_1(1), \dots, x_1(2^{n-1}-1))$ — соответственно векторы четных и нечетных компонент исходного вектора \mathbf{X} . Тогда, замечая, что $\omega_n^2 = \omega_{n-1}$, из (2.24) получаем:

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

где $\tilde{y}_0(k)$ — k -й элемент ДПФ (2.24) (размерности $N = 2^{n-1}$) вектора \mathbf{X}_0 , а $\tilde{y}_1(k)$ — k -й элемент для ДПФ (2.24) 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:

$$\begin{cases} \tilde{y}(k) = \tilde{y}_0(k) + \omega_n^k \tilde{y}_1(k) \\ \tilde{y}(k+2^{n-1}) = \tilde{y}_0(k) - \omega_n^k \tilde{y}_1(k) \\ k = 0, 1, \dots, 2^{n-1}-1 \end{cases} \quad (2.25)$$

При помощи (2.25) мы выразили коэффициенты $\tilde{Y} = (\tilde{y}(0), \tilde{y}(1), \dots, \tilde{y}(2^n - 1))$ ДПФ (2.24) размерности 2^n через коэффициенты ДПФ размерности 2^{n-1} , $\tilde{Y}_0 = (\tilde{y}_0(0), \tilde{y}_0(1), \dots, \tilde{y}_0(2^{n-1} - 1))$ и $\tilde{Y}_1 = (\tilde{y}_1(0), \tilde{y}_1(1), \dots, \tilde{y}_1(2^{n-1} - 1))$, которые получены из векторов X_0 и 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{Y}_0 и \tilde{Y}_1 размерности 2^{n-1} согласно (2.24) 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 (2.25). Итого, вычисление ДПФ размерности 2^n с использованием (2.25) 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 (2.24). Граф вычислений 8-точечного ДПФ по описанной схеме приведен на рис. 2.6.

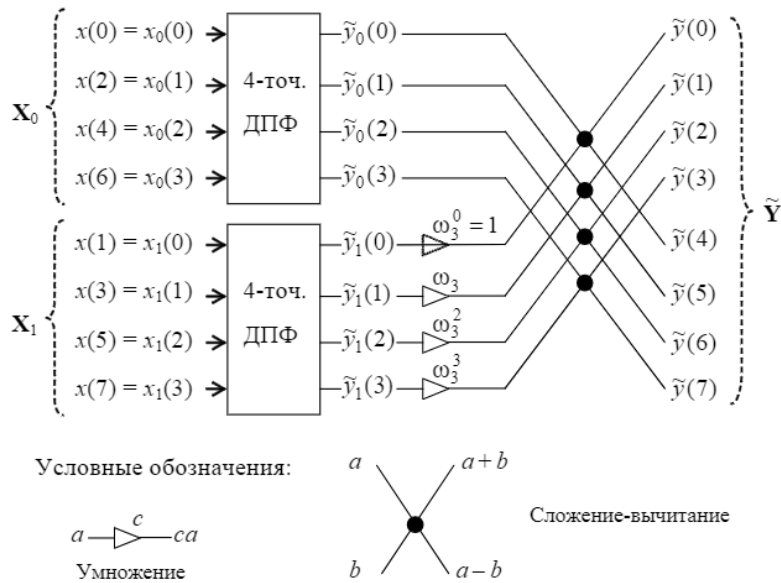


Рис. 2.6. Граф вычислений ДПФ, определяемый формулой (2.25), на примере преобразования размерности $N = 8$, $\omega_3 = \exp(-i\pi/4)$

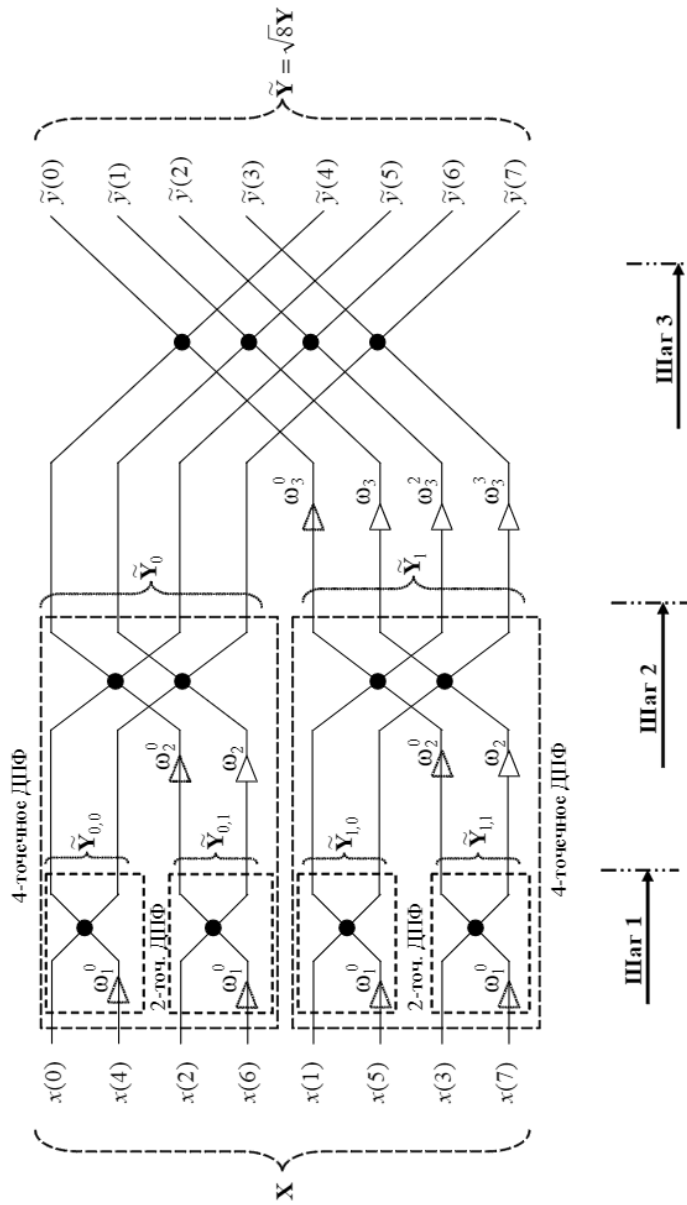
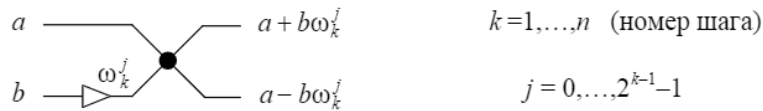


Рис. 2.7. Граф быстрого вычисления 8-точечного ДПФ (обозначения см. рис. 2.6)

Рис. 2.8. Элементарные операции алгоритма БПФ на k -ом шаге

Формулу (2.25) можно рекуррентно применять и далее для вычисления ДПФ меньших размерностей. Например, дальнейшее использование формулы (2.25) для ДПФ размерности $N = 8$ дает полный граф вычислений, который изображен на рис. 2.7.

В общем случае при размерности преобразования $N = 2^n$ процедура сведения ДПФ к двум ДПФ меньшей размерности рекуррентно выполняется n раз. В соответствующем алгоритме *быстрого преобразования Фурье* (БПФ) имеем n шагов (см. также рис. 2.7). На k -м шаге 2^{n-k} раз выполняются вычисления, аналогичные (2.25), результатом которых являются 2^{n-k} «частичных» ДПФ размерности 2^k , выполненных над некоторыми подмножествами отсчетов исходного вектора X . Алгоритм БПФ, основанный на рекуррентном использовании формулы (2.25), называется алгоритмом БПФ с *прореживанием по времени*, так как (см. также рис. 2.6) вычисление ДПФ размерности N сводится к обработке векторов X_0, X_1 , полученных «прореживанием» (выбором каждого второго отсчета, только четных или только нечетных номеров) вектора X .

Как видно из формулы (2.25) и примеров графов вычислений, приведенных на рис. 2.6 и 2.7, базовой операцией БПФ на k -м шаге является так называемая «бабочка», см. рис. 2.8. Каждый шаг требует (см. (2.25)) около $N/2 = 2^{n-1}$ комплексных умножений, итого в алгоритме БПФ имеем $\approx n2^{n-1}$ комплексных умножений вместо $\approx 2^{2n}$ при непосредственной реализации ДПФ по формуле (2.24). С точки зрения реализации на вычислительных устройствах комплексное умножение является 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 выигрыш примерно в 200 раз. Оценку количества вычислительных операций, необходимых для реализации БПФ, можно дополнительно уточнить.

Пример 2.8. Определить количество операций комплексного умножения M_C и комплексного сложения A_C , необходимых для реализации БПФ размерности $N = 2^n$.

◀ Так как k -й шаг ($k = 1, \dots, n$) рассматриваемого алгоритма БПФ состоит в 2^{n-k} -кратном выполнении комплексных умножений на числа $\{\omega_k^j = \exp(-i2\pi j/2^k)\}_{j=0}^{2^{k-1}-1}$ (см. также рис. 2.7, 2.8), где при $j=0$ имеем умножение на единицу, то на k -м шаге алгоритма необходимо выполнить $M_k = 2^{n-k}(2^{k-1} - 1) = 2^{n-1} - 2^{n-k}$ комплексных умножений. Тогда общее количество необходимых в БПФ комплексных умножений получаем равным

$$M_C = \sum_{k=1}^n M_k = n2^{n-1} - \sum_{k=1}^n 2^{n-k} = n2^{n-1} - 2^n + 1 = (n-2)2^{n-1} + 1.$$

Считая операции сложения и вычитания однотипными и замечая, что на каждом шаге $k = 1, \dots, n$ требуется выполнить 2^{n-1} сложений и столько же вычитаний (по числу «бабочек», проходимых на каждом шаге), получаем, что общее количество операций типа сложение/вычитание в алгоритме БПФ

$$A_C = n2^n.$$

Если в формуле (2.24) исключить умножения на $\omega_n^0 = 1$, то для непосредственного выполнения ДПФ по (2.24) требуется $(N-1)^2 = (2^n - 1)^2$ комплексных умножений и $N(N-1) = 2^n(2^n - 1)$ комплексных сложений или вычитаний. ▶

Таким образом, реализация БПФ размерности N требует порядка $O(N \log N)$ арифметических операций, в то время как непосредственная реализация ДПФ по формуле (2.23) или (2.24) — порядка $O(N^2)$ операций.

Упражнение. Определите, какое количество операций вещественного умножения M_R и вещественного сложения (вычитания) A_R необходимо для реализации БПФ.

Упражнение. Используя соотношение (2.25) и рис. 2.7, постройте граф вычислений БПФ размерности $N = 16$.

Отметим, что при обработке вещественных сигналов получаемые дискретные спектры обладают двойной избыточностью

данных (см. свойство 1° ДПФ), что можно использовать для дополнительного ускорения вычислений за счет использования схемы совмещенного ДПФ (см. свойство 6°).

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

$$Y = \frac{1}{\sqrt{N}} W X, \quad X = \frac{1}{\sqrt{N}} \bar{W} Y, \quad (2.26)$$

где

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

Тогда алгоритм БПФ можно трактовать как представление матрицы ДПФ из (2.26) в виде произведения *слабозаполненных* (т. е. состоящих в основном из нулевых элементов) матриц: $W = W_n W_{n-1} \dots W_1$, где каждая матрица W_j соответствует j -му шагу алгоритма БПФ

$$\sqrt{N} Y = \tilde{Y} = \underbrace{W_n W_{n-1} \dots W_1}_{\substack{\text{шаг } n-1 \\ \dots \\ \text{шаг } n}} X$$

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

Пример 2.9. 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, $\sqrt{8} Y = \tilde{Y} = W_3 W_2 W_1 X$ и, соответственно, три шага алгоритма БПФ (см. рис. 2.7). Структура матрицы W_3 соответствует последнему шагу алгоритма БПФ; в использованных ранее обозначениях: $\tilde{Y} = W_3 \begin{pmatrix} \tilde{Y}_0 \\ \tilde{Y}_1 \end{pmatrix}$, где $\tilde{Y} = \sqrt{8} Y$, а векторы \tilde{Y}_0 и \tilde{Y}_1 есть соответственно результаты 4-точечных ненормированных ДПФ (2.24), выполненных над четными X_0 и нечетными X_1 компонентами вектора X . Несложно убедиться, что граф рис. 2.6 (см. также (2.25)) соответствует следующей матричной операции умножения:

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

где $\omega_3 = \exp(-i\pi/4)$. В свою очередь (см. рис. 2.7 и соотношения (2.25))

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

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

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

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

Аналогично, по графу рис. 2.7 получаем структуру матрицы W_1 :

$$\begin{pmatrix} \tilde{Y}_{0,0} \\ \tilde{Y}_{0,1} \\ \tilde{Y}_{1,0} \\ \tilde{Y}_{1,1} \end{pmatrix} = \begin{pmatrix} \tilde{y}_{0,0}(0) \\ \tilde{y}_{0,0}(1) \\ \tilde{y}_{0,1}(0) \\ \tilde{y}_{0,1}(1) \\ \tilde{y}_{1,0}(0) \\ \tilde{y}_{1,0}(1) \\ \tilde{y}_{1,1}(0) \\ \tilde{y}_{1,1}(1) \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}}_{W_1} \underbrace{\begin{pmatrix} \tilde{y}_{0,0,0} = x(0) \\ \tilde{y}_{0,0,1} = x(4) \\ \tilde{y}_{0,1,0} = x(2) \\ \tilde{y}_{0,1,1} = x(6) \\ \tilde{y}_{1,0,0} = x(1) \\ \tilde{y}_{1,0,1} = x(5) \\ \tilde{y}_{1,1,0} = x(3) \\ \tilde{y}_{1,1,1} = x(7) \end{pmatrix}}_{\tilde{X}},$$

где вектор-столбец

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

получен из вектора

$$X = (x(0), x(1), x(2), x(3), x(4), x(5), x(6), x(7))^T$$

перестановкой пар компонент, индексы которых образуют *двоично-инверсную пару* (k, k^*) : $X \xrightarrow{x(k) \leftrightarrow x(k^*)} \tilde{X}$. Двоично-инверсным для целого числа $k \in \{0, 1, \dots, 2^n - 1\}$ будем называть число k^* , которое получается в результате обратной записи k в двоичном представлении. То есть если $k = \sum_{j=0}^{n-1} k_j 2^j = k_{n-1}k_{n-2} \dots k_0$, $k_j \in \{0, 1\}$, 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 интерпретировать в виде матричного умножения, $\tilde{X} = CX$. Для рассматриваемого нами примера 8-точечного преобразования

$$\tilde{X} = \begin{pmatrix} x(0) \\ x(4) \\ x(2) \\ x(6) \\ x(1) \\ x(5) \\ x(3) \\ x(7) \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_C \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \\ x(4) \\ x(5) \\ x(6) \\ x(7) \end{pmatrix}.$$

Обобщая результаты примера 2.9 на случай размерности ДПФ $N = 2^n$ для произвольного числа $n \in \mathbb{N}$, матрицу W ненормированного преобразования $\tilde{Y} = WX$, которая определяет алгоритм БПФ с прореживанием по времени, можем записать в следующем виде:

$$W = W_n \cdot W_{n-1} \cdot \dots \cdot W_1 \cdot C, \quad (2.27)$$

где $C = \{c_{k,j}\}_{k,j=0}^{2^n-1}$ — симметрическая матрица перестановок элементов с двоично-инверсными номерами: $c_{k,j} = \begin{cases} 1, & \text{если } k = j^* \\ 0, & \text{если } k \neq j^* \end{cases}$. Матрица W_k в (2.27) имеет блочный вид:

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

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

$$\mathbf{B}_k^{(j)} = \begin{pmatrix} 0 \dots 0 & 1 & 0 & 0 \dots 0 & 1 & 0 & 0 & \dots & 0 & 0 \dots 0 \\ 0 \dots 0 & 0 & 1 & 0 \dots 0 & 0 & \omega_k & 0 & \dots & 0 & 0 \dots 0 \\ 0 \dots 0 & 0 & 0 & 1 \dots 0 & 0 & 0 & \omega_k^2 & \dots & 0 & 0 \dots 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 \dots 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & \omega_k^{2^{k-1}-1} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \dots 0 & 1 & 0 & 0 \dots 0 & -1 & 0 & 0 & \dots & 0 & \dots \\ 0 \dots 0 & 0 & 1 & 0 \dots 0 & 0 & -\omega_k & 0 & \dots & 0 & 0 \dots 0 \\ 0 \dots 0 & 0 & 0 & 1 \dots 0 & 0 & 0 & -\omega_k^2 & \dots & 0 & 0 \dots 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 \dots 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & -\omega_k^{2^{k-1}-1} & 0 \dots 0 \end{pmatrix}$$

$\underbrace{\hspace{10em}}_{j \cdot 2^k \text{ нулевых столбцов}} \quad \underbrace{\hspace{5em}}_{2^{k-1} \text{ столбцов}} \quad \underbrace{\hspace{5em}}_{2^{k-1} \text{ столбцов}} \quad \underbrace{\hspace{10em}}_{2^n - (j+1) \cdot 2^k \text{ нулевых столбцов}}$

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

1. Цикл по матрицам W_k , $k = 1, \dots, n$.
2. Цикл по блокам $B_k^{(j)}$ матрицы W_k , $j = 0, \dots, 2^{n-k} - 1$.
3. Цикл по строкам блока $B_k^{(j)}$, $l = 0, \dots, 2^{k-1} - 1$:

$$\begin{cases} y(j2^k + l) = x(j2^k + l) + \omega_k^l \cdot x(j2^k + l + 2^{k-1}) \\ y(2^{k-1} + j2^k + l) = x(j2^k + l) - \omega_k^l \cdot x(j2^k + l + 2^{k-1}) \end{cases}$$

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

Описанный способ вычисления БПФ позволяет экономно использовать память компьютера, так как получаемые после выполнения «бабочки» значения $y(j2^k + l)$ и $y(j2^k + 2^{k-1} + l)$ можно сохранять на месте элементов $x(j2^k + l)$ и $x(j2^k + 2^{k-1} + l)$, которые больше не будут участвовать в вычислениях. Поэтому достаточно выделить память для хранения только одного вектора размерности N , при этом результирующий вектор \tilde{Y} будет получен в результате последовательного замещения пар компонент входного вектора X .

Алгоритм БПФ с прореживанием по времени можно реализовать, исключив операции перестановок элементов исходного вектора X с двоично-инверсными номерами. Рассмотрим это на примере 8-точечного ДПФ. Если схему сведения одного ДПФ к двум

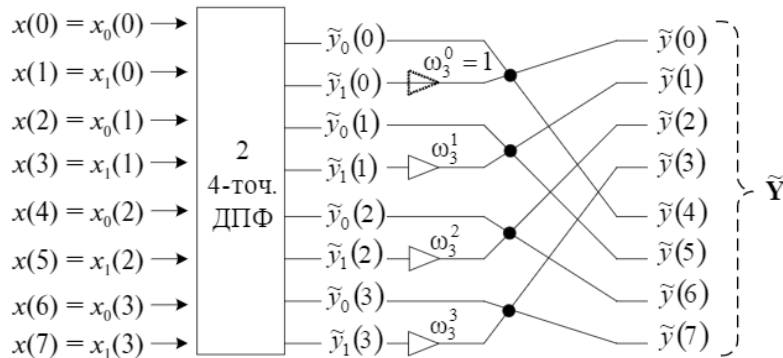


Рис. 2.9. Вариант графа вычислений ДПФ, определяемый формулой (2.25), на примере преобразования размерности $N = 8$. Обозначения см. рис. 2.5

ДПФ меньшей размерности (рис. 2.6) представить в несколько ином виде (рис. 2.9) и рекуррентно применить ее далее, то получим полный граф вычислений, приведенный на рис. 2.10.

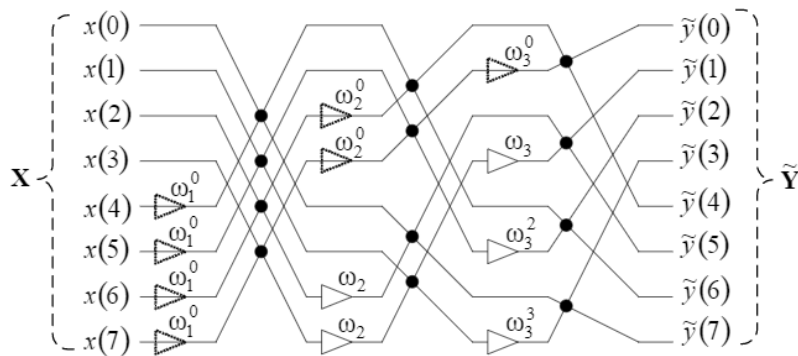


Рис. 2.10. Граф вычислений БПФ (алгоритм с прореживанием по времени) без перестановок элементов входного вектора с двоично-инверсными индексами, на примере преобразования размерности $N=8$. Обозначения см. рис. 2.6

Отметим, что исключение операций перестановок элементов входного вектора достигается ценой двойного увеличения объема памяти, необходимой при реализации БПФ, так как для векторов X и \tilde{Y} приходится выделять уже различные участки памяти компьютера.

Пример 2.10. 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: $W = W_3 W_2 W_1$ и, соответственно, три шага алгоритма БПФ. По аналогии с тем, как были получены матрицы-сомножители в примере 2.9 по графу рис. 2.7, получаем по графу рис. 2.10 следующие структуры матриц факторизованного представления:

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

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

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

Вновь можно заметить, что матрицы-сомножители факторизованного представления матрицы ДПФ имеют блочную структуру. Обобщая результаты примера 2.10 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

$$\mathbf{W} = \mathbf{W}_n \cdot \mathbf{W}_{n-1} \cdot \dots \cdot \mathbf{W}_1, \tag{2.28}$$

где

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

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

$$\mathbf{B}_k^{(j)}(\pm) = \left(\begin{array}{cccc} \underbrace{0 \ 0 \ 0}_{j \cdot 2^{n-k+1} \text{ нулевых столбцов}} & \underbrace{\begin{array}{ccc} 1 & 0 & 0 \dots 0 \\ 0 & 1 & 0 \dots 0 \\ 0 & 0 & 1 \dots 0 \\ \vdots & \vdots & \vdots \end{array}}_{2^{n-k} \text{ столбцов}} & \underbrace{\begin{array}{ccc} \pm \omega_k^j & 0 & 0 \dots 0 \\ 0 & \pm \omega_k^j & 0 \dots 0 \\ 0 & 0 & \pm \omega_k^j \dots 0 \\ \vdots & \vdots & \vdots \end{array}}_{2^{n-k} \text{ столбцов}} & \underbrace{0 \ 0 \ 0}_{2^n - (j+1) \cdot 2^{n-k+1} \text{ нулевых столбцов}} \end{array} \right) \left. \vphantom{\begin{array}{c} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{array}} \right\} 2^{n-k} \text{ строк}$$

Фактически, вторая форма алгоритма БПФ с прореживанием по времени, определяемая разложением (2.28), получается из первой формы (2.27) путем определенных перестановок строк в матрицах факторизации. Вычисление БПФ без перестановки элементов входного вектора \mathbf{X} также реализуется посредством трех вложенных циклов.

1. Цикл по матрицам \mathbf{W}_k , $k = 1, \dots, n$.
2. Цикл по блокам $\mathbf{B}_k^{(j)}(\pm)$ матрицы \mathbf{W}_k , $j = 0, \dots, 2^{k-1} - 1$.
3. Цикл по строкам блоков $\mathbf{B}_k^{(j)}(\pm)$, $l = 0, \dots, 2^{n-k} - 1$:

$$\begin{cases} y(j2^{n-k} + l) = x(j2^{n-k+1} + l) + \omega_k^j \cdot x(j2^{n-k+1} + l + 2^{n-k}) \\ y(2^{n-1} + j2^{n-k} + l) = x(j2^{n-k+1} + l) - \omega_k^j \cdot x(j2^{n-k+1} + l + 2^{n-k}) \end{cases}$$

Вновь под вектором \mathbf{X} на k -м шаге при $k = 2, \dots, n$ здесь понимается вектор \mathbf{Y} , полученный на $(k-1)$ -м шаге. В данной схеме вычислений

для хранения векторов X и Y требуется уже выделение отдельных массивов памяти.

Кратко остановимся на построении быстрого алгоритма обратного ДПФ. Все проведенные рассуждения для прямого ДПФ (2.24) можно в точности повторить, заменив $\omega_j = \exp(-2\pi i/2^j)$ на $\omega_j = \exp(2\pi i/2^j)$. Верными останутся и соотношения (2.25), и примеры графов вычислений, приведенные на рис. 2.6–2.10. Однако, имея вычислительную процедуру нахождения ДПФ $\tilde{Y} = WX$, можно осуществить и умножение на матрицу \bar{W} , необходимое при ОДПФ (см. (2.26)), поскольку $\sqrt{N}\bar{X} = W\tilde{Y}$. Реализация процедуры $B = \bar{W}A$ производится путем вычисления матричного произведения $\bar{B} = \bar{W}A$ и сводится к следующей последовательности действий.

1. $U \leftarrow \bar{A}$. (Сформировать вектор U из комплексно сопряженных компонент вектора A .)
2. $V = WU$. (Используя имеющуюся процедуру реализации ДПФ, выполнить умножение матрицы на вектор.)
3. $B \leftarrow \bar{V}$. (Сформировать вектор B из комплексно сопряженных компонент вектора V .)

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

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

2.8. Алгоритм БПФ с прореживанием по частоте

Ненормированное ДПФ (2.24) можно записать 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 из второй суммы множитель $\omega_n^{2^{n-1}k} = e^{-\frac{2\pi i}{2^n} 2^{n-1}k} = e^{-\pi i k} = (-1)^k$, 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:

$$\begin{cases} \tilde{y}(2k) = \sum_{j=0}^{2^{n-1}-1} (x(j) + x(j+2^{n-1})) \omega_{n-1}^{kj} \\ \tilde{y}(2k+1) = \sum_{j=0}^{2^{n-1}-1} (x(j) - x(j+2^{n-1})) \omega_n^j \omega_{n-1}^{kj} \\ k = 0, \dots, 2^{n-1} - 1 \end{cases} \quad (2.29)$$

Видим, что четные компоненты вектора $\tilde{\mathbf{Y}}$ (первая строка системы (2.29)) представляют собой ДПФ размерности $N/2 = 2^{n-1}$, выполненное над вектором

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

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

$$\mathbf{X}_1 = ((x(0) - x(2^{n-1})) \omega_n^0, (x(1) - x(2^{n-1} + 1)) \omega_n^1, \dots, (x(2^{n-1} - 1) - x(2^n - 1)) \omega_n^{2^{n-1}-1}).$$

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 рис. 2.11.

Формулы (2.29) определяют схему вычислений БПФ с прореживанием по частоте, которую можно рекурсивно применять далее для ДПФ векторов $\tilde{\mathbf{Y}}_0 = (\tilde{y}(0), \tilde{y}(2), \dots, \tilde{y}(2^n - 2)) = (\tilde{y}_0(0), \tilde{y}_0(1), \dots, \tilde{y}_0(2^{n-1} - 1))$ и $\tilde{\mathbf{Y}}_1 = (\tilde{y}(1), \tilde{y}(3), \dots, \tilde{y}(2^n - 1)) = (\tilde{y}_1(0), \tilde{y}_1(1), \dots, \tilde{y}_1(2^{n-1} - 1))$. Так, для случая $N = 8$ полный граф вычислений БПФ будет иметь вид, приведенный на рис. 2.12. Как и в алгоритме с прореживанием 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 (см. рис. 2.12).

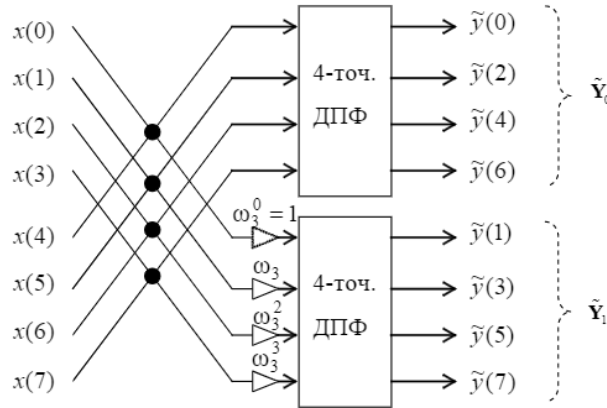


Рис. 2.11. Граф вычислений ДПФ, определяемый формулой (2.29), на примере преобразования размерности $N = 8$ (обозначения см. на рис. 2.6)

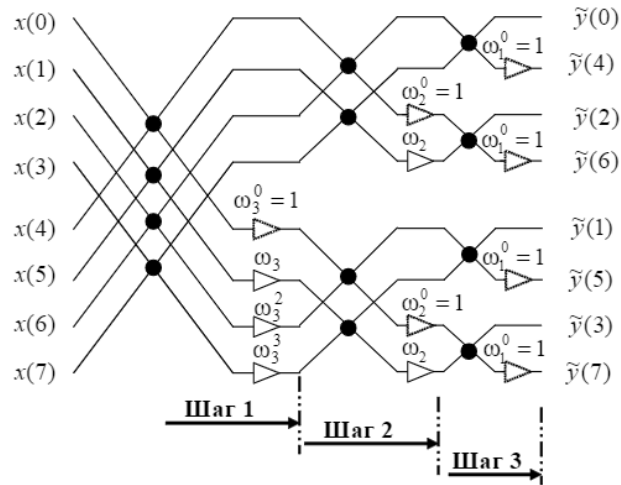


Рис. 2.12. Граф 8-точечного БПФ, алгоритм с прореживанием по частоте (обозначения см. на рис. 2.6)

Упражнение. Используя соотношение (2.29) и рис. 2.12, постройте граф вычислений БПФ размерности $N = 16$ для алгоритма с прореживанием по частоте.

Как видно из формулы (2.29) и примеров графов вычислений, приведенных на рис. 2.11 и 2.12, базовой операцией БПФ на j -м шаге для алгоритма с прореживанием по частоте является «бабочка», которая имеет вид, приведенный на рис. 2.13 (сравните с рис. 2.8).

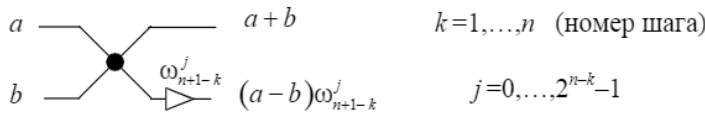


Рис. 2.13. Элементарные операции алгоритма БПФ с прореживанием по частоте на j -м шаге

Фактически (см. рис. 2.7 и 2.12) алгоритм с прореживанием по частоте имеет граф, который получается при проходе в обратном направлении графа алгоритма БПФ с прореживанием по времени с заменой узловых операций рис. 2.8 на «бабочки» рис. 2.13. В матричном виде факторизация матрицы ДПФ, определяемая алгоритмом с прореживанием по частоте, может быть получена непосредственно из представления (2.27) (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 $\mathbf{C} = \mathbf{C}^T$ представляет собой матрицу перестановок элементов с двоично-инверсными индексами, а матрица \mathbf{W}_j определяется по (2.27), $j = 1, \dots, n$.

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

Первая, рассмотренная выше, позволяет производить вычисления с использованием $\approx N$ ячеек памяти компьютера, 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 $\approx 2N$ ячеек памяти компьютера — соответствующая факторизация матрицы ДПФ задается соотношением:

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

где матрицы \mathbf{W}_j ($j = 1, \dots, n$) определены в представлении (2.28).

Пример 2.11. Представить матрицу ДПФ размерности $N = 8$ в факторизованном виде, соответствующем алгоритму БПФ с прореживанием по частоте без перестановок элементов выходного вектора.

◀ Воспользуемся результатами и обозначениями примера 2.10. Получаем:

$$\begin{aligned} \mathbf{W} = \mathbf{W}_1^T \mathbf{W}_2^T \mathbf{W}_3^T = & \\ = & \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & \omega_2 & 0 & 0 & 0 & -\omega_2 & 0 \\ 0 & 0 & 0 & \omega_2 & 0 & 0 & 0 & -\omega_2 \end{pmatrix} \times \\ & \times \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \omega_3 & 0 & 0 & 0 & -\omega_3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \omega_3^2 & 0 & 0 & 0 & -\omega_3^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \omega_3^3 & 0 & 0 & 0 & -\omega_3^3 \end{pmatrix}. \end{aligned}$$

Граф вычислений алгоритма БПФ с прореживанием по частоте, который не требует перестановок элементов вектора-спектра, получается при обратном проходе графа соответствующего (без перестановок) алгоритма с прореживанием по времени, с заменой узловых операций рис. 2.8 на операции рис. 2.13. Например, для случая

$N = 8$ получим полный граф вычислений, приведенный на рис. 2.14 (сравните с рис. 2.10).

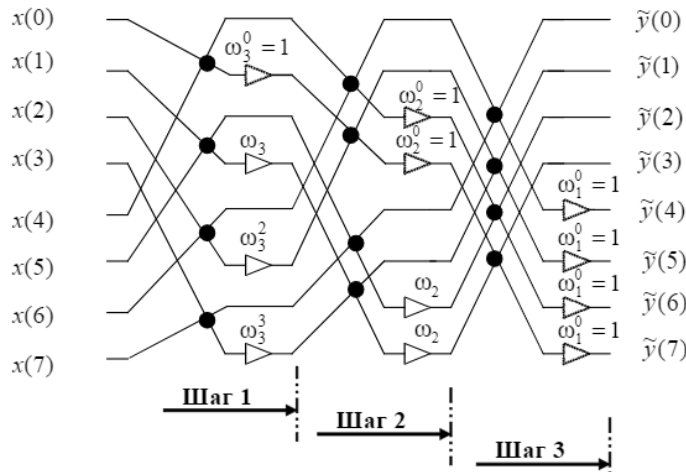


Рис. 2.14. Граф 8-точечного БПФ, алгоритм с прореживанием по частоте, не требующий перестановок элементов вектора-спектра (обозначения см. рис. 2.6)

Вычисление БПФ по алгоритму с прореживанием по частоте сводится к выполнению в точности того же количества вычислительных операций, что и реализация БПФ по алгоритму с прореживанием по времени.

Упражнение. Воспользовавшись результатами примера 2.9, выпишите факторизованное представление для матрицы ДПФ размерности $N = 8$, которое определяет алгоритм БПФ с прореживанием по частоте с перестановками компонент выходного вектора.

2.9. 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:

$$y_k = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x_j w_k(j/2^n), \quad (2.30)$$

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

где $w_k(t)$ — функции системы Уолша (см. раздел 1.3), а $n \in \mathbb{N}$.

Замечание. Если x_j — это отсчеты некоторой функции $f(t)$, определенной на отрезке $t \in [0; 1]$, т. е. $x_j = f(j/2^n)$, то (2.30) можно рассматривать как формулу для приближенного вычисления коэффициентов Фурье c_k по системе Уолша для 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, $Y = WX$, где X и Y — векторы-столбцы, а матрица ДПУ

$$W = \left\{ \begin{array}{l} \underbrace{w_{k,j}}_{\text{эл-т матрицы}} = \frac{1}{\sqrt{2^n}} \underbrace{w_k(j/2^n)}_{\text{отсчет ф. Уолша}} \end{array} \right\}_{k,j=0}^{2^n-1}.$$

Лемма 2.2. Матрица ДПУ — симметрическая, т. е. $W = W^T$.

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

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

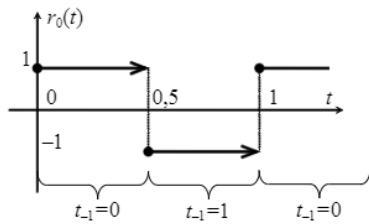


Рис. 2.15

где $\forall m \ t_m \in \{0; 1\}$, $t_m = [t/2^m] \bmod 2$. Тогда для $t \geq 0$ функцию $r_0(t)$ системы Радемахера (определение системы см. в разделе 1.3) можно задать следующим образом (см. рис. 2.15):

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

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

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

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. (1.11)) при $k = 0, \dots, 2^n - 1$ имеем:

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

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

Пример 2.12. Найти матрицу ДПУ для преобразования размерности $N = 2^2$.

◀ При доказательстве леммы 2.2 мы получили общее выражение:

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

В нашем случае $n = 2$, поэтому для $j = 2j_1 + j_0$, $k = 2k_1 + k_0$ имеем:

$$\mathbf{W} = \left(w_{k,j} = \frac{1}{2} (-1)^{j_0 k_1 + j_1 k_0} \right)_{k,j=0}^3. \quad \text{Получаем: } \mathbf{W} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}. \quad \blacktriangleright$$

Теорема 2.5. Матрица ДПУ — ортогональна: $\mathbf{W}^{-1} = \mathbf{W}^T$.

◀ Обозначим $\mathbf{U} = \mathbf{W}^T \mathbf{W} = \mathbf{W} \mathbf{W}$ 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 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. (1.9)), то на этих же полуинтервалах Δ_m^n также $\forall l < 2^n$:

$w_l(t) = \prod_{k=0}^{n-1} (r_k(t))^{l_k} = \text{const}$. Поэтому последнее выражение для $u_{k,j}$ можем переписать в виде:

$$u_{k,j} = \sum_{m=0}^{2^n-1} \int_{\Delta_m^n} \underbrace{w_k(t)w_j(t)}_{w_k(\frac{m}{2^n})w_j(\frac{m}{2^n})=\text{const}} dt = \int_0^1 w_k(t)w_j(t)dt = \langle w_k(t), w_j(t) \rangle = \begin{cases} 1, & \text{при } k = j \\ 0, & \text{при } k \neq j \end{cases}$$

(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 (2.30): $\mathbf{X} = \mathbf{W}\mathbf{Y}$, или

$$x_k = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} y_j w_k(j/2^n), \quad k = 0, 1, \dots, 2^n - 1.$$

Упражнение. Найдите ДПУ каждой из приведенных ниже конечных последовательностей длины $N = 4$:

$$\text{а) } x(n) = \tilde{\delta}(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}; \quad \text{б) } x(n) = \tilde{\delta}(n - n_0), \quad 1 \leq n_0 \leq 3.$$

Если не учитывать окончательной нормировки ($N = 2^n$ умножений на множитель $1/\sqrt{2^n}$), то ДПУ по формуле (2.30) сводится примерно к $N^2 = 2^{2n}$ операциям сложения (вычитания) чисел. Однако ДПУ также обладает быстрыми алгоритмами вычислений, позволяющими реализовать преобразование за счет вычисления $\sim N \log N$ сложений (вычитаний). Идеи построения таких алгоритмов имеют много общего с принципами построения алгоритмов БПФ.

Рассмотрим один из возможных алгоритмов вычисления быстрого преобразования Уолша (БПУ). Обозначая $y(k) = y_k$, $x(j) = x_j$, на основании полученных при доказательстве леммы 2.2 соотношений для ДПУ (2.30) имеем:

$$y(k) = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x(j) w_k(j/2^n) = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x(j) \prod_{m=0}^{n-1} (-1)^{j_{n-1-m} k_m}.$$

Запишем в последней сумме индекс $j = 2l + j_0$, где $j_0 \in \{0; 1\}$, $l = 0, \dots, 2^{n-1} - 1$. Тогда в двоичном представлении чисел l и j имеем $j_{m+1} = l_m$ ($m = 0, \dots, n-2$) и выражение для ДПУ можем переписать как

$$y(k) = \frac{1}{\sqrt{2^n}} \sum_{l=0}^{2^{n-1}-1} \sum_{j_0=0}^1 \left(x(2l + j_0) (-1)^{j_0 k_{n-1}} \prod_{m=0}^{n-2} (-1)^{l_{n-2-m} k_m} \right) =$$

$$= \frac{1}{\sqrt{2^n}} \sum_{l=0}^{2^{n-1}-1} (x(2l) + x(2l+1)) (-1)^{k_{n-1}} \prod_{m=0}^{n-2} (-1)^{l_{n-2-m} k_m}, \quad k = 0, \dots, 2^n - 1.$$

Для $k = 0, \dots, 2^{n-1} - 1$, т.е. когда $k_{n-1} = 0$, из последнего выражения получаем

$$y(k) = \frac{1}{\sqrt{2^n}} \sum_{l=0}^{2^{n-1}-1} (x(2l) + x(2l+1)) \prod_{m=0}^{n-2} (-1)^{l_{n-2-m} k_m},$$

а для $k = 2^{n-1}, \dots, 2^n - 1$, т.е. когда $k_{n-1} = 1$, имеем:

$$y(k) = \frac{1}{\sqrt{2^n}} \sum_{l=0}^{2^{n-1}-1} (x(2l) - x(2l+1)) \prod_{m=0}^{n-2} (-1)^{l_{n-2-m} k_m}.$$

Учитывая, что $\prod_{m=0}^{n-2} (-1)^{l_{n-2-m} k_m} = w_k(l/2^{n-1}) = w_l(k/2^{n-1})$, окончательно запишем:

$$\begin{cases} y(k) = \frac{1}{\sqrt{2^n}} \sum_{l=0}^{2^{n-1}-1} (x(2l) + x(2l+1)) w_k(l/2^{n-1}) = \frac{1}{\sqrt{2}} y_0(k), \\ y(k + 2^{n-1}) = \frac{1}{\sqrt{2^n}} \sum_{l=0}^{2^{n-1}-1} (x(2l) - x(2l+1)) w_k(l/2^{n-1}) = \frac{1}{\sqrt{2}} y_1(k), \\ k = 0, \dots, 2^{n-1} - 1, \end{cases}$$

где векторы размерности 2^{n-1} : $\mathbf{Y}_0 = (y_0(0), y_0(1), \dots, y_0(2^{n-1}-1))$, $\mathbf{Y}_1 = (y_1(0), y_1(1), \dots, y_1(2^{n-1}-1))$ представляют собой ДПУ (2.30) векторов $\mathbf{X}_0 = (x(0) + x(1), x(2) + x(3), \dots, x(2^n - 2) + x(2^n - 1))$ и $\mathbf{X}_1 = (x(0) - x(1), x(2) - x(3), \dots, x(2^n - 2) - x(2^n - 1))$ соответственно.

Таким образом, мы свели вычисление ДПУ размерности $N = 2^n$ к вычислению двух ДПУ размерностей 2^{n-1} . Если обозначить матрицу ДПУ размерности $N = 2^n$ как $\mathbf{W}(2^n)$, то полученная схема вычислений определяет следующую структуру матрицы ДПУ:

$$\mathbf{W}(2^n) = \frac{1}{\sqrt{2}} \left(\begin{array}{ccc|ccc} & & & 0 & \dots & 0 \\ & & & \vdots & & \vdots \\ & & \mathbf{W}(2^{n-1}) & \vdots & & \vdots \\ & & & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & & & \\ \vdots & & \vdots & \mathbf{W}(2^{n-1}) & & \\ 0 & \dots & 0 & & & \end{array} \right) \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \ddots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \\ \hline 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & -1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \ddots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix}. \quad (2.31)$$

Применяя формулу (2.31) рекуррентно для матриц-блоков $W(2^{n-1})$ и далее для $W(2^{n-2}), \dots, W(2)$, получим, что матрица ДПУ размерности $N = 2^n$ может быть представлена в виде произведения n слабозаполненных матриц, в каждой строке и каждом столбце которых имеется по два ненулевых элемента (1 или -1). Например,

$$W(4) = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

В общем случае умножение на каждую из n матриц факторизованного представления реализуется $N = 2^n$ операциями сложения (вычитания) чисел, таким образом, всего для вычисления БПУ необходимо выполнить $n2^n$ сложений (вычитаний).

Описанный алгоритм БПУ также можно представить в виде графа. Для размерности преобразования $N = 8$ такой граф изображен на рис. 2.16.

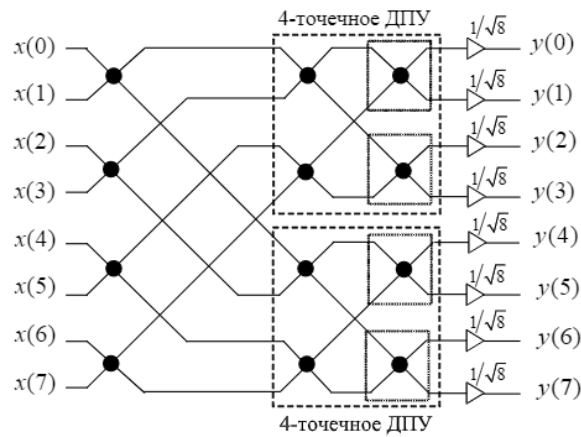


Рис. 2.16. Граф 8-точечного БПУ. Обозначения см. на рис. 2.6

Упражнение. Запишите факторизованное представление для матрицы ДПУ размерности $N = 8$.

Отметим, что наряду с ДПУ часто рассматривается *дискретное преобразование Адамара* (или *Уолша — Адамара*), отличающееся только способом нумерации строк матрицы преобразования. Вектор, полученный в результате дискретного преобразования Адамара, представляет собой вектор ДПУ (2.30), в котором переставлены местами пары компонент с двоично-инверсными индексами.

2.10. 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:

$$y_k = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x_j h_k(j/2^n), \quad (2.32)$$

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

где $h_k(t)$ — функции системы Хаара (см. раздел 1.3), а $n \in \mathbb{N}$.

Замечание. Выражение (2.32) также (см. замечание к (2.30)) можно рассматривать как формулу для приближенного вычисления 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 — k -й коэффициент Фурье разложения функции $f(t)$ по системе Хаара.

Пример 2.13. Найдем матрицу ДПХ для преобразования размерности $N = 2^2$.

◀ 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:

$$\mathbf{H} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}. \quad \blacktriangleright$$

Теорема 2.6. Матрица $\mathbf{H} = \left\{ \frac{1}{\sqrt{2^n}} h_k(j/2^n) \right\}_{k,j=0}^{2^n-1}$, 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}):

$u_{k,j} = \frac{1}{2^n} \sum_{m=0}^{2^n-1} h_k\left(\frac{m}{2^n}\right) h_j\left(\frac{m}{2^n}\right)$. 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:

$$u_{k,j} = \sum_{m=0}^{2^n-1} \int_{\Delta_m^n} h_k(t) h_j(t) dt = \int_0^1 h_k(t) h_j(t) dt = \langle h_k(t), h_j(t) \rangle = \begin{cases} 1, & \text{при } k = j \\ 0, & \text{при } k \neq j \end{cases}$$

(в силу ортонормированности системы Хаара, см. теорему 1.11). ▶

Следствие. Обратное к ДПХ (2.32) преобразование записывается в виде:

$$x_j = \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} y_k h_k(j/2^n), \quad j = 0, 1, \dots, 2^n-1.$$

Упражнение. Найдите ДПХ каждой из приведенных ниже конечных последовательностей длины $N = 4$:

$$\text{а) } x(n) = \tilde{\delta}(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}; \quad \text{б) } x(n) = \tilde{\delta}(n - n_0), \quad 1 \leq n_0 \leq 3.$$

Рассмотрим построение алгоритма вычисления быстрого преобразования Хаара (БПХ). Для этого вновь переобозначим: $y(k) = y_k$, $x(j) = x_j$.

Прежде всего, заметим, что для $k = 0, \dots, 2^{n-1}-1$ функции $h_k(t) = \text{const}$ при $t \in \Delta_l^{n-1} = \left[\frac{l}{2^{n-1}}, \frac{l+1}{2^{n-1}} \right)$, $l = 0, \dots, 2^{n-1}-1$ (см. главу 1: пример 1.12 и свойство 1° системы функций Хаара). Поэтому для

$k = 0, \dots, 2^{n-1} - 1$ выражение (2.32) можем представить следующим образом:

$$\begin{aligned} y(k) &= \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x(j) h_k(j/2^n) = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} \left(x(2j) h_k \left(\frac{2j}{2^n} \right) + x(2j+1) h_k \left(\frac{2j+1}{2^n} \right) \right) = \\ &= \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^{n-1}-1} (x(2j) + x(2j+1)) h_k(j/2^{n-1}) = \frac{1}{\sqrt{2}} y_0(k), \end{aligned}$$

где $\mathbf{Y}_0 = (y_0(0), y_0(1), \dots, y_0(2^{n-1} - 1))$ – вектор, полученный в результате ДПХ (2.32) размерности 2^{n-1} из вектора $\mathbf{X}_0 = (x_0 + x_1, x_2 + x_3, \dots, x_{2^{n-2}} + x_{2^{n-1}})$.

Для второй половины индексов $k = 2^{n-1}, \dots, 2^n - 1$ в формуле (2.32), вводя обозначение $k = 2^{n-1} + l$ ($l = 0, \dots, 2^{n-1} - 1$), имеем для функций системы Хаара (см. определение (1.12)):

$$h_{2^{n-1}+l}(x) = \begin{cases} 2^{-\frac{n-1}{2}} & \text{при } x \in \Delta_{2l}^n \\ -2^{-\frac{n-1}{2}} & \text{при } x \in \Delta_{2l+1}^n \\ 0 & \text{иначе} \end{cases}.$$

Поэтому из формулы (2.32) получаем для $k = 2^{n-1} + l$ ($l = 0, \dots, 2^{n-1} - 1$):

$$y(2^{n-1} + l) = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} x(j) h_{2^{n-1}+l}(j/2^n) = \frac{1}{\sqrt{2}} (x(2l) - x(2l+1)) = \frac{1}{\sqrt{2}} y_1(l),$$

где $\mathbf{Y}_1 = (y_1(0), y_1(1), \dots, y_1(2^{n-1} - 1)) = (x_0 - x_1, x_2 - x_3, \dots, x_{2^{n-2}} - x_{2^{n-1}})$.

Таким образом, с точностью до множителя $1/\sqrt{2}$ первая половина компонент вектора \mathbf{Y} ДПХ (2.26) размерности $N = 2^n$ представляет собой ДПХ (2.32), выполненное над вектором $\mathbf{X}_0 = (x_0 + x_1, x_2 + x_3, \dots, x_{2^{n-2}} + x_{2^{n-1}})$, а вторая половина компонент представляет собой вектор разностей $\mathbf{X}_1 = (x_0 - x_1, x_2 - x_3, \dots, x_{2^{n-2}} - x_{2^{n-1}})$. Если обозначить матрицу ДПХ размерности $N = 2^n$ как $\mathbf{H}(2^n)$, то установленная схема вычислений определяет следующую структуру матрицы ДПХ:

$$\mathbf{H}(2^n) = \frac{1}{\sqrt{2}} \left(\begin{array}{c|ccc} \mathbf{H}(2^{n-1}) & 0 & \dots & 0 \\ & \vdots & \ddots & \vdots \\ & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & 0 & 1 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{array} \right) \cdot \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \\ \hline 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & -1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix}. \quad (2.33)$$

Применяя формулу (2.33) рекуррентно для матриц-блоков $\mathbf{H}(2^{n-1})$ и далее для $\mathbf{H}(2^{n-2}), \dots, \mathbf{H}(2)$, получим, что матрица ДПХ размерности $N = 2^n$ может быть представлена в виде произведения n слабозаполненных матриц: $\mathbf{H}(2^n) = \mathbf{H}_n(2^n) \mathbf{H}_{n-1}(2^n) \dots \mathbf{H}_1(2^n)$. Так (см. также пример 2.13),

$$\mathbf{H}(4) = \begin{pmatrix} 1/2 & 1 & 1 & 1 & 1 \\ 1/2 & 1 & 1 & -1 & -1 \\ 1/\sqrt{2} & 1 & -1 & 0 & 0 \\ 1/\sqrt{2} & 0 & 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1/2 & 1 & 1 & 0 & 0 \\ 1/2 & 1 & -1 & 0 & 0 \\ 1/\sqrt{2} & 0 & 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Упражнение. Запишите факторизованное представление для матрицы ДПХ размерности $N = 8$.

В общем случае умножение на матрицу $\mathbf{H}_k(2^n)$ факторизованного представления ($k = 1, \dots, n$) реализуется 2^{n+1-k} операциями сложения (вычитания) чисел, таким образом, всего для вычисления БПХ необходимо выполнить $\sum_{k=1}^n 2^{n+1-k} = 2^{n+1} - 2$ сложений (вычитаний). Дополнительно $N = 2^n$ умножений требуется выполнить для окончательной нормировки коэффициентов преобразования.

Описанный алгоритм БПХ, определяемый рекуррентной формулой (2.33), также можно представить в виде графа. Для размерности преобразования $N = 8$ такой граф изображен на рис. 2.17.

Упражнение. Поясните, как может быть реализовано вычисление обратного ДПХ с использованием быстрого алгоритма.

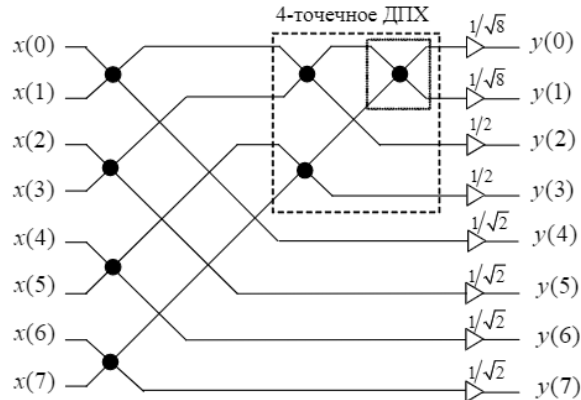


Рис. 2.17. Граф 8-точечного БПХ. Обозначения см. на рис. 2.6

Таким образом, все рассмотренные нами дискретные преобразования (ДПФ, ДПУ, ДПХ) имеют быстрые алгоритмы вычисления. С увеличением размерности преобразования N количество необходимых вычислительных операций при использовании быстрых алгоритмов для БПФ и БПУ имеет порядок $\sim N \log N$, а для БПХ $\sim N$. Причем ДПХ и ДПУ являются вещественными преобразованиями и реализуются только за счет операций сложения и вычитания чисел, не требуя операций умножения (за исключением окончательной нормировки коэффициентов, N умножений, что на практике требуется не всегда).

2.11. Некоторые применения дискретных ортогональных преобразований

Кратко рассмотрим некоторые применения дискретных ортогональных преобразований.

Пример 2.14. Приближенное вычисление первых коэффициентов ряда Фурье.

◀ Как было отмечено в данной главе ранее, выражения для коэффициентов рассмотренных дискретных преобразований (2.20), (2.30),

(2.32) с точностью до нормировочного множителя представляют собой приближенное вычисление по формуле прямоугольников интегралов, определяющих первые коэффициенты $c_k = \langle f(t), g_k(t) \rangle$ ряда Фурье $f(t) = \sum_k c_k g_k(t)$ по соответствующим системам функций $\{g_k(t)\}$ (комплексных экспонент, Уолша, Хаара). ►

Пример 2.15. Вычисление свертки векторов с использованием БПФ.

◀ Для вектора-свертки $\mathbf{C} = (c_0, c_1, \dots, c_{2N-1})$ векторов $\mathbf{A} = (a_0, a_1, \dots, a_{N-1})$ и $\mathbf{B} = (b_0, b_1, \dots, b_{N-1})$ была определена формула (2.22), выписав которую по отдельности для каждой компоненты, можем подсчитать количество умножений, необходимых для вычисления всего вектора \mathbf{C} . Получаем $2(1+2+\dots+(N-1))+N = N^2$ умножений (вещественных или комплексных, в зависимости от характера векторов \mathbf{A} и \mathbf{B}), см. таблицу.

| Компонента свертки, формула вычисления | Количество 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 |
| $c_{2N-1} = 0$ | 0 |

Положим, что $N = 2^n$. Используя свойство 5° ДПФ, можно вычислить вектор-свертку следующим образом.

1. Сформируем $2N$ -компонентные векторы $\tilde{\mathbf{A}} = (a_0, \dots, a_{N-1}, \underbrace{0, \dots, 0}_N)$, $\tilde{\mathbf{B}} = (b_0, \dots, b_{N-1}, \underbrace{0, \dots, 0}_N)$, и с использованием алгоритма БПФ получим векторы $\hat{\mathbf{A}} = (\hat{a}_0, \hat{a}_1, \dots, \hat{a}_{2N-1})$ — ДПФ вектора $\tilde{\mathbf{A}}$ и $\hat{\mathbf{B}} = (\hat{b}_0, \hat{b}_1, \dots, \hat{b}_{2N-1})$ — ДПФ вектора $\tilde{\mathbf{B}}$. Для $N = 2^n$ каждое вычисление двух БПФ размерности $2N = 2^{n+1}$ потребует (см. раздел 2.7) 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 $\hat{\mathbf{A}}$ и $\hat{\mathbf{B}}$, $\forall k: \hat{c}_k = \sqrt{2N} \hat{a}_k \hat{b}_k$. 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 количество 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 . ►

Упражнение. Положим, что для некоторого вычислительного устройства, реализующего вещественные арифметические операции, выполнение операции умножения требует намного больше времени, чем операция сложения (вычитания). По приблизительному количеству необходимых операций вещественного умножения оцените размерность $N = 2^n$ векторов \mathbf{X} , \mathbf{Y} , для которых использование БПФ при вычислении свертки (2.22) дает выигрыш во времени вычислений, если:

- а) векторы комплекснозначные; б) векторы вещественные.

Пример 2.16. 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. Например, во многих случаях высокочастотные компоненты ДПФ (2.18), (2.20) по абсолютной величине намного меньше низкочастотных составляющих. Если вместо вектора \mathbf{X} кодировать вектор меньшей размерности, образованный из \mathbf{Y} отбрасыванием малых 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 $\|\mathbf{X} - \check{\mathbf{X}}\|_E$, то в силу ортогональности используемого преобразования, так как $\mathbf{Y} - \check{\mathbf{Y}} = \mathbf{F}(\mathbf{X} - \check{\mathbf{X}})$, имеем: $\|\mathbf{X} - \check{\mathbf{X}}\|_E^2 = \|\mathbf{Y} - \check{\mathbf{Y}}\|_E^2 = \sum_{k \in \Theta} |y_k|^2$, 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 $\check{\mathbf{Y}}$, то такие именно преобразования используются для сжатия данных чаще всего. Причем важным (нередко решающим) фактором, определяющим возможность использования того или иного ортогонального преобразования, является наличие у него быстрого алгоритма вычисления. ►

Подробнее вопросы анализа эффективности использования дискретных ортогональных преобразований для сжатия данных с потерями информации будут рассмотрены в шестой главе.

2.12. 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 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 варианты выбора 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. 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 $q = (x_{\max} - x_{\min})/N$: $\Delta_j = [x_{\min} + jq; x_{\min} + (j+1)q)$, $j = 0, \dots, N-1$. 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: $|\hat{X} - X| \leq q/2$. Часто эту ошибку характеризуют некоторыми средними значениями, например, используется величина

$$E = M((\hat{X} - X)^2), \quad (2.34)$$

где M означает математическое ожидание. Известно, что при достаточно малой величине дискрета квантования q ошибку квантования $\varepsilon = \hat{X} - X$ 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 $D(\varepsilon) = E = q^2/12$.

Оптимальное квантование Ллойда — Макса

Вновь положим, что случайная величина непрерывного типа X принимает значения из некоторого 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 $\Delta_j = [t_j; t_{j+1})$, $j = 0, \dots, N-1$; $t_0 = x_{\min}$, $t_N = x_{\max}$. 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 (2.34) как функции данных параметров:

$$E = g(t_1, \dots, t_{N-1}; d_0, \dots, d_{N-1}) = M[(X - Q(X))^2] = \int_{t_0}^{t_N} (x - Q(x))^2 f_X(x) dx = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} (x - d_j)^2 f_X(x) dx. \quad (2.35)$$

Необходимым условием минимума функции $g(\dots)$ из (2.35) 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 $t_j - d_{j-1} = d_j - t_j$, получаем:

$$t_j = \frac{d_{j-1} + d_j}{2}, \quad j = 1, \dots, N-1. \quad (2.36)$$

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 (2.35) по переменным $\{d_j\}_{j=0}^{N-1}$ получаем:

$$\frac{\partial g(\dots)}{\partial d_j} = -2 \int_{t_j}^{t_{j+1}} (x - d_j) f_X(x) dx = 0.$$

Отсюда



$$d_j = \int_{t_j}^{t_{j+1}} x f_X(x) dx \bigg/ \int_{t_j}^{t_{j+1}} f_X(x) dx, \quad (2.37)$$

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

Заметим, что величины d_j из (2.37) 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 (2.36), (2.37) определяет *квантователь Ллойда — Макса*. В общем случае решение системы уравнений (2.36), (2.37) 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. (2.36).

Отметим также, что все рассуждения, которые привели к построению системы уравнений (2.36), (2.37), 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 2.17. 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 (2.37) получаем:

$$d_j = \int_{t_j}^{t_{j+1}} \frac{x}{b-a} dx \bigg/ \int_{t_j}^{t_{j+1}} \frac{1}{b-a} dx = \frac{t_{j+1}^2 - t_j^2}{2(t_{j+1} - t_j)} = \frac{t_{j+1} + t_j}{2}, \quad j = 0, \dots, N-1.$$

С учетом (2.36) и условий $t_0 = a$, $t_N = b$ для параметров оптимального квантователя Макса получаем окончательно следующую систему уравнений, которая является линейной и легко разрешима:

$$\begin{cases} t_0 = a; & t_N = b; \\ t_j = \frac{d_{j-1} + d_j}{2}, & j = 1, \dots, N-1; \\ d_j = \frac{t_j + t_{j+1}}{2}, & j = 0, \dots, N-1. \end{cases}$$

Решение системы — очевидно, оно определяет рассмотренный выше равномерный квантователь, который для равномерно распределенной случайной величины является оптимальным в смысле минимума ошибки (2.34). ►

Функции плотности вероятности, используемые для моделирования распределений квантуемой величины X , часто обладают свойством симметрии, т. е. $f_X(m+t) = f_X(m-t)$, где $m = M[X]$. В этом случае пороги и уровни оптимального квантователя также оказываются расположенными симметрично относительно математического ожидания m :

$$m - t_k = t_{N-k} - m, \quad k = 0, \dots, N; \quad m - d_n = d_{N-1-n} - m, \quad n = 0, \dots, N-1.$$

Это свойство помогает упростить нахождение параметров квантователя для симметричных распределений.

Упражнение. Найти параметры оптимального квантователя Ллойда — Макса для квантования с $N = 2$ уровнями случайной величины X , имеющей следующую функцию плотности распределения вероятностей:

$$f_X(x) = \begin{cases} \frac{3}{2}x^2, & x \in [-1; 1] \\ 0, & x \notin [-1; 1] \end{cases}.$$

Итак, процесс квантования отсчетов дискретного сигнала можно рассматривать как преобразование случайной величины непрерывного типа X в случайную величину дискретного типа $\hat{X} = Q(X)$. Трактруя уровни квантования как символы некоторого алфавита $\{d_0, \dots, d_{N-1}\}$, 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 пятой главе.

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 элементы, 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 $X \in \mathbb{R}^M$ ставится в соответствие некоторый вектор V из того же пространства \mathbb{R}^M . Множество всех N возможных значений кодовых векторов $C = \{V_0, V_1, \dots, V_{N-1}\}$ называется *кодовой книгой*, а количество векторов N — объемом (или размером) кодовой книги. После векторного квантования передаче или хранению вместо исходного вектора X подвергается индекс j вектора $V_j = \hat{X} = Q(X)$. 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 не 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 большим потенциалом с точки зрения эффективного представления данных. Покажем это на простом примере.

Пример 2.18. Пусть случайный 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.

плотность распределения $f_{x,y}(x, y) = \begin{cases} 1/2, & (x, y) \in G \\ 0, & (x, y) \notin G \end{cases}$. 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 компонент. 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 2.17, 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 плотность 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 $\hat{Y} \in \{-1/2; 1/2\}$.

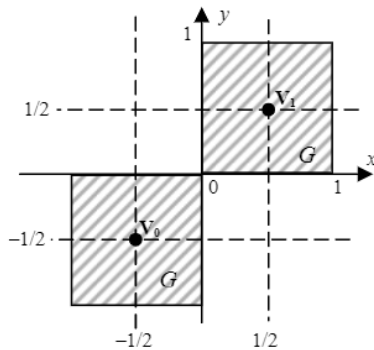


Рис. 2.18

Вместе с тем очевидно (см. рис. 2.18), 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 $C = \{V_0 = (-1/2; -1/2), V_1 = (1/2; 1/2)\}$. Объем кодовой книги $N = 2$, при этом ошибка векторного квантования совпадает с ошибкой скалярного квантования. ▶

Ошибку, возникающую в результате векторного квантования, характеризуют некоторой нормой $\|X - V_m\|$. Для каждого входного вектора X векторный квантователь Q ставит в соответствие выходной вектор $V_{m^*} = Q(X)$, который выбирается из кодовой книги $C = \{V_0, V_1, \dots, V_{N-1}\}$ по критерию максимальной близости:

$\|X - V_m\| = \min_{m=0,1,\dots,N-1} \|X - V_m\|$ (в случае скалярного квантования это определялось условием (2.36)). Таким образом, квантователь (функция Q) полностью определяется кодовой книгой C . Поэтому построение оптимального векторного квантователя для заданного объема кодовой книги N можно понимать как отыскание таких векторов $C = \{V_0, V_1, \dots, V_{N-1}\}$, которые дают минимальное значение математического ожидания ошибки квантования:

$$e(Q) = e(V_0, V_1, \dots, V_{N-1}) = M(\|X - Q(X)\|) \rightarrow \min_{V_0, V_1, \dots, V_{N-1}}.$$

В качестве векторной нормы чаще всего используется евклидова норма, а кодовая книга ищется по критерию

$$E(Q) = E(V_0, V_1, \dots, V_{N-1}) = M(\|X - Q(X)\|^2) \rightarrow \min_{V_0, V_1, \dots, V_{N-1}}. \quad (2.38)$$

Соответствующий квантователь называют *обобщенным квантователем Ллойда — Макса*.

Задание кодовой книги $C = \{V_0, V_1, \dots, V_{N-1}\}$ определяет разбиение пространства $\mathbb{R}^M = \bigcup_{j=0}^{N-1} w_j$ на области

$$w_j = \left\{ X \in \mathbb{R}^M \mid \|X - V_j\| = \min_{k=0,1,\dots,N-1} \|X - V_k\| \right\},$$

где для оптимального квантователя Ллойда — Макса, как можно показать, кодовые векторы $V_j \in w_j$ являются *центроидами* соответствующих ячеек, т. е. представляют собой условные математические ожидания: $V_j = M(X \mid X \in w_j)$. (В скалярном случае данное свойство принимает вид (2.37).) На рис. 2.19 приведен пример разбиения пространства \mathbb{R}^2 .

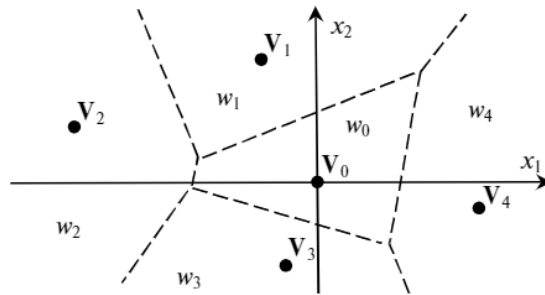


Рис. 2.19. Пример разбиения пространства \mathbb{R}^2 , определяемый кодовой книгой $C = \{V_0, V_1, V_2, V_3, V_4\}$

При известной функции плотности совместного распределения $f(x_1, \dots, x_M)$ компонент обрабатываемого вектора \mathbf{X} поиск оптимального векторного квантователя по критерию (2.38) означает минимизацию функции

$$E(\mathbf{V}_0, \mathbf{V}_1, \dots, \mathbf{V}_{N-1}) = \int \dots \int_{\mathbb{R}^M} \|\mathbf{X} - Q(\mathbf{X})\|^2 f(x_1, \dots, x_M) dx_1 \dots dx_M. \quad (2.39)$$

Обратим внимание на то, что для $M=1$ выражение (2.39) переходит в частный случай — формулу (2.35), 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 (2.39), 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) в варианте, соответствующем фиксированному размеру кодовой книги N . 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 $C = \{\mathbf{V}_k\}_{k=0}^{N-1}$. Положить среднюю ошибку квантования $E_0 = +\infty$.

Шаг 2. Построить разбиение $\Omega = \{\omega_k\}_{k=0}^{N-1}$ множества обучающих векторов A , где ячейки разбиения определяются по кодовой книге $C = \{\mathbf{V}_k\}_{k=0}^{N-1}$ следующим образом:

$$\omega_k = \left\{ \mathbf{a}_m \in A \mid \|\mathbf{a}_m - \mathbf{V}_k\| = \min_{j=0, \dots, N-1} \|\mathbf{a}_m - \mathbf{V}_j\| \right\}.$$

Шаг 3. Вычислить средний квадрат ошибки квантования:

$$E = \frac{1}{L} \sum_{k=0}^{N-1} \sum_{\mathbf{a}_m \in \omega_k} \|\mathbf{a}_m - \mathbf{V}_k\|^2.$$

Шаг 4. Проверка возможности останова. Если $(E_0 - E)/E < \varepsilon$, то прекратить вычисления, иначе перейти на шаг 5.

Шаг 5. Найти новые кодовые векторы усреднением по ячейкам:

$$\mathbf{V}_k = \frac{1}{|\{\mathbf{a}_m \in \omega_k\}|} \sum_{\mathbf{a}_m \in \omega_k} \mathbf{a}_m, \quad k = 0, \dots, N-1,$$

где $|\{\mathbf{a}_m \in \omega_k\}|$ — количество векторов обучающей последовательности A , попавших в ячейку ω_k .

Положить $E_0 = E$ и перейти на шаг 2.

Как видно из приведенного описания алгоритма LBG, его идея состоит в следующем¹. 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 $C = \{\mathbf{V}_k\}_{k=0}^{N-1}$ (шаг 2). Затем кодовые векторы заменяются усредненными векторами, найденными для классов $\{\omega_k\}_{k=0}^{N-1}$ (шаг 5). Процедура повторяется до тех пор, пока изменение ошибки квантования E обучающей последовательности A не станет малым (шаги 3, 4). Полученное разбиение множества обучающих векторов $A = \bigcup_{k=0}^{N-1} \omega_k$ на подмножества $\{\omega_k\}_{k=0}^{N-1}$ определяет итоговую кодовую книгу $C = \{\mathbf{V}_k\}_{k=0}^{N-1}$ и соответствующее разбиение пространства $\mathbb{R}^M = \bigcup_{k=0}^{N-1} \omega_k$.

К сожалению, результаты работы алгоритма LBG во многом зависят от того, насколько удачна была выбрана на первом шаге начальная кодовая книга C . Алгоритм не дает гарантии, что для заданной последовательности обучающих векторов итоговая ошибка (см. шаг 3)

$$E(\mathbf{V}_0, \dots, \mathbf{V}_{N-1}) = \frac{1}{L} \sum_{k=0}^{N-1} \sum_{\mathbf{a}_m \in \omega_k} \|\mathbf{a}_m - \mathbf{V}_k\|^2 \quad (2.40)$$

достигнет своего глобального, а не локального минимума. Поэтому на практике алгоритм запускается несколько раз со случайной инициализацией кодовой книги, и затем из полученных вариантов выбирают тот, который дал на обучающей последовательности векторов минимальное значение ошибки (2.40).

¹ Отметим, что полностью аналогичный алгоритму LBG подход — метод *K-средних* — используется в анализе данных для выделения и группировки близких по некоторой метрике векторов обучающей выборки.

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. По этой причине 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 2.18), проявляющегося при скалярном квантовании и кодировании статистически зависимых компонент обрабатываемых данных, чаще используются другие приемы, которые будут подробнее рассмотрены в шестой главе.

ГЛАВА 3

ОПИСАНИЕ И АНАЛИЗ ЛИНЕЙНЫХ ДИСКРЕТНЫХ СИСТЕМ

3.1. Нормирование временной и частотной осей

В предыдущей главе мы рассмотрели вопросы преобразования аналогового сигнала в дискретный: $f(t) \rightarrow \{f(n\Delta t)\}_{n=-\infty}^{\infty}$, где $T = \Delta t$ — период дискретизации сигнала, определенный, например, в соответствии с критерием (2.7). Однако мы можем перейти к «нормированному» времени $s = t/T$ и рассмотреть сигнал $f_0(s) = f(sT)$. Тогда точкам $t_n = nT$ оригинального масштаба времени будут соответствовать целые числа $s_n = n$ на оси нормированного времени и $f_0(n) = f(nT)$. То есть дискретный сигнал $f_0(n)$ представляет собой отсчеты аналогового сигнала $f_0(s)$, взятые с шагом дискретизации $\Delta s = 1$ (что соответствует и единичной частоте дискретизации $1/\Delta s = 1$). Оригинальный масштаб оси времени t всегда можно получить по реальному значению периода дискретизации $T = \Delta t$ в результате денормировки $t = sT$.

Нормирование оси времени влечет изменения и в частотной области, которые определяются свойством 3° преобразования Фурье (см. раздел 1.5). По спектру (2.9)

$$\tilde{S}_0(\eta) = \sum_{n \in \mathbb{Z}} f_0(n) e^{-2\pi i \eta n}$$

нормированного сигнала $f_0(n)$ спектр (2.9) оригинального дискретного сигнала $f(nT)$ всегда можно выразить в результате денормировки в частотной области:

$$\tilde{S}(v) = \sum_{n \in \mathbb{Z}} \underbrace{f(nT)}_{f_0(n)} e^{-2\pi i v n T} T = T \tilde{S}_0(vT).$$

Таким образом, нормировка оси времени привела и к определенной нормировке оси частот: спектр $\tilde{S}_0(\eta) = \tilde{S}_0(\eta + 1)$ имеет единичный период, а нормированные частоты η связаны с реальными v как $\eta = vT = v/f_d$. Например, частоте дискретизации $v = f_d = 1/T$ соответствует нормированная частота $\eta = 1$, а для частоты Найквиста имеем: $v = f_d/2 \leftrightarrow \eta = 1/2$.

Далее, где это возможно, мы обычно будем рассматривать дискретные сигналы на нормированных осях времени и частот. Получить, при необходимости, оригинальный масштаб представления сигнала (или спектра) всегда можно денормированием по рассмотренным выше правилам. Кроме того, в данной главе мы будем подразумевать, что рассматриваемые дискретные сигналы (последовательности) $x(n)$ таковы, что $x(n) = 0$ при $n < 0$.

3.2. Z-преобразование

Определение. Пусть $\{x(n)\}_{n=-\infty}^{\infty}$ — последовательность комплексных чисел, причем $x(n) = 0$ при $n < 0$, и существует верхний предел $\lim_{n \rightarrow \infty} \sqrt[n]{|x(n)|} = r < \infty$. Z-преобразованием данной последовательности назовем следующую функцию $X(z)$ комплексной переменной z :

$$X(z) = \begin{cases} \varphi(z) = \sum_{n=0}^{\infty} x(n)z^{-n} \text{ для } |z| > r \\ \text{аналитическое продолжение функции } \varphi(z) \text{ для } |z| \leq r \end{cases}. \quad (3.1)$$

Будем использовать обозначение $X(z) = Z\{x(n)\}$.

Поясним определение. Функция $\varphi(z)$, фигурирующая в (3.1), представляет собой сумму ряда Лорана специального вида: $\sum_{n=0}^{\infty} x(n)z^{-n}$. В общем случае ряд Лорана имеет следующий вид:

$$\sum_{n=-\infty}^{\infty} c_n (z - z_0)^n, \quad (3.2)$$

где z_0 — точка комплексной плоскости, c_n — комплексные числа. Напомним следующую теорему [17].

Теорема 3.1. Если коэффициенты ряда Лорана (3.2) таковы, что $\lim_{n \rightarrow \infty} \sqrt[n]{|c_{-n}|} = r < R = 1 / \lim_{n \rightarrow \infty} \sqrt[n]{|c_n|}$, то ряд (3.2) 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 функцией¹, а коэффициенты ряда выражаются с помощью контурных интегралов:

$$c_n = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz, \quad (3.3)$$

$n \in \mathbb{Z}$, 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 3.1 ряд $\sum_{n=0}^{\infty} x(n)z^{-n}$ в (3.1) абсолютно сходится в области $r < |z|$ ($R = \infty$), где его сумма $\varphi(z)$ 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 3.1. Найти Z-преобразование последовательности $x(n) = n$, $n = 0, 1, \dots$

◀ Для $|z| > \lim_{n \rightarrow \infty} \sqrt[n]{|n|} = 1$ имеем:

$$X(z) = \sum_{n=0}^{\infty} n z^{-n} = -z \sum_{n=0}^{\infty} \frac{d}{dz} z^{-n} = -z \frac{d}{dz} \left(\sum_{n=0}^{\infty} z^{-n} \right) = -z \frac{d}{dz} \left(\frac{z}{z-1} \right) = \frac{z}{(z-1)^2}.$$

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. $\forall z \in \mathbb{C}: Z\{x(n)\} = X(z) = z/(z-1)^2$. ▶

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: $u(n) = x(n)*y(n)$.

¹ Функцию $f(z)$ называем *аналитической* в некоторой области D , если она комплексно дифференцируема в этой области.

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:

$$u(n) = \sum_{k=-\infty}^{\infty} x(k)y(n-k) = \sum_{k=0}^n x(k)y(n-k) = \sum_{k=0}^n y(k)x(n-k) = \sum_{k=-\infty}^{\infty} y(k)x(n-k).$$

Соответствие между последовательностью $\{x(n)\}_{n=0}^{\infty}$ и ее Z -образом $X(z) = Z\{x(n)\}$ обозначим так: $x(n) \leftrightarrow X(z)$. Укажем на ряд свойств 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:

$$\sum_{n=0}^{\infty} x(n-N)z^{-n} = \sum_{k=-N}^{\infty} x(k)z^{-k-N} = z^{-N} \sum_{k=0}^{\infty} x(k)z^{-k} = z^{-N} X(z). \quad \blacktriangleright$$

3°. Опережающий сдвиг последовательности на $M \geq 0$ отсчетов:

$$x(n+M) \leftrightarrow z^M (X(z) - x(0) - x(1)z^{-1} - \dots - x(M-1)z^{-(M-1)}).$$

$$\blacktriangleleft \sum_{n=0}^{\infty} x(n+M)z^{-n} = \sum_{k=M}^{\infty} x(k)z^{-k+M} = z^M \left(\sum_{k=0}^{\infty} x(k)z^{-k} - \sum_{k=0}^{M-1} x(k)z^{-k} \right). \quad \blacktriangleright$$

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 3.1.

$$\sum_{n=0}^{\infty} nx(n)z^{-n} = -z \sum_{n=0}^{\infty} x(n) \frac{d}{dz} z^{-n} = -z \frac{d}{dz} \left(\sum_{n=0}^{\infty} x(n)z^{-n} \right) = -z \frac{d}{dz} X(z). \quad \blacktriangleright$$

6°. Z -преобразование свертки: $x(n)*y(n) \leftrightarrow X(z)Y(z)$.

◀ Используя доказанное выше свойство 2°, получаем:

$$\begin{aligned} Z\{x(n)*y(n)\} &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^n x(k)y(n-k) \right) z^{-n} = \sum_{k=0}^{\infty} x(k) \sum_{n=0}^{\infty} y(n-k)z^{-n} = \\ &= \sum_{k=0}^{\infty} x(k)z^{-k} Y(z) = X(z)Y(z). \quad \blacktriangleright \end{aligned}$$

Упражнение. Убедитесь в справедливости приведенных в табл. 3.1 выражений для Z -преобразований ряда последовательностей $x(n)$.

Таблица 3.1. Z-преобразования некоторых последовательностей*

| Последовательность $x(n)$ | Z-преобразование $X(z)$ |
|---|--|
| $\tilde{\delta}(n-n_0) = \begin{cases} 1, & n = n_0 \\ 0, & n \neq n_0 \end{cases}$ | z^{-n_0} |
| a^n | $z/(z-a)$ |
| n | $z/(z-1)^2$ |
| n^2 | $z(z+1)/(z-1)^3$ |
| $\cos(\beta n)$ | $\frac{z(z-\cos\beta)}{z^2-2z\cos\beta+1}$ |
| $\sin(\beta n)$ | $\frac{z\sin\beta}{z^2-2z\cos\beta+1}$ |

* Везде считается, что $x(n) = 0$ при $n < 0$.

Из теоремы 3.1 и формулы (3.3) непосредственно следует, что элементы последовательности $x(n)$ могут быть восстановлены по ее Z-образу $X(z)$ при помощи формулы:

$$x(n) = \frac{1}{2\pi i} \oint_{\gamma} X(z) z^{n-1} dz, \quad (3.4)$$

где γ — произвольный замкнутый контур, лежащий в области аналитичности функции $X(z)$ и охватывающий все ее особые точки.

Напомним следующую теорему о вычетах, которую удобно использовать для вычисления интегралов (3.4) при обращении Z-преобразования.

Теорема 3.2. 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 особые точки $\{z_k\}_{k=1}^N$, то

$$\frac{1}{2\pi i} \oint_{\gamma} f(z) dz = \sum_{k=1}^N \text{Res}[f(z), z_k], \quad (3.5)$$

где $\text{Res}[f(z), z_k]$ — это вычет функции $f(z)$ в точке z_k .

Таким образом (см. (3.4)), для обращения Z -преобразования в формулу (3.5) 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 (3.2), $f(z) = \sum_{k=-\infty}^{\infty} c_k (z - z_0)^k$, то (см. (3.3)) 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 (3.2) $c_{-m} \neq 0$, а $\forall j > m: c_{-j} = 0$), 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 $u(z_0) \neq 0$, $v(z_0) = 0$, $v'(z_0) \neq 0$, то $c_{-1} = \text{Res}[f(z), z_0] = u(z_0)/v'(z_0)$.

Пример 3.2. 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 (3.4) и (3.5), получаем:

$$\begin{aligned} x(n) &= \frac{1}{2\pi i} \oint_{\gamma} \frac{z^n}{z^2 + 1} dz = \text{Res}\left[\frac{z^n}{z^2 + 1}, z = i\right] + \text{Res}\left[\frac{z^n}{z^2 + 1}, z = -i\right] = \\ &= \lim_{z \rightarrow i} \frac{z^n (z - i)}{(z + i)(z - i)} + \lim_{z \rightarrow -i} \frac{z^n (z + i)}{(z + i)(z - i)} = \frac{i^n - (-i)^n}{2i} = \frac{e^{i\frac{\pi n}{2}} - e^{-i\frac{\pi n}{2}}}{2i} = \sin \frac{\pi n}{2}. \quad \blacktriangleright \end{aligned}$$

Упражнение. Найдите последовательность $x(n)$ по ее Z -образу $X(z) = z/(z^2 - 9)^2$.

Математический аппарат Z -преобразования часто оказывается удобным для решения разностных уравнений.

Определение. Линейным разностным уравнением N -го порядка с постоянными коэффициентами называется уравнение вида

$$\sum_{k=0}^N a_k x(n+k) = f(n), \quad n = 0, 1, \dots, \quad (3.6)$$

где $\{a_k\}_{k=0}^N$, $\{f(n)\}_{n=0}^{\infty}$ — заданные числа (вообще говоря, комплексные), $\{x(n)\}_{n=0}^{\infty}$ — искомая последовательность (решение уравнения).

Разностные уравнения являются дискретными аналогами дифференциальных уравнений. Для того чтобы уравнение (3.6) имело единственное решение, необходимо задать, например, начальные условия, т. е. значения N первых элементов искомой последовательности: $x(0) = x_0, \dots, x(N-1) = x_{N-1}$. Не ограничивая общности, можно положить $a_N = 1$. Тогда оставшиеся члены последовательности можно выразить через найденные ранее, записав (3.6) 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 (3.6) можно с использованием свойства 3° Z-преобразования. Покажем это на примере.

Пример 3.3. Решить разностное уравнение $x(j+2) - 2x(j+1) + x(j) = -4$ с условиями: $x(0) = 0, x(1) = 2$.

◀ Так как (см. свойство 3° Z-преобразования)

$$Z\{x(j+2)\} = z^2 \left(X(z) - x(0) - \frac{x(1)}{z} \right) = z^2 \left(X(z) - \frac{2}{z} \right),$$

$$Z\{x(j+1)\} = z(X(z) - x(0)) = zX(z), \quad Z\{-4\} = \sum_{n=0}^{\infty} (-4)z^{-n} = \frac{-4z}{z-1},$$

то для исходного разностного уравнения получаем следующий Z-образ:

$$z^2 \left(X(z) - \frac{2}{z} \right) - 2zX(z) + X(z) = -\frac{4z}{z-1}, \quad \text{откуда} \quad X(z) = \frac{2z^2 - 6z}{(z-1)^3}.$$

Обращая Z-преобразование при помощи формул (3.4), (3.5), находим:

$$x(n) = \text{Res} \left[\frac{2z^2 - 6z}{(z-1)^3} z^{n-1}, z=1 \right] = \frac{1}{2} \lim_{z \rightarrow 1} \frac{d^2((2z-6)z^n)}{dz^2} = 2n(2-n). \quad \blacktriangleright$$

Упражнение. Решите уравнение $x(n+2) - 5x(n+1) + 6x(n) = 0$ с условиями: $x(0) = 1, x(1) = 2$.

3.3. Линейные дискретные фильтры

Определение. *Линейным дискретным фильтром (или линейной дискретной системой) L назовем устройство, которое преобразует входную последовательность $x(n)$ в выходную $y(n) = L\{x(n)\}$ по правилу, определяемому для любого целого значения n следующим разностным уравнением:*

$$\sum_{m=0}^M a_m y(n-m) = \sum_{k=0}^N b_k x(n-k), \quad (3.7)$$

где a_m, b_k — некоторые постоянные коэффициенты. Вновь полагаем, что $x(n) = y(n) = 0$ при $n < 0$. Число $K = \max(M, N)$ назовем *порядком системы (фильтра)*.

Таким образом, параметрами системы являются наборы коэффициентов $\{a_m\}_{m=0}^M, \{b_k\}_{k=0}^N$, причем мы будем рассматривать только системы с вещественными коэффициентами. Постоянство параметров характеризует свойство *стационарности* системы. Линейность системы, алгоритм работы которой определяется уравнением (3.7), очевидна. Действительно, если пары последовательностей $\{x_1(n), y_1(n)\}$ и $\{x_2(n), y_2(n)\}$ удовлетворяют уравнению (3.7), 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 (3.7) делим на a_0), уравнение (3.7) можно переписать в другом виде:

$$y(n) = -\sum_{m=1}^M a_m y(n-m) + \sum_{k=0}^N b_k x(n-k), \quad (3.8)$$

что задает в виде рекуррентной формулы некоторую расчетную процедуру нахождения очередного элемента выходной последовательности по уже известным элементам входной и выходной последовательностей. Если в уравнении (3.8) есть хотя бы один коэффициент $a_m \neq 0$, то фильтр называется *рекурсивным*, в противном случае — *нерекурсивным*.

Уравнение (3.8) в явном виде определяет алгоритм формирования отклика ЛДФ $y(n)$ как вычисление взвешенной суммы,

в которую входят текущее $x(n)$ и N предшествующих $x(n-k)$ значений входного сигнала, а также M предыдущих значений выходного сигнала $y(n-m)$. Линейные дискретные системы, у которых значение выходного сигнала $y(n)$ в момент дискретного времени n не зависит от еще не поступивших на вход системы отсчетов $x(n+1)$, $x(n+2)$, ..., называются *каузальными* (или физически реализуемыми). Рассматриваемые в данной главе системы, описываемые уравнением типа (3.8), таковыми являются.

Пример 3.4. 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 $n < 0$.

◀ $y(0) = 1$, $y(1) = ay(0) + x(1) = a$, $y(2) = ay(1) + x(2) = a^2$, ..., $y(n) = a^n$, ... ▶

Пример 3.5. Найти реакцию нерекурсивного фильтра, описываемого разностным уравнением $y(n) = x(n) + bx(n-1)$, на входное воздействие $x(n) = \tilde{\delta}(n)$. (Вновь полагаем $y(n) = 0$ для $n < 0$.)

◀ $y(0) = 1$, $y(1) = b$, $\forall n > 1$: $y(n) = 0$. ▶

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 обычно используется обозначение $h(n)$, т. е. $h(n) = y(n)|_{x(n)=\tilde{\delta}(n)}$.

В примерах 3.4 и 3.5 как раз и были найдены импульсные характеристики некоторых фильтров. В примере 3.4 получена *бесконечная ИХ* ($\forall M \exists m > M: h(m) \neq 0$), а в примере 3.5 рассмотрен фильтр с *конечной ИХ* ($\exists M \forall m > M: h(m) = 0$). Фильтры с конечной ИХ называются 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 в виде

$$x(n) = \sum_{m=0}^{\infty} x(m) \underbrace{\tilde{\delta}(n-m)}_{0 \text{ при } n \neq m}.$$

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 $x(n) = \sum_{m=0}^{\infty} x(m)\delta(n-m)$ получаем равной $y(n) = \sum_{m=0}^{\infty} x(m)h(n-m)$. Таким образом, по известной ИХ отклик фильтра находится с помощью дискретной свертки:

$$y(n) = \sum_{m=0}^{\infty} x(m)h(n-m) = \sum_{m=0}^n x(m)h(n-m) = \sum_{m=0}^n h(m)x(n-m). \quad (3.9)$$

Пример 3.6. По заданной ИХ $h(n)$ найти реакцию фильтра на входное воздействие $x(n) = 1$.

◀ По формуле (3.9) получаем $y(n) = \sum_{m=0}^n h(m)$ — данный выходной сигнал (отклик на $x(n) = 1$) называют также *переходной характеристикой*. ▶

Другой важнейшей характеристикой фильтра является передаточная функция.

Определение. *Передаточной функцией (ПФ)* линейной дискретной системы называется отношение Z -образов выходной и входной последовательностей: $H(z) = Y(z)/X(z)$.

Пример 3.7. 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 (3.8): $Y(z) = -\sum_{m=1}^M a_m Y(z)z^{-m} + \sum_{k=0}^N b_k X(z)z^{-k}$. Отсюда

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^N b_k z^{-k}}{1 + \sum_{m=1}^M a_m z^{-m}}. \quad (3.10)$$

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 (3.8), можно сразу записать ПФ в виде (3.10). По известной ПФ отклик фильтра на произвольное входное воздействие можно определить,

найдя сначала Z -образ входной последовательности $X(z) = Z\{x(n)\}$. Тогда $Y(z) = H(z)X(z)$ и $y(n) = Z^{-1}\{Y(z)\}$.

Пример 3.8. Найти ПФ для фильтра из примера 3.4.

$$\blacktriangleleft H(z) = \frac{1}{1+(-az^{-1})} = \frac{z}{z-a} \blacktriangleright$$

Между передаточной функцией и импульсной характеристикой существует очевидная связь. Так как ПФ не зависит от входного воздействия, возьмем $x(n) = \tilde{\delta}(n)$: $H(z) = Z\{h(n)\}/Z\{\tilde{\delta}(n)\}$. Поскольку $Z\{\tilde{\delta}(n)\} = 1$, то $H(z) = Z\{h(n)\}$ – ПФ представляет собой Z -преобразование ИХ.

Пример 3.9. Для ЛДФ, описываемого уравнением

$$y(n) = -2y(n-1) - y(n-2) + x(n) + x(n-1),$$

найти:

- передаточную функцию фильтра,
- импульсную характеристику фильтра,
- отклик фильтра на входное воздействие $x(n) = 3^n$ с использованием уравнения свертки (3.9), а также посредством обращения Z -образа отклика фильтра.

◀ а) Сравнивая заданное разностное уравнение с уравнением в общем виде (3.8), находим ненулевые коэффициенты ЛДФ: $a_1 = 2$, $a_2 = 1$, $b_0 = 1$, $b_1 = 1$. Поэтому с учетом (3.10)

$$H(z) = \frac{1+z^{-1}}{1+2z^{-1}+z^{-2}} = \frac{z}{z+1}.$$

$$\text{б) } h(n) = Z^{-1}\{H(z)\} = \frac{1}{2\pi i} \oint_{\gamma} H(z)z^{n-1} dz = \text{Res} \left[\frac{z^n}{z+1}, z = -1 \right] = (-1)^n.$$

в) Используя формулу свертки (3.9), 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

$$y(n) = Z^{-1} \left\{ \frac{z^2}{(z+1)(z-3)} \right\} = \text{Res} \left[\frac{z^{n+1}}{(z+1)(z-3)}, z = -1 \right] + \\ + \text{Res} \left[\frac{z^{n+1}}{(z+1)(z-3)}, z = 3 \right] = \frac{(-1)^{n+1}}{-4} + \frac{3^{n+1}}{4} = \frac{(-1)^n + 3^{n+1}}{4}. \quad \blacktriangleright$$

Упражнение. Найдите отклик фильтра, заданного передаточной функцией $H(z) = \frac{1}{4(z-1)(z-0,5)}$, на входное воздействие $x(n) = 0,5^n$.

Из формулы (3.10) следует важный для последующего рассмотрения факт: ПФ ЛДФ может быть представлена в виде отношения некоторых алгебраических многочленов неотрицательной степени: $H(z) = P(z)/Q(z)$ (для этого необходимо умножить числитель и знаменатель ПФ (3.10) на величину $z^{\max(M,N)}$). Поэтому передаточные функции ЛДФ, определяемых уравнением (3.7) или (3.8), имеют конечное число нулей — точек комплексной плоскости, в которых $P(z) = 0$, и полюсов — точек, где $Q(z) = 0$.

Если известны нули и полюсы передаточной функции $H(z)$, то ее можно представить в виде $H(z) = K \prod_{j=1}^J (z - z_{z_j}) / \prod_{l=1}^L (z - z_{p_l})$, где — нули, $\{z_{p_l}\}_{l=1}^L$ — полюсы ПФ, а K — некоторый масштабный (вещественный) коэффициент. Поскольку коэффициенты ПФ (3.10) — вещественные, то комплекснозначные полюсы (или нули) ПФ могут встречаться только комплексно сопряженными парами.

Таким образом, задав нули и полюсы передаточной функции, ее можно определить с точностью до некоторого вещественного множителя-константы K .

3.4. Соединения и структурные схемы фильтров

Найдем выражения для передаточных функций систем, полученных в результате различных способов соединения двух фильтров.

1. Последовательное соединение фильтров (рис. 3.1). Очевидно, что $Y(z) = H_2(z)V(z) = H_2(z)H_1(z)X(z)$, т.е. передаточная функция $H(z) = H_2(z)H_1(z)$.

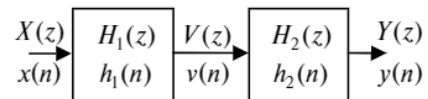


Рис. 3.1. Последовательное соединение фильтров

2. Параллельное соединение фильтров (рис. 3.2). Очевидно,

$$Y(z) = H_1(z)X(z) + H_2(z)X(z).$$

Отсюда передаточная функция системы: $H(z) = H_1(z) + H_2(z)$.

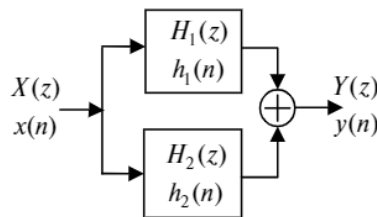


Рис. 3.2. Параллельное соединение фильтров

3. Соединение с обратной связью (рис. 3.3). Выполняя несложные преобразования, получаем:

$$Y(z) = H_1(z)(X(z) + U(z)) = H_1(z)(X(z) + H_2(z)Y(z)).$$

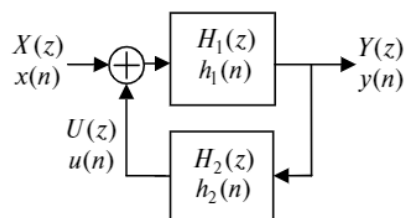


Рис. 3.3. Соединение с обратной связью

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 (3.8) и (3.10).

Схемы ЛДФ будем строить из элементов, выполняющих задержку последовательности на один отсчет, т.е. работающих по правилу $y(n) = x(n-1)$ и имеющих импульсную характеристику $h(n) = \delta(n-1)$ и передаточную функцию $H(z) = z^{-1}$, а также из элементов-умножителей: $y(n) = cx(n)$, для которых $h(n) = c\delta(n)$ и $H(z) = c$ (см. рис. 3.4). Элементы-сумматоры будем обозначать символом «+» или « Σ ».

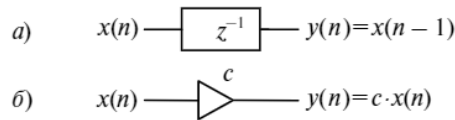


Рис. 3.4. Элементы структурных схем фильтров: а) элемент задержки, б) умножитель

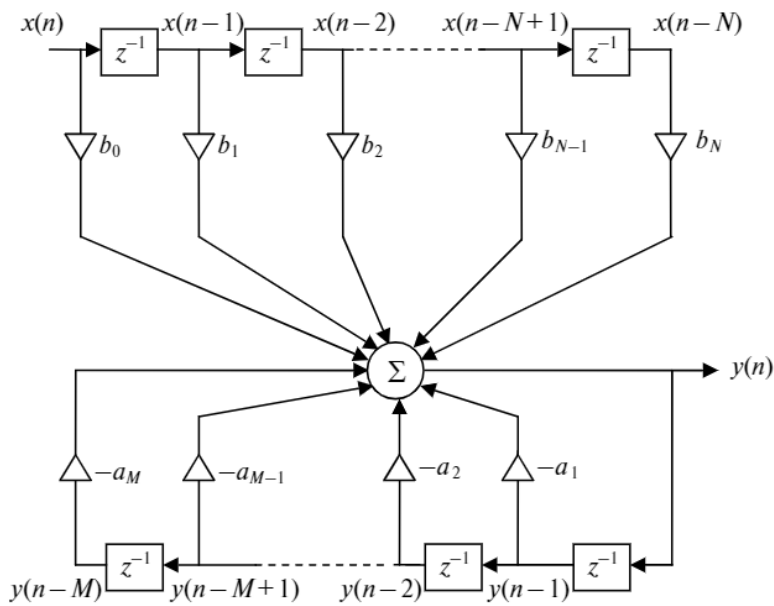


Рис. 3.5. Прямая форма структурной схемы ЛДФ

Прямая форма структурной схемы ЛДФ реализуется непосредственно по формуле (3.8), 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 соответственно, см. рис. 3.5.

Пример 3.10. Найти передаточную функцию для фильтра, имеющего следующую структурную схему (рис. 3.6).

$$\leftarrow H(z) = \frac{1 + 3z^{-1} + 2z^{-2}}{1 + 0,4z^{-1}} = \frac{z^2 + 3z + 2}{z^2 + 0,4z} \rightarrow$$

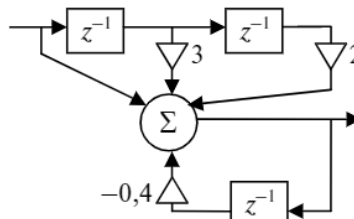


Рис. 3.6. К условиям примера 3.10

Прямая каноническая форма — это форма структурной схемы, содержащая минимальное количество элементов задержки. Для ее получения представим передаточную функцию (3.10) как результат последовательного соединения фильтров, соответствующего схеме рис. 3.7:

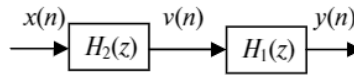


Рис. 3.7. К получению прямой канонической формы

$$H(z) = \frac{1}{1 + \underbrace{\sum_{m=1}^M a_m z^{-m}}_{H_2(z)}} \cdot \underbrace{\sum_{k=0}^N b_k z^{-k}}_{H_1(z)}$$

Обозначим $V(z) = H_2(z)X(z)$, $Y(z) = H_1(z)V(z)$, что соответствует структурной схеме последовательного соединения. 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, изображенная на рис. 3.8, где $L = \max(N, M)$.

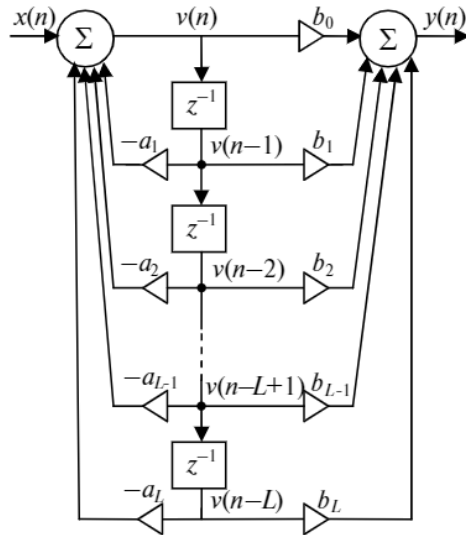


Рис. 3.8. Прямая каноническая форма структурной схемы ЛДФ

Пример 3.11. Построить прямую каноническую форму структурной схемы для фильтра из примера 3.10.

◀ Прodelайте самостоятельно, результат приведен на рис. 3.9. ▶

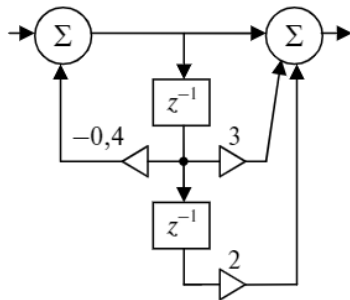


Рис. 3.9. К решению примера 3.11

Если в разностном уравнении (3.8) отсутствует зависимость отсчета выходной последовательности $y(n)$ от ее предыдущих отсчетов $y(n-k)$ ($k \geq 1$), т.е. $y(n) = \sum_{k=0}^N b_k x(n-k)$, $H(z) = \sum_{k=0}^N b_k z^{-k}$, 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 (3.8) 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 3.12. Найти передаточную функцию, импульсную характеристику и упростить структурную схему фильтра, изображенную на рис. 3.10.

◀ Убедитесь, что $H(z) = 1$, $h(n) = \delta(n)$. ▶

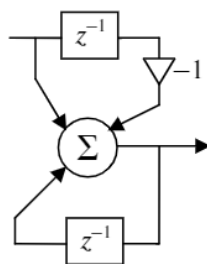


Рис. 3.10. К условиям примера 3.12

Пример 3.12 носит искусственный характер; при рассмотрении рекурсивных фильтров считается, что они имеют бесконечную ИХ.

Синтезировать КИХ-фильтр по заданной ИХ $h(n) = \{h(0), h(1), \dots, h(N), 0, 0, \dots\}$ очень просто: для этого нужно 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 количеством элементов задержки и умножителей. Подробнее эти вопросы будут рассмотрены в четвертой главе.

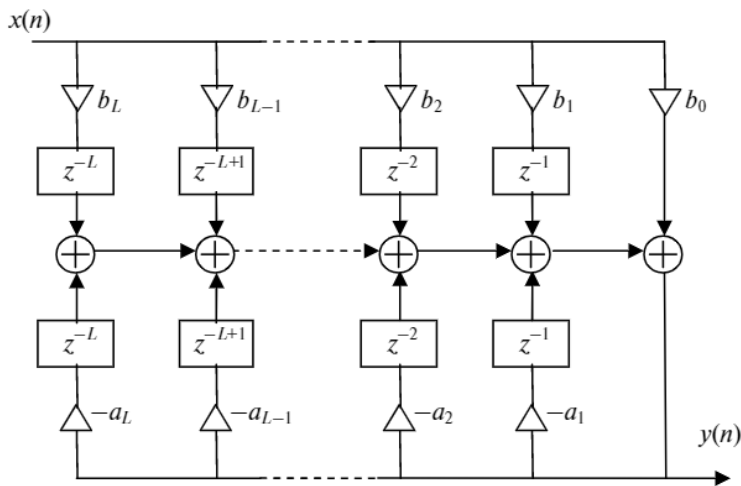


Рис. 3.11. Промежуточная структура для получения транспонированной формы

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. 3.11, где $L = \max(M, N)$. Эта структурная схема, очевидно, непосредственно реализует уравнение (3.8) и может быть представлена как преобразование прямой формы, при котором меняются местами операции умножения и задержки, а общий сумматор заменяется цепочкой сумматоров. В полученной схеме рис. 3.11 можно поменять местами операции суммирования и задержки, в итоге получим транспонированную форму фильтра, изображенную на рис. 3.12.

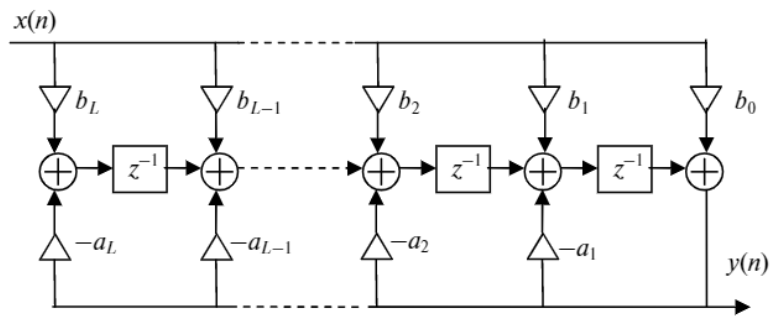


Рис. 3.12. Транспонированная форма структурной схемы ЛДФ

Упражнение. Постройте самостоятельно транспонированную форму для фильтра из примера 3.10.

Упражнение. Постройте прямую, прямую каноническую, транспонированную формы фильтра, заданного импульсной характеристикой $h(n) = n2^{-n}$. *Указание:* сначала найдите ПФ (3.10).

Если линейная дискретная система имеет высокий порядок, то помимо приведенных выше структурных схем фильтров часто оказывается удобным рассматривать и использовать при синтезе дискретных фильтров структуры, полученные в результате последовательного (каскадного) или параллельного соединения более простых структур — звеньев первого или второго порядка.

Для построения каскадной структуры передаточную функцию системы сначала нужно представить в виде разложения

на множители $H(z) = H_0(z) \prod_{k=1}^K H_k(z)$, где $H_k(z)$ представляет собой для $k \geq 1$ передаточную функцию рекурсивного звена первого или второго порядка:

$$H_k(z) = \frac{b_{k,0} + b_{k,1}z^{-1}}{1 + a_{k,1}z^{-1}} \quad \text{— первый порядок,}$$

$$H_k(z) = \frac{b_{k,0} + b_{k,1}z^{-1} + b_{k,2}z^{-2}}{1 + a_{k,1}z^{-1} + a_{k,2}z^{-2}} \quad \text{— второй порядок.}$$

Число рекурсивных звеньев K определяется порядком M многочлена, стоящего в знаменателе передаточной функции (3.10) как $K = [(M+1)/2]$ (квадратными скобками обозначена операция взятия целой части). Если степень N полинома в числителе ПФ (3.10) $N > M$, то в разложении ПФ на множители присутствует также нерекурсивное звено с ПФ вида $H_0(z) = \sum_{m=0}^{N-M} c_m z^{-m}$ (когда нерекурсивного звена нет, полагаем $H_0(z) = 1$). Общая каскадная структура системы представляется в виде последовательного соединения звеньев с передаточными функциями $H_k(z)$, $k = 0, \dots, K$ (см. рис. 3.13).

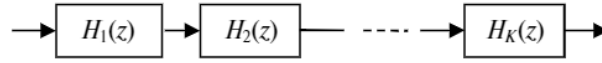


Рис. 3.13. Общий вид структуры каскадной реализации системы (нерекурсивное звено отсутствует)

Для реализации системы в виде параллельной структуры передаточную функцию $H(z)$ (3.10) нужно представить в виде разложения

$$H(z) = \sum_{m=0}^{N_p} c_m z^{-m} + \sum_{k=1}^K H_k(z), \quad (3.11)$$

где каждое слагаемое $H_k(z)$ вновь представляет собой звено первого или второго порядка, но на этот раз частного вида:

$$H_k(z) = \frac{b_{k,0}}{1 + a_{k,1}z^{-1}} \quad \text{— первый порядок,}$$

$$H_k(z) = \frac{b_{k,0} + b_{k,1}z^{-1}}{1 + a_{k,1}z^{-1} + a_{k,2}z^{-2}} \quad \text{— второй порядок.}$$

Как и прежде, $K = [(M+1)/2]$, а значение N_p находится для $N \geq M$ по функции (3.10) как $N_p = N - M$ (если $N < M$, то первая сумма

$H_0(z) = \sum_{m=0}^{N_p} c_m z^{-m}$ в выражении (3.11), определяющая нерекурсивное звено, отсутствует). Соответствующая формуле (3.11) параллельная структура реализации системы отражена на рис. 3.14.

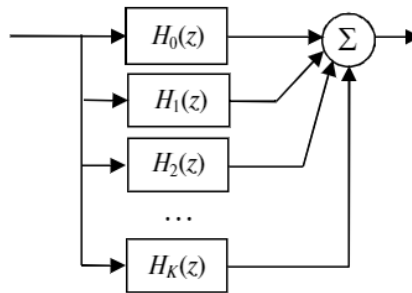


Рис. 3.14. Общий вид структуры параллельной реализации системы (нерекурсивное звено присутствует)

Пример 3.13. Дискретный фильтр реализован в виде последовательного соединения двух звеньев с передаточными функциями

$$H_1(z) = \frac{1 - z^{-1}}{1 + 0,5z^{-1}} \quad \text{и} \quad H_2(z) = \frac{z^{-1} + 2z^{-2}}{1 + z^{-1} + z^{-2}}.$$

Реализовать данный фильтр в виде параллельной структуры.

◀ Найдем передаточную функцию всей системы:

$$H(z) = H_1(z)H_2(z) = \frac{z^{-1} + z^{-2} - 2z^{-3}}{(1 + 0,5z^{-1})(1 + z^{-1} + z^{-2})} = \frac{z^{-1} + z^{-2} - 2z^{-3}}{1 + 1,5z^{-1} + 1,5z^{-2} + 0,5z^{-3}}.$$

Имеем $N = M = 3$, поэтому $K = [(M + 1)/2] = 2$, а передаточная функция нерекурсивного звена имеет нулевой порядок и соответствует множителю, т. е. $H_0(z) = c_0$. Представим передаточную функцию как структуру рис. 3.14 с множителем $c_0 = b_3/a_3 = -2/0,5 = -4$ и двумя параллельными звеньями, тогда имеем для $H(z)$:

$$\frac{z^{-1} + z^{-2} - 2z^{-3}}{1 + 1,5z^{-1} + 1,5z^{-2} + 0,5z^{-3}} = -4 + \frac{b_{1,0}}{1 + 0,5z^{-1}} + \frac{b_{2,0} + b_{2,1}z^{-1}}{1 + z^{-1} + z^{-2}}.$$

Приводя выражение в правой части к общему знаменателю и приравнявая коэффициенты при одинаковых степенях в числителях дробей полученного равенства, находим: $b_{1,0} = 6$, $b_{2,0} = -2$, $b_{2,1} = 2$.

Таким образом, параллельная структура, определяемая рис. 3.14, состоит из множителя на коэффициент $c_0 = -4$, звена

первого порядка и звена второго порядка с передаточными функциями соответственно

$$H_1(z) = \frac{6}{1 + 0,5z^{-1}} \quad \text{и} \quad H_2(z) = \frac{-2 + 2z^{-1}}{1 + z^{-1} + z^{-2}}. \quad \blacktriangleright$$

3.5. Устойчивость ЛДФ

Определение. ЛДФ называется *устойчивым*, если для любой ограниченной входной последовательности $x(n)$ ($\forall n: |x(n)| \leq C_1 < \infty$) выходная последовательность также является ограниченной, т.е. $\forall n: |y(n)| \leq C_2 < \infty$.

В *неустойчивой* системе при некоторых ограниченных входных сигналах выходной сигнал неограниченно возрастает.

Теорема 3.3. Необходимым и достаточным условием устойчивости ЛДФ является сходимость ряда

$$\sum_{n=0}^{\infty} |h(n)| \leq C < \infty. \quad (3.12)$$

◀ Докажем сначала достаточность условия (3.12) для устойчивости ЛДФ. Если выполняется (3.12), то для любого входного воздействия, такого, что $|x(n)| \leq C_1 < \infty$, получаем отклик ЛДФ $\forall n$:

$$|y(n)| = \left| \sum_{k=0}^n x(n-k)h(k) \right| \leq \sum_{k=0}^n |x(n-k)||h(k)| \leq C_1 \sum_{k=0}^n |h(k)| \leq C_1 C < \infty.$$

То есть выходная последовательность ограничена, фильтр устойчив.

Докажем необходимость условия (3.12) для устойчивости ЛДФ. Допустим противное: фильтр устойчив, но (3.12) не выполняется, последовательность частичных сумм $S_j = \sum_{n=0}^j |h(n)|$ неограниченно возрастает с увеличением индекса j , т.е. $\forall C > 0 \exists J = J(C): \sum_{n=0}^J |h(n)| > C$. Выбрав такое целое число J по произвольному сколь 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

$$y(J) = \sum_{n=0}^J x(n)h(J-n) = \sum_{n=0}^J |h(J-n)| = \sum_{n=0}^J |h(n)| > C.$$

Таким образом, для сколь 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 (3.12) может не выполняться, неверно. ▶

Следствие 1. Необходимое условие устойчивости ЛДФ: $\lim_{n \rightarrow \infty} h(n) = 0$.

Следствие 2. КИХ-фильтры всегда устойчивы.

Пример 3.14. При каких значениях параметра a фильтр из примера 3.4 является устойчивым?

◀ При решении примера 3.4 было установлено, что для данного фильтра ИХ $h(n) = a^n$. Условие (3.12) принимает вид: $\sum_{n=0}^{\infty} |a|^n < \infty$. Отсюда $|a| < 1$. ▶

Вместо проверки условия (3.12) часто оказывается удобнее использовать другой критерий устойчивости.

Теорема 3.4. 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 на основании теоремы 3.3 ограничена:

$$|H(z)| = \left| \sum_{n=0}^{\infty} h(n)z^{-n} \right| \leq \sum_{n=0}^{\infty} |h(n)| |z^{-n}| \leq \sum_{n=0}^{\infty} |h(n)| \leq C < \infty.$$

Следовательно, в области $|z| \geq 1$ у передаточной функции нет полюсов.

Достаточность. В соответствии с (3.4) 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 алгебраических 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 $\gamma = \{z \mid |z| = r\}$, где $r = \frac{1+r_1}{2} < 1$. На этой окружности $Q(z) \neq 0$, ПФ 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 $z = re^{i\varphi}$, для ИХ получаем следующую оценку:

$$|h(n)| = \left| \frac{1}{2\pi i} \oint_{|z|=r} H(z)z^{n-1} dz \right| = \left| \frac{1}{2\pi i} \int_0^{2\pi} H(re^{i\varphi}) ir^n e^{i\varphi n} d\varphi \right| \leq \frac{r^n}{2\pi} \int_0^{2\pi} M |e^{i\varphi n}| d\varphi = Mr^n.$$

Поскольку $r < 1$, получаем для условия (3.12):

$$\sum_{n=0}^{\infty} |h(n)| \leq M \sum_{n=0}^{\infty} r^n = \frac{M}{1-r} < \infty,$$

т.е. ЛДФ устойчив. ►

Пример 3.15. Определить по передаточной функции, при каких значениях параметра a фильтр из примера 3.4 является устойчивым.

◀ Ранее найдено, что $H(z) = z/(z-a)$ (см. пример 3.8). Единственная особая точка передаточной функции — полюс первого порядка $z = a$. Требование теоремы 3.4 приводит к уже установленному ранее в примере 3.14 условию устойчивости фильтра: $|a| < 1$. ►

3.6. 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 $z = e^{i\omega}$.

Так как передаточная функция есть Z -преобразование ИХ, то ЧХ 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 на основании теоремы 3.3 $\sum_{k=0}^{\infty} |h(k)e^{-i\omega k}| = \sum_{k=0}^{\infty} |h(k)| \leq C < \infty$. Как несложно видеть, ЧХ $K(\omega) = H(e^{i\omega})$ имеет период 2π , причем $K(-\omega) = \overline{K(\omega)}$ (в силу вещественного характера рассматриваемых ИХ $h(n)$). Исходя из данных свойств ЧХ, для ее описания на всей оси частот достаточно задать $K(\omega)$ только для значений нормированной циклической частоты

$\omega \in [0; \pi]$ — эта частотная область называется *основной полосой частот*.

Подстановка переменной $z = e^{i\omega}$ в выражение для ПФ означает, что частотная характеристика представляет собой значения ПФ $H(z)$, взятые на единичной окружности комплексной плоскости, см. рис. 3.15. На рисунке приведен также пример карты нулей и полюсов ПФ, т. е. их расположение на комплексной z -плоскости: символами «o» принято обозначать положение нулей, а «*» — полюсов ПФ.

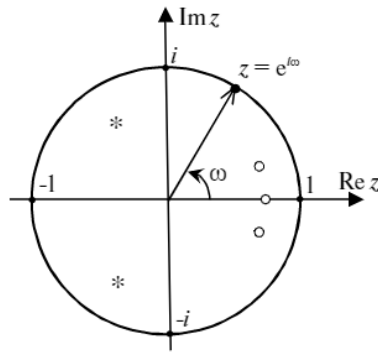


Рис. 3.15. Геометрическая иллюстрация связи ЧХ и ПФ: $K(\omega) = H(e^{i\omega})$. Пример карты нулей и полюсов ПФ

Как отмечалось ранее в разделе 3.3, для рассматриваемых нами систем ПФ (3.10) имеет вид дробно-рациональной функции. Поэтому карта (расположение) нулей и полюсов ПФ определяет ее с точностью до постоянного множителя. Кроме того, если на окружности $|z| = 1$ ПФ $H(z)$ не имеет полюсов (это всегда верно для устойчивых систем), то ЧХ $K(\omega) = H(e^{i\omega})$ будет также и непрерывной по ω функцией.

Вводя определение ЛДС при помощи уравнения (3.7) (или вытекающего из него соотношения (3.8)), мы указали, что рассматриваемые нами системы имеют вещественные значения коэффициентов $\{a_m\}_{m=1}^M, \{b_k\}_{k=0}^N$. Поэтому если входным воздействием является комплексный сигнал $\tilde{x}(n) = x(n) + iv(n)$ (где $x(n), v(n)$ — вещественные последовательности), то в силу линейности ЛДС отклик такой системы дает комплексный сигнал

$$\tilde{y}(n) = L\{\tilde{x}(n)\} = L\{x(n)\} + iL\{v(n)\} = y(n) + iu(n).$$

Здесь вещественный сигнал $y(n) = L\{x(n)\}$ соответствует отклику системы на воздействие $x(n)$, а вещественный сигнал $u(n) = L\{v(n)\}$ — реакции на воздействие $v(n)$.

Рассмотрим в качестве входной последовательности отсчеты комплексного гармонического сигнала $\tilde{x}(n) = Ae^{i\omega n}$ ($n \geq 0$), где A — некоторое комплексное число (комплексная амплитуда гармонического колебания). Тогда найденный по формуле свертки (3.9) отклик ЛДФ получаем равным

$$\tilde{y}(n) = \sum_{k=0}^n h(k) \underbrace{Ae^{i\omega(n-k)}}_{\tilde{x}(n-k)} = \underbrace{Ae^{i\omega n}}_{\tilde{x}(n)} \underbrace{\sum_{k=0}^n h(k)e^{-i\omega k}}_{K_n(\omega) \xrightarrow{n \rightarrow \infty} K(\omega)} \xrightarrow{n \rightarrow \infty} \tilde{x}(n)K(\omega). \quad (3.13)$$

То есть в установившемся режиме (при значениях $n \gg 1$) на выходе ЛДФ получаем гармоническое колебание той же частоты ω , а его амплитуда и фаза определяются ЧХ, которую можно также записать в показательной форме: $K(\omega) = |K(\omega)|e^{i\varphi(\omega)}$. 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) $\varphi(\omega) = \arg K(\omega)$. Рассмотрим это на примере.

Пример 3.16. Для значения параметра $a = 1/\sqrt{3}$ определить отклик фильтра из примера 3.4 в установившемся режиме при подаче на вход последовательности $x(n) = \{1, 0, -1, 0, 1, 0, -1, 0, \dots\}$.

◀ Для данного фильтра ПФ $H(z) = \frac{z}{z-a}$, откуда $K(\omega) = 1 / \left(1 - \frac{1}{\sqrt{3}}e^{-i\omega}\right)$. Заданное входное воздействие: $x(n) = \cos(\pi n/2) = \operatorname{Re}(e^{i\pi n/2})$. 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 $\tilde{y}(n) = K(\pi/2)\tilde{x}(n)$ и для $n \gg 1$ имеем:

$$y(n) = \operatorname{Re} \left(\frac{e^{i\frac{\pi n}{2}}}{1 - \frac{1}{\sqrt{3}}} \right) = \operatorname{Re} \left(\frac{e^{i\frac{\pi n}{2}}}{1 + \frac{i}{\sqrt{3}}} \right) = \operatorname{Re} \left(\frac{e^{i\frac{\pi n}{2}}}{\frac{2}{\sqrt{3}}e^{i\frac{\pi}{6}}} \right) =$$

$$= \operatorname{Re} \left(\underbrace{e^{i\frac{\pi n}{2}}}_{\tilde{x}(n)} \underbrace{\frac{\sqrt{3}}{2} e^{-i\frac{\pi}{6}}}_{K(\pi/2)} \right) = \frac{\sqrt{3}}{2} \cos \left(\frac{\pi n}{2} - \frac{\pi}{6} \right).$$

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 сдвиг фазы выходного сигнала определяется ФЧХ $\varphi(\pi/2) = \arg K(\pi/2) = -\pi/6$. ▶

Результат примера 3.16 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

$$y(n) = \operatorname{Re}\{A e^{i\omega n} |K(\omega)| e^{i\varphi(\omega)}\} = |A| |K(\omega)| \cos(\omega n + \gamma + \varphi(\omega)). \quad (3.14)$$

Использование формулы (3.14) возможно лишь для достаточно больших значений индекса n , когда можно полагать (см. (3.13)), 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 абсолютно сходится ряд $|K(\omega)| = \left| \sum_{k=0}^{\infty} h(k) e^{-i\omega k} \right| \leq \sum_{k=0}^{\infty} |h(k)| < \infty$, поэтому ИХ также можно считать конечной, выбрав такое достаточно большое число n , что с некоторой требуемой точностью $K_n(\omega) \approx K(\omega)$. Для неустойчивых фильтров формула (3.14) неприменима. Таким образом, физический смысл ЧХ — комплексный коэффициент передачи гармонических колебаний в установившемся режиме.

При прохождении через фильтр дискретный сигнал (не обязательно гармоническое колебание) некоторым образом изменяется. Записав общее уравнение преобразования сигнала линейной системой в Z -области, $Y(z) = H(z)X(z)$, выполним подстановку $z = e^{i2\pi v}$. Тогда получим следующую связь спектров $\tilde{S}_y(v) = \sum_{n=0}^{\infty} y(n) e^{-i2\pi v n} = Y(e^{i2\pi v})$ и $\tilde{S}_x(v) = \sum_{n=0}^{\infty} x(n) e^{-i2\pi v n} = X(e^{i2\pi v})$ (выходного и входного сигналов):

$$\tilde{S}_y(v) = K_f(v) \tilde{S}_x(v),$$

где $K_f(v) = H(e^{i2\pi v}) = \sum_{n=0}^{\infty} h(n) e^{-i2\pi v n}$ представляет собой ЧХ фильтра с аргументом в терминах линейной частоты $v = \omega/2\pi$. Характеризуя

в частотной области изменение сигнала фильтром, выделяют *амплитудные* и *фазовые* искажения, связанные с изменениями амплитудного и фазового спектров сигнала, которые определяются АЧХ и ФЧХ системы.

Учитывая, что $h(n) = 0$ при $n < 0$, и переходя к линейной частоте $\nu = \omega/2\pi$, рассмотрим ЧХ $K_f(\nu) = \sum_{k=-\infty}^{+\infty} h(k)e^{-2\pi i\nu k}$. Данное выражение представляет собой ряд Фурье в комплексной форме (1.15) для функции единичного периода, где коэффициенты Фурье (1.16): $c_k = h(-k)$. Как видим, нормированная ЧХ $K_f(\nu)$ представляет собой частотный спектр дискретного сигнала — ИХ $h(n)$ для единичного интервала дискретизации $T = 1$, см. раздел 2.4 и формулу (2.9). Частота Найквиста $f_D/2$ (см. (2.7)) соответствует нормированной частоте $\nu = 1/2$ ($\omega = \pi$), а период нормированной ЧХ в соответствии со свойством периодичности спектра дискретного сигнала будет единичным для $K_f(\nu)$ и равным 2π для $K(\omega)$. Нормированные частоты не имеют размерности; реальная линейная частота измеряется в герцах (Гц), а циклическая — в радианах в секунду (рад/с).

Итак, для линейных нормированных частот основной полосой сигнала является область $\nu \in [0; 1/2]$, что для ненормированных (оригинальных) линейных частот соответствует полосе $\omega \in [0; f_D/2]$.

Пример 3.17. В основной полосе частот фильтра задана следующая

ЧХ: $K_f(\nu) = \begin{cases} 0, & \text{при } 0 \leq \nu < 1/4 \\ 1, & \text{при } 1/4 \leq \nu \leq 1/2 \end{cases}$. Найти ИХ данного фильтра.

◀ Так как $K_f(-\nu) = \overline{K_f(\nu)}$ и $K_f(\nu) = K_f(\nu + 1)$, то на периоде ЧХ имеем:

$$K_f(\nu) = \sum_{k=-\infty}^{+\infty} h(k)e^{-2\pi i\nu k} = \begin{cases} 0, & \text{при } \nu \in (-1/4; 1/4) \\ 1, & \text{при } \nu \in [-1/2; -1/4] \cup [1/4; 1/2] \end{cases}$$

Компоненты ИХ находим как коэффициенты (1.16) ряда Фурье (1.15):

$$h(k) = c_{-k} = \int_{-1/2}^{1/2} K_f(\nu)e^{i2\pi\nu k} d\nu = 2 \int_{1/4}^{1/2} \cos(2\pi\nu k) d\nu,$$

где последний переход осуществлен в силу четности ЧХ $K_f(\nu)$. Вычисляя полученный определенный интеграл, находим ИХ:

$$h(k) = \begin{cases} 1/2 & \text{при } k = 0 \\ -\frac{\sin(\pi k/2)}{\pi k} & \text{при } k = \pm 1, \pm 2, \dots \end{cases}$$

Напомним, что ИХ представляет собой отклик системы на единичное входное воздействие $x(n) = \delta(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$, поэтому рассмотренный в примере фильтр является *физически нереализуемым*, так как $h(k) \neq 0$ при $k < 0$ (отклик на выходе системы не может появиться ранее входного воздействия). ►

Рассмотренный в примере 3.17 фильтр является *идеальным фильтром верхних частот*: нижние частоты фильтр подавляет, верхние — пропускает. В терминах значений АЧХ $|K_f(v)|$ в основной полосе нормированных частот $v \in [0; 1/2]$ определим следующие типы идеальных дискретных фильтров.

Определения. *Идеальным фильтром нижних частот (ФНЧ)* назовем фильтр, имеющий в основной (нормированной) полосе частот следующую АЧХ:

$$|K_f(v)| = \begin{cases} 1, & \text{при } 0 \leq v < v_c \\ 0, & \text{при } v_c \leq v \leq 1/2 \end{cases}$$

Идеальным фильтром верхних частот (ФВЧ) назовем фильтр с АЧХ

$$|K_f(v)| = \begin{cases} 0, & \text{при } 0 \leq v < v_c \\ 1, & \text{при } v_c \leq v \leq 1/2 \end{cases}$$

идеальным полосовым фильтром (ПФ) — фильтр с АЧХ

$$|K_f(v)| = \begin{cases} 0, & \text{при } 0 \leq v < v_{c1} \\ 1, & \text{при } v_{c1} \leq v < v_{c2} \\ 0, & \text{при } v_{c2} \leq v \leq 1/2 \end{cases}$$

идеальным режекторным (или полосно-заграждающим) — с АЧХ

$$|K_f(v)| = \begin{cases} 1, & \text{при } 0 \leq v < v_{c1} \\ 0, & \text{при } v_{c1} \leq v < v_{c2} \\ 1, & \text{при } v_{c2} \leq v \leq 1/2 \end{cases}$$

Область частот, в которой АЧХ идеальных фильтров равна единице (в общем случае некоторой положительной константе), называют *полосой пропускания* фильтра, а область частот, в которой АЧХ равна нулю, — *полосой подавления*. Граничная частота v_c , разделяющая полосу пропускания и полосу подавления, называется *частотой среза*.

Упражнение. С учетом свойств периодичности, $K_f(v) = K_f(v+1)$, и сопряженной симметрии ЧХ, $K_f(-v) = \overline{K_f(v)}$, нарисуйте графики АЧХ для определенных выше идеальных дискретных фильтров.

Идеальные фильтры физически нереализуемы, так как для них всегда $h(k) \neq 0$ при $k < 0$. АЧХ $|K_f(v)|$ реальных фильтров отличны от прямоугольных, и, чтобы подчеркнуть это, схематично АЧХ ФНЧ можно изобразить, например, так, как это отражено на рис. 3.16. Способы задания спецификаций частотных характеристик фильтров и вопросы синтеза структур фильтров будут рассмотрены в четвертой главе.

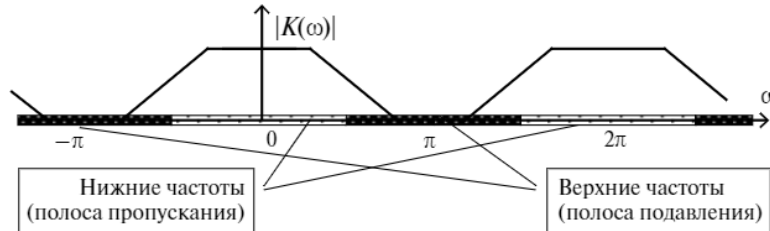


Рис. 3.16. Схематичный вид нормированной АЧХ фильтра нижних частот

Определение. Пусть в некоторой полосе частот $\omega \in \Omega$ ФЧХ имеет вид

$$\varphi(\omega) = \arg K(\omega) = \alpha\omega + c + 2\pi t, \quad (3.15)$$

где α, c — некоторые вещественные константы, а целое число t выбирается так, чтобы в соответствии с определением аргумента комплексного числа выполнялось условие $-\pi < \varphi(\omega) \leq \pi$. Тогда говорят, что в полосе частот Ω ФЧХ является *линейной*.

Замечание. Наличие слагаемого $2\pi t$ в выражении (3.15) связано только с необходимостью приведения значения ФЧХ к диапазону $-\pi < \varphi(\omega) \leq \pi$. Если это требование снять, то соответствующая *непрерывная фаза* $\Phi(\omega) = \alpha\omega + c$ будет иметь вид линейной функции, что и объясняет приведенное определение.

Пример 3.18. Для фильтра, заданного передаточной функцией $H(z) = \frac{c}{2} + cz^{-1} + \frac{c}{2}z^{-2}$, где $c > 0$, найти АЧХ и ФЧХ, определить

отклик $y(n)$ в установившемся режиме на входное воздействие $x(n) = \sin\left(\frac{\pi}{3}n - \frac{\pi}{6}\right)$.

◀ 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 $|K(\omega)| = c(1 + \cos \omega) = 2c \cos^2(\omega/2)$ приведен на рис. 3.17. Данный ЛДФ является фильтром нижних частот. График линейной ФЧХ $\varphi(\omega) = \arg K(\omega) = \arg e^{-i\omega}$ постройте самостоятельно, помня, что $-\pi < \varphi(\omega) \leq \pi$.

В установившемся режиме отклик фильтра на гармоническое входное воздействие $x(n) = \operatorname{Im} e^{i\left(\frac{\pi}{3}n - \frac{\pi}{6}\right)}$ с циклической частотой $\omega = \pi/3$ находим как

$$\begin{aligned} y(n) &= \operatorname{Im} \left(K\left(\frac{\pi}{3}\right) e^{i\left(\frac{\pi}{3}n - \frac{\pi}{6}\right)} \right) = \operatorname{Im} \left(\frac{3}{2} c e^{i\left(-\frac{\pi}{3} + \frac{\pi}{3}n - \frac{\pi}{6}\right)} \right) = \\ &= \frac{3}{2} c \sin\left(\frac{\pi n}{3} - \frac{\pi}{2}\right) = -\frac{3c}{2} \cos \frac{\pi n}{3}. \end{aligned}$$

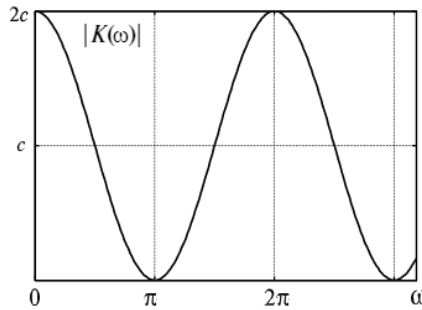


Рис. 3.17. АЧХ фильтра из примера 3.18

Упражнение. Для фильтра, заданного ПФ $H(z) = \frac{c}{2} - cz^{-1} + \frac{c}{2} z^{-2}$, $c > 0$, найти АЧХ и ФЧХ, определить отклик $y(n)$ в установившемся режиме на входное воздействие $x(n) = \cos\left(\frac{\pi}{3}n - \frac{\pi}{6}\right)$ и изобразить структурную схему фильтра. Является ли ФЧХ фильтра линейной?

Линейная ФЧХ (3.15) гарантирует, что гармонические колебания $x(n) = |A| \cos(\omega n + \gamma)$ любой частоты при прохождении через



фильтр имеют одинаковые временной сдвиг α и фазовый сдвиг c . На выходе системы в соответствии с (3.14) получается сигнал

$$y(n) = |A||K(\omega)| \cos(\omega n + \gamma + \alpha\omega + c + 2\pi m) = |A||K(\omega)| \cos(\omega(n + \alpha) + \gamma + c).$$

Различия наблюдаются только в амплитуде колебаний, изменение которой определяется АЧХ. Величина α определяет временную задержку выходного сигнала относительно входного, взятую с противоположным знаком. Свойство линейности ФЧХ часто является необходимым требованием, предъявляемым к синтезируемому фильтру.

В общем случае непрерывная фаза $\Phi(\omega) \in (-\infty; \infty)$ связана с ФЧХ как $\Phi(\omega) = \varphi(\omega) + 2\pi m$, где целое число m выбирается так, чтобы в точках разрыва $\varphi(\omega)$ со «скачком» значения на 2π обеспечить непрерывность функции $\Phi(\omega)$.

Определение. Назовем *групповым временем задержки* (ГВЗ) функцию $\tau(\omega) = -d\Phi(\omega)/d\omega$.

Несложно видеть, что для линейной ФЧХ $d\Phi(\omega)/d\omega = \alpha = \text{const}$. Поэтому линейность ФЧХ характеризуется тем, насколько близко ГВЗ к константе. В точках непрерывности функции $\tau(\omega) = -d\Phi(\omega)/d\omega$ значение ГВЗ $\tau(\omega_0)$ для некоторой частоты ω_0 определяет временную задержку спектральных составляющих сигнала с частотами ω , близкими к ω_0 , для которых можно считать, что $\tau(\omega) \approx \tau(\omega_0)$. Это приближенное равенство будет выполняться тем точнее, чем ближе ФЧХ к линейной.

Упражнение. Найдите ГВЗ для фильтра из примера 3.18.

3.7. Минимально-фазовые системы. Фазовые звенья

Определение. Для линейной дискретной системы с ПФ $H(z)$ *обратной* назовем систему, имеющую ПФ $\tilde{H}(z)$ такую, что $\tilde{H}(z)H(z) = 1$.

Очевидно, что $\tilde{H}(z) = 1/H(z)$, и для ЧХ обратной системы также имеем $\tilde{K}(\omega) = 1/K(\omega)$.

Как было установлено ранее в разделе 3.3, передаточная функция систем, определяемых разностным уравнением (3.8), может быть представлена в виде отношения полиномов: $H(z) = K \prod_{j=1}^J (z - z_{c_j}) / \prod_{l=1}^L (z - z_{*l})$, где $\{z_{c_j}\}_{j=1}^J$ — нули, $\{z_{*l}\}_{l=1}^L$ — полюсы ПФ, а K — некоторый масштабный коэффициент. Задав множество нулей и полюсов ПФ, мы можем определить ее с точностью до вещественного множителя K . Для такого класса систем, очевидно, обратная система имеет следующую ПФ:

$$\tilde{H}(z) = \frac{1}{K} \prod_{l=1}^L (z - z_{*l}) / \prod_{j=1}^J (z - z_{c_j}).$$

Таким образом, полюсы ПФ обратной системы являются нулями исходной ПФ, поэтому для устойчивости обратной системы (см. теорему 3.4) необходимо и достаточно, чтобы нули ПФ оригинальной системы также лежали на комплексной плоскости внутри единичного круга. Требование расположения не только полюсов, но и нулей ПФ дискретной системы внутри круга $|z| < 1$ часто возникает также и из других соображений, связанных с условиями, накладываемыми на фазочастотные характеристики. Поясним это на примере.

Пример 3.19. Показать, что системы с передаточными функциями $H_1(z) = 1 - bz^{-1}$ и $H_2(z) = b - z^{-1}$, где b — вещественный параметр, $0 < b < 1$, имеют одинаковые АЧХ $|K_1(\omega)| = |K_2(\omega)|$, и в основной полосе частот $\omega \in [0; \pi]$ выполняется неравенство: $0 \leq \arg K_1(\omega) \leq \arg K_2(\omega)$.

◀ Несложно видеть, что представленные системы имеют один полюс первого порядка в точке $z = 0$. Первая система имеет ноль внутри единичного круга в точке $z = b$, вторая — вне единичного круга в точке $z = 1/b$. Для ЧХ систем имеем:

$$K_1(\omega) = H_1(e^{i\omega}) = 1 - be^{-i\omega} = 1 - b \cos \omega + i b \sin \omega,$$

$$K_2(\omega) = H_2(e^{i\omega}) = b - e^{-i\omega} = b - \cos \omega + i \sin \omega.$$

Отсюда получаем одинаковые АЧХ (убедитесь):

$$|K_1(\omega)| = \sqrt{(1 - b \cos \omega)^2 + b^2 \sin^2 \omega} = \sqrt{1 + b^2 - 2b \cos \omega},$$

$$|K_2(\omega)| = \sqrt{(b - \cos \omega)^2 + \sin^2 \omega} = \sqrt{1 + b^2 - 2b \cos \omega}.$$

Сравним ФЧХ двух систем в основной полосе частот $\omega \in [0; \pi]$. Как было показано в предыдущем разделе 3.6, в установившемся режиме гармоническое колебание $x(n) = \cos(\omega n + \gamma)$, поданное на вход ЛДС с частотной характеристикой $K(\omega)$, преобразуется



в выходной сигнал $y(n) = |K(\omega)| \cos(\omega n + \gamma + \varphi(\omega))$, где $|K(\omega)|$ — АЧХ, $\varphi(\omega) = \arg K(\omega)$ — ФЧХ системы, см. (3.14). Чем ближе к нулю величина $|\varphi(\omega)| = |\arg K(\omega)|$, тем меньшие фазовые искажения приобретает сигнал, тем меньше сдвиг выходного сигнала $y(n)$ относительно входного $x(n)$.

Для первого из рассмотренных фильтров в основной полосе частот $0 \leq \omega \leq \pi$ имеем:

$$\varphi_1(\omega) = \arg K_1(\omega) = \operatorname{arctg} \left(\frac{b \sin \omega}{1 - b \cos \omega} \right), \quad \text{причем } 0 \leq \varphi_1(\omega) < \pi/2,$$

так как для $0 < b < 1$ в основной полосе частот $b \sin \omega \geq 0$ и $1 - b \cos \omega > 0$.

Для второго фильтра при $0 < b < 1$ действительная часть ЧХ принимает отрицательные значения в основной полосе частот для $b < \cos \omega$ и положительные значения при $b > \cos \omega$. Поэтому, обозначив $\Omega = \arccos b$, имеем:

$$\varphi_2(\omega) = \arg K_2(\omega) = \begin{cases} \pi + \operatorname{arctg} \left(\frac{\sin \omega}{b - \cos \omega} \right), & \text{при } 0 \leq \omega < \Omega \\ \operatorname{arctg} \left(\frac{\sin \omega}{b - \cos \omega} \right), & \text{при } \Omega < \omega \leq \pi \end{cases}.$$

Доопределим полученное выражение «по непрерывности», а именно, положим $\varphi_2(\Omega) = \lim_{\omega \rightarrow \Omega} \varphi_2(\omega) = \pi/2$.

Очевидно, что для $0 \leq \omega \leq \Omega$: $\pi \geq \varphi_2(\omega) \geq \pi/2$, поэтому $\varphi_2(\omega) > \varphi_1(\omega)$. Для частотной полосы $\Omega < \omega < \pi$ также получим $\varphi_2(\omega) > \varphi_1(\omega)$, поскольку функция арктангенса является монотонно возрастающей и при $0 < b < 1$:

$$\frac{\sin \omega}{b - \cos \omega} > \frac{b \sin \omega}{1 - b \cos \omega} = \frac{\sin \omega}{\frac{1}{b} - \cos \omega}.$$

При $\omega = \pi$ имеем $\varphi_2(\pi) = \varphi_1(\pi) = 0$.

Таким образом, для основной полосы частот $0 \leq \omega \leq \pi$: $0 \leq \arg K_1(\omega) \leq \arg K_2(\omega)$. Поскольку $K(-\omega) = \overline{K(\omega)}$, то для частот $-\pi \leq \omega < 0$ аргумент меняет знак на противоположный и $\arg K_2(\omega) \leq \arg K_1(\omega) \leq 0$. ►

Две рассмотренные в примере 3.19 системы с вещественными нулями $z = b$ (первая) и $z = 1/b$ (вторая), $0 < b < 1$, имеют одинаковые АЧХ, но первая система дает меньшее искажение фазы выходного гармонического сигнала, т. к. $|\varphi_2(\omega)| > |\varphi_1(\omega)|$.

Приведенный пример показывает, что для простейших систем с одинаковой АЧХ, имеющих в передаточной функции единственный ноль, меньшие фазовые искажения обеспечивают те системы, у которых ноль ПФ расположен внутри единичного круга $|z| < 1$. Данное наблюдение поясняет смысл следующего определения.

Определение. *Минимально-фазовой* назовем систему с ПФ вида (3.10), все полюсы и нули которой лежат в области $|z| < 1$.

Помимо вещественных нулей или полюсов, ПФ систем, определяемых уравнением (3.8) с вещественными коэффициентами, может также иметь и комплексно-сопряженные пары нулей (полюсов).

Пример 3.20. Показать, что системы с передаточными функциями $H_{11}(z) = (1 - Cz^{-1})(1 - \bar{C}z^{-1})$ и $H_{22}(z) = (C - z^{-1})(\bar{C} - z^{-1})$, где комплексное число C таково, что $0 < |C| < 1$, имеют одинаковые АЧХ $|K_{11}(\omega)| = |K_{22}(\omega)|$. Какая из систем является минимально-фазовой? ◀ Обозначим $b = |C|$, $0 < b < 1$; $\varphi = \arg C$, $-\pi < \varphi \leq \pi$. Тогда $C = be^{i\varphi}$ и для частотных характеристик имеем:

$$K_{11}(\omega) = (1 - be^{-i(\omega-\varphi)})(1 - be^{-i(\omega+\varphi)}),$$

$$K_{22}(\omega) = (be^{i\varphi} - e^{-i\omega})(be^{-i\varphi} - e^{-i\omega}) = (b - e^{-i(\omega+\varphi)})(b - e^{-i(\omega-\varphi)}).$$

Используя обозначения из рассмотренного выше примера 3.19 $K_1(\omega) = 1 - be^{-i\omega}$ и $K_2(\omega) = b - e^{-i\omega}$, получаем представления:

$$K_{11}(\omega) = K_1(\omega - \varphi)K_1(\omega + \varphi), \quad K_{22}(\omega) = K_2(\omega + \varphi)K_2(\omega - \varphi).$$

В примере 3.19 показано, что $|K_1(\omega)| = |K_2(\omega)|$, поэтому $|K_{11}(\omega)| = |K_{22}(\omega)|$.

Два нуля ПФ $H_{11}(z)$, C и \bar{C} , лежат внутри единичного круга $|z| < 1$, поэтому первая система является минимально-фазовой. Нули ПФ $H_{22}(z)$, $1/C$ и $1/\bar{C}$, лежат в области $|z| > 1$, потому вторая система является *неминимально-фазовой*. ▶

Отметим, что для ПФ рассматриваемых нами систем, определяемых уравнением (3.8) с вещественными коэффициентами, верно соотношение $H(\bar{z}) = \overline{H(z)}$, непосредственно вытекающее из (3.10), и АЧХ $|K(\omega)|$ может быть найдена через ПФ $H(z)$ на основании следующей связи:

$$|K(\omega)|^2 = \underbrace{H(e^{i\omega})}_{K(\omega)} \overbrace{H(e^{i\omega})}^{\overline{K(\omega)}} = H(e^{i\omega})H(e^{-i\omega}) = H(z)H(1/z)|_{z=e^{i\omega}}. \quad (3.16)$$

Из данного соотношения следует, что если две системы имеют такие передаточные функции $H(z)$ и $G(z)$, что $H(z) = G(1/z)$, то эти системы имеют одинаковые АЧХ, так как $H(z)H(1/z) = G(1/z)G(z)$. Соотношение $H(z) = G(1/z)$ означает, что каждому нулю (полюсу) z_k ПФ $H(z)$ соответствует ноль (полюс) ПФ $G(z)$: $z_k^* = 1/z_k$. Поэтому если одна из этих двух систем с одинаковыми АЧХ является минимально-фазовой, то вторая будет неминимально-фазовой. В примерах 3.19 и 3.20 мы наблюдали выполнение данного свойства для простейших нерекурсивных звеньев первого и второго порядков.

Как отмечено ранее, расположение нулей и полюсов определяет ПФ с точностью до постоянного нормирующего множителя. Поэтому неминимально-фазовую систему можно преобразовать к минимально-фазовой с той же АЧХ, заменив нули $\{z_j\}$ исходной ПФ, лежащие вне области $|z| < 1$, на их обратные значения $\{1/z_j\}$. Исключение составляют неминимально-фазовые системы, имеющие нули ПФ на окружности $|z| = 1$ — такие нули после замены на $1/z$ остаются на единичной окружности, меняя знак аргумента комплексного числа на противоположный.

Пример 3.21. Покажем, что замена одного нуля единичной кратности z_k ПФ $H(z)$ на $z_k^* = 1/z_k$, сохраняющая АЧХ системы, соответствует умножению ПФ $H(z)$ на множитель $I_k(z) = \frac{z_k z - 1}{z - z_k}$.

◀ Рассмотрим ПФ $U(z) = H(z)I_k(z) = H(z)\frac{z_k z - 1}{z - z_k}$. Очевидно, что один ноль функции $H(z)$ сокращается с полюсом z_k функции $I_k(z)$. Вместе с тем функция $U(z)$ получит от $I_k(z)$ дополнительный ноль в точке $z_k^* = 1/z_k$. Так как $|U(e^{i\omega})| = |H(e^{i\omega})||I_k(e^{i\omega})|$ и на основании (3.16) $|I_k(e^{i\omega})|^2 = I_k(z)I_k(1/z)|_{z=e^{i\omega}} = 1$ (убедитесь), то АЧХ полученной системы сохраняется: $|U(e^{i\omega})| = |H(e^{i\omega})|$. ▶

Система из примера 3.21, определяемая ПФ $I_k(z)$, обладает постоянной АЧХ для всех значений частоты ω . При прохождении через такую систему гармоническое колебание любой частоты не изменяет своей амплитуды, возможны только изменения фазы.

Определение. *Все пропускающей системой* (или *фазовым звеном*) назовем линейную дискретную систему с АЧХ $|K(\omega)| = \text{const} > 0$.

Простейшие фазовые звенья первого и второго порядков с АЧХ $|P(e^{j\omega})| = 1$ можно получить, воспользовавшись результатами решения примеров 3.19 и 3.20. Фазовое звено первого порядка имеет ПФ вида

$$P(z) = \frac{b - z^{-1}}{1 - bz^{-1}} = \frac{bz - 1}{z - b},$$

где для устойчивости системы вещественное число b должно быть таким, что $|b| < 1$. Для фазового звена второго порядка имеем

$$P(z) = \frac{(C - z^{-1})(\bar{C} - z^{-1})}{(1 - Cz^{-1})(1 - \bar{C}z^{-1})} = \frac{(Cz - 1)(\bar{C}z - 1)}{(z - C)(z - \bar{C})},$$

где для устойчивости системы комплексное число C должно удовлетворять условию $|C| < 1$. Следствием требования устойчивости рассмотренных фазовых звеньев первого и второго порядков является то, что данные звенья не являются минимально-фазовыми, так как каждому полюсу z_k ПФ $P(z)$ внутри круга $|z| < 1$ соответствует ноль ПФ в точке $z_k^* = 1/z_k^*$, $|z_k^*| > 1$.

При последовательном соединении различных фазовых звеньев, очевидно, также получается все пропускающая система, ПФ которой в общем виде можно представить как

$$\begin{aligned} P(z) &= K \prod_{k=1}^{M_r} \frac{b_k - z^{-1}}{1 - b_k z^{-1}} \prod_{j=1}^{M_c} \frac{(C_j - z^{-1})(\bar{C}_j - z^{-1})}{(1 - C_j z^{-1})(1 - \bar{C}_j z^{-1})} = \\ &= K \prod_{k=1}^{M_r} \frac{b_k z - 1}{z - b_k} \prod_{j=1}^{M_c} \frac{(C_j z - 1)(\bar{C}_j z - 1)}{(z - C_j)(z - \bar{C}_j)}, \end{aligned} \quad (3.17)$$

где определяющие нули и полюсы параметры $\{b_k\}_{k=1}^{M_r}$ являются вещественными, а $\{C_j\}_{j=1}^{M_c}$ — комплексными числами; K — вещественный множитель.

Применяя к каждому лежащему в области $|z| > 1$ нулю ПФ $H(z)$ любой неминимально-фазовой системы прием, рассмотренный в примере 3.21, для ПФ можно получить представление $H(z) = G(z)P(z)$, где $G(z)$ — ПФ минимально-фазовой системы¹, причем $|H(z)| = |G(z)|$, а $P(z)$ — ПФ (3.17) все пропускающей системы.

¹ В данном случае допускаем наличие нулей на окружности $|z| = 1$.



Упражнение. Для системы, имеющей ПФ

$$H(z) = \frac{\left(1 - \frac{3}{2}e^{i\frac{\pi}{4}}z^{-1}\right)\left(1 - \frac{3}{2}e^{-i\frac{\pi}{4}}z^{-1}\right)}{1 - \frac{1}{3}z^{-1}} = \frac{\left(z - \frac{3}{2}e^{i\frac{\pi}{4}}\right)\left(z - \frac{3}{2}e^{-i\frac{\pi}{4}}\right)}{z\left(z - \frac{1}{3}\right)},$$

найти разложение в виде $H(z) = G(z)P(z)$, где $G(z)$ — ПФ минимально-фазовой системы, $P(z)$ — ПФ всепропускающей системы (3.17).

Поскольку всепропускающие системы не изменяют АЧХ, они используются для коррекции фазовых искажений сигналов.

3.8. Нахождение отклика фильтра с использованием БПФ

Данный метод основан на свойстве 5° ДПФ, см. раздел 2.6. 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 $x(n)$ представляет собой конечную последовательность длины L : $x(L-1) \neq 0$ и $x(n) = 0$ при $n \geq L$. Тогда отклик фильтра (см. (3.9))

$$y(n) = \sum_{m=0}^{M-1} h(m)x(n-m) = \sum_{m=0}^{L-1} h(n-m)x(m) \quad (3.18)$$

будет иметь длительность $M+L-1$ отсчетов, 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 2.15: 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 (3.18), а остальные компоненты будут равны нулю:

$$Y = \left(\underbrace{y(0), y(1), \dots, y(L-1)}_L, \underbrace{y(L), \dots, y(L+M-2)}_{M-1}, 0, 0, \dots, 0 \right).$$

Использование описанной процедуры в вычислительном плане может быть более эффективным, чем непосредственная реализация формулы свертки. По сравнению с оценками из примера 2.15, выигрыш может быть более существенным. Действительно, фильтр можно задать не вектором ИХ $(h(0), \dots, h(M-1))$, а сразу отсчетами ЧХ $K_f(v) = \sum_{k=0}^{M-1} h(k)e^{-2\pi i v k}$, взятыми в точках $v_j = j/(2N)$, $j = 0, \dots, 2N-1$:

$$K_f(v_j) = \sum_{k=0}^{M-1} h(k)e^{-\frac{2\pi i}{2N} j k} = \sum_{k=0}^{2N-1} h(k)e^{-\frac{2\pi i}{2N} j k}$$

(здесь учтено, что $h(n) = 0$ при $n \geq M$). Правая часть последнего равенства с точностью до нормировочного множителя представляет собой ДПФ импульсной характеристики $\mathbf{H} = \underbrace{(h(0), h(1), \dots, h(M-1), 0, 0, \dots, 0)}_{2N \text{ элементов}}$:

$$\hat{h}_j = \frac{1}{\sqrt{2N}} \sum_{k=0}^{2N-1} h(k)e^{-\frac{2\pi i}{2N} j k}, \quad j = 0, 1, \dots, 2N-1.$$

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 вычислений.

Вычислительная эффективность использования БПФ тем выше, чем больше размерность используемого преобразования. Однако чаще всего обработать весь дискретный сигнал целиком оказывается 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, то последовательность

$x(n)$ обрабатывается поблочно. Выбирая из входного сигнала блоки по L отсчетов $X_j = \{x(jL), \dots, x((j+1)L-1)\}$, $j = 0, 1, \dots$, далее последовательно обрабатывают каждый из них по описанной выше схеме на основе БПФ. Входной сигнал можно представить в виде $x(n) = \sum_j X_j$, и в силу линейности фильтра его отклик также представляет собой сумму $y(n) = \sum_j Y_j$, где

$$Y_j = \left\{ \underbrace{y_j(jL), \dots, y_j((j+1)L-1)}_L, \underbrace{y_j((j+1)L), \dots, y_j((j+1)L+M-2)}_{M-1}, 0, \dots, 0 \right\}$$

— реакция фильтра на входное воздействие X_j .

Таким образом, хотя входная последовательность разделена на непересекающиеся блоки X_j , выходные блоки Y_j перекрываются. Задав исходно отклик фильтра нулевой последовательностью $y(n) = 0$, далее необходимо прибавлять компоненты очередного полученного блока Y_j ($j = 0, 1, \dots$) к тем элементам выходной последовательности $y(n)$, которые были сформированы ранее, см. рис. 3.18.

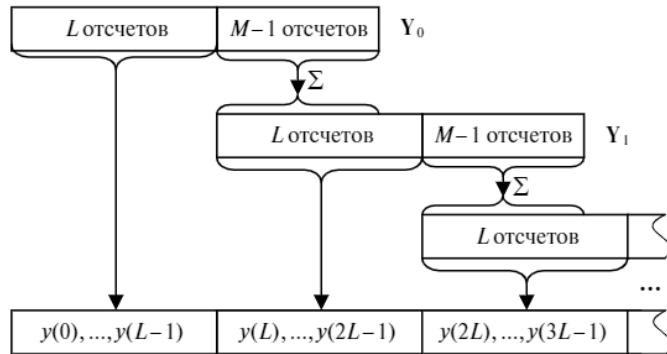


Рис. 3.18. Вычисление отклика ЛДФ при поблочной обработке входного сигнала

3.9. Многоскоростная обработка сигналов

Термин *многоскоростная обработка* применяется прежде всего, когда речь идет о преобразовании частоты дискретизации сигнала (повышении или понижении) в соответствующих *многоскоростных* системах ЦОС.

Уменьшение частоты дискретизации в целое число раз

Понижение частоты дискретизации называют также *децимацией* сигнала. Понижение частоты дискретизации $f_d = 1/T$ в M (целое число) раз реализовать очень просто: для этого необходимо прорежить дискретный сигнал $x(nT)$, оставив только отсчеты $x(nMT)$ и отбросив остальные. При этом нужно помнить, что для предотвращения появления эффекта наложения частот может потребоваться предварительная НЧ-фильтрация исходного сигнала $x(nT)$ с полным подавлением частот выше половины новой частоты дискретизации $f_{\text{new}}/2 = f_{\text{old}}/(2M)$. Спектральный эффект от прореживания сигнала вполне ожидаем, на рис. 3.19 приведен пример изменения спектра в результате уменьшения частоты дискретизации: сплошной линией показан амплитудный спектр соответствующего аналогового сигнала, пунктирной — амплитудный спектр дискретного сигнала.

Понижение частоты дискретизации целесообразно проводить в тех случаях, когда дополнительная НЧ-фильтрация для предотвращения наложения частот не требуется. Очевидно, что меньшая частота дискретизации снижает требования к аппаратуре ЦОС.

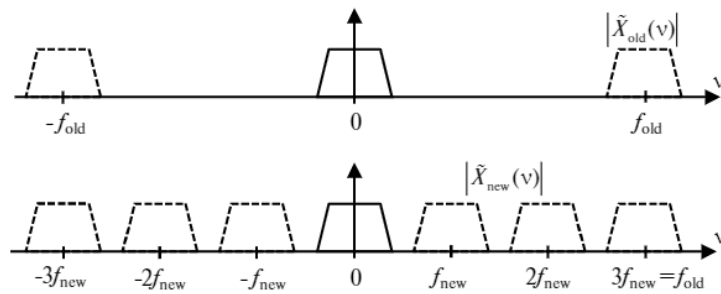


Рис. 3.19. Изменение спектра дискретного сигнала при уменьшении частоты дискретизации в 3 раза: сверху — амплитудный спектр исходного дискретного сигнала $x_{\text{old}}(n) = x(nT)$, внизу — сигнала $x_{\text{new}}(n) = x(3nT)$. Сплошная линия — часть амплитудного спектра, соответствующая исходному аналоговому сигналу $x(t)$

Повышение частоты дискретизации в целое число раз

Для повышения частоты дискретизации $f_d = 1/T$ заданного дискретного сигнала $x_{\text{old}}(n) = x(nT)$ в L раз необходимо сформировать последовательность $x_{\text{new}}(n) = x(nT/L)$, элементы которой представляют собой взятые в L раз чаще отсчеты того же самого аналогового сигнала $x(t)$, по которому была получена последовательность $x_{\text{old}}(n) = x(nT)$, см. рис. 3.20.

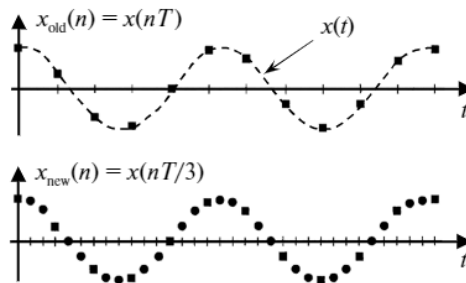


Рис. 3.20. Пример повышения частоты дискретизации в 3 раза

Так как предполагается, что последовательность $x_{\text{old}}(n) = x(nT)$ была сформирована с соблюдением условий теоремы отсчетов и аналоговый сигнал $x(t)$ может быть *точно* восстановлен по дискретному сигналу $x_{\text{old}}(n) = x(nT)$ с использованием формулы Котельникова (2.5), то решение задачи формирования последовательности $x_{\text{new}}(n) = x(nT/L)$ по заданной последовательности $x_{\text{old}}(n) = x(nT)$ является очевидным. На основании формулы (2.5) получаем:

$$x_{\text{new}}(n) = x(nT/L) = \sum_{k=-\infty}^{\infty} x_{\text{old}}(k) \frac{\sin\left(\pi\left(\frac{n}{L} - k\right)\right)}{\pi\left(\frac{n}{L} - k\right)}. \quad (3.19)$$

Очевидно, что $x_{\text{new}}(nL) = x_{\text{old}}(n)$, а остальные отсчеты новой последовательности можно вычислить по (3.19).

Однако практическая реализация такого подхода с точки зрения организации вычислений достаточно сложна. Тем не менее формула (3.19) демонстрирует основную идею, используемую при интерполяции дискретных сигналов.

Пример 3.22. Покажем, что вычисления по формуле (3.19) можно интерпретировать как дискретную свертку последовательности $h(n) = \frac{\sin(\pi n/L)}{\pi n/L}$ (где $h(0) = 1$) и последовательности

$$x_{\text{int}}(k) = \begin{cases} x_{\text{old}}(k/L), & \text{если } k \bmod L = 0 \\ 0, & \text{если } k \bmod L \neq 0 \end{cases}$$

полученной вставкой $(L-1)$ нулей после каждого элемента $x_{\text{old}}(n)$.

◀ С учетом введенных обозначений формулу (3.19) можем переписать:

$$\begin{aligned} x_{\text{new}}(n) &= \sum_{k \in \mathbb{Z}} x_{\text{int}}(kL) \frac{\sin(\pi(n-kL)/L)}{\pi(n-kL)/L} = \\ &= \sum_{k \in \mathbb{Z}} x_{\text{int}}(k) \frac{\sin(\pi(n-k)/L)}{\pi(n-k)/L} = \sum_{k \in \mathbb{Z}} x_{\text{int}}(k) h(n-k). \end{aligned} \quad \blacktriangleright$$

Упражнение. Покажите, что в основной полосе нормированных частот $0 \leq \nu \leq 1/2$ фильтр с ИХ $h(n) = \frac{\sin(\pi n/L)}{\pi n/L}$ ($L \geq 2$) имеет следующую ЧХ:

$$K(\nu) = \begin{cases} L, & \text{при } 0 \leq \nu < 1/2L \\ 0, & \text{при } 1/2L \leq \nu \leq 1/2 \end{cases},$$

т. е. этот фильтр обладает идеальной прямоугольной ЧХ ФНЧ с частотой среза (полного подавления) $f_s = 1/(2L)$ и коэффициентом усиления L в полосе пропускания $0 \leq \nu < 1/(2L)$. *Указание:* учитывая, что нормированная ЧХ есть функция единичного периода, найдите по заданной ЧХ отсчеты ИХ $h(n)$ как коэффициенты ряда Фурье (по аналогии со схемой решения примера 3.17).

Как видим, ИХ $h(n) = \frac{\sin(\pi n/L)}{\pi n/L}$ соответствует ЧХ идеального фильтра нижних частот (физически нереализуемого, так как $h(n) \neq 0$ при $n < 0$). Поэтому реализация вычислений по формуле (3.19) сводится к добавлению $(L-1)$ нулей после каждого отсчета в дискретный сигнал $x_{\text{old}}(n)$ и последующей НЧ-фильтрации полученной последовательности с частотой отсечки (полного подавления) $f_{\text{old}}/2 = f_{\text{new}}/(2L)$ и коэффициентом усиления L в полосе пропускания. Такая схема и используется на практике, конечно, с другими, физически реализуемыми (уже неидеальными) ФНЧ, которые проектируются исходя из требований конкретных приложений.



Два этапа реализации дискретной интерполяции сигнала при повышении частоты интерполяции (вставка нулей и НЧ-фильтрация) можно наглядно проиллюстрировать как преобразование спектров $X(v)$ соответствующих сигналов, см. рис. 3.21, где приведен пример преобразования сигналов и их спектров при четырехкратном повышении частоты дискретизации. Поскольку идеальный НЧ-фильтр реализовать невозможно, после НЧ-фильтрации лишние копии спектра подавляются не полностью (рис. 3.21*в*), а в сохраняемые копии также могут быть внесены некоторые искажения. Вследствие этого вместо точных значений (3.19) интерполяция дает некоторую ошибку, которая будет тем меньше, чем ближе ЧХ применяемого фильтра НЧ к идеальной.

Применяемый ФНЧ в полосе пропускания должен обеспечивать коэффициент усиления, равный L , для того чтобы компенсировать уменьшение амплитуды сигнала $x_{\text{new}}(n)$, вызванное добавлением нулевых отсчетов на этапе формирования последовательности $x_{\text{int}}(n)$.

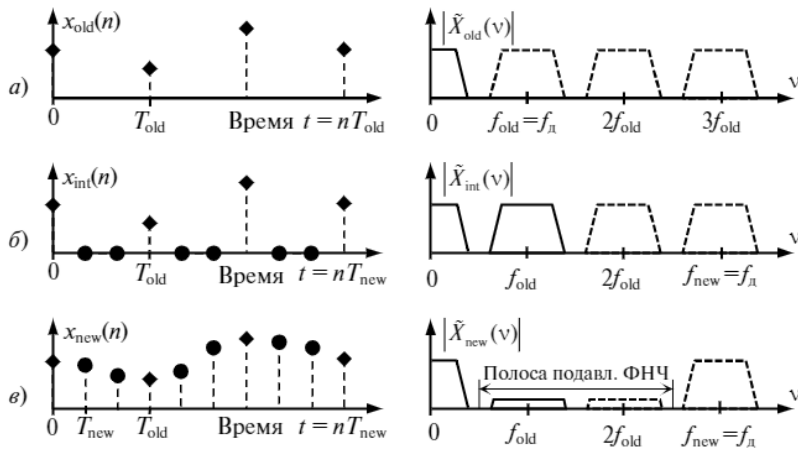


Рис. 3.21. Преобразование спектра при интерполяции дискретного сигнала (сплошной линией обозначен амплитудный спектр аналогового сигнала в области допустимых частот): *а*) оригинальный дискретный сигнал, *б*) промежуточный сигнал, разреженный нулями, *в*) интерполированный сигнал после низкочастотной фильтрации

Изменение частоты дискретизации на рациональный множитель

Используя рассмотренную выше схему повышения частоты дискретизации с множителем L и последующее понижение частоты дискретизации в M раз, можно изменить частоту дискретизации в любое (рациональное) число раз $k = L/M$. Например, чтобы получить частоту дискретизации, большую в 1,75 раза исходной, нужно взять $L = 7$, $M = 4$.

Обозначая операцию вставки $L - 1$ нулей после каждого отсчета последовательности как « $\uparrow L$ », а прореживание последовательности в M раз как « $\downarrow M$ », схему изменения (увеличения или уменьшения) частоты дискретизации можно представить в виде, изображенном на рис. 3.22а. Два ФНЧ в схеме рис. 3.22а, обрабатывающие сигналы на промежуточной частоте $f_{\text{пч}} = L/T$, можно объединить в один ФНЧ и упростить схему преобразования частоты дискретизации так, как это отражено на рис. 3.22б.

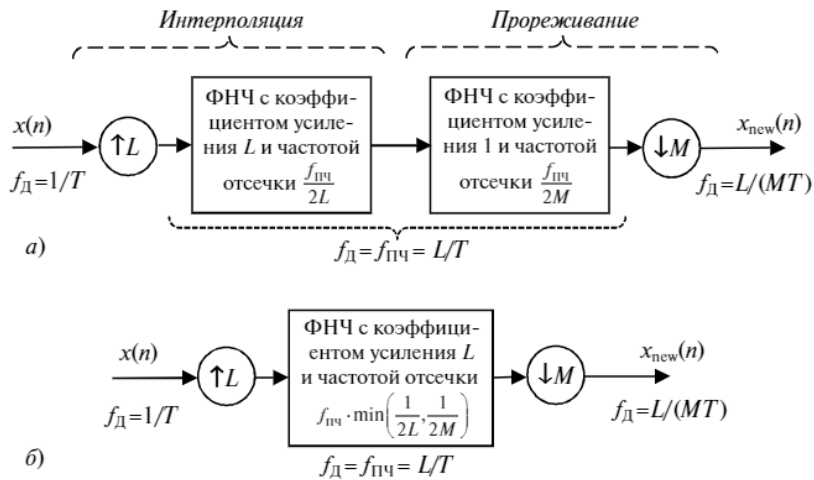


Рис. 3.22. Преобразование частоты дискретизации в спектра при интерполяции дискретного сигнала в $k = L/M$ раз: а) непосредственная реализация системы передискретизации, б) упрощенная схема реализации системы передискретизации



Пример 3.23. Сформулировать требования к ФНЧ и предложить схему передискретизации сигнала, при которой новая частота дискретизации $f_{\text{new}} = 0,8f_{\text{old}}$.

◀ Очевидно, что $L = 4$, $M = 5$. Для схемы обработки, определяемой рис. 3.22б, ФНЧ имеет рабочую частоту $f_{\text{ПЧ}} = 4f_{\text{old}}$, частоту полного подавления (отсечки) $f_{\text{ПЧ}}/10 = 0,4f_{\text{old}}$ и коэффициент усиления в полосе пропускания, равный 4. ▶

Упражнение. Сформулировать требования к ФНЧ и привести схему передискретизации сигнала по схеме рис. 3.22б, при которой новая частота дискретизации $f_{\text{new}} = 1,2f_{\text{old}}$.

Рассмотренные нами выше способы преобразования частоты дискретизации в вычислительном плане весьма неэффективны. Например, если требуется выполнить изменение частоты дискретизации в $4/3$ раза, то, вставив после каждого отсчета исходной последовательности по 3 нуля, в ходе вычислений свертки при последующей НЧ фильтрации 3 из 4 необходимых умножений будут умножениями на 0. Причем в получаемых в результате вычисления свертки данных 2 из 3 элементов не нужны, так как будут отброшены при прореживании. Более эффективные способы изменения частоты дискретизации (т.е. передискретизации) сигнала основаны на *полифазном* представлении дискретных сигналов.

3.10. Изменение частоты дискретизации сигналов с помощью полифазных фильтров

Пусть последовательность $y(n)$ получена в результате прореживания в M раз последовательности $x(n)$, т.е. $y(n) = x(nM)$. Найдем связь Z -образов $Y(z) = Z\{y(n)\}$ и $X(z) = Z\{x(n)\}$.

Сначала заметим, что $I(n) = \sum_{k=0}^{M-1} e^{\frac{2\pi i}{M}kn} = \begin{cases} M, & \text{при } n \bmod M = 0 \\ 0, & \text{при } n \bmod M \neq 0 \end{cases}$
(убедитесь самостоятельно, воспользовавшись формулой для суммы геометрической прогрессии). Поэтому

$$\begin{aligned}
 Y(z) &= \sum_{n=0}^{\infty} x(nM)z^{-n} = \frac{1}{M} \sum_{n=0}^{\infty} x(n) \left(\sum_{k=0}^{M-1} e^{\frac{2\pi i kn}{M}} \right) z^{-\frac{n}{M}} = \\
 &= \frac{1}{M} \sum_{k=0}^{M-1} \sum_{n=0}^{\infty} x(n) \left(z e^{-i2\pi k} \right)^{\frac{n}{M}} = \frac{1}{M} \sum_{k=0}^{M-1} X \left(z^{\frac{1}{M}} e^{-i\frac{2\pi k}{M}} \right). \quad (3.20)
 \end{aligned}$$

Полученное соотношение позволяет установить эквивалентность схем обработки сигналов, приведенных на рис. 3.23.



Рис. 3.23. Эквивалентные схемы обработки с понижающей дискретизацией

Пример 3.24. Показать, что схемы обработки сигналов на рис. 3.23а, б эквивалентны.

◀ Используя (3.20), для схемы рис. 3.23б имеем: $X_b(z) = X(z)H(z^M)$,

$$\begin{aligned}
 Y(z) &= \frac{1}{M} \sum_{k=0}^{M-1} X_b \left(z^{\frac{1}{M}} e^{-i\frac{2\pi k}{M}} \right) = \frac{1}{M} \sum_{k=0}^{M-1} X \left(z^{\frac{1}{M}} e^{-i\frac{2\pi k}{M}} \right) H \left(z e^{-i2\pi k} \right) = \\
 &= \frac{H(z)}{M} \sum_{k=0}^{M-1} X \left(z^{\frac{1}{M}} e^{-i\frac{2\pi k}{M}} \right),
 \end{aligned}$$

что соответствует выражению для преобразования сигнала по схеме рис. 3.23а. ▶

В вычислительном плане из двух эквивалентных схем обработки сигналов, представленных на рис. 3.23, предпочтительной является схема на рис. 3.23а. Действительно, при обработке по схеме рис. 3.23б после нахождения отклика $x_b(n)$ фильтра, имеющего передаточную функцию $G(z) = H(z^M)$, каждые $M-1$ из M отсчетов отклика $x_b(n)$ отбрасываются, т. е. оказываются напрасно вычисленными. Если же, как это определяет схема обработки сигнала на рис. 3.23а, сначала M -кратно сократить объем обрабатываемых данных в результате прореживания и уже после этого проводить обработку фильтром с передаточной функцией $H(z)$, то лишних вычислений не будет.

Упражнение. Покажите, что если имеется некоторый фильтр с ИХ $h(n)$ и ПФ $H(z) = Z\{h(n)\}$, то переход к фильтру с ПФ $G(z) = H(z^M)$ означает преобразование ИХ, состоящее в добавлении $M-1$ нулей после каждого отсчета $h(n)$:

$$g(n) = \begin{cases} h(n/M), & \text{при } n \bmod M = 0 \\ 0, & \text{при } n \bmod M \neq 0 \end{cases}$$

Упражнение. Докажите эквивалентность схем обработки рис. 3.24. *Указание:* сначала найдите связь Z -образов последовательностей $x_b(n)$ и $x(n)$, $y(n)$ и $x_a(n)$.



Рис. 3.24. Эквивалентные схемы с повышающей дискретизацией

Полифазное представление дискретного сигнала

Рассмотрим фильтр с импульсной характеристикой $h(n)$. Выберем некоторое целое число $K \geq 2$ и представим передаточную функцию этого фильтра в следующем виде:

$$H(z) = \sum_{n=0}^{\infty} h(n)z^{-n} = \sum_{k=0}^{K-1} \sum_{m=0}^{\infty} h(mK+k)z^{-(mK+k)} = \sum_{k=0}^{K-1} z^{-k} \sum_{m=0}^{\infty} \underbrace{h(mK+k)}_{\text{обозначим } e_k(m)} z^{-mK}.$$

Назовем *полифазными компонентами* ИХ $h(n)$ последовательности

$$\{e_k(m) = h(mK+k)\}_{m=0}^{\infty}$$

($k = 0, \dots, K-1$). Рассматривая каждую последовательность $\{e_k(m)\}_{m=0}^{\infty}$ как ИХ некоторого фильтра, имеющего ПФ

$$E_k(z) = \sum_{m=0}^{\infty} e_k(m)z^{-m},$$

получаем *полифазное представление* для ПФ исходного фильтра

$$H(z) = \sum_{k=0}^{K-1} z^{-k} E_k(z^K). \quad (3.21)$$

Форме представления ПФ (3.21) соответствуют эквивалентные структурные схемы фильтров на рис. 3.25, которые очевидным образом преобразуются в структуры, приведенные на рис. 3.26 и 3.27.

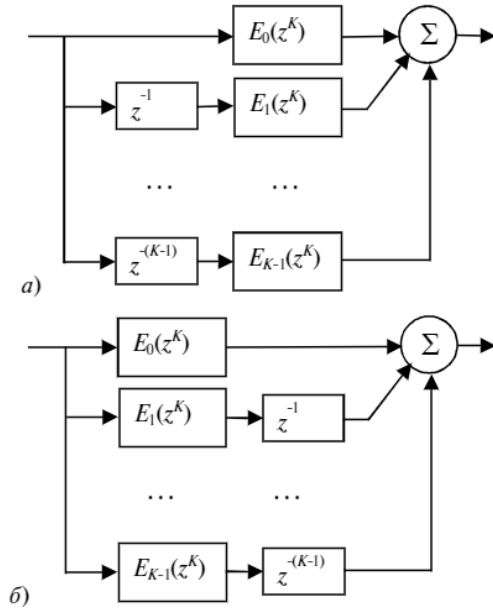


Рис. 3.25. Структурные схемы эквивалентных фильтров, соответствующие представлению передаточной функции системы в виде (3.21)

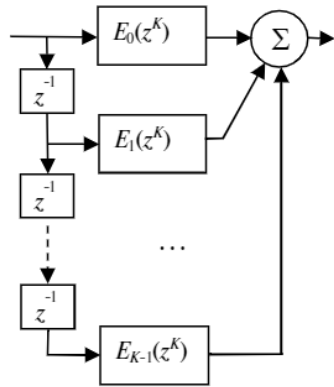


Рис. 3.26. Преобразованная структурная схема рис. 3.25a

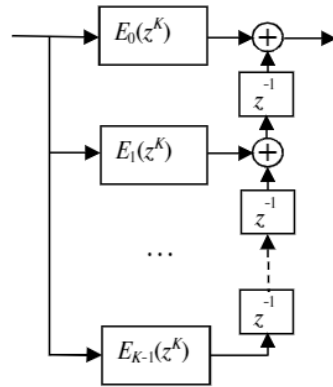


Рис. 3.27. Преобразованная структурная схема рис. 3.25б

Используя представление (3.21) для передаточной функции и соответствующие ему структуры реализации фильтра на рис. 3.26 и 3.27, можно построить схемы децимации и интерполяции сигналов, которые являются более эффективными в вычислительном плане по сравнению с рассмотренными ранее в разделе 3.9.

Полифазная реализация прореживающих систем

В простейшей схеме понижения частоты дискретизации, описанной в разделе 3.9, дискретный сигнал подвергается НЧ-фильтрации (с целью избежания наложения частот) и затем прореживается, см. рис. 3.28.

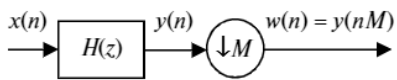


Рис. 3.28. Простейшая структура реализации прореживания сигнала

Фильтрацию сигнала в схеме рис. 3.28 заменим на параллельную обработку сдвинутых подпоследовательностей по схеме рис. 3.26. С учетом прореживания получим схему обработки, представленную на рис. 3.29. Применяя далее тождественное преобразование полученной схемы в соответствии с правилом рис. 3.23, из рис. 3.29 получаем структуру, приведенную на рис. 3.30.

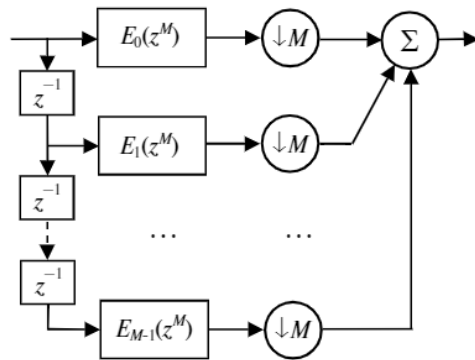


Рис. 3.29. Параллельная структура прореживающего фильтра на основе полифазного разложения ИХ

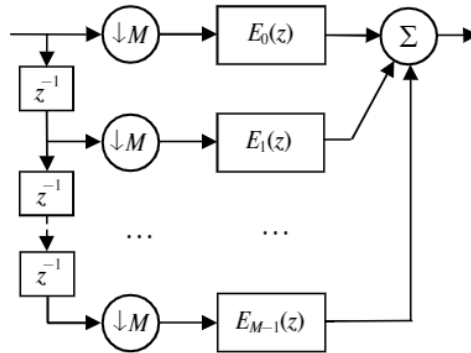


Рис. 3.30. Параллельная структура прореживающего фильтра на основе полифазных фильтров, эквивалентная структуре рис. 3.29

Пример 3.25. Пусть ФНЧ имеет ИХ $h(n) = \{h(0) = 1/\sqrt{2}, h(1) = 1/\sqrt{2}\}$.

Для структуры рис. 3.30 при $M = 2$ найдите ПФ $E_0(z)$ и $E_1(z)$.

◀ Для $e_k(n) = h(nM + k)$ имеем $e_0(n) = h(2n) = \{1/\sqrt{2}, 0, 0, \dots\}$, $e_1(n) = h(2n+1) = \{1/\sqrt{2}, 0, 0, \dots\}$. Отсюда $E_0(z) = Z\{h_0(n)\} = 1/\sqrt{2}$, $E_1(z) = Z\{h_1(n)\} = 1/\sqrt{2}$. Полифазные фильтры представляют собой обычные умножители на величину $1/\sqrt{2}$. Изобразите самостоятельно полученную структуру системы рис. 3.30. ▶

Оценим вычислительную сложность реализации прореживания по схемам рис. 3.28 и 3.30 по количеству необходимых умножений. Предположим, что входной сигнал $x(n)$ поступает с частотой один отсчет в единицу времени, а конечная ИХ $h(n)$ исходного фильтра рис. 3.28 имеет длину N отсчетов. Тогда обработка сигналов по схеме рис. 3.28 требует N умножений в единицу времени, необходимых для вычисления свертки. В схеме рис. 3.30 каждый фильтр E_k имеет ИХ примерно из (не более!) N/M отсчетов, а обработка сигналов проводится на частоте, в M раз меньшей. То есть каждый из M фильтров E_k рис. 3.30 производит N/M^2 умножений в единицу времени, всего для всей структуры рис. 3.30 получаем N/M необходимых умножений в единицу времени. Итого, имеем общее сокращение количества умножений при замене схемы фильтрации рис. 3.28 на схему рис. 3.30 в M раз.

Полифазная реализация интерполирующих систем

Простейшая структура для интерполяции сигнала, описанная в разделе 3.9 и использующая ФНЧ, приведена на рис. 3.31.

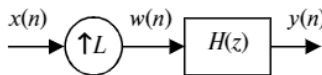


Рис. 3.31. Простейшая структура реализации интерполяции сигнала

Используя структуру рис. 3.27, реализацию схемы интерполяции можно осуществить при помощи фильтра рис. 3.32, который, в свою очередь, на основе эквивалентного преобразования рис. 3.24 может быть трансформирован в структуру, приведенную на рис. 3.33.

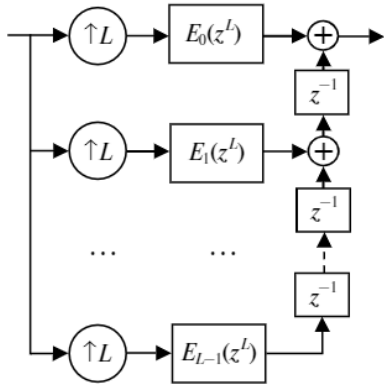


Рис. 3.32. Параллельная структура интерполяционного фильтра на основе полифазного разложения

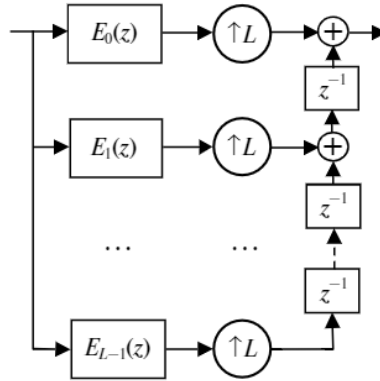


Рис. 3.33. Параллельная структура интерполяционного фильтра на основе полифазного разложения, эквивалентная рис. 3.31

Упражнение. Пусть ФНЧ имеет ИХ $h(n) = \{1/4, 1/2, 1/4\}$. Изобразите структуру рис. 3.33 для $L = 2$ и найдите передаточные функции $E_0(z)$, $E_1(z)$.

Несложно показать (проделайте это самостоятельно), что реализация интерполяционного фильтра по схеме рис. 3.33 требует в L раз меньшее количество умножений в единицу времени по сравнению со схемой на рис. 3.31.

3.11. Эффекты квантования в цифровых системах

В цифровых системах источниками ошибок квантования являются:

- аналого-цифровое преобразование — квантование отсчетов дискретных сигналов;
- умножение цифровых сигналов, результат которого может быть округлен или усечен в связи с ограниченностью разрядности представления чисел;
- квантование величин коэффициентов $\{a_m\}$ и $\{b_k\}$, определяющих разностное уравнение или ПФ системы (см. (3.8), (3.10)).

Мы будем рассматривать только равномерное квантование, когда проквантованное значение \hat{X} находится по правилу $\hat{X} = q \text{round}(X/q)$ (квантование с округлением) или $\hat{X} = q[X/q]$ (квантование с отбрасыванием дробной части числа, или квантование с усечением), где число q представляет собой шаг (дискрет) квантования.

Равномерное квантование отсчетов дискретного сигнала с округлением, что имеет отношение к ошибкам аналого-цифрового преобразования сигналов, было рассмотрено в разделе 2.12. Далее мы обобщим полученные ранее результаты на остальные источники ошибок квантования.

Для ошибки равномерного квантования $\varepsilon = \hat{X} - X$: $|\hat{X} - X| \leq q/2$ (квантование с округлением) и $|\hat{X} - X| \leq q$ (квантование с усечением). Эту ошибку удобно характеризовать средними значениями, например величиной (2.34): $E = M((\hat{X} - X)^2)$. При достаточно малой величине дискрета квантования q ошибку квантования $\varepsilon = \hat{X} - X$ можно считать случайной величиной, подчиняющейся равномерному закону распределения на отрезке $[-q/2; q/2]$ (при квантовании с округлением) или $[-q; 0]$ (при квантовании с усечением), вне зависимости от того, какой закон распределения имеет исходная случайная величина X . Для равномерного закона распределения $\varepsilon \sim R(-q/2; q/2)$ (квантование с округлением) математическое ожидание $M(\varepsilon) = 0$, а дисперсия ошибки квантования, совпадающая с величиной (2.34): $D(\varepsilon) = E = q^2/12$.



Пример 3.26. Найти средний квадрат ошибки квантования (2.34) для равномерного квантования с усечением.

◀ Поскольку $\varepsilon = (\hat{X} - X) \sim R(-q; 0)$, то $E = M(\varepsilon^2) = \int_{-q}^0 \frac{1}{q} x^2 dx = q^2/3$. ▶

Обозначив квантуемую последовательность как $X(n)$, а проквантованную — как $\hat{X}(n)$, для последовательности ошибок квантования $\varepsilon(n) = \hat{X}(n) - X(n)$ относительно всех возможных источников ошибок будем считать верными следующие допущения.

- Случайные величины $\varepsilon(n)$ имеют равномерное распределение $\varepsilon \sim R(-q/2; q/2)$ (при квантовании с округлением) или $\varepsilon \sim R(-q; 0)$ (при квантовании с усечением).
- Последовательность $\varepsilon(n)$ является стационарным процессом, эргодическим относительно математического ожидания.
- Элементы последовательности $\varepsilon(n)$ попарно некоррелированы.
- Последовательность $\varepsilon(n)$ не коррелирована с исходной последовательностью $X(n)$.

Далее будем рассматривать представление проквантованного числа в формате с фиксированной запятой, $x = \pm 0, b_1 b_2 \dots b_B$, тогда шаг равномерного квантования определяется весом младшего значащего разряда: $q = 2^{-B}$, где B — номер (после запятой) последнего разряда в двоичном представлении проквантованного числа.

Шум аналого-цифрового преобразования (АЦП)

Положим, что проквантованные с округлением отсчеты $\hat{X}(n)$ дискретного сигнала представлены в нормализованной цифровой форме формата с фиксированной запятой, когда все B значащих двоичных разрядов $\{b_k\}_{k=1}^B$ относятся к дробной части числа: $\hat{X}(n) = \pm 0, b_1 b_2 \dots b_B$. Шаг квантования в этом случае $q = 2^{-B}$, а дисперсия шума квантования АЦП $\varepsilon_\Lambda = \hat{X} - X$: $D(\varepsilon_\Lambda) = q^2/12 = 2^{-2B}/12$.

Мощность шума квантования принято выражать в логарифмическом масштабе как $P_\Lambda = 10 \lg D(\varepsilon_\Lambda)$ (дБ). Получим

$$P_\Lambda = 10 \lg \frac{2^{-2B}}{12} = -20B \lg 2 - 10 \lg 12 \approx -(6,02B + 10,79) \text{ дБ.}$$

Видим, что добавление одного значащего бита в АЦП влечет уменьшение мощности шума примерно на 6 дБ.

Пример 3.27. Пусть линейная система имеет ИХ $h(n)$. Предположим, что все арифметические операции, необходимые для определения отклика фильтра, реализуются точно. Определить дисперсию шума на выходе системы, вызванного шумом АЦП входного сигнала.

◀ Входной сигнал системы представим в виде: $\hat{X}(n) = X(n) + \varepsilon_\Lambda(n)$. В силу линейности системы для отклика на ее выходе, обусловленного шумом АЦП, получаем: $\varepsilon_{\Lambda, \text{вых}}(n) = \sum_{k=0}^{\infty} \varepsilon_\Lambda(n-k)h(k)$. Для дисперсии шума на выходе в силу некоррелированности входного шума имеем:

$$\begin{aligned} \sigma_{\Lambda, \text{вых}}^2 &= D(\varepsilon_{\Lambda, \text{вых}}(n)) = D\left(\sum_{k=0}^{\infty} \varepsilon_\Lambda(n-k)h(k)\right) = \sum_{k=0}^{\infty} D(\varepsilon_\Lambda(n-k)h(k)) = \\ &= \sum_{k=0}^{\infty} h^2(k)D(\varepsilon_\Lambda(n-k)) = D(\varepsilon_\Lambda(n)) \sum_{k=0}^{\infty} h^2(k) = \frac{q^2}{12} \sum_{k=0}^{\infty} h^2(k), \end{aligned}$$

где $q = 2^{-B}$ — дискрет равномерного квантования с округлением АЦП.

Тогда для мощности шума в логарифмических единицах получим $P_{\Lambda, \text{вых}} = 10 \lg D(\varepsilon_{\Lambda, \text{вых}}(n)) = P_\Lambda + 10 \lg \sum_{k=0}^{\infty} h^2(k)$ (дБ). ▶

Обобщая результаты решения примера 3.27, отметим, что если на вход линейного дискретного устройства поступает некоррелированный шум с дисперсией на входе σ_X^2 , то на выходе системы данный сигнал порождает шумовой отклик с дисперсией

$$\sigma_Y^2 = \sigma_X^2 \sum_{k=0}^{\infty} h^2(k). \quad (3.22)$$

Величина (3.22) определяется только значением дисперсии шума на входе и импульсной характеристикой системы.

В предположении, что все арифметические операции, необходимые для определения отклика фильтра, реализуются точно, оценим теперь максимальную амплитуду сигнала-ошибки на выходе системы, который вызван шумом АЦП входного сигнала. Имеем:

$$|\varepsilon_{\Lambda, \text{вых}}(n)| \leq \sum_{k=0}^{\infty} |\varepsilon_\Lambda(n-k)| |h(k)| \leq \max_j |\varepsilon_\Lambda(j)| \sum_{k=0}^{\infty} |h(k)| = \frac{q}{2} \sum_{k=0}^{\infty} |h(k)|.$$

Видим, что оценка максимальной амплитуды сигнала на выходе системы также определяется только импульсной характеристикой и свойствами входного шумового сигнала.

Собственный шум цифровой системы

Собственный шум цифровой системы — это выходной шум, вызванный эффектами квантования результатов умножения в системе. Действительно, если сомножители представлены B значащими двоичными разрядами (не считая знакового), то их произведение имеет уже $2B$ разрядов, и округление результата произведения до B значащих битов означает выполнение квантования с шагом $q = 2^{-B}$.

Оценка собственного шума проводится в три этапа.

1. Составляется линейная модель системы, учитывающая появление шумов квантования в точках выполнения операций умножения.
2. Вычисляются выходные реакции системы на каждый собственный шумовой сигнал.
3. Находятся итоговые оценки шума системы, учитывающие вклад каждой составляющей (каждого источника) шума.

Линейная модель умножителя как источника шума квантования показана на рис. 3.34. Если сомножители (отсчет сигнала $x(n)$, масштабирующий коэффициент a) и результат умножения представлены B значащими двоичными разрядами в дробной части чисел, то ошибка квантования считается равномерно распределенной случайной величиной $\varepsilon_y \sim R(-q_y/2; q_y/2)$, где $q_y = 2^{-B}$. Тогда для шума умножителя верны оценки:

$$m_y = M(\varepsilon_y) = 0, \quad \sigma_y^2 = D(\varepsilon_y) = q_y^2/12, \quad E_y = \max|\varepsilon_y| \leq q_y/2.$$

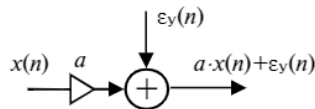


Рис. 3.34. Линейная модель умножителя

Пример 3.28. Для линейной дискретной системы, описываемой ПФ

$$H(z) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1}},$$

$|a_1| < 1$, построить линейную модель с источниками шумов квантования.

◀ Прямая форма реализации заданного фильтра приведена на рис. 3.35а. Линейная модель системы, учитывающая собственные шумы, связанные с умножителями, приведена на рис. 3.35б. Внешним шумовым сигналом указан также шум АЦП входного сигнала, обозначенный $\epsilon_\Lambda(n)$. ▶

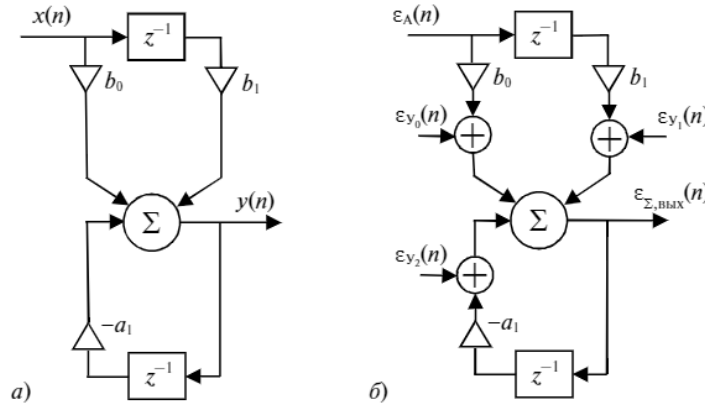


Рис. 3.35. Прямая форма схемы системы из примера 3.28 (а) и ее линейная модель с указанием собственных и внешних источников шумов квантования (б)

Найдем вклад каждого связанного с умножителями источника собственного шума в выходной шум системы. В силу линейности рассматриваемых систем выходной шум может быть представлен как суперпозиция шумов квантования, созданных каждым из собственных (внутренних) источников в системе:

$$\epsilon_{y, \text{вых}}(n) = \sum_k \epsilon_{y_k, \text{вых}}(n). \tag{3.23}$$

Для определения k -й составляющей выходного шумового сигнала $\epsilon_{y_k, \text{вых}}(n)$ необходимо рассмотреть «частичную» ИХ $h_k(n)$, характеризующую реакцию системы на единичное входное воздействие, приложенное в точке нахождения соответствующего источника шума $\epsilon_{y_k}(n)$. Тогда

$$\varepsilon_{y_k, \text{вых}}(n) = \sum_{m=0}^{\infty} h_k(m) \varepsilon_{y_k}(n-m).$$

Так как для шумовых составляющих $\varepsilon_{y_k}(n)$ принимается модель некоррелированной случайной последовательности с равномерным законом распределения $\varepsilon_{y_k} \sim R(-q_y/2; q_y/2)$, то получаем следующие оценки для выходных шумовых составляющих:

$$M(\varepsilon_{y_k, \text{вых}}(n)) = 0,$$

$$\sigma_{y_k, \text{вых}}^2 = D(\varepsilon_{y_k, \text{вых}}(n)) = D(\varepsilon_{y_k}(n)) \sum_{m=0}^{\infty} h_k^2(m) = \frac{q_y^2}{12} \sum_{m=0}^{\infty} h_k^2(m),$$

$$|\varepsilon_{y_k, \text{вых}}(n)| \leq \max_j |\varepsilon_{y_k}(j)| \sum_{m=0}^{\infty} |h_k(m)| = \frac{q_y}{2} \sum_{m=0}^{\infty} |h_k(m)|,$$

где предполагается, что все множители имеют B битов в представлении дробной части чисел и $q_y = 2^{-B}$.

Из последних соотношений на основании (3.23) и в предположении о взаимной некоррелированности шумов множителей следуют оценки для выходного собственного шума линейной системы. Так, для системы, имеющей L множителей, получаем:

$$M(\varepsilon_{y, \text{вых}}(n)) = 0, \quad (3.24)$$

$$\sigma_{y, \text{вых}}^2 = D(\varepsilon_{y, \text{вых}}(n)) = \frac{q_y^2}{12} \sum_{k=1}^L \sum_{m=0}^{\infty} h_k^2(m), \quad (3.25)$$

$$|\varepsilon_{y, \text{вых}}(n)| \leq \frac{q_y}{2} \sum_{k=1}^L \sum_{m=0}^{\infty} |h_k(m)|. \quad (3.26)$$

Пример 3.29. Для линейной дискретной системы из примера 3.28 получить оценку собственного шума (3.25).

◀ Система имеет три источника шума (множителя), см. рис. 3.35. Несложно видеть, что если подать в любой из точек приложения сигналов $\{\varepsilon_{y_k}(n)\}_{k=0}^2$ единичный импульс $x(n) = \tilde{\delta}(n) = \begin{cases} 1, & n=0 \\ 0, & n \neq 0 \end{cases}$, при отсутствии других сигналов, то на выходе получим последовательность $\tilde{h}(n) = (-a_1)^n$. Данная последовательность представляет собой ИХ, соответствующую части фильтра, обрабатывающей шумовые сигналы $\{\varepsilon_{y_k}(n)\}_{k=0}^2$. То есть $h_1(n) = h_2(n) = h_3(n) = (-a_1)^n$ и на основании соотношения (3.25) получаем: $\sigma_{y, \text{вых}}^2 = \frac{3q_y^2}{12} \sum_{m=0}^{\infty} (-a_1)^{2m} = \frac{q_y^2}{4(1-a_1^2)}$. ▶

Упражнение. Для линейной дискретной системы из примера 3.28 получить оценку собственного шума (3.26).

Полный выходной шум системы

В силу линейности рассматриваемых цифровых систем полный шум на выходе можно записать в виде:

$$\varepsilon_{\Sigma}(n) = \varepsilon_{\Lambda, \text{вых}}(n) + \varepsilon_{\Upsilon, \text{вых}}(n),$$

где, как и ранее, $\varepsilon_{\Lambda, \text{вых}}(n)$ — шум на выходе системы, вызванный ошибкой АЦП входного сигнала, а $\varepsilon_{\Upsilon, \text{вых}}(n)$ — собственный шум системы, связанный с квантованием чисел в умножителях. Поскольку две указанные составляющие выходного шума считаются независимыми, то

$$M(\varepsilon_{\Sigma}(n)) = M(\varepsilon_{\Lambda, \text{вых}}(n)) + M(\varepsilon_{\Upsilon, \text{вых}}(n)) = 0,$$

$$\sigma_{\Sigma}^2 = D(\varepsilon_{\Sigma}(n)) = D(\varepsilon_{\Lambda, \text{вых}}(n)) + D(\varepsilon_{\Upsilon, \text{вых}}(n)) = \sigma_{\Lambda, \text{вых}}^2 + \sigma_{\Upsilon, \text{вых}}^2, \quad (3.27)$$

$$|\varepsilon_{\Sigma}(n)| \leq |\varepsilon_{\Lambda, \text{вых}}(n)| + |\varepsilon_{\Upsilon, \text{вых}}(n)|. \quad (3.28)$$

Пример 3.30. Для линейной дискретной системы из примера 3.28 получить оценку дисперсии полного выходного шума (3.27).

◀ Обозначим: q_{Λ} — шаг квантования АЦП, q_{Υ} — шаг квантования умножителей. Фильтр из примера 3.28 имеет ИХ $h(0) = b_0$, а для $n \geq 1$

$$h(n) = b_0(-a_1)^n + b_1(-a_1)^{n-1} = (-a_1)^n \left(b_0 - \frac{b_1}{a_1} \right)$$

(убедитесь). Используя результат решения примера 3.27 и формулу (3.22), получаем для выходной составляющей шума АЦП входного сигнала:

$$\begin{aligned} \sigma_{\Lambda, \text{вых}}^2 &= \frac{q_{\Lambda}^2}{12} \sum_{k=0}^{\infty} h^2(k) = \frac{q_{\Lambda}^2}{12} \left(b_0^2 + \left(b_0 - \frac{b_1}{a_1} \right)^2 \sum_{k=1}^{\infty} (-a_1)^{2k} \right) = \\ &= \frac{q_{\Lambda}^2}{12} \left(b_0^2 + \left(b_0 - \frac{b_1}{a_1} \right)^2 \frac{a_1^2}{1-a_1^2} \right) = \frac{q_{\Lambda}^2}{12(1-a_1^2)} \left(b_0^2(1-a_1^2) + a_1^2 \left(b_0 - \frac{b_1}{a_1} \right)^2 \right). \end{aligned}$$

Окончательно, воспользовавшись результатом решения примера 3.29, имеем:

$$\sigma_{\Sigma}^2 = \sigma_{A, \text{вых}}^2 + \sigma_{Y, \text{вых}}^2 = \frac{1}{4(1-a_1^2)} \left(\frac{q_A^2}{3} \left(b_0^2(1-a_1^2) + a_1^2 \left(b_0 - \frac{b_1}{a_1} \right)^2 \right) + q_Y^2 \right). \quad \blacktriangleright$$

Упражнение. Представив структуру фильтра с ПФ из примера 3.28 в прямой канонической форме (см. раздел 3.4), получите оценку собственного шума системы (3.25) и сравните ее с результатом решения примера 3.29.

Упражнение. Для линейной дискретной системы из примера 3.28 получите оценку полного выходного шума (3.28).

Эффекты квантования коэффициентов цифровой системы

В цифровых системах коэффициенты разностных уравнений (передаточных функций) $\{a_m\}$ и $\{b_k\}$, определяющие множители в структурных схемах, также должны принимать квантованные значения. Точность представления коэффициентов определяется разрядностью умножителей, и погрешность этого представления равна половине веса младшего разряда двоичного представления числа.

Очевидно, что замена коэффициентов фильтра проквантованными значениями вносит погрешность и искажает частотную и импульсную характеристики фильтра, приводит к смещению нулей и полюсов ПФ.

Пример 3.31. Для рекурсивной системы с передаточной функцией $H(z) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}}$, где $a_1 = -1,6$ и $a_2 = 0,6425$, исследовать влияние на эту ПФ округления коэффициентов фильтра до одной (десятичной) значащей цифры после запятой.

◀ У исходной ПФ $H(z) = \frac{1}{1 - 1,6z^{-1} + 0,6425z^{-2}}$ имеется два комплексно сопряженных полюса: $z_{1,2} = 0,8 \pm 0,05i$. Система устойчива, так как $|z_1| = |z_2| = \sqrt{0,6425} < 1$. Округление коэффициентов фильтра до одной значащей цифры после запятой влечет изменение только одного коэффициента: $a_2 = 0,6425 \rightarrow 0,6$. Полученная ПФ $H(z) = \frac{1}{1 - 1,6z^{-1} + 0,6z^{-2}}$ имеет действительные полюсы: $z_1 = 0,6$

и $z_2 = 1$. Таким образом, после округления одного из коэффициентов система стала неустойчивой, так как один из полюсов лежит на окружности $|z| = 1$. ►

Как видим, квантование коэффициентов фильтра может даже привести к потере устойчивости системы.

На практике для учета искажений характеристик, связанных с конечной разрядностью двоичного представления коэффициентов системы, требуется проводить дополнительный анализ характеристик системы после квантования коэффициентов умножителей. Напомним также, что квантование коэффициентов в любом случае не может привести к потере системой устойчивости, если используется нерекурсивная структурная схема (трансверсальный фильтр).

Эффекты квантования коэффициентов системы в значительной степени обусловлены типом фильтра и его структуры, видом конкретной передаточной функции.

Предельные циклы цифровых систем

Если начиная с некоторого момента дискретного времени n_0 на вход системы перестает поступать возбуждение $x(n)$, т. е. $\forall n > n_0 \quad x(n) = 0$, то отклик устойчивой системы должен асимптотически стремиться к нулю: $\lim_{n \rightarrow \infty} y(n) = 0$. Однако из-за ошибок округления чисел в умножителях и возможных переполнений в сумматорах даже для устойчивых БИХ-систем это может не выполняться.

Пример 3.32. Для системы, описываемой разностным уравнением $y(n) = a y(n-1) + x(n)$ и использующей 3-битовое представление дробной части чисел в умножителе, найти реакцию на входное возбуждение $x(n) = 7\tilde{\delta}(n)/8$ (где $\tilde{\delta}(n)$ – единичный импульс) для $a = -0,5$.

◀ Прежде всего заметим, что система является устойчивой и имеет единственный полюс $z = a = -0,5$. Обозначим операцию округления числа u до трех значащих битов после запятой $Q(u)$, а число, имеющее двоичное представление с тремя битами после запятой \hat{u} , т. е. $\hat{u} = Q(u)$. Тогда алгоритм работы исследуемой системы будет описываться следующим уравнением: $\hat{y}(n) = Q(-0,5\hat{y}(n-1)) + x(n)$.



В момент дискретного времени $n = 0$ на вход системы подается импульс, который запишем в двоичной форме представления: $x(0) = 7/8 = 0,111_{(2)}$. Формирование выходного сигнала в соответствии с алгоритмом работы системы отражено в следующей таблице, где значения чисел представлены в двоичном виде.

| | | | | | | | | |
|--------------------|-------|---------|-------|--------|--------|---------|--------|-----|
| n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | ... |
| $-0,5\hat{y}(n-1)$ | 0 | -0,0111 | 0,010 | -0,001 | 0,0001 | -0,0001 | 0,0001 | ... |
| $\hat{y}(n)$ | 0,111 | -0,100 | 0,010 | -0,001 | 0,001 | -0,001 | 0,001 | ... |

Таким образом, начиная с момента дискретного времени $n = 3$ на выходе системы возникают незатухающие колебания $\pm 0,001_{(2)} = \pm 1/8$. ►

Эффект, наблюдаемый в примере 3.32, относится к *предельным циклам при нулевом входном воздействии* и обусловлен нелинейностью квантования в умножителях. Такой эффект может наблюдаться как при квантовании с округлением (рассмотренном в примере), так и при квантовании с усечением. Предельные циклы, т. е. незатухающие сигналы на выходе теоретически устойчивых систем в отсутствие входного возбуждения, могут возникать также в результате переполнения в сумматорах.

Предельные циклы систем ЦОС характеризуют ситуацию, когда перестают работать использовавшиеся нами ранее гипотезы о равномерном распределении ошибок и некоррелированности шума квантования и квантуемого сигнала. Это имеет место в рассмотренном примере 3.32, когда квантуемая величина дискретна, с шагом значений всего лишь в два раза меньшим дискрета квантования.

Обычно увеличение разрядности сумматоров и умножителей ведет к уменьшению амплитуды паразитных выходных колебаний и уменьшению вероятности их появления в отсутствие сигналов на входе, поэтому увеличение разрядности устройств является основным средством борьбы с предельными циклами.

Отметим, что возникновение предельных циклов принципиально невозможно в нерекурсивных КИХ-фильтрах, в структуре которых нет обратной связи. В отсутствие предельных циклов заключается одно из важных достоинств КИХ-фильтров.

Упражнение. В условиях примера 3.32 определить отклик системы для значения параметра $a = 0,5$.

3.12. Согласованный дискретный фильтр

Во многих практических приложениях возникает необходимость определения факта наличия или отсутствия заданного «эталонного» сигнала $u(t)$ в обрабатываемом сигнале $X(t)$, искаженном шумом. При этом эталонный сигнал может иметь неизвестную задержку времени появления t_0 и уровень β , т. е. фактически необходимо определить наличие сигнала $\beta u(t-t_0)$ на фоне шума. Например, $u(t)$ может представлять собой излучаемый радиолокационный сигнал, а $\beta u(t-t_0)$ — отраженный от цели принимаемый сигнал. Для выявления эталонного сигнала на фоне помех используют *согласованные* фильтры. Сформулируем задачу синтеза согласованного фильтра для случая дискретных сигналов.

Пусть $u(n) = \{u(0), u(1), \dots, u(N-1)\}$ — заданный эталонный сигнал конечной длительности, $X_m(n)$ — отсчеты дискретного белого шума, представляющие собой попарно некоррелированные случайные величины с дисперсией σ_x^2 и нулевым математическим ожиданием $m_x = 0$. На вход фильтра поступает случайный сигнал

$$X(n) = u(n) + X_m(n).$$

Необходимо синтезировать такой линейный дискретный фильтр, чтобы получить на его выходе отклик $Y(n) = v(n) + Y_m(n)$, для которого в момент дискретного времени $n = N-1$ обеспечивается максимальное отношение «мощностей» полезного сигнала и шума:

$$k_m = v^2(N-1) / M(Y_m^2(N-1)), \quad (3.29)$$

при этом полезный сигнал и шум определяются по формулам свертки:

$$v(N-1) = \sum_{n=0}^{N-1} u(N-1-n)h(n), \quad (3.30)$$

$$Y_m(N-1) = \sum_{n=0}^{N-1} X_m(N-1-n)h(n). \quad (3.31)$$

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 (3.30) и (3.31), 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 (3.31), в силу некоррелированности и нулевого математического ожидания отсчетов, имеем на основании формулы (3.22):

$$M(Y_m^2(N-1)) = D(Y_m(N-1)) = \sigma_x^2 \sum_{n=0}^{N-1} h^2(n). \quad (3.32)$$

Подставим полученное выражение (3.32) и соотношение (3.30) в формулу (3.29), получаем:

$$k_m = \frac{\left(\sum_{n=0}^{N-1} u(N-1-n)h(n) \right)^2}{\sigma_x^2 \sum_{n=0}^{N-1} h^2(n)}. \quad (3.33)$$

Нужно найти такую оптимальную ИХ $h(n) = \{h(0), h(1), \dots, h(N-1)\}$, которая обращает в максимум выражение (3.33).

Легко видеть, что для различных значений множителя α семейство ИХ $\{\alpha h(n)\}$ дает одно и то же значение k_m (3.33). Поэтому проведем нормировку искомой ИХ, потребовав выполнения условия $\sum_{n=0}^{N-1} h^2(n) = 1$ — тогда дисперсия (мощность) шумовой составляющей на выходе фильтра будет той же, что и во входном сигнале, как это следует из (3.32). Поиск безусловного максимума функции (3.33) сводится в таком случае к задаче об условном максимуме выражения, стоящего в числителе дроби (3.33), т.е. значения $v^2(N-1)$ (см. (3.30)). 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 $v(N-1)$ (3.30). Окончательно получаем следующую задачу условной минимизации:

$$F(h(0), h(1), \dots, h(N-1)) = v(N-1) = \sum_{n=0}^{N-1} u(N-1-n)h(n) \rightarrow \max,$$

при условии $\sum_{n=0}^{N-1} h^2(n) = 1$.

Составим для данной задачи функцию Лагранжа

$$L(h(0), h(1), \dots, h(N-1)) = \sum_{n=0}^{N-1} u(N-1-n)h(n) + \lambda \left(1 - \sum_{n=0}^{N-1} h^2(n)\right),$$

и из условий

$$\frac{\partial L(h(0), h(1), \dots, h(N-1))}{\partial h(m)} = u(N-1-m) - 2\lambda h(m) = 0$$

найдем её критическую точку $\mathbf{h} = (h(0), h(1), \dots, h(N-1))$:

$$h(m) = \frac{1}{2\lambda} u(N-1-m); \quad m = 0, 1, \dots, N-1.$$

Множитель $\lambda = \pm \frac{1}{2} \sqrt{\sum_{k=0}^{N-1} u^2(k)}$ находится здесь из принятого требования нормировки ИХ: $\sum_{n=0}^{N-1} h^2(n) = 1$.

Найденная критическая точка \mathbf{h} является максимумом функции Лагранжа, если её матрица вторых производных отрицательно определена:

$$\left[\frac{\partial^2 L(h(0), h(1), \dots, h(N-1))}{\partial h(m) \partial h(k)} \right]_{m, k=1}^{N-1} = \begin{bmatrix} -2\lambda & 0 & \dots & 0 \\ 0 & -2\lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -2\lambda \end{bmatrix} < 0.$$

В соответствии с критерием Сильвестра это требование выполняется, если принять $\lambda = \frac{1}{2} \sqrt{\sum_{k=0}^{N-1} u^2(k)} > 0$.

Окончательно получаем, что оптимальная ИХ, обращающая в максимум отношение сигнал/шум (3.29) или (3.33), имеет вид:

$$h(j) = \alpha u(N-1-j), \quad (3.34)$$

$j = 0, 1, \dots, N-1$, где

$$\alpha = \left(\sum_{k=0}^{N-1} u^2(k) \right)^{-1/2}. \quad (3.35)$$

Полученная ИХ (3.34) 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 (3.34) можно выбрать отличным от значения из (3.35). Несложно видеть, что при $\alpha > 0$ в равенстве (3.30) $v(N-1) > 0$, а для $\alpha < 0$ имеем $v(N-1) < 0$.

Пример 3.33. Определить, как связаны спектр (2.9) дискретного эталонного сигнала $\tilde{S}_u(v)$ и частотная характеристика $K(v)$ согласованного дискретного фильтра.

◀ Используя соотношения (3.34) и (3.35), найдем ЧХ согласованного фильтра:

$$\begin{aligned} K(v) &= \sum_{k=0}^{N-1} h(k)e^{-i2\pi vk} = \alpha \sum_{k=0}^{N-1} u(N-1-k)e^{-i2\pi vk} = \alpha \sum_{k=0}^{N-1} u(k)e^{-i2\pi v(N-1-k)} = \\ &= \alpha e^{-i2\pi v(N-1)} \sum_{k=0}^{N-1} u(k)e^{i2\pi vk} = \alpha e^{-i2\pi v(N-1)} \tilde{S}_u(v), \end{aligned}$$

где $\tilde{S}_u(v) = \sum_{k=0}^{N-1} u(k)e^{i2\pi vk}$ — спектр дискретного эталонного сигнала. ▶

Упражнение. Убедитесь самостоятельно, что передаточная функция согласованного фильтра может быть представлена в виде: $H(z) = \alpha z^{-(N-1)} U(1/z)$, где $U(z) = Z\{u(n)\} = \sum_{n=0}^{N-1} u(n)z^{-n}$, а параметр α определен по формуле (3.35).

Выбрав для согласованного фильтра с ИХ (3.34) значение масштабирующего множителя α в соответствии с (3.35), согласно (3.32) имеем: $\sigma_Y^2 = M(Y_m^2) = M(X_m^2) = \sigma_X^2$, а для квадрата полезного сигнала в момент $n = N - 1$ дискретного времени (см. (3.30), (3.34), (3.35)) получим:

$$v^2(N-1) = 1/\alpha^2 = \sum_{k=0}^{N-1} u^2(k) \quad (3.36)$$

— 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 (3.30), (3.34), (3.35), получаем:

$$\begin{aligned} v(J) &= \sum_{k=0}^{N-1} h(k)u(J-k) = \sum_{k=-\infty}^{\infty} h(k)u(J-k) = \frac{1}{\alpha} \sum_{k=-\infty}^{\infty} h(k)h(N-1-J+k) = \\ &= \frac{1}{2\alpha} \left(\sum_{k=-\infty}^{\infty} (h(N-1-J+k) - h(k))^2 + \underbrace{\sum_{k=-\infty}^{\infty} h^2(N-1-J+k)}_1 + \underbrace{\sum_{k=-\infty}^{\infty} h^2(k)}_1 \right) = \\ &= \frac{1}{2\alpha} \left(2 - \sum_{k=-\infty}^{\infty} (h(N-1-J+k) - h(k))^2 \right) \leq \frac{1}{\alpha} = \sqrt{E_u}. \end{aligned}$$

Таким образом,

$$\max_{n \in \mathbb{Z}} v(n) = v(N-1) = \sqrt{E_u}. \quad (3.37)$$

Как отмечалось ранее, согласованные фильтры применяются для обнаружения заданных сигналов на фоне помех в различных практических приложениях: радиолокации, распознавании образов и др.

Рассмотрим задачу, для решения которой используется согласованный фильтр. Эта задача состоит в том, чтобы определить, подан ли на вход фильтра эталонный сигнал $u(n)$, искаженный аддитивными помехами (некоррелированным шумом): $X(n) = u(n) + X_m(n)$, или же входная последовательность не содержит эталона: $X(n) = X_m(n)$.

Положим, что шум имеет нулевое математическое ожидание. При наличии на входе фильтра эталонного сигнала $u(n)$ математическое ожидание выходного сигнала $Y(N-1) = v(N-1) + Y_m(N-1)$ при выполнении условий (3.34), (3.35) равно

$$M(Y(N-1)) = v(N-1) = \sqrt{E_u},$$

для дисперсии, с учетом (3.32), получаем

$$D(Y(N-1)) = D(v(N-1) + Y_m(N-1)) = D(Y_m(N-1)) = \sigma_x^2,$$

а отношение (3.29) принимает значение $k_m = E_u / \sigma_x^2$.

При отсутствии эталонного сигнала во входной последовательности, когда $X(n) = X_m(n)$, получаем:

$$M(Y(N-1)) = M(Y_m(N-1)) = 0, \quad D(Y(N-1)) = D(Y_m(N-1)) = \sigma_x^2.$$

Таким образом, в момент дискретного времени $n = N-1$ дисперсия сигнала $Y(N-1)$ на выходе согласованного фильтра при выполнении условия (3.35) всегда равна дисперсии шума входного сигнала σ_x^2 . Математическое ожидание отклика фильтра $Y(N-1)$ равно нулю в случае отсутствия эталонного сигнала на входе и значению $\sqrt{E_u}$ если эталон $u(n)$ присутствует во входном сигнале.

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 $T \in (0; \sqrt{E_u})$ оказывают влияние два противодействующих 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 (необнаружения) эталона при его наличии. Очевидно, что для снижения вероятностей различного рода ошибок энергия эталонного сигнала E_u (3.36) должна быть по возможности большей и превосходить мощность шума σ_x^2 .

Пример 3.34. Для нормального некоррелированного шума $X_{ш}(n)$ на входе согласованного фильтра определить значение порога T , при котором вероятность пропуска (необнаружения) эталонного сигнала равна вероятности ложного обнаружения.

◀ В силу композиционной устойчивости нормального закона распределения выходной шум (3.31) также будет иметь нормальный закон, причем при выполнении условия (3.35) имеем $\sigma_Y^2 = \sigma_x^2$. Как отмечено выше, при наличии эталонного сигнала на входе $M(Y(N-1)) = \sqrt{E_u}$, при его отсутствии — $M(Y(N-1)) = 0$. Соответствующие графики функций плотности распределения вероятностей случайной величины $Y(N-1)$ приведены на рис. 3.36. Из рисунка видно, что вероятность необнаружения эталонного сигнала при его наличии на входе фильтра (площадь криволинейной трапеции S_1) равна вероятности ложного обнаружения эталонного сигнала при его отсутствии на входе фильтра (площади криволинейной трапеции S_2), если порог срабатывания детектора на выходе согласованного фильтра принять равным $T = \sqrt{E_u}/2$.

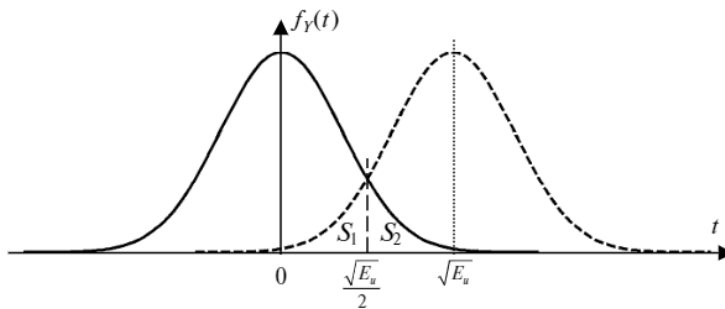


Рис. 3.36. Графики функций плотности распределения случайной величины $Y(N-1)$ — отклика согласованного фильтра: сплошная линия — эталонный сигнал на входе отсутствует, пунктирная линия — эталонный сигнал на входе присутствует

Вероятности данных ошибок будут тем меньше, чем меньше дисперсия входного шума σ_x^2 и больше энергия эталонного сигнала E_u . ►

Другая возникающая на практике задача состоит в том, чтобы в искаженной шумом входной последовательности

$$X(n) = \beta u(n - M) + X_{\text{ш}}(n)$$

с неизвестными уровнем $\beta > 0$ и моментом времени M начала эталонного сигнала определить «местоположение» эталона во входной последовательности, т.е. значение числа M . Математическое ожидание $M(Y(n)) = \beta v(n - M)$ отклика $Y(n) = \beta v(n - M) + Y_{\text{ш}}(n)$ согласованного фильтра принимает максимальное значение в момент n дискретного времени, который находится (см. (3.37)) из условия $n - M = N - 1$:

$$M(Y(M + N - 1)) = \beta v(M + N - 1) = \beta \sqrt{E_u} = \beta \max_n v(n),$$

где E_u — энергия эталона (3.36). Поэтому для определения числа M необходимо найти в выходной последовательности момент J дискретного времени с максимальным значением отклика: $Y(J) = \max_n Y(n)$, и положить $M = J - N + 1$. Из-за влияния шумов найденное так значение M может оказаться неверным. Вновь для повышения достоверности результата желательно, чтобы энергия $\beta^2 E_u$ полезного сигнала, поступающего на вход фильтра, была как можно большей.

Обе описанные задачи могут ставиться совместно, когда нужно 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 шум. Тогда, если для некоторого порога T и индекса J выполняется условие $Y(J) > T$, то принимается решение о том, что эталонный сигнал присутствует, а его первый отсчет приходится на момент дискретного времени $M = J - N + 1$.

При рассмотрении задач выделения эталона на фоне шума мы считали сигнал $u(n)$ заданным. Для повышения достоверности и точности определения момента появления эталона используются специально синтезированные эталонные сигналы.

Пример 3.35. Из эталонных сигналов $u_1(n) = \{1, 1, 1, -1, -1, 1, -1\}$ и $u_2(n) = \{1, 1, 1, 1, 1, 1, 1\}$ одинаковой длительности $N = 7$ и энергии $E_u = 7$ выбрать тот, который позволяет более достоверно определять момент появления эталонного сигнала на фоне шумов.

◀ Найдем отклик согласованного с сигналом $u_1(n)$ фильтра, имеющего ИХ $h_1(n) = u_1(6 - n)$ (см. (3.34) при $\alpha = 1$). Используя формулу свертки $y_1(j) = \sum_{k=0}^j u_1(k)h_1(j-k) = \sum_{k=0}^j u_1(k)u_1(6-j+k)$, получаем для $j = 0, 1, 2, \dots$:

$$y_1(j) = \{-1, 0, -1, 0, -1, 0, 7, 0, -1, 0, -1, 0, -1, 0, 0, \dots\}.$$

Аналогично, для $y_2(j) = \sum_{k=0}^j u_2(k)u_2(6-j+k)$, $j = 0, 1, 2, \dots$, находим:

$$y_2(j) = \{1, 2, 3, 4, 5, 6, 7, 6, 5, 4, 3, 2, 1, 0, 0, \dots\}.$$

В отсутствие шумов оба согласованных фильтра дают максимальное значение отклика $y_1(6) = y_2(6) = E_u = 7$. Однако первый из этих фильтров будет в меньшей степени подвержен влиянию шума на точность нахождения максимума отклика фильтра. Например, если шумовая составляющая на выходе согласованного фильтра принимает значения $Y_{\text{ш}} \in [-3; 3]$, то для отклика первого согласованного фильтра $Y_1(j) = y_1(j) + Y_{\text{ш}}$ максимальное значение по-прежнему будет приходиться на отсчет с номером $j = 6$: $Y_1(6) \in [4; 10]$; $\forall j \neq 6: Y_1(j) \in [-4; 3]$. При таком же шуме $Y_{\text{ш}} \in [-3; 3]$ для второго согласованного фильтра максимальное значение выходного сигнала $Y_2(j) = y_2(j) + Y_{\text{ш}}$ возможно для любого значения $j = 1, 2, \dots, 11$ и достоверное определение момента появления эталонного сигнала становится невозможным. ▶

Обобщая результаты примера 3.35, можно отметить, что среди эталонных сигналов с одинаковыми длительностью и энергией более пригодным для выделения на фоне помех является тот сигнал, который дает большее превышение отклика согласованного фильтра в момент дискретного времени $N - 1$ по сравнению со значениями отклика в другие моменты времени. Для синтеза соответствующих сигналов разработаны специальные методы, изучение которых выходит за рамки нашего рассмотрения.

Упражнение. Из двух эталонных сигналов $\{1, -1, 1, 1, -1, -1\}$ и $\{1, -1, 1, -1, 1, -1\}$ выберите тот, который предпочтительнее использовать для определения момента появления эталонного сигнала на фоне шумов.

3.13. Линейная дискретная система как генератор случайных сигналов. Полюсная модель сигналов

Ранее мы рассмотрели некоторые примеры анализа случайных воздействий на линейные дискретные системы. Обобщим полученные результаты. Пусть входным воздействием линейной дискретной системы с ИХ $h(n)$ является стационарный в широком смысле случайный процесс $X(n)$, для которого известно математическое ожидание $m_X = M(X(n))$ и ковариационная функция $K_X(k) = \text{cov}(X(n-k), X(n)) = R_X(k) - m_X^2$, где использовано обозначение $R_X(k) = M(X(n-k)X(n))$. Функцию $R_X(k)$, представляющую собой второй начальный момент, в технической литературе часто называют корреляционной (или автокорреляционной), хотя такой термин является не совсем корректным¹. Для отклика системы

$$Y(n) = \sum_{k=0}^{\infty} X(n-k)h(k), \quad (3.38)$$

очевидно, имеем

$$m_Y = M(Y(n)) = m_X \sum_{k=0}^{\infty} h(k).$$

Если отсчеты входного сигнала попарно некоррелированы ($X(n)$ представляет собой дискретный белый шум), то дисперсия выходного сигнала (3.38) определяется соотношением (3.22):

$$\sigma_Y^2 = \sigma_X^2 \sum_{k=0}^{\infty} h^2(k).$$

При прохождении через ЛДС преобразование сигнала $X(n)$, отсчеты которого являются коррелированными, удобнее характеризовать в терминах спектральных характеристик.

¹ В математической литературе автокорреляционной называют обычно функцию $R_X(k) = K_X(k)/\sigma_X^2 = K_X(k)/K_X(0)$.

Пример 3.36. Показать, что случайный сигнал $Y(n)$ на выходе ЛДС, полученный в результате преобразования (3.38) из стационарного случайного процесса $X(n)$, имеет спектр мощности

$$\tilde{s}_Y(\omega) = |H(e^{i\omega})|^2 \tilde{s}_X(\omega), \quad (3.39)$$

где $\tilde{s}_Y(\omega) = \sum_{n=-\infty}^{\infty} K_Y(n) e^{-i\omega n}$, $\tilde{s}_X(\omega) = \sum_{n=-\infty}^{\infty} K_X(n) e^{-i\omega n}$ – спектры мощности сигналов $Y(n)$ и $X(n)$ соответственно, а $K(\omega) = H(e^{i\omega}) = \sum_{n=-\infty}^{\infty} h(n) e^{-i\omega n}$ – частотная характеристика ЛДС.

◀ Так как ковариационная функция (КФ) $K_X(k) = \text{cov}(X(n-k), X(n))$ не зависит от значения $m_X = M(X(n))$, для упрощения выкладок положим $m_X = 0$. Тогда $K_X(k) = R_X(k) = M(X(n-k)X(n))$ и $m_Y = M(Y(n)) = m_X \sum_{k=0}^{\infty} h(k) = 0$, $K_Y(k) = R_Y(k) = M(Y(n-k)Y(n))$.

Найдем КФ $K_Y(k)$ дискретного случайного процесса $Y(n)$. Воспользовавшись (3.38), имеем:

$$\begin{aligned} K_Y(k) &= M(Y(n-k)Y(n)) = M\left(\sum_{m=0}^{\infty} X(n-k-m)h(m) \sum_{j=0}^{\infty} X(n-j)h(j)\right) = \\ &= \sum_{m=0}^{\infty} h(m) \sum_{j=0}^{\infty} h(j) M(X(n-k-m)X(n-j)) = \sum_{m=0}^{\infty} h(m) \sum_{j=0}^{\infty} h(j) K_X(j-k-m). \end{aligned}$$

Спектром мощности случайного процесса является, по определению, преобразование Фурье его КФ (см. раздел 1.8). По дискретной КФ $K_Y(k)$ находим спектр мощности

$$\begin{aligned} \tilde{s}_Y(\omega) &= \sum_{k=-\infty}^{\infty} K_Y(k) e^{-i\omega k} = \sum_{m=0}^{\infty} h(m) \sum_{j=0}^{\infty} h(j) \sum_{k=-\infty}^{\infty} K_X(j-k-m) e^{-i\omega k} = \\ &= \sum_{m=0}^{\infty} h(m) e^{i\omega m} \sum_{j=0}^{\infty} h(j) e^{-i\omega j} \sum_{k=-\infty}^{\infty} K_X(j-k-m) e^{-i\omega(k+m-j)} = \\ &= \overline{K(\omega)} K(\omega) \sum_{k=-\infty}^{\infty} K_X(-k) e^{-i\omega k} = |K(\omega)|^2 \sum_{k=-\infty}^{\infty} K_X(k) e^{-i\omega k} = |H(e^{i\omega})|^2 \tilde{s}_X(\omega), \end{aligned}$$

где предпоследний переход осуществлен на основании четности ковариационной функции: $K_X(-k) = K_X(k)$. ▶

Упражнение. Пусть $y(n) = \sum_{k=0}^{\infty} x(n-k)h(k)$, $n \in \mathbb{Z}$. Покажите, что для спектра дискретного сигнала $\tilde{S}_y(\omega) = \sum_{n=-\infty}^{\infty} y(n) e^{-i\omega n}$ (см. также определение (2.9) в терминах линейной частоты) справедливо соотношение:

$$\tilde{S}_y(\omega) = H(e^{i\omega}) \tilde{S}_x(\omega), \quad (3.40)$$

где $\tilde{S}_x(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{-i\omega n}$ – спектральная плотность сигнала $x(n)$,
 $K(\omega) = H(e^{i\omega}) = \sum_{n=0}^{\infty} h(n)e^{-i\omega n}$ – ЧХ некоторого фильтра с ИХ $\{h(n)\}_{n=0}^{\infty}$.

Из формулы (3.40) следует, что при подаче на вход ЛДС единичного импульса $x(n) = \tilde{\delta}(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$, имеющего спектр $\tilde{S}_x(\omega) = 1$, выходной сигнал $y(n)$ будет иметь спектр, совпадающий с ЧХ системы: $\tilde{S}_y(\omega) = H(e^{i\omega})$. Если же входное воздействие ЛДС моделируется случайным сигналом $X(n)$ с постоянным спектром мощности $\tilde{S}_x(\omega) = 1$ (белым шумом), то спектр мощности случайного сигнала на выходе, как это следует из (3.39), будет совпадать с квадратом АЧХ системы: $\tilde{S}_y(\omega) = |H(e^{i\omega})|^2 = H(z)H(z^{-1})|_{z=e^{i\omega}}$. Указанные свойства ЧХ используются для моделирования сигналов: в ряде приложений оказывается удобно описывать сигнал как отклик ЛДС на единичный импульс (для детерминированной модели), или на входное возбуждение системы белым шумом (для стохастической модели). При таком подходе ЛДС выступает в роли генератора модельного сигнала, спектр которого задан ЧХ ЛДС.

Рассматривая дискретный сигнал $x(n)$ на некотором интервале наблюдения при $n = 0, 1, \dots, L-1$, для моделирования этого сигнала необходимо так подобрать параметры ЛДС, чтобы ее ИХ хорошо аппроксимировала реализацию моделируемого сигнала на данном интервале. Параметрами, определяющими модель сигнала, для рассматриваемых нами ЛДС являются коэффициенты разностного уравнения (3.8) (передаточной функции (3.10)) $\{a_m\}_{m=1}^M, \{b_k\}_{k=0}^N$. Для стохастической модели при этом предполагается, что на рассматриваемом интервале наблюдения $n = 0, 1, \dots, L-1$ сигнал $X(n)$ является стационарным.

Наибольшее распространение для параметрического моделирования сигналов получила так называемая *полюсная* модель, в которой нерекурсивная часть ЛДС содержит только один коэффициент $b_0 \neq 0$ и передаточная функция (3.10) принимает вид

$$H(z) = \frac{b_0}{1 + \sum_{m=1}^M a_m z^{-m}}, \tag{3.41}$$

а соответствующее разностное уравнение (3.8) записывается как

$$y(n) = b_0 x(n) - \sum_{m=1}^M a_m y(n-m). \quad (3.42)$$

Такая модель называется полюсной (в оригинале *all-pole*), поскольку ПФ (3.41) с точностью до некоторого масштабного множителя целиком определяется расположением своих полюсов $\{p_k\}_{k=1}^M$ – нулей (возможно, кратных) многочлена

$$Q(z) = 1 + \sum_{m=1}^M a_m z^{-m} = K \prod_{k=1}^M (1 - p_k z^{-1}),$$

стоящего в знаменателе (3.41).

Упражнение. Показать, что полюсная система второго порядка с ПФ

$$H(z) = \frac{1}{1 - 2a(\cos \omega_0)z^{-1} + a^2 z^{-2}},$$

где $0 < a < 1$ и $0 < \omega_0 < \pi$, имеет ИХ, представляющую собой затухающее дискретное гармоническое колебание, а именно:

$$h(n) = \frac{a^n}{\sin \omega_0} \sin(\omega_0(n+1)).$$

Изобразите карту нулей и полюсов этой системы (аналогично рис. 3.15).

Полюсное описание используется, например, для параметрического моделирования сигналов при сжатии данных в кодах речи — *вокодерах*. Речевой сигнал разбивается при этом на фрагменты (*кадры*), в которых он аппроксимируется ИХ полюсной ЛДС. Например, в стандартизованном вокодере LPC-10 порядок полинома знаменателя ПФ (3.41) $M = 10$, при этом количество отсчетов сигнала в кадре $L = 120$ [43]. Для локального описания сигнала в каждом кадре используются свои параметры модели (3.41), дающие оптимальную (обычно в среднеквадратичном смысле) аппроксимацию. Поскольку при формировании декодированного речевого сигнала $y(n)$ его отсчеты формируются по предыдущим восстановленным в соответствии с формулой (3.42), то процесс декодирования можно интерпретировать как построение линейного прогноза значения сигнала $y(n)$ по уже известным отсчетам $y(n-1), \dots, y(n-M)$, полученным ранее. Поэтому описанный способ параметрического представления сигналов называют также *кодированием с линейным предсказанием* (linear prediction coding — LPC).

Полюсные ЛДС используются также и для моделирования случайных процессов. Дискретный случайный процесс $Y(n)$, представляющий собой отклик системы (3.42) при входном возбуждении в виде белого шума, называют *авторегрессионным*.

Пример 3.37. Найти спектр мощности и ковариационную функцию сигнала, генерируемого системой, работа которой описывается разностным уравнением $y(n) = ay(n-1) + bx(n)$, при ее возбуждении белым шумом.

◀ ПФ фильтра имеет вид $H(z) = \frac{b}{1 - az^{-1}} = b \frac{z}{z - a}$, что соответствует простейшей полюсной модели (с единственным параметром — полюсом $z = a$ и коэффициентом $b_0 = b$). При подаче на вход белого шума, имеющего спектр мощности $\tilde{s}_x(\omega) = C$, для генерируемого выходного случайного сигнала имеем в соответствии с (3.39) и (3.16) спектр мощности

$$\begin{aligned} \tilde{s}_y(\omega) &= |H(e^{j\omega})|^2 C = H(z)H(z^{-1}) \Big|_{z=e^{j\omega}} C = \frac{be^{j\omega}}{e^{j\omega} - a} \cdot \frac{be^{-j\omega}}{e^{-j\omega} - a} C = \\ &= \frac{Cb^2}{(a - \cos\omega)^2 + \sin^2\omega} = \frac{Cb^2}{1 - 2a\cos\omega + a^2}. \end{aligned}$$

Убедитесь, что полученному спектру мощности $\tilde{s}_y(\omega)$ соответствует ковариационная функция $K_Y(n) = \frac{Cb^2}{1 - a^2} a^{|n|}$, $|a| < 1$, т.е.

$$\tilde{s}_y(\omega) = \sum_{n=-\infty}^{\infty} K_Y(n) e^{-j\omega n} = \frac{Cb^2}{1 - a^2} \sum_{n=-\infty}^{\infty} a^{|n|} e^{-j\omega n} = \dots = \frac{Cb^2}{1 - 2a\cos\omega + a^2}$$

(опущенные преобразования проделайте самостоятельно). ▶

Упражнение. Покажите, что значение коэффициента корреляции $\rho_Y(k) = \text{cov}(Y(n), Y(n+k)) / \sigma_Y^2$ отсчетов случайного сигнала, генерируемого системой из примера 3.37: $\rho_Y(k) = a^{|k|}$.

Определение. Дискретным марковским процессом первого порядка (или простой цепью Маркова) называется случайная дискретная последовательность $Y(n)$, $n \in \mathbb{Z}$, для которой закон распределения случайной величины $Y(n)$ полностью определяется реализацией предыдущего состояния $Y(n-1) = y(n-1)$. Таким образом, простая цепь Маркова полностью описывается следующей функцией условного распределения: $F_Y(y_n | y_{n-1}) = P\{Y(n) < y_n | Y(n-1) = y_{n-1}\}$.

Случайный сигнал $Y(n)$ из примера 3.37 является простым частным случаем дискретного марковского процесса.

Упражнение. Пусть авторегрессионный случайный процесс $Y(n)$ представляет собой отклик полюсной ЛДС из примера 3.37 на возбуждение белым шумом $X(n)$, т.е. $Y(n) = aY(n-1) + bX(n)$, и известна функция распределения $F_X(x) = P\{X(n) < x\}$. Покажите, что закон распределения $Y(n)$ полностью определяется реализацией состояния $Y(n-1) = y(n-1)$ (т.е. соответствующий дискретный случайный процесс – марковский), причем функция распределения $F_Y(y_n | y_{n-1}) = P\{Y(n) < y_n | Y(n-1) = y_{n-1}\}$ имеет вид:

$$F_Y(y_n | y_{n-1}) = F_X\left(\frac{y_n - ay_{n-1}}{b}\right).$$

Марковские процессы представляют собой полезную модель для описания сигналов, часто используемую в приложениях.

3.14. Фильтр Калмана

Положим, что сигнал $Y(n)$ можно представить моделью марковского процесса, который возникает на выходе полюсной системы первого порядка, возбужденной дискретным белым шумом $G(n)$ с дисперсией σ_G^2 и нулевым математическим ожиданием (см. пример 3.37):

$$Y(n) = aY(n-1) + G(n), \quad (3.43)$$

где для устойчивости системы $-1 < a < 1$. После прохождения по некоторому каналу сигнал (3.43) получает амплитудное искажение, определяемое множителем c , и аддитивное наложение белого шума $V(n)$ с нулевым средним и дисперсией σ_V^2 . Результат воздействия помех на сигнал $Y(n)$ дает следующую модель искаженного сигнала:

$$X(n) = cY(n) + V(n). \quad (3.44)$$

Пример 3.38. Покажите, что для модели (3.44) при $k \neq 0$ коэффициент корреляции

$$\rho_X(k) = \frac{\text{cov}(X(n), X(n+k))}{\sigma_X^2} = a^{|k|} \left(1 - \frac{\sigma_V^2}{\sigma_X^2}\right),$$

где $\sigma_X^2 = D(X(n))$.

◀ В силу некоррелированности белого шума $V(n)$ с отсчетами сигнала $Y(n)$, с учетом того, что $M[V(n)] = M[Y(n)] = M[X(n)] = 0$, из (3.44) получаем (убедитесь самостоятельно, проделав опущенные выкладки):

$$K_X(k) = \text{cov}(X(n), X(n+k)) = M[(cY(n)+V(n))(cY(n+k)+V(n+k))] = \dots \\ \dots = c^2 \text{cov}(Y(n), Y(n+k)) = c^2 K_Y(k) = c^2 \sigma_Y^2 \rho_Y(k),$$

где обозначено: $\sigma_Y^2 = D(Y(n))$, $\rho_Y(k) = K_Y(k)/\sigma_Y^2$.

Поскольку (см. 3.44) $\sigma_X^2 = c^2 \sigma_Y^2 + \sigma_V^2$, то

$$\rho_X(k) = \frac{K_X(k)}{\sigma_X^2} = \frac{\sigma_X^2 - \sigma_V^2}{\sigma_X^2} \rho_Y(k) = \rho_Y(k) \left(1 - \frac{\sigma_V^2}{\sigma_X^2}\right) = a^{|k|} \left(1 - \frac{\sigma_V^2}{\sigma_X^2}\right),$$

где последний переход выполнен на основании полученного при решении примера 3.37 выражения для ковариационной функции $K_Y(k)$, т.к. $\rho_Y(k) = K_Y(k)/K_Y(0)$. Видим, что в сигнале $X(n)$ коэффициент корреляции отсчетов $\rho_X(k)$ ослабевает по сравнению с $\rho_Y(k)$ тем сильнее, чем больше дисперсия шума σ_V^2 , искажающего полезный сигнал $Y(n)$ в соответствии с моделью (3.44). ▶

По доступным значениям $X(n)$ будем искать оценку для полезного сигнала $Y(n)$ в виде

$$\hat{Y}(n) = b(n)\hat{Y}(n-1) + k(n)X(n), \quad (3.45)$$

что соответствует структурной схеме простейшего рекурсивного фильтра, представленной на рис. 3.37. Предполагается, что в процессе работы фильтра коэффициенты $b(n)$, $k(n)$ могут изменяться.

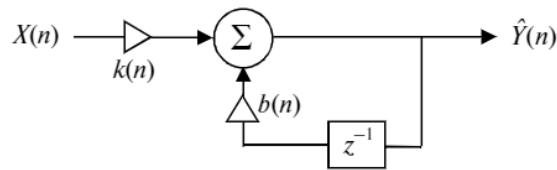


Рис. 3.37. Структурная схема рекурсивного фильтра первого порядка для формирования оценки (3.45)

Обозначим $E(n) = \hat{Y}(n) - Y(n) = b(n)\hat{Y}(n-1) + k(n)X(n) - Y(n)$ и найдем оптимальную с точки зрения минимизации среднеквадратичной ошибки

$$e(n) = M[E^2(n)] = M\left[\left(b(n)\hat{Y}(n-1) + k(n)X(n) - Y(n)\right)^2\right] \quad (3.46)$$

оценку (3.45). Для этого частные производные от $e(n)$ по искомым параметрам $b(n)$ и $k(n)$ приравняем нулю:

$$\frac{\partial e(n)}{\partial b(n)} = 2M\left[\left(b(n)\hat{Y}(n-1) + k(n)X(n) - Y(n)\right)\hat{Y}(n-1)\right] = 0, \quad (3.47)$$

$$\frac{\partial e(n)}{\partial k(n)} = 2M\left[\left(b(n)\hat{Y}(n-1) + k(n)X(n) - Y(n)\right)X(n)\right] = 0. \quad (3.48)$$

С учетом введенного для $E(n)$ обозначения можно записать равенства (3.47) и (3.48) в эквивалентной форме соответственно:

$$M[E(n)\hat{Y}(n-1)] = 0, \quad M[E(n)X(n)] = 0.$$

Тогда (см. (3.45)) $\forall n$:

$$M[E(n)\hat{Y}(n)] = b(n)M[E(n)\hat{Y}(n-1)] + k(n)M[E(n)X(n)] = 0. \quad (3.49)$$

Представим (3.47) в виде

$$b(n)M[\hat{Y}(n-1)\hat{Y}(n-1)] = M[(Y(n) - k(n)X(n))\hat{Y}(n-1)]. \quad (3.50)$$

Для левой части равенства (3.50) имеем

$$\begin{aligned} M[\hat{Y}(n-1)\hat{Y}(n-1)] &= M\left[\left(\hat{Y}(n-1) - Y(n-1) + Y(n-1)\right)\hat{Y}(n-1)\right] = \\ &= \underbrace{M[E(n-1)\hat{Y}(n-1)]}_0 + M[Y(n-1)\hat{Y}(n-1)] = M[Y(n-1)\hat{Y}(n-1)], \end{aligned}$$

где последний переход осуществлен в силу справедливости (3.49) для любого значения n . В правой части равенства (3.50), после подстановки (3.44), получаем

$$\begin{aligned} M[(Y(n) - k(n)X(n))\hat{Y}(n-1)] &= M\left[(Y(n)(1 - k(n)c) - k(n)V(n))\hat{Y}(n-1)\right] = \\ &= (1 - k(n)c)M[Y(n)\hat{Y}(n-1)] - k(n)M[V(n)\hat{Y}(n-1)] = \\ &= (1 - ck(n))M[Y(n)\hat{Y}(n-1)], \end{aligned}$$

так как $M[V(n)\hat{Y}(n-1)] = 0$ в силу некоррелированности отсчета белого шума $V(n)$ с предшествующим значением отклика фильтра $\hat{Y}(n-1)$. В результате проделанных преобразований получаем из равенства (3.50):

$$b(n)M[Y(n-1)\hat{Y}(n-1)] = (1 - ck(n))M[Y(n)\hat{Y}(n-1)]. \quad (3.51)$$

Проводя подстановку (3.43), в правой части (3.51) имеем

$$M[Y(n)\hat{Y}(n-1)] = aM[Y(n-1)\hat{Y}(n-1)] + M[G(n)\hat{Y}(n-1)],$$

причем $M[G(n)\hat{Y}(n-1)] = 0$, так как текущее значение белого шума $G(n)$ некоррелировано с предшествующим значением отклика фильтра $\hat{Y}(n-1)$. Поэтому получаем из (3.51) окончательно:

$$b(n)M[Y(n-1)\hat{Y}(n-1)] = a(1 - ck(n))M[Y(n-1)\hat{Y}(n-1)],$$

откуда находим выражение для оптимального (в смысле минимума (3.46)) значения коэффициента $b(n)$ из оценки (3.45):

$$b(n) = a(1 - ck(n)). \quad (3.52)$$

Подстановка выражения (3.52) в (3.45) приводит к следующему соотношению:

$$\hat{Y}(n) = a\hat{Y}(n-1) + k(n)(X(n) - ac\hat{Y}(n-1)). \quad (3.53)$$

Уравнение (3.53) определяет прогнозное значение $\hat{Y}(n)$ полезного сигнала $Y(n)$ в виде линейного предсказания, формируемого из предыдущего прогнозного значения $\hat{Y}(n-1)$ и текущего входного значения зашумленного сигнала $X(n)$. Устройство, реализующее алгоритм работы в соответствии с (3.53), представляет собой рекурсивный предсказатель первого порядка и называется *скалярным фильтром Калмана*. Несложно убедиться, что такой фильтр имеет структуру, изображенную на рис. 3.38.

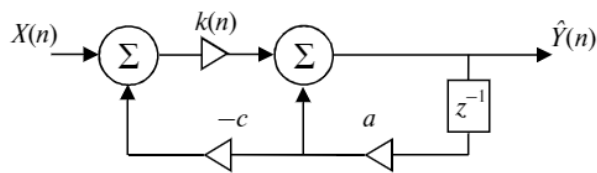


Рис. 3.38. Структурная схема скалярного фильтра Калмана

Проанализируем алгоритм работы данного фильтра. Предыдущая оценка $\hat{Y}(n-1)$ сигнала, умноженная на коэффициент a , т.е. $a\hat{Y}(n-1)$, дает предсказание очередного отсчета $Y(n)$, см. уравнение модели сигнала (3.43). После дополнительного умножения на c

получаем предсказание $\hat{X}(n) = ac\hat{Y}(n-1)$ сигнала $X(n)$ на выходе канала передачи, см. (3.44). Затем ошибка прогноза, т.е. разница между реальным текущим отсчетом $X(n)$ и предсказанием $\hat{X}(n)$, с адаптивно подстраиваемым «коэффициентом доверия» $k(n)$ прибавляется к предсказанной оценке $a\hat{Y}(n-1)$ очередного отсчета $Y(n)$, в результате чего получается текущая оценка $\hat{Y}(n)$ этого отсчета.

Упражнение. Найдите выражение для передаточной функции фильтра, структура которого приведена на рис. 3.38. Сколько особых точек (и какого типа) имеет ПФ этого фильтра? *Указание:* найдите Z -образ уравнения (3.53) и выразите из него передаточную функцию.

Определим рекуррентное правило изменения параметра $k(n)$. Запишем ошибку прогноза фильтра $E(n) = \hat{Y}(n) - Y(n)$, воспользовавшись найденным выражением (3.53), и затем проведем подстановки (3.44) и (3.43):

$$\begin{aligned} E(n) &= a\hat{Y}(n-1) + k(n)(X(n) - ac\hat{Y}(n-1)) - Y(n) = \\ &= a\hat{Y}(n-1) + k(n)cY(n) + k(n)V(n) - k(n)ac\hat{Y}(n-1) - Y(n) = \\ &= a(1 - ck(n))\hat{Y}(n-1) - (1 - ck(n))Y(n) + k(n)V(n) = \\ &= a(1 - ck(n))\hat{Y}(n-1) - a(1 - ck(n))Y(n-1) - (1 - ck(n))G(n) + k(n)V(n) = \\ &= a(1 - ck(n))E(n-1) - (1 - ck(n))G(n) + k(n)V(n). \end{aligned} \quad (3.54)$$

В полученном выражении (3.54) случайные величины $E(n-1)$ (ошибка прогноза в предшествующий момент дискретного времени), $G(n)$ и $V(n)$ (независимые отсчеты белых шумов в текущий момент времени) попарно некоррелированы. Поэтому для среднеквадратичной ошибки (3.46) $e(n) = M[E^2(n)]$ имеем:

$$e(n) = a^2(1 - ck(n))^2 e(n-1) + (1 - ck(n))^2 \sigma_G^2 + k^2(n)\sigma_V^2, \quad (3.55)$$

где $\sigma_G^2 = M[G^2(n)]$, $\sigma_V^2 = M[V^2(n)]$ – дисперсии шумов с нулевыми средними.

С другой стороны,

$$e(n) = M[E^2(n)] = M[E(n)(\hat{Y}(n) - Y(n))] = M[E(n)\hat{Y}(n)] - M[E(n)Y(n)].$$

Учитывая (3.49) и выражая $Y(n)$ из (3.44), находим далее:

$$e(n) = -M[E(n)Y(n)] = -\frac{1}{c}M[E(n)(X(n) - V(n))] = \frac{1}{c}M[E(n)V(n)],$$

где последний переход осуществлен на основании (3.48). Воспользовавшись затем для $E(n)$ представлением (3.54), в силу попарной некоррелированности случайных величин $E(n-1)$, $G(n)$ и $V(n)$, имеем:

$$e(n) = k(n)\sigma_V^2/c. \tag{3.56}$$

Приравняв правые части выражений (3.55) и (3.56), после несложных преобразований (проделайте их самостоятельно), получаем:

$$k(n) = \frac{c(a^2e(n-1) + \sigma_G^2)}{c^2(a^2e(n-1) + \sigma_G^2) + \sigma_V^2}. \tag{3.57}$$

Формулы (3.56) и (3.57) определяют итерационный алгоритм адаптации фильтра Калмана. Для текущего значения $k(n-1)$ этот алгоритм определяет по формуле (3.56) параметр $e(n-1)$, который, в свою очередь, используется для нахождения следующего значения $k(n)$ в соответствии с (3.57). Коэффициент $k(n)$ обновляется с обработкой каждого очередного отсчета входного сигнала $X(n)$. То есть, подставив (3.56) в (3.57), окончательно получаем следующее правило для итерационного определения коэффициента $k(n)$ фильтра Калмана:

$$k(n) = \frac{a^2\sigma_V^2k(n-1) + c\sigma_G^2}{ca^2\sigma_V^2k(n-1) + c^2\sigma_G^2 + \sigma_V^2}. \tag{3.58}$$

Пример 3.39. Показать, что при подаче на вход фильтра Калмана стационарного сигнала (3.44), полученного из (3.43), итерационный процесс нахождения коэффициента $k(n)$ по формуле (3.58) сходится, т. е. $\exists K = \lim_{n \rightarrow \infty} k(n)$. Как следует задавать начальное значение $k(0)$?

◀ Правило (3.58) задает поиск решения уравнения $x = f(x)$, где

$$f(x) = \frac{a^2\sigma_V^2x + c\sigma_G^2}{ca^2\sigma_V^2x + c^2\sigma_G^2 + \sigma_V^2},$$

методом простых итераций. Напомним, что сходимость итерационного поиска $x_n = f(x_{n-1})$, $n = 1, 2, \dots$ гарантирована, если выбранное начальное приближение x_0 и точка $x^* = K$ — решение уравнения $x = f(x)$ — лежат внутри некоторого отрезка Ω такого, что $\forall x \in \Omega: |f'(x)| < 1$. Условие для производной исследуемой функции

$$|f'(x)| = f'(x) = a^2\sigma_V^4 / (ca^2\sigma_V^2x + c^2\sigma_G^2 + \sigma_V^2)^2 < 1$$

эквивалентно требованию

$$\frac{1}{f'(x)} = \frac{(ca^2\sigma_v^2x + c^2\sigma_G^2 + \sigma_v^2)^2}{a^2\sigma_v^4} > 1.$$

Это неравенство справедливо, в силу условия $|a| < 1$, если

$$\frac{(ca^2\sigma_v^2x + c^2\sigma_G^2 + \sigma_v^2)^2}{\sigma_v^4} > 1 \Leftrightarrow \left| a^2cx + c^2\frac{\sigma_G^2}{\sigma_v^2} + 1 \right| > 1.$$

Последнее условие выполняется, если $cx > 0$. Поэтому в качестве начального приближения $x_0 = k(0)$ для поиска оптимального коэффициента $K = \lim_{n \rightarrow \infty} k(n)$ можно выбрать любое число того же знака, что и c , например $x_0 = k(0) = \text{sign } c$. ►

Приведенный пример иллюстрирует сходимость процесса итерационного уточнения оптимального параметра $k(n)$ калмановского фильтра при стационарном входном воздействии. Однако основное применение фильтров Калмана связано с обработкой сигналов, характеристики которых могут меняться. Чаще всего изменения касаются уровня шума, т. е. величины дисперсии шумовой помехи σ_v^2 — при этом, как несложно видеть, структуру фильтра (рис. 3.38) изменять не потребуется, нужно лишь внести корректировку в алгоритм адаптации, определяемый расчетной формулой (3.58).

Пример 3.40. Используя фильтр Калмана, определяемый соотношениями (3.53) и (3.58), провести обработку сигнала $X(n)$, который описывается моделью (3.43)–(3.44) с параметрами $a = 0,97$, $c = 1$, $\sigma_G^2 = 1$, $\sigma_v^2 = 25$.

◀ Положив $Y(0) = 0$, далее по (3.43) рекуррентно сформируем массив, моделирующий случайный дискретный сигнал $Y(n)$. При генерации случайных величин $G(n)$ будем использовать стандартный нормальный закон. График, соединяющий отсчеты полезного сигнала $Y(n)$, отображен на рис. 3.39 сплошной кривой. Дискретный сигнал $X(n)$, полученный по (3.44) добавлением нормального шума $V(n)$ с нулевым средним и дисперсией $\sigma_v^2 = 25$, обозначен на рис. 3.39 пунктиром. Результат применения фильтра Калмана к последовательности $X(n)$ дает оценку полезного сигнала $\hat{Y}(n)$, которая на рис. 3.39 изображена штриховой кривой. ►

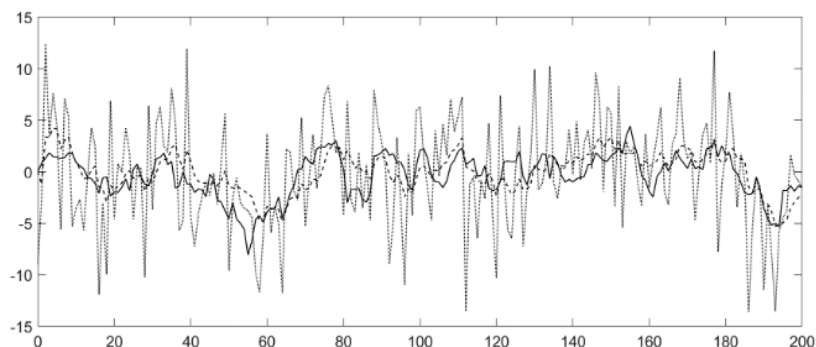


Рис. 3.39. Иллюстрация к примеру 3.40

Рассмотренные нами структура (рис. 3.38) и алгоритм работы (3.58) рекурсивного предсказателя первого порядка (скалярного фильтра Калмана) основаны на простейшей полюсной модели сигнала, которая определяет единственный полюс $z = a$ генерирующей сигнал ЛДС. Авторегрессионные модели более высокого порядка для оцениваемого сигнала влекут и повышение порядка рекурсивных фильтров, обеспечивающих итерационное приближение к его оптимальной среднеквадратичной оценке [20]. Соответствующие системы выходят за рамки нашего изучения и приводят к *векторным фильтрам Калмана*. Как векторные, так и скалярные фильтры Калмана нашли широчайшее применение в радиолокации, системах управления, навигации, технического зрения и во многих других приложениях, где необходимо проводить оценивание и экстраполяцию сигналов по результатам их последовательных измерений в условиях помех (шума).

Фильтры Калмана имеют настраиваемые в процессе работы параметры (в простейшем случае скалярного фильтра — это коэффициент $k(n)$). Фильтры с регулируемыми параметрами называются *адаптивными*. Другие примеры и приложения адаптивной фильтрации будут приведены в заключительных разделах следующей главы.

ГЛАВА 4

ВВЕДЕНИЕ В МЕТОДЫ СИНТЕЗА ЦИФРОВЫХ ФИЛЬТРОВ

4.1. Этапы разработки цифровых фильтров

В широком смысле под термином *цифровой фильтр* понимают любую систему, преобразующую входной цифровой сигнал $x(n)$ в выходной $y(n)$ по некоторому правилу. Например, для рассматриваемых нами линейных дискретных систем (фильтров) это правило определяется соотношением (3.8).

Часто термин «фильтр» как в аналоговой, так и в цифровой технике понимается и в более узком смысле, как устройство, преобразующее входной сигнал в выходной с подавлением или усилением частотных составляющих сигнала из заданной полосы частот. Когда необходимо подчеркнуть, что речь идет именно о таком устройстве, будем называть его *частотно-избирательным фильтром* (или просто частотным фильтром). Задача разработки таких фильтров наиболее часто встречается на практике, и ее рассмотрению будет уделено основное внимание в данной главе.

Процесс разработки цифровых фильтров, алгоритм работы которых описывается уравнением (3.8), можно разделить на следующие этапы.

1. Определение требований к фильтру (задание спецификации).
2. Выбор типа (КИХ, БИХ) и определение коэффициентов фильтра.
3. Выбор структурной формы реализации (параллельная, каскадная, транспонированная и пр.).

4. Анализ влияния ошибок, вызванных конечной разрядностью представления данных.
5. Программная, аппаратная, или программно-аппаратная реализация фильтра.

Указанное выделение этапов проектирования в достаточной степени условно. Так, в требованиях к фильтру может быть сразу задана требуемая аппаратная платформа и соответствующая ей архитектура реализации (например, в виде каскадных звеньев), что полностью предопределяет выбор структуры фильтра (этап 3). Часто указанная схема проектирования используется итерационно, например, после выполнения этапа 4 может оказаться, что из-за влияния ошибок квантования заданные в спецификации характеристики фильтра не достигнуты; тогда требуется повторить выполнение этапов 2 и 3, скорректировав принятые ранее на этих этапах решения. Несмотря на сделанные оговорки, приведенная выше поэтапная схема разработки цифровых фильтров полезна для наглядного описания процесса их проектирования.

Этап задания спецификации фильтра, с которого начинается процесс проектирования, представляет собой формулировку условий задачи, которую далее и предстоит решать проектировщику. Спецификация требований обычно включает в себя: задание частоты дискретизации, цифрового формата отсчетов входного и выходного сигнала; требования к частотной характеристике и быстродействию фильтра; требования к аппаратной платформе реализации фильтра (разрядность используемых регистров памяти, тактовая частота и другие характеристики используемого процессора ЦОС). В качестве аппаратной платформы может выступать обычный персональный компьютер, тогда результатом проектирования фильтра будет являться компьютерная программа.

Требуемые характеристики фильтров очень часто определяются в частотной области (в основной полосе частот). Для частотно-избирательных фильтров (нижних частот, верхних частот, полосовых, режекторных) спецификации чаще всего задаются в виде допусков отклонения от АЧХ соответствующих идеальных фильтров, которые являются физически не реализуемым (см. раздел 3.6). Пример такой спецификации для полосового фильтра приведен

на рис. 4.1: в полосе пропускания АЧХ должна удовлетворять условию $1 - \delta_p \leq |K(\omega)| \leq 1 + \delta_p$, а в полосе подавления $|K(\omega)| \leq \delta_s$. Допустимые отклонения АЧХ чаще выражают в логарифмических единицах (децибелах): $A_s = -20 \lg \delta_s$ — затухание в полосе подавления и $A_p = 20 \lg(1 + \delta_p)$ — неравномерность в полосе пропускания.

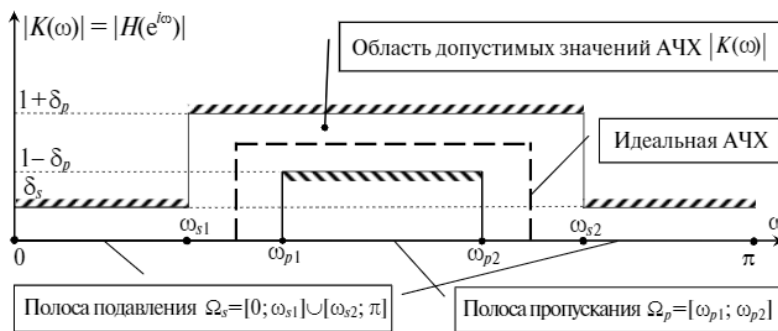


Рис. 4.1. Пример задания спецификации АЧХ полосового фильтра

Так же, как и в предыдущей главе, рассматривая частотные характеристики фильтров, обычно мы будем использовать нормированные частоты, преобразуя в спектральной области реальную линейную частоту f в нормированную линейную $v = f/f_d$, где f_d — частота дискретизации сигнала. Нормированной циклической частотой тогда является $\omega = 2\pi v$.

При формулировке исходных требований к фазовой характеристике обычно указывается, важна ли ее близость к линейной ФЧХ (в полосе пропускания). Однако для ряда специальных фильтров — *фазовых корректоров* — именно спецификация ФЧХ определяет основные требования к фильтру. Фазовые корректоры реализуются посредством всепропускающих систем (см. раздел 3.7).

Второй этап проектирования представляет собой собственно *синтез фильтра* и заключается в нахождении коэффициентов уравнения (3.8) по заданной спецификации фильтра. При этом первая задача, которая встает перед разработчиком — выбор типа фильтра: КИХ или БИХ. Если для построения согласованного фильтра (см. раздел 3.12) выбор структуры в виде нерекурсивного КИХ-

фильтра является практически очевидным, то при решении задачи синтеза частотно-избирательного фильтра ответ на вопрос о выборе КИХ- или БИХ-типа фильтра является в общем случае не столь однозначным. Укажем на наиболее важные особенности, которые следует учитывать при выборе типа разрабатываемого фильтра.

- Для КИХ-фильтров возможно получить строго линейную фазовую характеристику, что важно для многих приложений; БИХ-фильтры имеют нелинейную ФЧХ, особенно на краях полос пропускания.
- Нерекурсивные фильтры всегда имеют КИХ и всегда устойчивы; БИХ-фильтры имеют рекурсивную структуру и вследствие этого подвержены таким негативным влияниям эффектов квантования коэффициентов фильтров, как потеря устойчивости и предельные циклы (см. раздел 3.11).
- При одинаковых требованиях к спецификациям АЧХ частотно-избирательных фильтров, аппроксимирующим идеальные, БИХ-фильтры обычно требуют меньшего числа коэффициентов (и, соответственно, меньше умножений и регистров памяти для реализации), чем КИХ-фильтры. Меньшее число арифметических операций снижает и фактор влияния ошибок округления и переполнения, связанных с конечной разрядностью умножителей и сумматоров.
- Для КИХ-фильтров разработаны методы синтеза, которые применимы для аппроксимации произвольных АЧХ, в то время как методы синтеза БИХ-фильтров ориентированы преимущественно на аппроксимацию идеальных (имеющих прямоугольную АЧХ) частотно-избирательных фильтров.

Методы синтеза цифровых КИХ- и БИХ-фильтров существенно различаются.

Как было отмечено в предыдущей главе, реализация БИХ-фильтров возможна только с использованием рекурсивных структурных схем. Среди методов синтеза БИХ-фильтров следует выделить следующие два основных: *метод инвариантного преобразования импульсной характеристики* и *метод билинейного преобразования*. Оба они основаны на определенных трансформациях методов синтеза аналоговых фильтров.



Для нахождения коэффициентов КИХ-фильтров также существуют несколько методов, из которых выделим три: *метод на основе дискретизации частотной характеристики, оконный (или взвешивания) и оптимизационный метод.*

Нужно отметить, что указанные выше методы, которые будут рассмотрены далее в данной главе, иллюстрируют концептуально наиболее важные подходы к синтезу цифровых рекурсивных и нерекурсивных фильтров и не охватывают все используемые в инженерной практике методы и технологии проектирования цифровых фильтров.

Кратко идеологию выбора типа частотно-избирательного фильтра на втором этапе проектирования можно резюмировать следующим правилом, которое, конечно, в ряде случаев может иметь исключения.

- Использовать БИХ-фильтры, когда необходимо обеспечить максимальную производительность, что особенно важно для приложений реального времени и для маломощных вычислительных аппаратных средств реализации (или очень высоких рабочих частот фильтра).
- Использовать КИХ-фильтры, когда важны требования к линейности ФЧХ, или же требуемая АЧХ заметно отличается от стандартных идеальных АЧХ частотно-избирательных фильтров, или для реализации фильтра имеются избыточные вычислительные мощности.

Отметим, что современные процессоры ЦОС часто имеют встроенные функции, реализующие нерекурсивные фильтры. Для многих приложений это упрощает (фактически предопределяет) выбор именно КИХ-фильтра для дальнейшего синтеза структуры.

Третий этап проектирования — выбор структурной схемы реализации фильтра. Главными факторами выбора здесь являются требования к программно-аппаратной платформе реализации, которые связаны также с областью приложений разрабатываемого фильтра и могут быть весьма специфичными. Основные структурные схемы реализации цифровых фильтров рассмотрены в разделе 3.4. Общие правила выбора структурных схем реализации фильтров сформулировать сложно. Однако можно отметить, что для

реализации КИХ-фильтров чаще всего используется прямая форма, которая для нерекурсивных фильтров называется также *трансверсальной*. Кроме того, если ИХ имеет достаточно большую длительность, вместо прямой формы фильтра, непосредственно реализующего операцию свертки, может быть применена схема вычислений с использованием БПФ, как это было описано в разделе 3.8.

Для реализации БИХ-фильтров могут применяться все формы структурных схем, рассмотренные в разделе 3.4, однако чаще используются параллельная и каскадная структуры, которые порождают системы, менее чувствительные к влиянию конечной разрядности умножителей и сумматоров на искажения характеристик фильтра, подобные рассмотренным в разделе 3.11. Прямая форма пригодна для БИХ-фильтров в основном тогда, когда ошибками, вызванными конечной разрядностью умножителей и сумматоров, можно пренебречь — это имеет место, например, при реализации фильтра на базе высокопроизводительных (по отношению к требуемой скорости обработки сигнала) многоразрядных вычислительных систем.

На четвертом этапе проектирования проводится анализ влияния различного рода ошибок, вызванных конечной разрядностью представления чисел в фильтре. Пример соответствующего анализа, с указанием причин возникновения ошибок и их видов, был приведен в разделе 3.11. Так как на втором и третьем этапах проектирования фильтра его коэффициенты и обрабатываемые данные рассматриваются как числа, имеющие неограниченную (очень высокую) точность представления, то после учета искажений характеристик фильтра, вызванных конечной разрядностью представления чисел, полученные параметры фильтра могут не удовлетворить заданным спецификациям. В этом случае потребуется пересмотреть решения, принятые на втором и третьем этапах проектирования.

Во многих случаях возможности современных вычислительных средств (высокая разрядность представления и обработки данных при достаточно высоком быстродействии) делают влияние конечной разрядности представления чисел на характеристики фильтра пренебрежимо малым.

Заключительный, пятый этап проектирования состоит в подборе необходимого процессора ЦОС и устройств интерфейса, если фильтр реализуется аппаратно и конкретная вычислительная структура не была задана в спецификации. Если не требуется обработка сигналов в режиме реального времени (например, обрабатываются предварительно записанные на носители данных цифровые сигналы), фильтр может быть реализован в виде программы для универсального компьютера. Спектр приложений ЦОС необычайно широк, поэтому возможно множество различных вариантов реализации цифровых фильтров под конкретные практические задачи.

Правила проектирования фильтров, используемых на этапах 3–5, в значительной мере определяются конкретными приложениями и требованиями к программно-аппаратной платформе их реализации. Вместе с тем, для решения задачи синтеза фильтра на втором этапе проектирования разработано немало общих универсальных методов. В данной главе мы рассмотрим основные методы синтеза фильтров, прежде всего частотно-избирательных.

4.2. КИХ-фильтры с линейной фазой. Синтез КИХ-фильтров методом частотной выборки

КИХ-фильтр можно представить в виде *трансверсальной* структуры, изображенной на рис. 4.2, которая получается непосредственно из прямой (прямой канонической) формы, см. раздел 3.4. Трансверсальный фильтр представляет собой чаще всего используемую на практике нерекурсивную структурную схему.

В общем случае ЧХ трансверсального фильтра порядка N имеет, очевидно, следующий вид:

$$K(\omega) = \sum_{k=0}^N h(k)e^{-i\omega k} = \sum_{k=0}^N b_k e^{-i\omega k},$$

так как ИХ $h(k) = b_k, k = 0, 1, \dots, N$.

Как отмечалось ранее, во многих случаях разрабатываемый фильтр должен обладать близкой к линейной ФЧХ. Такое

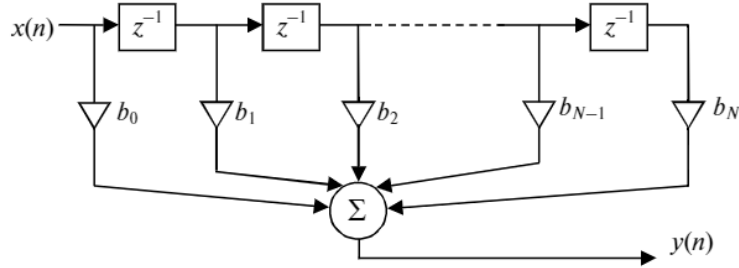


Рис. 4.2. Структурная схема трансверсального фильтра с ИХ $h(k) = b_k, k = 0, 1, \dots, N$

требование возникает вследствие того, что линейная ФЧХ вида (3.15) обеспечивает одинаковую временную задержку при прохождении через фильтр колебаний различных частот (см. раздел 3.6). Существует четыре типа КИХ-фильтров, обеспечивающих линейные ФЧХ.

КИХ-фильтры с линейной фазой I типа. К первому типу относятся фильтры четного порядка $N = 2M$ с импульсной характеристикой $\{h(n) = b_n\}_{n=0}^{2M}$, симметричной относительно центрального отсчета: $h(k) = h(N - k), k = 0, \dots, M$. Для ЧХ таких КИХ-фильтров имеем

$$\begin{aligned} K(\omega) &= \sum_{k=0}^N h(k) e^{-i\omega k} = e^{-i\omega M} \sum_{k=-M}^M h(M+k) e^{-i\omega k} = \\ &= e^{-i\omega M} \left(h(M) + \sum_{k=1}^M (h(M+k) e^{-i\omega k} + h(M-k) e^{i\omega k}) \right) = \\ &= e^{-i\omega M} \left(h(M) + 2 \sum_{k=1}^M h(M-k) \cos(\omega k) \right), \end{aligned}$$

так как $h(M-k) = h(M+k)$. Вводя обозначения $a(0) = h(M)$, $a(k) = 2h(M-k)$, окончательно получаем:

$$K(\omega) = e^{-i\omega M} \sum_{k=0}^M a(k) \cos(\omega k). \quad (4.1)$$

Несложно видеть, что трансверсальный симметричный фильтр четного порядка с ЧХ (4.1) имеет в полосе пропускания линейную ФЧХ. Действительно, представим (4.1) в виде $K(\omega) = A(\omega) e^{-i\omega M}$, где вещественный множитель $A(\omega) = \sum_{k=0}^M a(k) \cos(\omega k)$. Так как в полосе пропускания АЧХ $|K(\omega)| = |A(\omega)| \neq 0$ (см. рис. 4.1), то непрерывная

вещественная функция $A(\omega)$ не меняет своего знака, поэтому будем считать $A(\omega) > 0$ (при необходимости сменим знаки всех коэффициентов $a(k)$ на противоположные). Тогда в полосе пропускания АЧХ $|K(\omega)| = A(\omega)$ и ФЧХ имеет вид (3.15), а именно: $\varphi(\omega) = \arg K(\omega) = -M\omega + 2\pi t$ (напомним, целое число t здесь выбирается из условия $-\pi < \varphi(\omega) \leq \pi$). Поэтому при прохождении через фильтр гармонические колебания любой частоты из полосы пропускания получают одинаковую задержку на $M = N/2$ отсчетов.

Вне полосы пропускания множитель $A(\omega)$ может принимать как положительные, так и отрицательные значения. Смена знака — умножение на $-1 = e^{\pm i\pi}$ — происходит в точках, где $A(\omega) = 0$, и приводит к скачку фазы на величину π . Поэтому при прохождении через фильтр I типа гармонические колебания с частотами ω , для которых $A(\omega) < 0$, помимо задержки на $M = N/2$ отсчетов получают и дополнительный сдвиг фазы на π . При этом ФЧХ для таких частот также имеет линейный вид: $\varphi(\omega) = -M\omega \pm \pi + 2\pi t$.

КИХ-фильтры с линейной фазой II типа — это системы нечетного порядка $N = 2M - 1$ с симметричной импульсной характеристикой $h(k) = h(N - k)$, $k = 0, 1, \dots, M - 1$. В этом случае ЧХ

$$K(\omega) = e^{-i\omega(M-\frac{1}{2})} \sum_{k=1}^M b(k) \cos\left(\omega\left(k - \frac{1}{2}\right)\right), \quad (4.2)$$

где $b(k) = 2h(M - k)$. При прохождении через такой фильтр гармонические колебания получают временную задержку на полуцелое число $N/2 = M - \frac{1}{2}$ отсчетов. Для частот ω , при которых $B(\omega) = \sum_{k=1}^M b(k) \cos(\omega(k - \frac{1}{2})) < 0$, гармонические колебания получают также начальную фазу π .

КИХ-фильтры с линейной фазой III типа — это системы четного порядка $N = 2M$, имеющие антисимметричную ИХ: $h(k) = -h(N - k)$, $k = 0, \dots, M$. Обратим внимание, что требование антисимметричности ИХ в данном случае влечет $h(M) = -h(M)$, поэтому для фильтров III типа всегда $h(M) = 0$. Для ЧХ можно получить следующее соотношение:

$$K(\omega) = i e^{-i\omega M} \sum_{k=1}^M c(k) \sin(\omega k), \quad (4.3)$$

где $c(k) = 2h(M - k)$. При прохождении через такой фильтр гармонические колебания получают временную задержку на $M = N/2$ отсчетов, а начальная фаза изменяется на $\pm\pi/2$ в зависимости от знака вещественного множителя $C(\omega) = \sum_{k=1}^M c(k)\sin(\omega k)$. Для ФЧХ имеем (убедитесь!): $\varphi(\omega) = -M\omega + \frac{\pi}{2}\text{sign } C(\omega) + 2\pi m$.

КИХ-фильтры с линейной фазой IV типа — это системы нечетного порядка $N = 2M - 1$ с антисимметричной импульсной характеристикой $h(k) = -h(N - k)$, $k = 0, 1, \dots, M - 1$.

Пример 4.1. Показать, что ЧХ КИХ-фильтра IV типа имеет вид:

$$K(\omega) = ie^{-i\omega(M-\frac{1}{2})} \sum_{k=1}^M d(k)\sin\left(\omega\left(k - \frac{1}{2}\right)\right), \quad (4.4)$$

где $d(k) = 2h(M - k)$.

$$\begin{aligned} \blacktriangleleft K(\omega) &= \sum_{m=0}^{2M-1} h(m)e^{-i\omega m} = \sum_{m=0}^{M-1} \left(h(m)e^{-i\omega m} + \underbrace{h(2M-1-m)}_{-h(m)} e^{-i\omega(2M-1-m)} \right) = \\ &= 2ie^{-i\omega(M-\frac{1}{2})} \sum_{m=0}^{M-1} h(m) \frac{e^{-i\omega m} e^{i\omega(M-\frac{1}{2})} - e^{i\omega m} e^{-i\omega(M-\frac{1}{2})}}{2i} = \\ &= 2ie^{-i\omega(M-\frac{1}{2})} \sum_{m=0}^{M-1} h(m)\sin\left(\omega\left(M - \frac{1}{2} - m\right)\right). \end{aligned}$$

Выполняя далее замену индекса суммирования m на $k = M - m$, получаем выражение (4.4). \blacktriangleright

Как видно из соотношения (4.4), при прохождении через КИХ-фильтр с линейной ФЧХ четвертого типа гармонические колебания получают задержку на полуцелое число $N/2 = M - \frac{1}{2}$ отсчетов. Кроме того, колебания частоты ω приобретают начальный фазовый сдвиг на $\pm\pi/2$, определяемый знаком функции $D(\omega) = \sum_{k=1}^M d(k)\sin(\omega(k - \frac{1}{2}))$.

Упражнение. Докажите справедливость соотношений (4.2) и (4.3).

За счет симметрии (антисимметрии) коэффициентов $\{b_k\}_{k=0}^N$ КИХ-фильтров с линейной ФЧХ количество умножителей в реализующих их структурах может быть снижено вдвое по сравнению с исходной структурой трансверсального фильтра, приведенной

на рис. 4.2. Например, уравнение системы (3.8) для фильтров II и IV типа нечетного порядка $N = 2M - 1$ можно преобразовать к виду:

$$y(n) = \sum_{k=0}^{2M-1} b_k x(n-k) = \sum_{k=0}^{M-1} b_k x(n-k) + \underbrace{b_{N-k}}_{\pm b_k} x(n-(N-k)) = \sum_{k=0}^{M-1} b_k (x(n-k) \pm x(n-N+k)),$$

где обозначение « \pm » означает знак «+» для фильтров типа II и «-» для фильтров типа IV. Пример структуры с сокращенным числом умножений, реализующей КИХ-фильтр с линейной фазой II типа при $N = 7$, приведен на рис. 4.3. Соответствующие структуры КИХ-фильтров с линейной фазой, в которых число умножителей сокращено вдвое, будем называть *приведенными*.

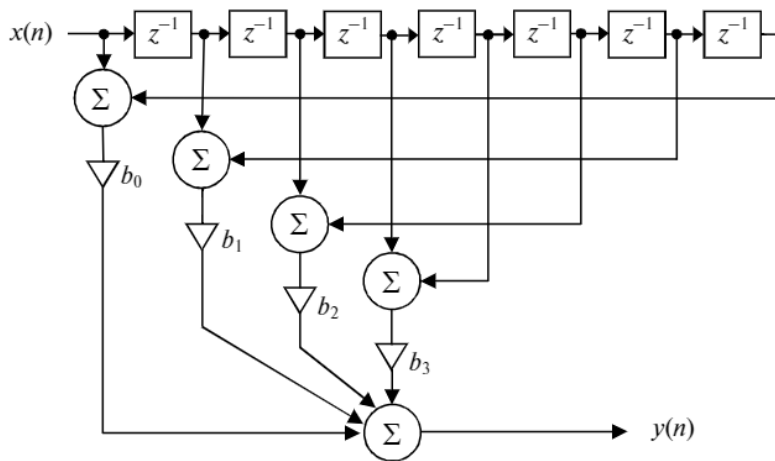


Рис. 4.3. Пример приведенной структуры седьмого порядка КИХ-фильтра с линейной фазой II типа

Упражнение. Изобразите приведенную структуру КИХ-фильтра с линейной фазой I типа для $N = 6$ (фильтр шестого порядка).

Среди четырех рассмотренных типов КИХ-фильтров с линейной ФЧХ фильтры первого типа наиболее универсальны и могут применяться для синтеза всех возможных вариантов частотно-избирательных фильтров (НЧ, ВЧ, полосовых, режекторных).

Для фильтров второго типа ЧХ (4.2) всегда равна нулю для максимальной (нормированной) верхней частоты сигнала $\omega = \pi$, поэтому такие фильтры непригодны для синтеза ФВЧ.

Фильтры третьего и четвертого типов имеют ЧХ ((4.3) и (4.4) соответственно), которые всегда равны нулю при $\omega = 0$, поэтому, очевидно, данные типы фильтров нельзя использовать для синтеза ФНЧ. Фильтры третьего типа к тому же не подходят и для синтеза ФВЧ, так как для нормированной частоты $\omega = \pi$ их ЧХ (4.3) всегда принимает нулевое значение.

Дополнительный сдвиг начальной фазы на $\pm\pi/2$, который придают колебаниям любой частоты КИХ-фильтры третьего и четвертого типов, оказывается полезным свойством для синтеза ряда специальных цифровых фильтров, таких как дифференциаторы и преобразователь Гильберта, которые будут рассмотрены далее в разделе 4.5.

Изучение методов синтеза частотно-избирательных КИХ-фильтров начнем с *метода на основе дискретизации ЧХ*, или *метода частотной выборки*, на примере фильтра с линейной фазой I типа, имеющего ЧХ (4.1).

Пусть задана требуемая АЧХ $K_D(\omega)$ некоторого частотно-избирательного фильтра. Коэффициенты фильтра $\{a(k)\}_{k=0}^M$ в (4.1) должны быть выбраны такими, чтобы АЧХ синтезируемого фильтра $|K(\omega)|$ хорошо аппроксимировала АЧХ $K_D(\omega)$. Для нахождения коэффициентов $\{a(k)\}_{k=0}^M$ приравняем значения требуемой АЧХ $K_D(\omega)$ значениям АЧХ $|K(\omega)|$ синтезируемого фильтра (4.1) в некоторых точках основной полосы частот $\omega_j \in [0; \pi]$, $j = 0, \dots, M$:

$$|K(\omega_j)| = \left| \sum_{k=0}^M a(k) \cos(\omega_j k) \right| = \left| h(M) + 2 \sum_{k=1}^M h(M-k) \cos(\omega_j k) \right| = K_D(\omega_j), \quad (4.5)$$

где $h(M) = a(0)$, $2h(M-k) = a(k)$. Равенствам (4.5), очевидно, будут удовлетворять коэффициенты $\{h(M-k)\}_{k=0}^M$, найденные из решения следующей системы линейных уравнений:

$$h(M) + 2 \sum_{k=1}^M h(M-k) \cos(\omega_j k) = K_D(\omega_j), \quad j = 0, \dots, M. \quad (4.6)$$

Тогда построенное с этими коэффициентами выражение

$$A(\omega) = h(M) + 2 \sum_{k=1}^M h(M-k) \cos(\omega k) \quad (4.7)$$

можно рассматривать как интерполяционную формулу приближения требуемой АЧХ $K_D(\omega)$, а точки ω_j задают узлы интерполяции, в которых $A(\omega_j) = K_D(\omega_j)$. Хотя вопрос выбора расположения узлов представляет собой отдельную задачу, в методе на основе дискретизации ЧХ используется самое простое равномерное расположение точек $\omega_j \in [0; \pi]$.

Пример 4.2. По АЧХ $K_D(\omega) = \begin{cases} 0, & \omega \in [0; \pi/2) \\ 1, & \omega \in [\pi/2; \pi) \end{cases}$ идеального фильтра верхних частот синтезировать трансверсальный фильтр с линейной фазой I типа порядков $N = 2M = 10$, $N = 2M = 22$ и построить графики соответствующих АЧХ $|K(\omega)| = |A(\omega)|$.

◀ Выберем $\omega_j = \frac{\pi(j+0,5)}{M+1}$, $j = 0, \dots, M$. Получаем из (4.6) систему:

$$h(M) + 2 \sum_{k=1}^M h(M-k) \cos \frac{\pi k(j+0,5)}{M+1} = K_D \left(\frac{\pi(j+0,5)}{M+1} \right), \quad j = 0, \dots, M.$$

Можно показать (на основании формул (6.13), (6.15), которые будут рассмотрены позднее), что решение полученной системы:

$$h(M-k) = \frac{1}{M+1} \sum_{j=0}^M K_D \left(\frac{\pi(j+0,5)}{M+1} \right) \cos \frac{\pi k(j+0,5)}{M+1}, \quad k = 0, \dots, M. \quad (4.8)$$

В нашем случае $K_D \left(\frac{\pi(j+0,5)}{M+1} \right) = \begin{cases} 0, & j = 0, 1, \dots, (M-1)/2 \\ 1, & j = (M+1)/2, \dots, M \end{cases}$ и для импульсной характеристики фильтра получаем из (4.8) выражение

$$h(M-k) = \frac{1}{M+1} \sum_{j=\frac{M+1}{2}}^M \cos \frac{\pi k(j+0,5)}{M+1}, \quad k = 0, \dots, M.$$

Отсюда, используя соотношение (4.7), находим АЧХ $|K(\omega)| = |A(\omega)|$. На рис. 4.4 приведены графики полученных АЧХ фильтров для случаев $M = 5$ (слева) и $M = 11$ (справа). ▶

Результаты решения примера 4.2 показывают, что увеличение порядка фильтра не уменьшило максимальное отклонение реальной АЧХ $|K(\omega)|$ от идеальной $K_D(\omega)$ вблизи ее точки разрыва $\omega = \pi/2$. Аппроксимация разрывной функции конечным тригонометрическим многочленом вида (4.7) всегда будет вызывать «всплески» вблизи точек разрыва, подобно наблюдаемому в примере 4.2 — это явление хорошо известно как *эффект Гиббса* и рассмотрено

в разделе 1.4. Как там было установлено, с увеличением порядка M полинома (4.7) точки максимума отклонения будут приближаться к точке разрыва аппроксимируемой функции, а величина максимального отклонения будет оставаться той же.

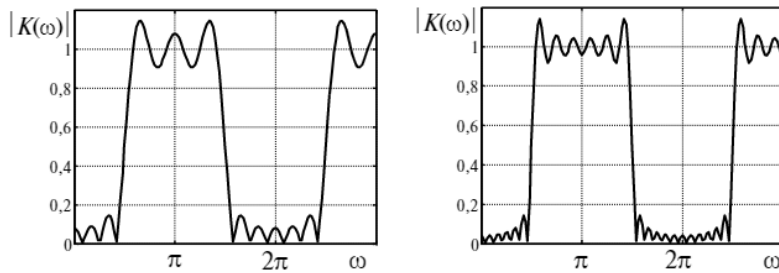


Рис. 4.4. Иллюстрации к результатам решения примера 4.2

Для более точного приближения конечными суммами тригонометрических рядов аппроксимируемая функция должна быть как можно более гладкой, т.е. быть непрерывной и иметь непрерывные производные до как можно более высокого порядка. Поэтому на практике при формировании требований к АЧХ определяют допустимый переходный участок частот, который разделяет полосы пропускания и подавления, как было изображено на рис. 4.1. Тогда можно подобрать более гладкую АЧХ $K_D(\omega)$ для ее последующей аппроксимации функцией $A(\omega)$ (4.7). Например, для узкополосного фильтра с полосой пропускания $\omega \in [\omega_0 - \delta; \omega_0 + \delta]$, где значение δ — мало, часто оказывается удобно выбирать АЧХ в виде $K_D(\omega) = \exp(-c(\omega - \omega_0)^2)$, $c > 0$. Однако конечная цель синтеза частотно-избирательного фильтра состоит не в поиске приемлемой функции $K_D(\omega)$, а в том, чтобы найти такие коэффициенты $\{b_k = h(k)\}_{k=0}^N$ фильтра минимально возможного порядка N , чтобы реальная АЧХ $|K(\omega)|$ не выходила за пределы заданных спецификацией границ, см. рис. 4.1. В методе на основе дискретизации ЧХ это достигается следующим образом.

Задавшись некоторым порядком фильтра N , равномерно расположим в основной полосе частот узлы $\omega_j \in [0; \pi]$, $j = 0, \dots, M$. Для значений $\omega_j \in \Omega_s$, которые попали в полосу подавления, приравняем

значения частотной характеристики к нулю: $K_D(\omega_j) = 0$; для узлов $\omega_j \in \Omega_p$, попавших в полосу пропускания, приравняем АЧХ константе — требуемому коэффициенту передачи — обычно единице: $K_D(\omega_j) = 1$. Значения АЧХ для узлов $\omega_j \in (\Omega_p \cup \Omega_s)$, попавших в переходную частотную область между полосами подавления и пропускания, считаем свободными переменными $\{x_j = K_D(\omega_j)\}$ и подбираем итерационно их значения так, чтобы в области пропускания Ω_p и в области подавления Ω_s минимизировать максимальное отклонение E реальной АЧХ от требуемой:

$$E = \min_{\{x_j = K_D(\omega_j)\}} \left(\max_{\omega \in (\Omega_p \cup \Omega_s)} |K_D(\omega) - A(\omega)| \right), \quad \text{где } K_D(\omega) = \begin{cases} 0, & \omega \in \Omega_s \\ 1, & \omega \in \Omega_p \end{cases}.$$

Для этого, варьируя значения $x_j = K_D(\omega_j)$ в узлах $\omega_j \in (\Omega_p \cup \Omega_s)$ из переходной полосы, находим решением системы (4.6) коэффициенты фильтра $\{b_k = h(k)\}_{k=0}^N$ и анализируем точность $\varepsilon = \max_{\omega \in (\Omega_p \cup \Omega_s)} |K_D(\omega) - A(\omega)|$ соответствующего приближения (4.7) требуемой АЧХ в полосах пропускания и подавления.

Поиск значений АЧХ в переходной области сводится к некоторой задаче оптимизации, решаемой численными методами. Если для полученного в результате такой оптимизации значения $E = E(N)$ хотя бы одно из требований $\max_{\omega \in \Omega_p} |A(\omega) - K_D(\omega)| \leq \delta_p$ или $\max_{\omega \in \Omega_s} |A(\omega)| \leq \delta_s$ оказывается не выполненным, то описанная процедура поиска коэффициентов $\{b_k = h(k)\}_{k=0}^N$ повторяется для большего значения порядка фильтра N .

Не останавливаясь на возможных методах численного решения соответствующей задачи оптимизации, проиллюстрируем возможность формирования требуемой АЧХ на несложном примере ФВЧ.

Пример 4.3. В условиях примера 4.2 рассмотреть возможность приближения идеальной АЧХ при отсутствии ограничений на значения $K_D(\omega)$ в двух ближайших к частоте $\omega = \pi/2$ точках интерполяции.

◀ Пусть M — нечетное число. Вновь, как и в примере 4.2, выберем $\omega_j = \frac{\pi(j+0,5)}{M+1}$, $j = 0, \dots, M$. Полагаем

$$K_D(\omega_j) = 0 \text{ (полоса подавления) для } j = 0, \dots, \frac{M-3}{2} \text{ и}$$

$$K_D(\omega_j) = 1 \text{ (полоса пропускания) для } j = \frac{M+3}{2}, \dots, M.$$

Если выбрать значения АЧХ в точках $\omega_{\frac{M+1}{2}} = \frac{\pi}{2} \left(\frac{M}{M+1} \right)$ и $\omega_{\frac{M+1}{2}} = \frac{\pi}{2} \left(\frac{M+2}{M+1} \right)$ равными соответственно $K_D(\omega_{\frac{M+1}{2}}) = 0,16$ и $K_D(\omega_{\frac{M+1}{2}}) = 0,84$, то, воспользовавшись решением (4.8) системы (4.6), получим КИХ-фильтры, графики АЧХ которых приведены на рис. 4.5 для значений $M = 5$ (слева) и $M = 11$ (справа). Сравните полученные АЧХ с графиками рис. 4.4. ►

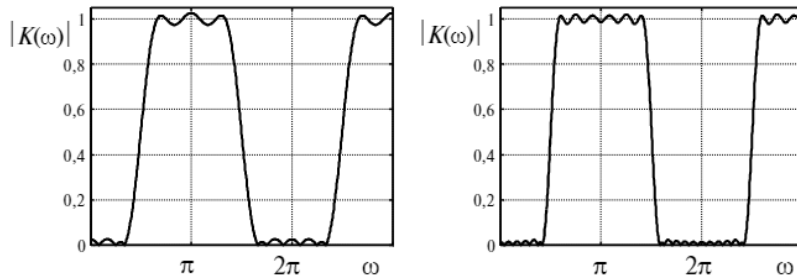


Рис. 4.5. Иллюстрации к результатам решения примера 4.3

Рассмотренный метод синтеза на основе дискретизации ЧХ можно применять для других видов частотно-избирательных фильтров (нижних частот, верхних частот, полосовых, режекторных).

4.3. Оконный метод синтеза КИХ-фильтров

Обычно при проектировании частотно-избирательных КИХ-фильтров принимается требование линейности ФЧХ, что, как мы видели в предыдущем разделе, достигается выбором симметричной или антисимметричной ИХ $\{h(k) = b_k\}_{k=0}^N$.

Требование к АЧХ формулируется в виде задания границ полос пропускания и подавления, а также неравномерности АЧХ в соответствующих частотных полосах, см. рис. 4.1. Например, для ФНЧ и ФВЧ эти требования определяются следующими параметрами в основной полосе частот:

$A_p = 20 \lg(1 + \delta_p)$ — неравномерность в полосе пропускания (дБ),
 $A_s = -20 \lg \delta_s$ — затухание в полосе подавления (дБ),
 f_p — граничная частота полосы пропускания,
 f_s — граничная частота полосы подавления.

Для полосовых и режекторных фильтров отличие в задании спецификации АЧХ будет состоять в том, что границы полос пропускания и подавления определяются двумя значениями частот каждая: $f_{p1} < f_{p2}$ и $f_{s1} < f_{s2}$. Пример задания спецификации АЧХ для ФНЧ отображен на рис. 4.6, где ширина переходной полосы $\Delta F = f_s - f_p$.

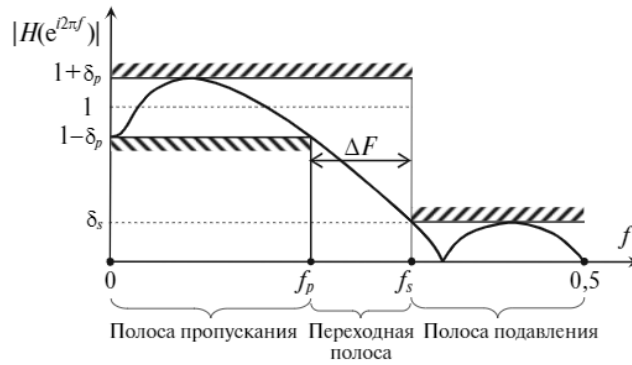


Рис. 4.6. Спецификация и пример АЧХ ФНЧ в области нормированных частот

Оконный метод, или *метод взвешивания импульсной характеристики*, предназначен прежде всего для синтеза частотно-избирательных КИХ-фильтров. Для пояснения сути метода сначала напомним, что (см. раздел 3.6) ИХ $h(k)$ и ЧХ $K(\omega)$ любого фильтра связаны соотношениями: $K(\omega) = \sum_{k=-\infty}^{+\infty} h(k)e^{-i\omega k}$, $\forall k$: $h(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K(\omega)e^{i\omega k} d\omega$.

Предположим, нам нужно разработать ФНЧ. В качестве начальной возьмем АЧХ идеального ФНЧ $K_{\text{НЧ}}(\omega)$ с частотой среза $f_c = (f_p + f_s)/2$, или для циклических частот $\omega_c = 2\pi f_c$. Такой идеальной АЧХ соответствует симметричная ИХ (см. также пример 3.17)

$$h_{\text{НЧ}}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K_{\text{НЧ}}(\omega)e^{i\omega k} d\omega = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{i\omega k} d\omega = \begin{cases} \omega_c/\pi, & k = 0 \\ \frac{\sin(\omega_c k)}{\pi k}, & k \neq 0 \end{cases}, \quad (4.9)$$

$h_{\text{НЧ}}(k) = h_{\text{НЧ}}(-k)$, физически нереализуемого БИХ-фильтра. Однако коэффициенты полученной ИХ (4.9) убывают по абсолютной величине с увеличением $|k|$, поэтому, выбрав некоторое число M и отбросив малые коэффициенты $h_{\text{НЧ}}(k)$, для которых $|k| > M$, БИХ (4.9) можно попытаться приблизить при помощи КИХ $\{h_{\text{НЧ}}(k)\}_{k=-M}^M$. После введения задержки на M отсчетов, не изменяющей АЧХ, фильтр с ИХ $\{\tilde{h}_{\text{НЧ}}(k) = h_{\text{НЧ}}(k - M)\}_{k=0}^{2M}$ становится физически реализуемым КИХ-фильтром с линейной ФЧХ первого типа. Если рассмотреть соответствующую АЧХ (см. также (4.1) и (4.7)),

$$|K(\omega)| = \left| \sum_{k=0}^{2M} \tilde{h}_{\text{НЧ}}(k) e^{-i\omega k} \right| = \left| \sum_{k=-M}^M h_{\text{НЧ}}(k) e^{-i\omega k} \right| = \left| h_{\text{НЧ}}(0) + 2 \sum_{k=1}^M h_{\text{НЧ}}(k) \cos(\omega k) \right|,$$

то становится ясно, что отклонение $\max_{\omega} |K_{\text{НЧ}}(\omega) - |K(\omega)||$ невозможно уменьшить увеличением числа M вследствие эффекта Гиббса, наблюдаемого вблизи точки разрыва $\omega = \omega_c$ идеальной ЧХ $K_{\text{НЧ}}(\omega)$. Подобное явление мы наблюдали также в примере 4.2 и отмечали далее, что сходимость ряда $K(\omega) = \sum_{k=-\infty}^{\infty} h(k) e^{-i\omega k}$ (скорость убывания коэффициентов $h(k)$) можно улучшить, если разрывную ЧХ идеального фильтра приблизить более гладкой функцией.

В оконном методе синтеза КИХ-фильтров скорость убывания коэффициентов ИХ по абсолютной величине повышают за счет умножения БИХ идеального частотно-избирательного фильтра $h_D(k)$ на некоторую неотрицательную симметричную последовательность $\{w(k) = w(-k)\}_{k=-M}^M$: $h_w(k) = w(k)h_D(k)$, $k = 0, \pm 1, \pm 2, \dots, \pm M$. Весовая функция $w(k)$ должна убывать с увеличением $|k|$. Такой подход позволяет уменьшить отклонение реальной АЧХ $|K_w(\omega)| = \left| \sum_{k=-M}^M h_w(k) e^{-i\omega k} \right|$ от идеальной в полосах пропускания и подавления. (В простейшем случае окно — прямоугольное и соответствует обычному усечению ИХ: $w(k) = 1$, $k = 0, \pm 1, \pm 2, \dots, \pm M$.)

Необходимые в оконном методе выражения для ИХ $h_D(k)$ разных типов идеальных частотно-избирательных фильтров несложно получить подобно тому, как это было проделано выше при нахождении ИХ (4.9) идеального ФНЧ.

Упражнение. Покажите, что для идеальных фильтров верхних частот (ВЧ), полосового и режекторного фильтров (П и Р соответственно) выражения для ИХ имеют вид:

$$h_{\text{ВЧ}}(k) = \begin{cases} 1 - (\omega_c/\pi), & k = 0 \\ \frac{-\sin(\omega_c k)}{\pi k}, & k \neq 0 \end{cases}, \quad (4.10)$$

$$h_{\text{П}}(k) = \begin{cases} (\omega_{c2} - \omega_{c1})/\pi, & k = 0 \\ \frac{\sin(\omega_{c2} k) - \sin(\omega_{c1} k)}{\pi k}, & k \neq 0 \end{cases}, \quad (4.11)$$

$$h_{\text{Р}}(k) = \begin{cases} 1 - (\omega_{c2} - \omega_{c1})/\pi, & k = 0 \\ \frac{\sin(\omega_{c1} k) - \sin(\omega_{c2} k)}{\pi k}, & k \neq 0 \end{cases}, \quad (4.12)$$

где для полосового и режекторного фильтров частоты среза, определяющие полосы пропускания (заграждения), таковы, что $\omega_{c1} < \omega_{c2}$.

Отметим, что все ИХ идеальных частотно-избирательных фильтров (4.9)–(4.12) являются симметричными, $h_D(k) = h_D(-k)$, и определяют физически не реализуемые БИХ-фильтры. В силу симметрии весового окна $\{w(k) = w(-k)\}_{k=-M}^M$ симметричными являются и КИХ-фильтры, полученные оконным методом из идеальных: $h_w(k) = w(k)h_D(k) = h_w(-k)$, $k = 0, \pm 1, \pm 2, \dots, \pm M$. После введения задержки ИХ на M отсчетов (не изменяющей АЧХ) фильтр с ИХ $\tilde{h}(k) = h_w(k - M) = w(k - M)h_D(k - M)$ становится физически реализуемым КИХ-фильтром с линейной фазой I типа, при этом его АЧХ определяется соотношением (убедитесь!)

$$|K_w(\omega)| = \left| \sum_{k=0}^{2M} \tilde{h}(k) e^{-i\omega k} \right| = \left| h_w(0) + 2 \sum_{k=1}^M h_w(k) \cos(\omega k) \right|. \quad (4.13)$$

Пример 4.4. Для ФНЧ с нормированной частотой среза $f_c = 1/4$ сравнить АЧХ КИХ-фильтров с линейной ФЧХ первого порядка $N = 2M = 18$ при использовании прямоугольного окна и окна Ханна:

$$w(k) = 0,5 + 0,5 \cos \frac{\pi k}{M+1}, \quad k = 0, \pm 1, \pm 2, \dots, \pm M.$$

◀ Выполним численные расчеты АЧХ по формуле, полученной из (4.13) переходом от нормированных циклических частот к линейным:

$$|K_w(\nu)| = \left| h_w(0) + 2 \sum_{k=1}^M h_w(k) \cos(2\pi\nu k) \right|,$$

положив $h_w(k) = h_{\text{ПЧ}}(k)$ для прямоугольного окна и $h_w(k) = w(k)h_{\text{ПЧ}}(k)$ для окна Ханна. Полученные графики приведены на рис. 4.7.

Как видно из примера, введение весовой функции Ханна существенно ослабляет пульсации АЧХ (в линейном масштабе рис. 4.7а они практически не заметны), но при том же порядке фильтра расширяет переходную полосу. ►

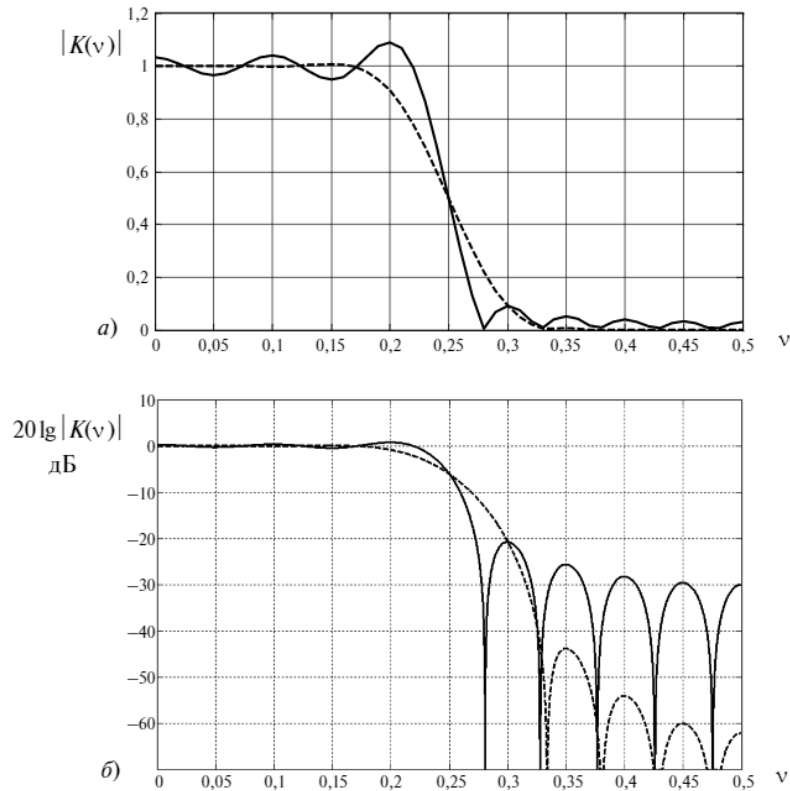


Рис. 4.7. АЧХ КИХ-фильтров с линейной фазой типа I порядка $N = 18$, синтезированных оконным методом по идеальному ФНЧ с нормированной линейной частотой среза $f_c = 0,25$ в линейном (а) и логарифмическом (б) масштабах по оси ординат: сплошная линия — прямоугольное окно, пунктирная линия — окно Ханна

Важной характеристикой окна $\{w(k)\}_{k=-M}^M$ является его спектр $W(\omega) = \sum_{k=-M}^M w(k)e^{-i\omega k}$, циклическая свертка которого с ЧХ идеального фильтра $K_D(\omega)$ определяет ЧХ синтезируемого фильтра:

$$K_w(\Omega) = \sum_{k=-M}^M h_w(k)e^{-i\Omega k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} K_D(\omega)W(\Omega - \omega)d\omega. \quad (4.14)$$

Упражнение. Докажите справедливость соотношения (4.14). *Указание:* используя интегральное представление (4.14) функции $K_w(\Omega)$, имеющей период 2π , разложите ее в ряд Фурье и убедитесь, что коэффициенты разложения получаются равными $h_w(k) = w(k)h_D(k)$.

С учетом того, что спектр $W(\omega)$ является вещественной четной функцией (убедитесь), формулу (4.14) можно записать в виде: $K_w(\Omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K_D(\omega)W(\omega - \Omega)d\omega$. Так как входящие в интеграл функции вещественные, он имеет несложную графическую интерпретацию, которая для идеального фильтра НЧ изображена на рис. 4.8.

Из рис. 4.8 видно, что ЧХ НЧ-фильтра $K_w(\Omega) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} W(\omega - \Omega)d\omega$ с точностью до нормировочного множителя $1/(2\pi)$ представляет собой площадь криволинейной трапеции, которая вырезана при помощи прямоугольной ЧХ идеального фильтра $K_D(\omega)$ из спектра оконной функции $W(\omega)$, сдвинутого по оси частот на величину Ω .

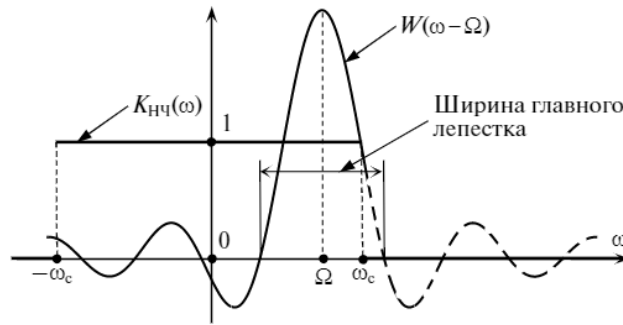


Рис. 4.8. Геометрическая интерпретация вычисления интеграла (4.14) на примере спектра $W(\omega)$ прямоугольной оконной функции

Формула (4.14) и ее графическая интерпретация поясняют влияние спектра окна $W(\omega)$ на ЧХ синтезируемого фильтра $K_w(\Omega)$. Так, неравномерность (пульсации) $K_w(\Omega)$ определяется соотношением между глобальным максимумом $W(0)$ и абсолютными величинами локальных экстремумов $W(\omega)$ в так называемых *боковых лепестках* (криволинейных трапециях, ограниченных осью абсцисс и графиком функции $W(\omega)$). Чем меньше амплитуда боковых лепестков, тем меньше отклонения реальной ЧХ $K_w(\Omega)$ от идеальной $K_D(\omega)$ в полосах пропускания и подавления. Ширина переходной зоны, разделяющей эти полосы, будет тем больше, чем шире *главный лепесток* спектра $W(\omega)$.

Таким образом, окна для синтеза фильтров должны выбираться такими, чтобы по сравнению с максимальным значением $W(0)$ главного лепестка спектра амплитуды боковых лепестков были как можно меньшими, а ширина главного лепестка была по возможности уже.

Выделим следующие получившие распространение оконные весовые функции (во всех случаях $k = 0, \pm 1, \pm 2, \dots, \pm N/2$ для четного числа $N = 2M$ и $k = 0, \pm 1, \pm 2, \dots, \pm (N-1)/2$ для нечетного числа $N = 2M - 1$; прочие весовые коэффициенты считаются равными нулю).

- Окно Ханна (или Хэннинга) $w(k) = 0,5 + 0,5 \cos \frac{2\pi k}{N+1}$.
- Окно Хэмминга $w(k) = 0,54 + 0,46 \cos \frac{2\pi k}{N+1}$.
- Окно Блэкмана $w(k) = 0,42 + 0,5 \cos \frac{2\pi k}{N+1} + 0,08 \cos \frac{4\pi k}{N+1}$.

Важнейшие характеристики данных весовых функций представлены в табл. 4.1. Уровень бокового лепестка выражен в децибелах как $20 \lg(A_1/A_0)$, где $A_0 = W(0)$ — уровень главного лепестка, A_1 — максимальное значение АЧХ $|W(\omega)|$ вне главного лепестка (соответствует локальному максимуму в ближайшем к главному боковом лепестке). Отметим, что все окна обеспечивают практически равную неравномерность АЧХ в полосе пропускания и в полосе подавления, т. е. $\delta_p \approx \delta_s$.

Таблица 4.1. Важнейшие характеристики некоторых весовых функций

| Тип окна | Уровень бокового лепестка, дБ | Ширина переходной полосы* ΔF | Неравномерность в полосе пропускания, дБ | Затухание в полосе подавления, дБ |
|---------------|-------------------------------|--------------------------------------|--|-----------------------------------|
| Прямоугольное | -13 | $\approx 0,9/N$ | 0,742 | 21 |
| Ханна | -31 | $\approx 3,1/N$ | 0,0546 | 44 |
| Хэмминга | -41 | $\approx 3,3/N$ | 0,0194 | 53 |
| Блэкмана | -57 | $\approx 5,5/N$ | 0,0017 | 74 |

* В единицах нормированных линейных частот.

Применение метода синтеза КИХ-фильтров на основе взвешивания ИХ рассмотрим на примере.

Пример 4.5. Синтезировать оконным методом КИХ-фильтр, удовлетворяющий следующим требованиям:

- полоса пропускания 150–250 Гц;
- ширина переходной полосы 50 Гц;
- неравномерность в полосе пропускания 0,1 дБ;
- затухание в полосе подавления 60 дБ;
- частота дискретизации 1 кГц.

◀ Найдем неравномерность АЧХ в полосе пропускания и в полосе подавления. Так как $A_p = 20 \lg(1 + \delta_p) = 0,1$ дБ, то $\delta_p = 0,0116$; $A_s = -20 \lg \delta_s = 60$ дБ, поэтому $\delta_s = 0,001$.

При синтезе фильтра необходимо обеспечить неравномерность АЧХ $\delta = \min(\delta_p, \delta_s) = 0,001$. Из приведенных в таблице 4.1 видов весовых функций данному требованию отвечает только окно Блэкмана. Число коэффициентов фильтра (порядок фильтра) оцениваем по таблице 4.1:

$$N \approx 5,5/\Delta F = 5,5/(50/1000) = 110.$$

Синтезируемый фильтр является полосовым, с полосами подавления 0–100 Гц и 300–500 Гц. Соответствующий идеальный полосовой фильтр для заданной ширины переходных полос в 50 Гц имеет частоты среза 125 Гц и 275 Гц, что соответствует нормированным циклическим частотам $\omega_{c1} = 2\pi \cdot (125/1000) = \pi/4$,

$\omega_{c2} = 2\pi \cdot (275/1000) = 11\pi/20$. Воспользовавшись формулой (4.11), численно находим взвешенную ИХ симметричного (физически не реализуемого) КИХ-фильтра: $h_w(k) = w(k)h_{\Pi}(k)$, где $w(k)$ — весовая функция Блэкмана, $k = 0, \pm 1, \pm 2, \dots, \pm 55$.

Вводя задержку на 55 отсчетов, получаем окончательно ИХ и набор коэффициентов физически реализуемого КИХ фильтра, отвечающего заданной спецификации: $b_k = \tilde{h}(k) = h_w(k - 55)$, $k = 0, 1, 2, \dots, 110$. ►

Упражнение. Используя какой-либо математический программный пакет, найдите численно коэффициенты КИХ-фильтра из примера 4.5 и постройте график АЧХ. Убедитесь, что заданная спецификация фильтра выполнена.

Основной недостаток рассмотренных нами выше оконных функций состоит в их неуниверсальности. Анализируя данные табл. 4.1, можно заметить, что улучшение характеристик неравномерности АЧХ влечет расширение ширины переходной полосы ΔF . Причем для каждого конкретного окна увеличение длины ИХ (порядка фильтра N) не улучшает неравномерность АЧХ, а лишь сокращает ширину переходной полосы. Для реализации заданных спецификаций фильтра приходится выбирать варианты из ограниченного набора окон, определяющих характеристики неравномерности АЧХ с достаточно большим дискретом. В результате в ряде случаев точность аппроксимации АЧХ может оказаться избыточной, причем платой за эту не всегда нужную точность оказывается увеличение количества отсчетов ИХ, т.е. усложнение фильтра. Указанные проблемы можно отчасти разрешить, используя в качестве весовой функции *окно Кайзера*. Данная весовая функция имеет дополнительный параметр β , который позволяет в определенных пределах задавать более гибкий баланс между шириной переходной полосы и неравномерностью АЧХ.

Для длины ИХ $2M+1$ (КИХ-фильтр порядка $N = 2M$) весовая функция окна Кайзера имеет вид:

$$w_k(k) = \begin{cases} I_0(\beta\sqrt{1-(k/M)^2})/I_0(\beta), & k = 0, \pm 1, \pm 2, \dots, \pm M; \\ 0, & |k| > M. \end{cases} \quad (4.15)$$

Здесь

$$I_0(x) = 1 + \sum_{n=1}^{\infty} \left(\frac{(x/2)^n}{n!} \right)^2 \quad (4.16)$$

— модифицированная функция Бесселя первого рода нулевого порядка¹.

Значения параметров $N = 2M$ и β , которые необходимо использовать для нахождения функции весового окна (4.15), вычисляются по эмпирическим формулам, установленным Кайзером, входными параметрами которых являются выраженная в децибелах неравномерность АЧХ $A = -20 \lg(\min(\delta_s, \delta_p))$ и ширина переходной полосы в единицах нормированных линейных частот ΔF . Параметр β задается по величине A следующим образом:

$$\beta = \begin{cases} 0,1102(A - 8,7), & A \geq 50; \\ 0,5842(A - 21)^{0,4} + 0,07886(A - 21), & 21 < A < 50; \\ 0, & A \leq 21. \end{cases} \quad (4.17)$$

Минимальный порядок фильтра определяется следующей формулой:

$$N = 2M \geq \begin{cases} \frac{A - 7,95}{14,36\Delta F}, & A > 21; \\ 0,9222/\Delta F, & A \leq 21. \end{cases} \quad (4.18)$$

Параметры окна Кайзера для некоторых значений неравномерности АЧХ приведены в табл. 4.2 (сравните с данными табл. 4.1).

Таблица 4.2. Характеристики некоторых вариантов окна Кайзера

| Ширина переходной полосы* ΔF | Неравномерность в полосе пропускания, дБ | Затухание в полосе подавления, дБ |
|--------------------------------------|--|-----------------------------------|
| $\approx 2,93/N$ ($\beta = 4,54$) | 0,0274 | 50 |
| $\approx 4,32/N$ ($\beta = 6,76$) | 0,00275 | 70 |
| $\approx 5,71/N$ ($\beta = 8,96$) | 0,000275 | 90 |

* В единицах нормированных линейных частот.

¹ Отметим, что ряд (4.16) достаточно быстро сходится, и 20–25 его первых слагаемых практически всегда обеспечивают требуемую точность расчетов окна Кайзера (4.15).

Пример 4.6. Описать процедуру синтеза КИХ-фильтра, удовлетворяющего спецификации из примера 4.5, с использованием окна Кайзера.

◀ Для данного фильтра $A = 60$ дБ, $\Delta F = 0,05$. Порядок фильтра находим по формуле (4.18): $N = 2M \geq \frac{60 - 7,95}{14,36 \times 0,05} = 72,49$. Примем $N = 74$.

Дополнительно необходимый для определения функции окна (4.15) параметр находим по формуле (4.17): $\beta = 0,1102(A - 8,7) \approx 5,65$.

Воспользовавшись формулой (4.11) с параметрами $\omega_{c1} = 2\pi(125/1000) = \pi/4$, $\omega_{c2} = 2\pi(275/1000) = 11\pi/20$, численно находим взвешенную ИХ симметричного КИХ-фильтра: $h_w(k) = w_k(k)h_{\Pi}(k)$, где $w_k(k)$ — весовая функция Кайзера (4.15) для $M = 37$, $\beta = 5,65$; $k = 0, \pm 1, \pm 2, \dots, \pm 37$.

Вводя задержку на 37 отсчетов, получаем окончательно ИХ и набор коэффициентов физически реализуемого КИХ-фильтра, отвечающего заданной спецификации: $b_k = \hat{h}(k) = h_w(k - 37)$, $k = 0, 1, 2, \dots, 74$.

Видим, что по сравнению с окном Блэкмана использование окна Кайзера позволило сократить число умножителей в структуре КИХ-фильтра примерно на треть. ▶

Основным достоинством рассмотренного нами оконного метода синтеза КИХ-фильтров является его простота. Метод не требует большого объема вычислений (даже при использовании более сложной оконной функции Кайзера). В отличие от других весовых функций, окно Кайзера можно «настраивать» под требуемое соотношение между шириной переходной полосы ΔF и неравномерностью δ АЧХ проектируемого фильтра, при этом общая закономерность для всех весовых функций остается общей: уменьшение δ ведет к расширению (увеличению) ΔF .

Главный недостаток оконного метода — его недостаточная гибкость. В частности, фильтры, синтезируемые методом взвешивания ИХ, имеют примерно одинаковую неравномерность АЧХ в полосе пропускания и в полосе подавления: $\delta_p \approx \delta_s$. Вместе с тем спецификация фильтра может иметь существенно различающиеся требования к этим неравномерностям. В результате приходится ориентироваться на более жесткие требования, и синтезированный

фильтр может с избытком перекрыть некоторую часть требований спецификации. Ценой такой избыточности будет более высокая (возможно, недопустимая) сложность аппаратной или программной реализации фильтра.

4.4. Синтез оптимальных КИХ-фильтров

По сравнению с рассмотренными в предыдущем разделе оконными методами синтеза на основе взвешивания ИХ более мощными с точки зрения оптимальности получаемых структур КИХ-фильтров, но и более сложными как в теоретическом, так и в вычислительном плане являются оптимизационные методы синтеза.

Говоря об оптимальности фильтров, прежде всего необходимо уточнить, по какому критерию и при каких ограничениях проводится оптимизация. Например, при синтезе КИХ-фильтров задача может быть сформулирована следующим образом: при заданной длине импульсной характеристики $N+1$ (заданном порядке фильтра N) найти такие коэффициенты симметричного (антисимметричного) фильтра с линейной фазой, чтобы ошибка аппроксимации заданной АЧХ $|K_D(\omega)|$ была минимальной. Если ошибка аппроксимации ε понимается в смысле среднеквадратичного отклонения на периоде ЧХ, когда

$$\varepsilon^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |K(\omega) - K_D(\omega)|^2 d\omega, \quad (4.19)$$

то решение такой задачи известно из теории рядов Фурье: в качестве отсчетов импульсной характеристики необходимо взять первые коэффициенты Фурье-разложения ЧХ $K_D(\omega)$, т. е. использовать рассмотренный в предыдущем разделе метод синтеза фильтра с прямоугольным весовым окном ИХ. Однако критерий минимизации ошибки (4.19) приводит к значительным пульсациям реальных ЧХ $K(\omega)$ вблизи точек разрыва аппроксимируемых ЧХ $K_D(\omega)$. Поэтому чаще при разработке фильтров минимизируют ошибку ε , которая понимается как *максимальное* отклонение реальной АЧХ от требуемой (в заданной частотной области Ω) — такая аппроксимация называется *равномерной*.

Рассмотрим КИХ-фильтры с линейной ФЧХ первого типа. Удобно представить ЧХ такого фильтра в виде $K(\omega) = A(\omega)e^{-i\omega M}$, где вещественный тригонометрический полином (см. (4.1))

$$A(\omega) = \sum_{k=0}^M a(k)\cos(\omega k). \quad (4.20)$$

Как и ранее, считаем, что спецификация синтезируемого фильтра определяется заданием границ полос пропускания Ω_p , подавления Ω_s и неравномерностей АЧХ δ_p и δ_s в этих полосах. Для заданного порядка M тригонометрического многочлена $A(\omega)$ будем искать такой вектор коэффициентов $\mathbf{a} = (a(0), a(1), \dots, a(M))$, который обеспечивает *наилучшее равномерное приближение* требуемой ЧХ $K_D(\omega)$ (вещественной функции) с минимальной ошибкой аппроксимации

$$\varepsilon = \min_{\mathbf{a} \in \mathbb{R}^{M+1}} \max_{\omega \in (\Omega_s \cup \Omega_p)} p(\omega) |A(\omega) - K_D(\omega)|. \quad (4.21)$$

Функция $p(\omega) > 0$ в выражении (4.21) позволяет по-разному масштабировать (взвешивать) ошибку аппроксимации для различных частот. (Если $\delta_p = \delta_s$, то полагаем $p(\omega) = 1$.)

Многочлен вида (4.20), который обеспечивает минимальную ошибку (4.21), будем называть оптимальным (наилучшим) многочленом равномерного приближения порядка M . Для его отыскания выполним замену переменной $x = \cos \omega$. Тогда основная полоса нормированных циклических частот $\arccos x = \omega \in [0; \pi]$ взаимно однозначно отобразится в отрезок $x \in [-1; 1]$, а (4.21) примет вид:

$$\varepsilon = \min_{\mathbf{a} \in \mathbb{R}^{M+1}} \max_{x \in (X_s \cup X_p)} p(\arccos x) |A(\arccos x) - K_D(\arccos x)|, \quad (4.22)$$

где $X_s \subset [-1; 1]$ и $X_p \subset [-1; 1]$ — множества (отрезки), в которые при замене переменной отобразились полосы подавления и пропускания $\Omega_s \subset [0; \pi]$ и $\Omega_p \subset [0; \pi]$ соответственно.

Пример 4.7. По спецификации ФНЧ, заданной в виде полос пропускания $\Omega_p = [0; \omega_p]$ и подавления $\Omega_s = [\omega_s; \pi]$, неравномерностей АЧХ δ_p и δ_s , сформулировать критерий оптимальности многочлена (4.20) в виде (4.22).

◀ Полосы пропускания и подавления отображаются соответственно:

$$\begin{aligned} \Omega_p = [0; \omega_p] &\xleftarrow[\omega = \arccos x]{x = \cos \omega} X_p = [x_p; 1], \quad x_p = \cos \omega_p; \\ \Omega_s = [\omega_s; \pi] &\xleftarrow[\omega = \arccos x]{x = \cos \omega} X_s = [-1; x_s], \quad x_s = \cos \omega_s. \end{aligned}$$

Идеальная АЧХ ФНЧ принимает вид:

$$K_D(\arccos x) = \begin{cases} 1, & x \in X_p \\ 0, & x \in X_s \end{cases}$$

Обозначим $\alpha = \delta_p/\delta_s$. Тогда весовая функция может быть выражена как

$$p(\arccos x) = \begin{cases} 1, & x \in X_p \\ \alpha, & x \in X_s \end{cases}$$

Полученная область ограничений на искомую функцию $P(x) = A(\arccos x)$ отображена на рис. 4.9. Если многочлен $A(\arccos x) = \sum_{k=0}^M a(k) \cos(k \arccos x)$, удовлетворяющий (4.22), даст ошибку равномерной аппроксимации $\varepsilon \leq \delta_p$, то спецификация фильтра будет выполнена. Иначе потребуются искать другой многочлен большего порядка M . ▶

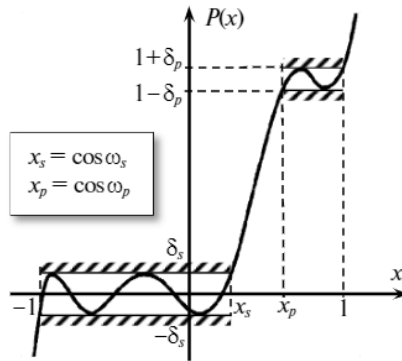


Рис. 4.9. Иллюстрация к примеру 4.7: область ограничений на искомую функцию $P(x)$ и ее примерный вид

Упражнение. По аналогии с рис. 4.9 изобразите допустимую область для функции $P(x) = A(\arccos x)$ по заданной спецификации полосового фильтра, имеющего две полосы подавления $\Omega_{s1} = [0; \omega_{s1}]$, $\Omega_{s2} = [\omega_{s2}; \pi]$ и одну полосу пропускания $\Omega_p = [\omega_{p1}; \omega_{p2}]$, $\omega_{s1} < \omega_{p1} < \omega_{p2} < \omega_{s2}$.

Поясним смысл замены переменной $x = \cos \omega$. После такой замены полином (4.20) принимает вид:

$$P(x) = A(\arccos x) = \sum_{k=0}^M a(k) \cos(k \arccos x). \quad (4.23)$$

Функции

$$T_n(x) = \cos(n \arccos x), \quad (4.24)$$

$n = 0, 1, \dots$, фигурирующие в выражении (4.23), носят название *полиномов Чебышева* и на отрезке $x \in [-1; 1]$ могут быть представлены в виде алгебраических многочленов порядка n : $T_n(x) = \sum_{j=0}^n c_j x^j$. Действительно, имеем:

$$T_0(x) = \cos 0 = 1, \quad T_1(x) = \cos(\arccos x) = x.$$

Далее, обозначая $\omega = \arccos x$, получаем:

$$\begin{aligned} T_{n-1}(x) &= \cos((n-1)\omega) = \cos(n\omega)\cos\omega + \sin(n\omega)\sin\omega, \\ T_{n+1}(x) &= \cos((n+1)\omega) = \cos(n\omega)\cos\omega - \sin(n\omega)\sin\omega. \end{aligned}$$

Складывая два последних равенства и учитывая, что $T_n(x) = \cos(n\omega)$, $x = \cos\omega$, получаем: $T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x)$. Отсюда рекуррентная формула для построения многочленов Чебышева для $n = 2, 3, \dots$ принимает вид:

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x). \quad (4.25)$$

Упражнение. Найдите выражение для полиномов Чебышева $T_2(x)$, $T_3(x)$, $T_4(x)$. Убедитесь, что $n \geq 1$ в многочленах $T_n(x) = \sum_{j=0}^n c_j x^j$: $c_n = 2^{n-1}$.

Очевидно, что в силу определения (4.24) $\forall x \in [-1; 1]: |T_n(x)| \leq 1$. Параметрическому изменению значения аргумента x в (4.23) в пределах $-1 \leq x \leq 1$ соответствует непрерывное убывание величины $t = \arccos x$ от π до 0. Поэтому функция (4.24), представленная как $\cos t$, $0 \leq t \leq \pi$, принимает нулевое значение $T_n(x) = 0$ при $-1 \leq x \leq 1$ ровно в n точках, в которых она меняет знак. Максимальные отклонения от нуля — значения $T_n(x) = \pm 1$ — многочлен (4.24) имеет в $n+1$ точках $\{x_m\}_{m=0}^{n+1}$ отрезка $x \in [-1; 1]$, включая его концы $x_0 = -1$ и $x_{n+1} = 1$, причем $T_n(x_{m+1}) = -T_n(x_m)$.

Упражнение. Используя представление (4.24), найдите точки $x \in [-1; 1]$, в которых $T_n(x) = 0$, $T_n(x) = -1$, $T_n(x) = 1$.

Итак, мы убедились, что замена переменной $x = \cos\omega$ преобразует тригонометрический многочлен $A(\omega)$ (4.20) в алгебраический $P(x)$ (4.23), который имеет те же коэффициенты $\mathbf{a} = (a(0), \dots, a(M))$ при многочленах Чебышева, что и коэффициенты при функциях

$\cos(\omega k)$ в (4.20). Поиск коэффициентов оптимального в смысле (4.21) многочлена (4.20), таким образом, можно свести к поиску такого алгебраического многочлена $P(x) = \sum_{k=0}^M a(k)T_k(x) = \sum_{k=0}^M a_k x^k$, что (см. (4.22), (4.23))

$$\varepsilon = \min_{a \in \mathbb{R}^{M+1}} \max_{x \in (X_s \cup X_p)} |p(\arccos x)(P(x) - K_D(\arccos x))|. \quad (4.26)$$

Поиск наилучшего равномерного приближения заданной функции алгебраическим многочленом представляет собой непростую, но хорошо изученную классическую задачу. В большинстве случаев аналитическое решение найти невозможно, и оно ищется численными методами, которые основаны на следующей *теореме Чебышева об альтернансе*.

Теорема 4.1. Пусть замкнутое множество $X \in \mathbb{R}$ есть объединение конечного числа непересекающихся отрезков на числовой оси, а $P_M(x) = \sum_{k=0}^M a_k x^k$ — многочлен степени M . Пусть, кроме того, $D(x)$ — непрерывная в области $x \in X$ функция, а $w(x)$ — положительная и непрерывная функция при $x \in X$. Обозначим: $E(x) = w(x)(P_M(x) - D(x))$,

$$\delta = \min_{\{a_k\}_{k=0}^M} \max_{x \in X} |E(x)|.$$

Тогда для того, чтобы полином $P_M(x)$ был (и притом единственным) многочленом наилучшего равномерного приближения на множестве X для функции $D(x)$ с весом $w(x)$, необходимо и достаточно, чтобы абсолютный максимум функции $E(x)$ достигался в $K \geq M+2$ принадлежащих множеству X точках *альтернанса*: $x_1 < x_2 < \dots < x_K$, в которых $\forall m = 1, \dots, K: |E(x_m)| = \delta, E(x_{m+1}) = -E(x_m)$.

Выполнив обратную замену переменных $\omega = \arccos x$, теорему 4.1 можно переформулировать, используя тригонометрический полином (4.20) вместо алгебраического многочлена $P_M(x) = \sum_{k=0}^M a_k x^k$, и вернуться к задаче (4.21) равномерного приближения требуемой АЧХ $K_D(\omega)$ тригонометрическим полиномом (4.20) на отрезках полос пропускания Ω_p и подавления Ω_s в области нормированных частот $\omega \in [0; \pi]$. Так как арккосинус является монотонно убывающей функцией, точки альтернанса $x_1 < x_2 < \dots < x_K$ перейдут в $\omega_K < \omega_{K-1} < \dots < \omega_1$, где $\omega_m = \arccos(x_m)$, $m = K, \dots, 1$. Основным результатом теоремы 4.1 останется прежним: полином (4.20) будет многочленом наилучшего равномерного приближения тогда

и только тогда, когда максимальное отклонение от аппроксимируемой функции будет достигаться не менее чем в $M+2$ частотах альтернанса, причем ошибка $E(\cos\omega_{m+1}) = -E(\cos\omega_m)$. На этом основополагающем свойстве многочлена наилучшего равномерного приближения основан численный метод его построения, известный как *метод Ремеза*. Приведем основные шаги реализующего его алгоритма численного решения задачи (4.21).

1. Задаем некоторую начальную сетку частот $\{\omega_m\}_{m=1}^{M+2} \in \Omega = \Omega_p \cup \Omega_s$ (например, расположив точки сетки равномерно в полосах пропускания и подавления).
2. Ищем такие коэффициенты полинома (4.20), чтобы для частот $\{\omega_m\}_{m=1}^{M+2}$ получить равнопульсирующее отклонение полинома от аппроксимируемой ЧХ: $p(\omega_m)(A(\omega_m) - K_D(\omega_m)) = (-1)^{m+1}\delta$. Для этого составляем и решаем систему из $M+2$ линейных уравнений относительно переменных — $(M+1)$ коэффициентов $\{a(k)\}_{k=0}^M$ полинома (4.20) и величины δ :

$$\sum_{k=0}^M a(k)\cos(\omega_m k) + \frac{(-1)^m}{p(\omega_m)}\delta = K_D(\omega_m), \quad m = 1, 2, \dots, M+2.$$
3. С имеющимися коэффициентами $\{a(k)\}_{k=0}^M$ на густой сетке частот $\{\omega_l\}_{l=1}^L \in \Omega$, $L \gg M$, вычисляем отклонения аппроксимирующего полинома (4.20) от заданной ЧХ $K_D(\omega)$: $\Delta_l = p(\omega_l)(A(\omega_l) - K_D(\omega_l))$, $l = 1, 2, \dots, L$.
4. Если $\varepsilon = \max_{l=1, \dots, L} |\Delta_l| \approx |\delta|$ (с приемлемой точностью, например, $0,99|\delta| \leq \varepsilon \leq 1,01|\delta|$), то считаем, что искомые коэффициенты $\{a(k)\}_{k=0}^M$ полинома (4.20) и ошибка аппроксимации (4.21) $\varepsilon = |\delta|$ найдены; алгоритм завершил работу. Иначе переходим на шаг 5.
5. Задаем новую сетку $\{\omega_m\}_{m=1}^{M+2} \in \Omega$, в которой обязательно должны быть те частоты, для которых при анализе на предыдущем шаге был достигнут максимум отклонения $\varepsilon = \max_{l=1, \dots, L} |\Delta_l| > |\delta|$, а также другие частоты $\{\omega_j\}$, для которых были получены самые большие отклонения $|\Delta_j| > |\delta|$. При этом знаки отклонений Δ_m в выбранных частотах сетки $\{\omega_m\}_{m=1}^{M+2}$ должны чередоваться¹. Затем выполняется переход на шаг 2.

¹ Например, новые узлы сетки можно искать в окрестности старых узлов так, чтобы знак отклонения сохранялся, но величина отклонения была большей.



Чем больше порядок полинома M используется в методе Ремеза, тем выше будет точность аппроксимации требуемой ЧХ. Полином минимального порядка, обеспечивающий выполнение спецификаций фильтра, является оптимальным.

Описанный оптимизационный метод синтеза КИХ-фильтров по заданной АЧХ имеет различные модификации и алгоритмические реализации, разработанные как самим Ремезом, так и другими исследователями. В частности, в зарубежной литературе часто упоминается алгоритм Паркса — Мак-Клиллана [29], который также является одним из вариантов реализации идеи метода Ремеза.

Мы рассмотрели, как свести задачу наилучшей равномерной аппроксимации (4.21) тригонометрическим полиномом (4.20) к классической задаче поиска наилучшего равномерного приближения алгебраическими многочленами (4.26) для случая КИХ-фильтров с линейной фазой I типа. Если вид тригонометрического полинома отличен от (4.20), как это имеет место для КИХ-фильтров с линейной фазой II–IV типов, то задачу поиска оптимального полинома равномерного приближения необходимо сначала переформулировать и свести к поиску наилучшего в смысле (4.21) тригонометрического полинома вида (4.20). Проиллюстрируем это на примере фильтра второго типа.

По аналогии с (4.21), построение оптимального КИХ-фильтра с линейной фазой II типа порядка $N = 2M - 1$ формулируется как поиск такого вектора коэффициентов $\mathbf{b} = (b(1), \dots, b(M))$, определяющего ЧХ (4.2), чтобы соответствующий тригонометрический полином

$$B(\omega) = \sum_{k=1}^M b(k) \cos\left(\omega\left(k - \frac{1}{2}\right)\right) \quad (4.27)$$

являлся решением задачи наилучшего равномерного приближения желаемой (вещественной) частотной характеристики $\tilde{K}_D(\omega)$ с заданной весовой функцией $\tilde{p}(\omega)$:

$$\varepsilon = \min_{\mathbf{b} \in \mathbb{R}^M} \max_{\omega \in (\Omega_s \cup \Omega_p)} \tilde{p}(\omega) |B(\omega) - \tilde{K}_D(\omega)|. \quad (4.28)$$

Для того чтобы преобразовать эту задачу к условиям (4.21), установим связь между тригонометрическими полиномами (4.20) из (4.1) и (4.27) из (4.2).

Пример 4.8. Показать, что для вещественного тригонометрического полинома (4.27) справедливо представление

$$B(\omega) = \sum_{k=1}^M b(k) \cos\left(\omega\left(k - \frac{1}{2}\right)\right) = \cos\frac{\omega}{2} \sum_{k=0}^{M-1} a(k) \cos(\omega k),$$

и выразить коэффициенты $b(k)$ через $a(k)$.

◀ Используя соотношение $\cos\alpha\cos\beta = \frac{1}{2}(\cos(\alpha+\beta) + \cos(\alpha-\beta))$, имеем:

$$\begin{aligned} B(\omega) &= \cos\frac{\omega}{2} \sum_{k=0}^{M-1} a(k) \cos(\omega k) = \frac{1}{2} \sum_{k=0}^{M-1} a(k) \left(\cos\omega\left(k - \frac{1}{2}\right) + \cos\omega\left(k + \frac{1}{2}\right) \right) = \\ &= \frac{1}{2} \left(\sum_{k=0}^{M-1} a(k) \cos\omega\left(k - \frac{1}{2}\right) + \sum_{k=1}^M a(k-1) \cos\omega\left(k - \frac{1}{2}\right) \right) = \\ &= \frac{1}{2} \left(a(0) \cos\frac{\omega}{2} + \sum_{k=1}^{M-1} (a(k) + a(k-1)) \cos\omega\left(k - \frac{1}{2}\right) + a(M-1) \cos\omega\left(M - \frac{1}{2}\right) \right) = \\ &= \sum_{k=1}^M b(k) \cos\left(\omega\left(k - \frac{1}{2}\right)\right), \end{aligned}$$

где $b(1) = a(0) + \frac{a(1)}{2}$, $b(k) = \frac{a(k) + a(k-1)}{2}$ для $k = 2, \dots, M-1$;

$b(M) = a(M-1)/2$. Таким образом, всегда можно следующим образом представить тригонометрический полином (4.27) через полином (4.20): $B(\omega) = A_{M-1}(\omega) \cos(\omega/2)$, где $A_{M-1}(\omega) = \sum_{k=0}^{M-1} a(k) \cos(\omega k)$. ▶

Воспользовавшись результатами решения примера 4.8, выражение (4.28) можем преобразовать к виду

$$\begin{aligned} \varepsilon &= \min_{\mathbf{b} \in \mathbb{R}^M} \max_{\omega \in (\Omega_s \cup \Omega_p)} \tilde{p}(\omega) |B(\omega) - \tilde{K}_D(\omega)| = \\ &= \min_{\mathbf{a} \in \mathbb{R}^M} \max_{\omega \in (\Omega_s \cup \Omega_p)} \tilde{p}(\omega) \left| A_{M-1}(\omega) \cos\frac{\omega}{2} - \tilde{K}_D(\omega) \right| = \\ &= \min_{\mathbf{a} \in \mathbb{R}^M} \max_{\omega \in (\Omega_s \cup \Omega_p)} p(\omega) |A_{M-1}(\omega) - K_D(\omega)|, \end{aligned} \quad (4.29)$$

где обозначено: $p(\omega) = \tilde{p}(\omega) \cos(\omega/2)$, $K_D(\omega) = \tilde{K}_D(\omega) / \cos(\omega/2)$. Изменив таким образом выражения для исходно заданных весовой функции $\tilde{p}(\omega)$ и требуемой ЧХ $\tilde{K}_D(\omega)$, мы свели поиск оптимального для (4.28) полинома вида (4.27) к уже рассмотренной задаче отыскания наилучшего в смысле (4.21) тригонометрического полинома вида (4.20). Найдя ее решение относительно коэффициентов $\mathbf{a} = (a(0), a(1), \dots, a(M-1))$ многочлена (4.20), коэффициен-



ты $\mathbf{b} = (b(1), \dots, b(M))$ тригонометрического полинома (4.27) затем выразим при помощи установленных в примере 4.8 соотношений.

В соответствии с условиями теоремы 4.1 весовая функция $p(\omega) = \tilde{p}(\omega)\cos(\omega/2)$ в (4.29) должна быть положительной и непрерывной в замкнутой области $\Omega = \Omega_p \cup \Omega_s$. Кроме того, для получения требуемой в (4.29) ЧХ $K_D(\omega) = \tilde{K}_D(\omega)/\cos(\omega/2)$ необходимо исключить деление на ноль. Поэтому замкнутая область $\Omega \subset [0; \pi]$ в (4.29) должна быть задана так, чтобы точка $\pi \notin \Omega$. Это не противоречит уже известным нам свойствам фильтров с линейной фазой II типа: так как ЧХ таких фильтров $K(\pi) \equiv 0$, то точка $\omega = \pi$ вместе с некоторой ее окрестностью должна быть исключена из области Ω , на которой ищется многочлен наилучшего равномерного приближения.

Подход, рассмотренный на примере приведения задачи поиска оптимального фильтра второго типа к задаче поиска оптимального фильтра первого типа, может быть распространен и на КИХ-фильтры третьего и четвертого типов. Для этого соответствующие тригонометрические полиномы в (4.3) и (4.4) нужно сначала выразить через (4.20) подобно тому, как это было проделано для полинома (4.27) в примере 4.8. Затем исходно заданные выражения для требуемой ЧХ $\tilde{K}_D(\omega)$ и весовой функции $\tilde{p}(\omega)$ следует преобразовать по аналогии с приведенным выше выводом соотношения (4.29).

Упражнение. Покажите, что для тригонометрических полиномов $C(\omega) = \sum_{k=1}^M c(k)\sin \omega k$ и $D(\omega) = \sum_{k=1}^M d(k)\sin(\omega(k - \frac{1}{2}))$ (см. (4.3) и (4.4)) можно найти такие коэффициенты $\tilde{c}(k)$ и $\tilde{d}(k)$, $k = 1, \dots, M$, что $C(\omega) = \sin \omega \sum_{k=0}^{M-1} \tilde{c}(k)\cos(\omega k)$, $D(\omega) = \sin \frac{\omega}{2} \sum_{k=0}^{M-1} \tilde{d}(k)\cos(\omega k)$ (см. (4.20)).

Указание: используйте схему решения, аналогичную приведенной в примере 4.8, и тождество $\sin \alpha \cos \beta = \frac{1}{2}(\sin(\alpha + \beta) + \sin(\alpha - \beta))$.

Применение оптимизационных методов синтеза КИХ-фильтров на сегодняшний день получило очень широкое распространение в связи с развитием и ростом производительности вычислительной техники, что практически свело на нет основной недостаток этих методов — вычислительную сложность. Результаты применения оптимизационных методов позволяют

разрабатывать такие КИХ-фильтры, сложность реализации которых (определяемая длиной ИХ или количеством коэффициентов-умножителей в структуре фильтра) оказывается значительно меньшей по сравнению с фильтрами, разработанными другими методами.

Важнейшим достоинством оптимизационных методов является то, что они весьма удобны для синтеза не только частотно-избирательных фильтров, но и фильтров с произвольной кусочно-непрерывной ЧХ $K_D(\omega)$.

4.5. Специальные КИХ-фильтры: преобразователь Гильберта и цифровой дифференциатор

Определение. Для вещественного аналогового сигнала $x(t)$, имеющего спектр $X(v) = \mathcal{F}\{x(t)\}$, преобразованием Гильберта назовем вещественный сигнал $\tilde{x}(t)$, спектр которого $\tilde{X}(v) = \mathcal{F}\{\tilde{x}(t)\}$ определяется по спектру $X(v)$ следующим образом:

$$\tilde{X}(v) = \begin{cases} -iX(v), & \text{при } v > 0 \\ 0, & \text{при } v = 0 \\ iX(v), & \text{при } v < 0 \end{cases} \quad (4.30)$$

Сигнал $\tilde{x}(t)$ будем называть сопряженным по Гильберту сигналу $x(t)$.

Пример 4.9. Найти преобразование Гильберта для сигнала $x(t) = A \cos(2\pi v_0 t - \gamma)$.

◀ В разделе 1.7 было установлено, что $\mathcal{F}\{e^{i2\pi v_0 t}\} = \delta(v - v_0)$. Используя свойство линейности преобразования Фурье, для спектра исходного сигнала $x(t) = A \cos(2\pi v_0 t - \gamma)$ получаем:

$$X(v) = \mathcal{F}\left\{A \frac{e^{i(2\pi v_0 t - \gamma)} + e^{-i(2\pi v_0 t - \gamma)}}{2}\right\} = A \frac{\delta(v - v_0)e^{-i\gamma} + \delta(v + v_0)e^{i\gamma}}{2}.$$

Тогда, на основании (4.30), спектр сопряженного по Гильберту сигнала:

$$\tilde{X}(v) = A \frac{-i\delta(v - v_0)e^{-i\gamma} + i\delta(v + v_0)e^{i\gamma}}{2} = A \frac{\delta(v - v_0)e^{-i\gamma} - \delta(v + v_0)e^{i\gamma}}{2i}.$$

Отсюда

$$\tilde{x}(t) = \mathcal{F}^{-1}\{\tilde{X}(v)\} = A \frac{e^{i(2\pi v_0 t - \gamma)} - e^{-i(2\pi v_0 t - \gamma)}}{2i} = A \sin(2\pi v_0 t - \gamma).$$

Таким образом, гармоническое колебание *любой* частоты в результате преобразования Гильберта получает фазовый сдвиг на $\pi/2$, так как $\sin \alpha = \cos\left(\alpha - \frac{\pi}{2}\right)$. ▶

Если интерпретировать преобразование Гильберта как обработку сигнала при помощи некоторого устройства, то ЧХ такого аналогового устройства, как следует из (4.30):

$$K(v) = \frac{\tilde{X}(v)}{X(v)} = \begin{cases} -i = e^{-i\pi/2}, & \text{при } v > 0 \\ 0, & \text{при } v = 0. \\ i = e^{i\pi/2}, & \text{при } v < 0 \end{cases} \quad (4.31)$$

Отсюда ФЧХ $\varphi(v) = \arg K(v) = \begin{cases} -\pi/2, & \text{при } v > 0 \\ \pi/2, & \text{при } v < 0 \end{cases}$ и АЧХ $|K(v)| = 1$ при $v \neq 0$. Назовем аналоговую систему, имеющую ЧХ (4.31), *преобразователем Гильберта*.

Покажем, что ИХ аналогового преобразователя Гильберта $h(t) = \mathcal{F}^{-1}\{K(v)\} = 1/\pi t$. Находя ЧХ, соответствующую данной ИХ, имеем:

$$K(v) = \int_{-\infty}^{+\infty} \frac{e^{-2\pi i v t}}{\pi t} dt = \int_{-\infty}^{+\infty} \frac{\cos(2\pi v t)}{\pi t} dt - i \int_{-\infty}^{+\infty} \frac{\sin(2\pi v t)}{\pi t} dt.$$

Действительная часть ЧХ представляет собой интеграл от нечетной функции в симметричных пределах и потому обращается в ноль. Интеграл, определяющий мнимую часть ЧХ, $I(v) = \int_{-\infty}^{+\infty} \frac{\sin(2\pi v t)}{\pi t} dt$, для любого $v > 0$ дает значение 1, причем $I(-v) = -I(v)$ и $I(0) = 0$. Поэтому для ЧХ получаем требуемое выражение (4.31).

Упражнение. Убедитесь, что $\forall v > 0: I(v) = \int_{-\infty}^{+\infty} \frac{\sin(2\pi v t)}{\pi t} dt = 1$.

Указание: покажите, что $S_v(u) = \int_{-\infty}^{+\infty} \frac{\sin(2\pi v t)}{\pi t} e^{-i2\pi u t} dt = \begin{cases} 1, & u \in (-v; v) \\ 0, & u \notin (-v; v) \end{cases}$.

Определив ИХ $h(t) = 1/\pi t$, преобразование Гильберта во временной области можно тогда представить в виде интегральной

свертки:

$$\tilde{x}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(u)}{t-u} du = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(t-u)}{u} du.$$

Заметим, что преобразователь Гильберта представляет собой физически нереализуемое устройство, так как $\forall t < 0 \quad h(t) \neq 0$, поэтому реализовать его на практике можно только приближенно.

Упражнение. Запишите ЧХ и найдите по ней ИХ устройства, выполняющего *обратное преобразование Гильберта*, т. е. восстановление (аналогового) сигнала $x(t)$ по сопряженному $\tilde{x}(t)$ сигналу. Является ли такое устройство физически реализуемым?

Определение. Комплексный сигнал $z(t) = x(t) + i\tilde{x}(t)$, у которого мнимая часть сигнала $\tilde{x}(t)$ сформирована по вещественному сигналу $x(t)$ посредством преобразования Гильберта, называется *аналитическим сигналом*.

Представив аналитический сигнал в показательной форме:

$$z(t) = x(t) + i\tilde{x}(t) = s(t)e^{i\Psi(t)} = s(t)\cos\Psi(t) + is(t)\sin\Psi(t),$$

функцию $s(t) = |z(t)| = \sqrt{x^2(t) + \tilde{x}^2(t)}$ назовем *огibaющей*,

$\Psi(t) = \arctg \frac{\tilde{x}(t)}{x(t)}$ — *мгновенной фазой*, а $\omega(t) = \Psi'(t) = \frac{\tilde{x}'(t)x(t) - \tilde{x}(t)x'(t)}{x^2(t) + \tilde{x}^2(t)}$

— *мгновенной циклической частотой* сигнала $x(t)$.

Упражнение. Используя соотношение (4.30) и свойство линейности преобразования Фурье, покажите, что если сигналу $x(t)$, имеющему спектральную плотность $X(\nu) = \mathcal{F}\{x(t)\}$, поставлен в соответствие аналитический сигнал $z(t) = x(t) + i\tilde{x}(t)$, то его спектральная плотность

$$Z(\nu) = \mathcal{F}\{z(t)\} = \begin{cases} 2X(\nu), & \nu > 0 \\ X(\nu), & \nu = 0 \\ 0, & \nu < 0 \end{cases}$$

Преобразователь Гильберта применяется, например, в радиотехнике для обработки *радиосигналов*, которые описываются функциями вида

$$x(t) = u(t)\cos(\omega_0 t + \varphi(t)) = u(t)\cos\Psi(t),$$

где ω_0 — несущая частота, а функция $u(t)$ и/или $\varphi(t)$ содержит передаваемую в радиосигнале информацию¹. Например, структурная схема устройства, выделяющего огибающую радиосигнала $u(t) \geq 0$, изображена на рис. 4.10.

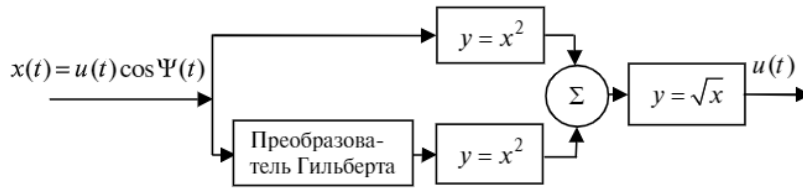


Рис. 4.10. Структурная схема устройства выделения огибающей радиосигнала

Для определения мгновенной фазы и циклической частоты радиосигнала также используются преобразователи Гильберта.

Определение. Преобразованием Гильберта дискретного сигнала $f(n) = x(nT)$, имеющего спектр $F(v) = T \sum_{n \in \mathbb{Z}} x(nT) e^{-2\pi i v n T}$ (см. (2.9)), назовем дискретный сигнал $\tilde{f}(n) = \tilde{x}(nT)$, спектральная плотность (2.9) которого

$$\tilde{F}(v) = \begin{cases} -iF(v), & \text{при } 0 < v < \frac{1}{2T} \\ 0, & \text{при } v = 0 \text{ и } v = \pm \frac{1}{2T}, \\ iF(v), & \text{при } \frac{-1}{2T} < v < 0 \end{cases} \quad \tilde{F}(v) = \tilde{F}\left(v + \frac{1}{\Delta t}\right).$$

Устройство, преобразующее сигнал $f(n) = x(nT)$ в $\tilde{f}(n) = \tilde{x}(nT)$, назовем *дискретным преобразователем Гильберта* (ДПГ).

Таким образом, отличие дискретного преобразования Гильберта от непрерывного преобразования, определяемого по (4.30), связано с тем, что спектр $F(v)$ дискретного сигнала $f(n) = x(nT)$ представляет собой периодическое продолжение спектра аналогового сигнала $X(v) = \mathcal{F}\{x(t)\}$ на всю ось частот с периодом $1/T$: $F(v) = \sum_{k \in \mathbb{Z}} X\left(v - \frac{k}{T}\right)$. При этом предполагается, что спектр исходного аналогового сигнала ограничен: $X(v) = 0$ при $|v| \geq 1/2T$. Как и

¹ Некоторые свойства спектров радиосигналов рассмотрены в разделе 2.5.

ранее, дискретный сигнал $\tilde{f}(n) = \tilde{x}(nT)$ называем сопряженным по Гильберту сигналу $f(n) = x(nT)$, а сигнал $z(nT) = x(nT) + i \cdot \tilde{x}(nT)$ называем аналитическим.

Упражнение. Покажите, что если сигналу $f(n) = x(nT)$, имеющему спектральную плотность $F(v)$, поставлен в соответствие аналитический сигнал $z(nT) = x(nT) + i \cdot \tilde{x}(nT)$, то его спектральная плотность

$$Z(v) = \begin{cases} 2F(v), & \text{при } 0 < v < \frac{1}{2T} \\ F(v), & \text{при } v = 0 \\ 0, & \text{при } \frac{-1}{2T} \leq v < 0 \end{cases}, \quad Z(v) = Z\left(v + \frac{1}{T}\right).$$

Вернемся к нормированным частотам, далее вновь положив $T = 1$.

Пример 4.10. Найти ИХ $h_T(n)$ ДПГ.

◀ Из определения дискретного преобразования Гильберта следует, что периодическая ЧХ ДПГ (для нормированных частот):

$$K(v) = \frac{\tilde{F}(v)}{F(v)} = \begin{cases} -i, & \text{при } 0 < v < 1/2 \\ 0, & \text{при } v = 0 \text{ и } v = \pm 1/2, \\ i, & \text{при } -1/2 < v < 0 \end{cases}, \quad K(v) = K(v + 1).$$

Найдем ИХ ДПГ, которая связана с коэффициентами Фурье (1.16) разложения в ряд (1.15) ЧХ $K(v)$ следующим образом: $h_T(n) = c_{-n}$. На основании формулы (1.16) имеем:

$$h_T(n) = \int_{-1/2}^{1/2} K(v) e^{2\pi i v n} dv = i \int_{-1/2}^0 e^{2\pi i v n} dv - i \int_0^{1/2} e^{2\pi i v n} dv = \dots = \frac{2 \sin^2 \frac{\pi n}{2}}{\pi n}, \quad (4.32)$$

при этом $h_T(0) = 0$. (Опущенные выкладки проделайте самостоятельно.) ▶

Очевидно, и в дискретном случае преобразователь Гильберта является физически нереализуемым устройством, т. е. на практике его можно реализовать только приближенно. Для этого с помощью формулы свертки запишем сопряженный по Гильберту сигнал как отклик ДПГ:

$$\tilde{f}(k) = \sum_{n=-\infty}^{\infty} h_T(n) f(k-n) \approx \tilde{f}_M(k) = \sum_{n=-M}^M w(n) h_T(n) f(k-n),$$

где $w(n)$ — некоторая весовая функция, например, окна Кайзера (см. раздел 4.3). Далее, вводя задержку на M отсчетов и обозначив $h(n) = w(n - M)h_T(n - M)$, имеем:

$$\tilde{f}(k) \approx \tilde{f}_M(k) = \sum_{n=-M}^M w(n)h_T(n)f(k - n) = \sum_{n=0}^{2M} h(n)f(k - n + M),$$

или
$$\tilde{f}(k - M) \approx \tilde{f}_M(k - M) = \sum_{n=0}^{2M} h(n)f(k - n) = y(k).$$

Таким образом, если ввести в рассмотрение КИХ-фильтр, имеющий импульсную характеристику $\{h(n) = w(n - M)h_T(n - M)\}_{n=0}^{2M}$, то отклик такого фильтра $y(n)$ на дискретный сигнал $f(n)$ будет представлять собой приближенное дискретное преобразование Гильберта сигнала $f(n)$, задержанное на M отсчетов: $y(n) \approx \tilde{f}(n - M)$. Очевидно, чем больше число M , тем точнее выполняется это приближенное равенство.

Пример 4.11. Изобразите структурную схему КИХ-фильтра, имеющего ИХ $\{h(n) = h_T(n - M)\}_{n=0}^{2M}$.

◀ В данном случае используется простейшее прямоугольное окно $w(n)$. Учтем, что $h_T(2k) = 0, k \in \mathbb{Z}$ (см. (4.32)). Получаем структурную схему, приведенную на рис. 4.11.

Возникающую задержку сопряженного по Гильберту сигнала необходимо учитывать при формировании аналитического сигнала, вводя такую же компенсационную задержку для исходного сигнала: $f(n - M)$. Для этого могут использоваться те же элементы

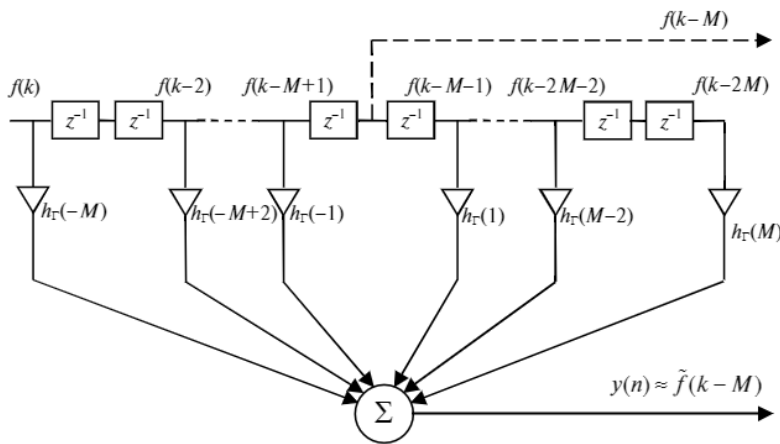


Рис. 4.11. Структурная схема фильтра с ИХ $\{h(n) = h_T(n - M)\}_{n=0}^{2M}$

задержки, которые имеются в структуре ДПГ — на рис. 4.11 линия формирования задержанного входного сигнала изображена пунктиром. ►

Упражнение. Если учесть, что $h_T(k) = -h_T(-k)$, то отклик фильтра, изображенного на рис. 4.11, можно записать в виде:

$$\tilde{f}(k - M) \approx y(k) = \sum_{n=0}^{\lfloor M/2 \rfloor} h_T(2n+1)(f(k - M - (2n+1)) - f(k - M + (2n+1))),$$

где квадратные скобки обозначают целую часть числа. Тогда в соответствующей структурной схеме фильтра появятся $\lfloor M/2 \rfloor$ вычитателей (устройств, вычисляющих разность двух чисел) и количество умножителей будет снижено вдвое, с M до $\lfloor M/2 \rfloor$. Изобразите соответствующую структурную схему ДПГ (см. также рис. 4.2 и 4.3).

Так как ИХ ДПГ (4.32) является антисимметричной, то для синтеза ДПГ используются КИХ-фильтры с линейной фазой III и IV типа. Для идеального ДПГ на всех частотах $|K(\omega)| = 1$ (за исключением отдельных точек $\omega = \pi k, k \in \mathbb{Z}$). В примере 4.11 была использована структура фильтра третьего типа — для этой структуры ЧХ (4.3) всегда обращается в ноль на границах основной полосы частот: $K(0) = K(\pi) = 0$.

Если для конкретного приложения ДПГ принципиально необходимо включение в рабочую полосу верхних частот, близких к нормированной частоте $\omega = \pi$ (частоте Найквиста), то для синтеза такого ДПГ лучше использовать КИХ-фильтры с линейной фазой IV типа, которые имеют ЧХ вида (4.4) и позволяют получить $K(\pi) \neq 0$. Однако следует иметь в виду, что использование КИХ-фильтра с линейной фазой IV типа для синтеза ДПГ приведет к вычислительно более сложной реализации по сравнению со структурой типа III. Сдвиг ИХ (4.32) на полуцелое число отсчетов (что имеет место в структурах фильтров четвертого типа) приведет к примерно двукратному увеличению числа необходимых для реализации ДПГ умножений в результате потери свойства ИХ $h(2k) = 0$, которое наследуют от ИХ вида (4.32) структуры III типа. Поэтому использование КИХ-фильтра с линейной фазой III типа обычно оказывается более предпочтительным в вычислительном плане.

Упражнение. Выпишите выражение (см. (4.32)) для $h_T\left(n - \left(M - \frac{1}{2}\right)\right)$.



Другим важным в практическом плане типом устройств, для реализации которых в цифровом виде используются структуры трансверсальных КИХ-фильтров, являются *дифференциаторы*. Для аналоговых сигналов дифференциатором называют устройство, которое преобразует входной сигнал $x(t)$ в выходной $y = x'(t)$. Частотная характеристика такого аналогового устройства с учетом свойства 9° интеграла Фурье (см. раздел 1.5) для циклических частот должна иметь вид $K(\omega) = i\omega$. Совершенно очевидно, что такое устройство, имеющее неограниченно возрастающую с увеличением частоты АЧХ $|K(\omega)| = |\omega|$, физически нереализуемо и представляет собой идеальную модель. На практике для аналогового дифференциатора необходимо определить некоторую ограниченную рабочую полосу частот, в которой ЧХ должна достаточно точно подчиняться закону $K(\omega) = i\omega$. Для цифровых дифференциаторов предполагается, что верхняя рабочая частота ЦД не превышает частоты Найквиста (в циклических или линейных нормированных частотах — это $\omega = \pi$ или $\nu = 0,5$ соответственно).

Определение. *Цифровым дифференциатором* (ЦД) назовем линейную дискретную систему, имеющую в линейных нормированных частотах ЧХ $K_d(\nu) = 2\pi i\nu$ (для $-0,5 < \nu < 0,5$), которая периодически продолжена на всю ось частот: $K_d(\nu) = K_d(\nu + 1)$.

Цифровые дифференциаторы, имеющие рабочую полосу, близкую или совпадающую с основной полосой частот $\omega \in [0; \pi]$, называются *широкополосными*. Если ширина рабочей полосы заметно уже основной полосы частот $\omega \in [0; \pi]$, то такие ЦД называют *узкополосными*. В зависимости от расположения рабочей полосы можно также разделить ЦД на низкочастотные, полосовые и высокочастотные.

Пример 4.12. Определить, какой тип КИХ-фильтров с линейной фазой пригоден для реализации цифрового дифференциатора.

◀ Прежде всего, найдем ИХ дифференциатора. Интегрируя по частям, находим (убедитесь):

$$h_d(n) = 2\pi i \int_{-1/2}^{1/2} \nu e^{2\pi i \nu n} d\nu = \begin{cases} \frac{\cos \pi n}{n} - \frac{\sin \pi n}{\pi n^2}, & n \neq 0 \\ 0, & n = 0 \end{cases}. \quad (4.33)$$

Полученная ИХ представляет собой нечетную функцию дискретного аргумента, поэтому для синтеза ЦД могут быть

применены антисимметричные трансверсальные фильтры с линейными ФЧХ III и IV типов. Импульсная характеристика (4.33) соответствует *идеальному* ЦД. Так же как и для ДПГ, реализовать ЦД возможно только приближенно.

В случае использования прямоугольного окна для структуры III типа с порядком фильтра $N = 2M$ ИХ (4.33) будет иметь вид

$$h(n) = h_d(n - M) = \begin{cases} \frac{\cos \pi(n - M)}{n - M}, & n \neq M \\ 0, & n = M \end{cases}. \quad (4.34)$$

При использовании КИХ-структуры IV типа с порядком фильтра $N = 2M - 1$ ИХ (4.33) получает задержку на полуцелое число отсчетов $(M - \frac{1}{2})$:

$$h(n) = h_d\left(n - \left(M - \frac{1}{2}\right)\right) = -\frac{\sin \pi\left(n - \left(M - \frac{1}{2}\right)\right)}{\pi\left(n - \left(M - \frac{1}{2}\right)\right)}. \quad (4.35)$$

Так как в линейных нормированных частотах ЧХ (4.3) КИХ-фильтра типа III допускает для верхней частоты обрабатываемого сигнала лишь нулевое значение $K(1/2) = 0$, в то время как для ЦД $K_d(1/2) = i\pi$, то на частотах, близких к $\nu = 1/2$, фильтр типа III будет плохо аппроксимировать ЧХ цифрового дифференциатора. Поэтому КИХ-структура фильтра типа III непригодна для реализации широкополосных и высокочастотных ЦД.

Реализовать широкополосные и высокочастотные дифференциаторы с линейной фазой можно только структурой КИХ-фильтра типа IV. Этот тип фильтра позволяет получить также и низкочастотные, и полосовые ЦД. Несложно заметить, что, по сравнению с (4.34), ИХ (4.35) быстрее убывает на краях с увеличением порядка фильтра. Поэтому усечение (возможно, с применением весовой функции) идеальной ИХ (4.33) с последующим сдвигом для получения ИХ (4.34) или (4.35) даст при одном и том же значении M более точную аппроксимацию ЧХ в случае (4.35).

Если полуцелая задержка на $(M - \frac{1}{2})$ отсчетов, которую дает сигналу КИХ-структура типа IV, не является каким-то ограничивающим фактором, то синтез ЦД по типу IV КИХ-фильтра с линейной фазой обычно дает лучшие результаты. ►



Упражнение. Используя компьютер, постройте графики ошибки аппроксимации АЧХ $|K_d(\nu)|$ идеального ЦД при использовании фильтров с ИХ (4.34) и (4.35) для $M = 5$, $M = 15$.

Точно так же, как и в случае ДПГ, при синтезе ЦД для усечения ИХ можно использовать различные (не только прямоугольные) весовые окна. Однако, как и для частотно-избирательных фильтров, применение оптимизационных методов позволяет получить более точную аппроксимацию идеальных ДПГ и ЦД КИХ-фильтрами меньшего порядка.

4.6. Основные характеристики аналоговых линейных систем и их связь с характеристиками ЛДС

Прежде чем переходить к рассмотрению методов синтеза цифровых БИХ-фильтров, необходимо напомнить некоторые положения из теории аналоговых фильтров. В широком смысле термин *аналоговый фильтр* означает любое устройство, которое преобразует по некоторому правилу входной аналоговый сигнал $x(t)$ в выходной $y(t)$. Мы рассмотрим такие фильтры, которые представляют собой линейные инвариантные во времени системы (ЛИВС). Преобразование сигнала $y(t) = L\{x(t)\}$ задается для них линейным оператором $L\{\dots\}$, причем параметры системы не меняются во времени (система обладает свойством *стационарности*), т. е. при задержке входного сигнала на величину t_0 имеем на выходе системы отклик $y(t - t_0) = L\{x(t - t_0)\}$.

Определения. Отклик $g(t)$ аналоговой ЛИВС на входное воздействие $\delta(t)$ (дельта-функцию) называется *импульсной характеристикой* этой системы: $g(t) = y(t)|_{x(t)=\delta(t)}$. *Каузальной* назовем ЛИВС, для которой $g(t) \equiv 0$ при $t < 0$.

Каузальные системы являются *физически реализуемыми*, так как для них реакция на выходе системы не может появиться ранее появления сигнала на входе. Как можно показать, при подаче на вход произвольного входного сигнала $x(t)$ такого, что $x(t) = 0$ при $t < 0$,

отклик на выходе каузальной системы может быть найден как интегральная свертка ИХ и входного воздействия:

$$y(t) = \int_0^t x(\tau)g(t - \tau)d\tau = \int_0^t g(\tau)x(t - \tau)d\tau. \quad (4.36)$$

Причем значения интегралов (4.36) в данном случае (т.е. для $x(t) = 0$ и $g(t) = 0$ при $t < 0$) останутся теми же, если пределы интегрирования будут выбраны от $-\infty$ до $+\infty$. Далее мы рассматриваем только каузальные аналоговые системы.

В теории аналоговых ЛИВС, помимо преобразования Фурье, важное место занимает преобразование Лапласа. Напомним, что если некоторая функция $f(t)$ является *оригиналом*, т.е. отвечает трем условиям:

1. $f(t) \equiv 0$ при $t < 0$,
2. $f(t)$ может иметь на каждом конечном отрезке $t \in [a; b] \subset [0; +\infty)$ лишь конечное число разрывов, причем только I рода,
3. $f(t)$ имеет при $t \rightarrow +\infty$ ограниченный показатель роста, т.е. $\exists \alpha \in \mathbb{R}, \exists M > 0, \forall t > 0: |f(t)| \leq Me^{\alpha t}$,

то ее преобразованием Лапласа, или *изображением* по Лапласу $F(p) = \mathcal{L}\{f(t)\}$, называется интеграл

$$F(p) = \int_0^{\infty} f(t)e^{-pt} dt, \quad (4.37)$$

который абсолютно сходится в области $\text{Re } p > \alpha$ и представляет собой в этой области аналитическую функцию. Восстановить оригинал $f(t)$ по заданному изображению $F(p)$, т.е. обратить преобразование Лапласа (4.37), можно при помощи интеграла Меллина:

$$f(t) = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} F(p)e^{pt} dp, \quad (4.38)$$

где $x > \alpha$. Особые точки $\{p_k\}$ подынтегральной функции в (4.38) лежат в области $\text{Re } p \leq \alpha$ и, если их число J конечно, для вычисления интеграла (4.38) часто оказывается удобнее использовать вычеты:

$$f(t) = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} F(p)e^{pt} dp = \sum_{k=1}^J \text{Res}[F(p)e^{pt}, p_k]. \quad (4.39)$$

Как можно видеть, при подстановке в функцию (4.37) комплексного аргумента $p = 2\pi i\nu$ или $p = i\omega$ мы получим в (4.37) «половину» интеграла Фурье для линейной или циклической частоты, см. (1.18). То есть если сигнал является оригиналом, то его спектр (1.18) представляет собой преобразование Лапласа $F(p)$ (4.37), рассматри-



ваемое на мнимой оси, а именно $S(v) = F(i2\pi v)$. В силу такой связи интегральных преобразований Лапласа и Фурье во многом общими являются и их свойства. В частности, преобразование Лапласа $Y(p) = \mathcal{L}\{y(t)\}$ интегральной свертки (4.36) представляет собой произведение изображений $G(p) = \mathcal{L}\{g(t)\}$ и $X(p) = \mathcal{L}\{x(t)\}$ (сравните со свойством 6° интеграла Фурье, см. раздел 1.5): $Y(p) = G(p)X(p)$.

Определение. *Передаточной функцией* $G(p)$ аналоговой ЛИВС называется отношение изображения по Лапласу выходного сигнала $y(t)$ к изображению входного сигнала $x(t)$: $G(p) = \mathcal{L}\{y(t)\} / \mathcal{L}\{x(t)\}$.

Передачная функция является характеристикой системы и не зависит от входного воздействия. Так как $\mathcal{L}\{\delta(t)\} = 1$, то ПФ представляет собой преобразование по Лапласу ИХ: $G(p) = \mathcal{L}\{g(t)\}$.

Рассматриваемые в теории аналоговых ЛИВС передачные функции являются *дробно-рациональными*, т.е. представимы в виде отношения алгебраических полиномов $P(p) = \sum_{k=0}^N \beta_k p^k$ и $Q(p) = \sum_{k=0}^M \alpha_k p^k$:

$$G(p) = \frac{\sum_{k=0}^N \beta_k p^k}{\sum_{k=0}^M \alpha_k p^k} = K \frac{\prod_{k=1}^N (p - c_k)}{\prod_{k=1}^M (p - p_k)}, \quad (4.40)$$

где $\{c_k\}_{k=1}^N$ — корни многочлена $P(p)$, $\{p_k\}_{k=1}^M$ — корни многочлена $Q(p)$, а K — некоторый множитель. *Порядком фильтра* называется величина $L = \max(M, N)$. Передачная функция вида (4.40) имеет нули в точках $\{c_k\}_{k=1}^N$ и особые точки — полюсы $\{p_k\}_{k=1}^M$. Нули и полюсы ПФ (4.40) определяют ее с точностью до константы-множителя K . Если коэффициенты $\{\beta_k\}_{k=0}^N$ и $\{\alpha_k\}_{k=0}^M$ многочленов $P(p)$ и $Q(p)$ — вещественные, то и константа K в (4.40) также является вещественной.

Для физически реализуемых систем порядок полинома числителя в (4.40) не превосходит порядка полинома знаменателя: $N \leq M$. *Правильной* называется дробно-рациональная функция вида (4.40) такая, что $N < M$. Если все полюсы $\{p_k\}_{k=1}^M$ правильной ПФ (4.40) являются простыми (кратных полюсов нет), то ее можно представить в виде следующего разложения на элементарные дроби

$$G(p) = \sum_{k=1}^M \frac{C_k}{p - p_k}, \quad (4.41)$$

где, очевидно, $C_k = \lim_{p \rightarrow p_k} (p - p_k)G(p)$. Как мы увидим далее, ПФ вида (4.41) часто встречается в теории аналоговых фильтров; воспользовавшись формулой (4.39), несложно установить, что соответствующая ИХ

$$g(t) = \sum_{k=1}^M C_k e^{p_k t}. \tag{4.42}$$

Определение. Аналоговая система называется *устойчивой*, если любое ограниченное входное воздействие $x(t)$ ($\exists C_1 < \infty \forall t: |x(t)| \leq C_1$) вызывает ограниченный отклик $y(t)$ ($\exists C_2 < \infty \forall t: |y(t)| \leq C_2$).

Если система является неустойчивой, то при подаче некоторого ограниченного входного воздействия $x(t)$ реакция системы $y(t)$ неограниченно возрастает по абсолютной величине. Для проверки устойчивости аналоговых систем используются следующие необходимые и достаточные условия (критерии устойчивости).

1. Система устойчива \Leftrightarrow для ИХ $g(t)$ абсолютно сходится интеграл $\int_0^{\infty} |g(t)| dt < \infty$.
2. Система устойчива \Leftrightarrow все полюсы $\{p_k\}_{k=1}^M$ ПФ (4.40) лежат в области $\text{Re } p_k < 0$ (левая полуплоскость комплексной плоскости).

Определение. *Частотной характеристикой* аналоговой ЛИВС называется спектр ИХ: $S_g(\nu) = G(i2\pi\nu) = \int_0^{\infty} g(t)e^{-2\pi i\nu t} dt$ (в линейных частотах) или $S(\omega) = G(i\omega) = \int_0^{\infty} g(t)e^{-i\omega t} dt$ (в циклических частотах). *Амплитудно-частотной характеристикой* (АЧХ) называется модуль ЧХ $|S_g(\nu)|$ (или $|S(\omega)|$), *фазочастотной характеристикой* (ФЧХ) — аргумент $\varphi(\nu) = \arg S_g(\nu)$ (или $\varphi(\omega) = \arg S(\omega)$) ЧХ.

Частотная характеристика определяет преобразование спектра сигнала при его прохождении через ЛИВС: из соотношения для изображений $Y(p) = G(p)X(p)$ при подстановке $p = i2\pi\nu$ имеем

$$\underbrace{Y(i2\pi\nu)}_{S_y(\nu) \text{ - спектр вых. сигнала}} = \underbrace{G(i2\pi\nu)}_{S_g(\nu) \text{ ЧХ}} \underbrace{X(i2\pi\nu)}_{S_x(\nu) \text{ - спектр вход. сигнала}}.$$

Если во временной области сигнал представляет собой комплексное гармоническое колебание $u(t) = Ae^{i2\pi\nu_0 t}$ (здесь

$A = |A|e^{i\gamma}$ — комплексная амплитуда), то для отклика системы с ИХ $g(t)$ по (4.36) получаем

$$v(t) = \int_0^t g(\tau) \underbrace{A e^{2\pi i \nu_0 (t-\tau)}}_{u(t-\tau)} d\tau = \underbrace{A e^{2\pi i \nu_0 t}}_{u(t)} \int_0^t \underbrace{g(\tau) e^{-2\pi i \nu_0 \tau}}_{S_g(\nu) \text{ при } t \rightarrow \infty} d\tau \underset{t \rightarrow \infty}{=} S_g(\nu_0) u(t).$$

Поэтому ЧХ представляет собой комплексный коэффициент передачи гармонического колебания в установившемся ($t \rightarrow \infty$) режиме. С использованием понятий АЧХ и ФЧХ представим ЧХ в показательной форме как $S_g(\nu) = |S_g(\nu)| e^{i\varphi(\nu)}$, тогда для вещественного гармонического воздействия $x(t) = |A| \cos(2\pi\nu_0 t + \gamma) = \text{Re}\{u(t)\}$ в установившемся режиме получим отклик аналогового фильтра с вещественной ИХ $g(t)$:

$$\begin{aligned} y(t) &= \int_0^t g(\tau) \underbrace{x(t-\tau)}_{\text{Re}\{u(t-\tau)\}} d\tau = \text{Re} \int_0^t g(\tau) \underbrace{A e^{2\pi i \nu_0 (t-\tau)}}_{u(t-\tau)} d\tau \underset{t \rightarrow \infty}{=} \text{Re} \{ S_g(\nu_0) u(t) \} = \\ &= \text{Re} \{ |S_g(\nu_0)| e^{i\varphi(\nu_0)} |A| e^{i\gamma} e^{2\pi i \nu_0 t} \} = |A| |S_g(\nu_0)| \cos(2\pi\nu_0 t + \gamma + \varphi(\nu_0)). \end{aligned}$$

Таким образом, при прохождении через ЛИВС гармоническое колебание частоты ν_0 изменяет амплитуду пропорционально величине АЧХ $|S_g(\nu_0)|$ и получает сдвиг фазы, определяемый значением ФЧХ $\varphi(\nu_0)$.

Понятия импульсной характеристики, передаточной функции, частотной характеристики переносятся с аналоговых ЛИВС на ЛДС с учетом специфики последних (определения ИХ, ПФ, ЧХ для ЛДС см. в главе 3). Рассмотрим взаимосвязь соответствующих характеристик аналоговых и дискретных линейных систем.

Сначала заметим, что если рассматривать аналоговую ИХ $g(t)$ лишь в дискретные моменты времени $\{t_n = nT\}_{n=0}^{\infty}$, где T — период дискретизации, то спектр (2.9) соответствующего дискретного сигнала $\{g_n = g(nT)\}_{n=0}^{\infty}$ можно представить (см. разделы 1.5, 2.4) как

$$\tilde{S}_g(\nu) = \mathcal{F} \left(T \sum_{n=0}^{\infty} g(nT) \delta(t - nT) \right) = T \sum_{n=0}^{\infty} g(nT) e^{-2\pi i \nu n T} = \sum_{m=-\infty}^{\infty} S_g \left(\nu - \frac{m}{T} \right), \quad (4.43)$$

где $S_g(\nu) = \mathcal{F}\{g(t)\} = \int_0^{\infty} g(t) e^{-2\pi i \nu t} dt$ — ЧХ оригинального аналогового фильтра.

Назвав импульсной характеристикой $h(n)$ ЛДС отклик на единичное воздействие $\tilde{\delta}(n) = \begin{cases} 1, n=0 \\ 0, n \neq 0 \end{cases}$ (см. раздел 3.3), будем

трактовать значения $h(n)$ как масштабированные отсчеты ИХ $g(t)$ некоторой аналоговой ЛИВС, а именно: $h(n) = Tg(nT)$, где T — период дискретизации. Тогда из ПФ дискретного фильтра $H(z) = Z\{h(n)\} = T \sum_{n=0}^{\infty} g(nT)z^{-n}$ спектр (4.43) можно получить как $\tilde{S}_g(v) = H(e^{2\pi i v T})$ — эта функция представляет собой ЧХ дискретного фильтра (см. раздел 3.6) с ИХ $h(n)$. Для циклических частот ЧХ соответствующего дискретного фильтра $\tilde{S}(\omega) = H(e^{i\omega T})$. Таким образом, для ЛДС с ИХ $h(n) = Tg(nT)$ частотной характеристикой является преобразование Фурье (в линейных или в циклических частотах) решетчатой функции $\tilde{g}(t) = \sum_{n=0}^{\infty} h(n)\delta(t - nT)$:

$$\tilde{S}_g(v) = \int_0^{\infty} \tilde{g}(t)e^{-2\pi i v t} dt \quad \text{или} \quad \tilde{S}(\omega) = \int_0^{\infty} \tilde{g}(t)e^{-i\omega t} dt.$$

Помня о рассмотренной выше связи интегральных преобразований Фурье и Лапласа и вытекающей из них связи между ЧХ и ПФ аналоговых ЛИВС, логично было бы назвать передаточной функцией ЛДС комплекснозначную функцию

$$\tilde{G}(p) = \mathcal{L}\{\tilde{g}(t)\} = \int_0^{\infty} \left(\sum_{n=0}^{\infty} h(n)\delta(t - nT) \right) e^{-pt} dt = \sum_{n=0}^{\infty} h(n)e^{-pTn}, \quad (4.44)$$

которая называется *дискретным преобразованием Лапласа* (ДПЛ) или *D-преобразованием* последовательности $\{h(n)\}_{n=0}^{\infty}$. Вместо (4.44), однако, в качестве ПФ используют *Z-преобразование* ИХ, $H(z) = \sum_{n=0}^{\infty} h(n)z^{-n}$, которое, очевидно, позволяет выразить функцию (4.44) как $\tilde{G}(p) = H(e^{pT})$. Использование вместо ДПЛ (4.44) *Z-преобразования* (3.1) является более удобным, поскольку в этом случае ПФ $H(z)$ принимает вид дробно-рациональной функции (3.10), которую проще анализировать, чем $\tilde{G}(p)$.

Переход от ДПЛ к *Z-преобразованию* означает замену переменных $z = e^{pT}$, определяющую некоторое отображение комплексной p -плоскости в комплексную z -плоскость. Поскольку $|z| = e^{(\text{Re } p)T}$ и период дискретизации $T > 0$, то точки левой полуплоскости из p -плоскости ($\text{Re } p < 0$) отображаются в z -области внутрь круга $|z| < 1$. Если же $\text{Re } p > 0$, то $|z| > 1$. Мнимая ось $p = i\omega$ отображается в единичную окружность $|z| = 1$, поэтому ЧХ ЛДС $\tilde{S}(\omega) = \tilde{G}(i\omega) = H(e^{i\omega T})$.

Эта функция является периодической, $\tilde{S}(\omega) = \tilde{S}\left(\omega + \frac{2\pi}{T}\right)$, и один ее период соответствует в z -области однократному прохождению замкнутого контура $|z| = 1$, а основной полосе циклических частот

$\omega \in [0; \pi/T]$ соответствует та часть окружности $|z| = 1$, которая приходится на верхнюю полуплоскость $\text{Im } z \geq 0$. Схематически отображение $p \rightarrow z = e^{pT}$ изображено на рис. 4.12.

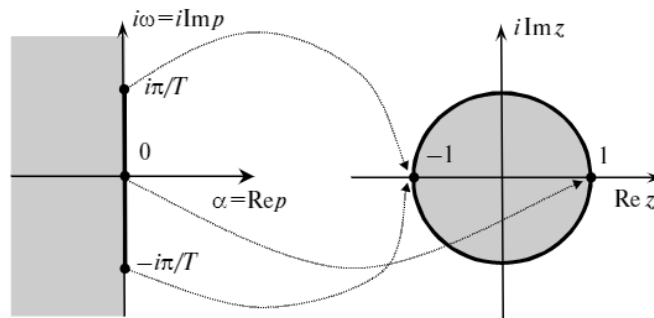


Рис. 4.12. Отображение комплексной плоскости $p = \alpha + i\omega \rightarrow z = e^{pT}$

4.7. Синтез БИХ-фильтров по аналоговым прототипам

Спецификацию требований (схему допусков) в случае частотно-избирательных БИХ-фильтров принято задавать в несколько ином виде по сравнению с тем, что рассматривалось ранее для КИХ-фильтров. Отличие состоит в определении параметра неравномерности в полосе пропускания δ_p . Данный параметр по-прежнему характеризует «размах» отклонений АЧХ в полосе пропускания, но в спецификациях БИХ-фильтров такие отклонения теперь могут быть только в меньшую сторону относительно максимального значения АЧХ. Общий вид спецификации АЧХ для полосового фильтра представлен на рис. 4.13 (сравните с заданием параметра δ_p для КИХ-фильтра на рис. 4.1 и 4.6). Неравномерность в полосе пропускания в децибелах для БИХ-фильтров определяется как $A_p = -20 \lg(1 - \delta_p)$.

Наиболее распространенные методы синтеза частотно-избирательных БИХ-фильтров основаны на использовании хорошо развитой теории расчета аналоговых фильтров. Различные методы

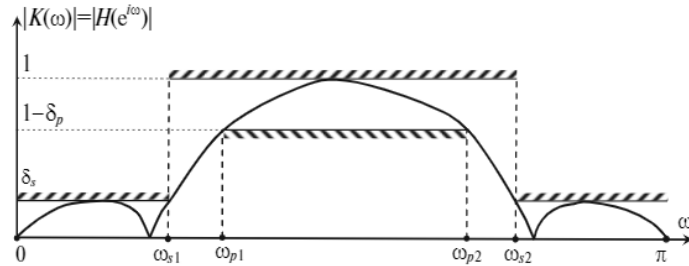


Рис. 4.13. Схема допусков и пример АЧХ для полосового БИХ-фильтра

синтеза цифровых БИХ-фильтров отличаются фактически тем, каким образом устанавливается соответствие (отображение) между синтезируемым дискретным фильтром и его «двойником» — аналоговым *фильтром-прототипом*. Мы рассмотрим далее два способа такого отображения, которые определяют *метод инвариантности импульсной характеристики* и *метод билинейного преобразования*. Однако прежде необходимо дать краткий обзор методов синтеза аналоговых фильтров.

В теории синтеза аналоговых частотно-избирательных (НЧ, ВЧ, полосовых, режекторных) фильтров требования к разрабатываемому фильтру сначала при помощи определенной замены переменной $\omega \rightarrow \Omega$ переводятся в требования к *нормированному НЧ фильтру-прототипу* (также аналоговому) с границей полосы пропускания на частоте $\Omega_p = 1$ рад/с, см. рис. 4.14. Для такого прототипа находятся полюсы и нули передаточной функции, обеспечивающей выполнение спецификации рис. 4.14, которые затем в результате процедуры

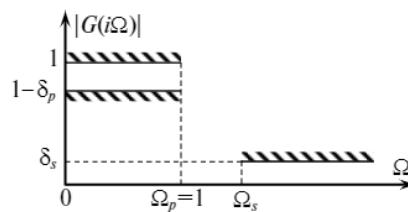


Рис. 4.14. Схема допусков для нормированного НЧ фильтра-прототипа



денормирования пересчитываются в нули и полюсы ПФ синтезируемого фильтра.

В этом разделе для ПФ синтезируемого аналогового фильтра будем использовать обозначение $H(p)$, для ПФ нормированного НЧ-прототипа — $G(s)$. Частотными характеристиками являются функции $H(i\omega)$ и $G(i\Omega)$. Приведем формулы, которые используются для преобразования различных частотно-избирательных фильтров в нормированный НЧ-прототип. Отметим, что все соответствующие отображения комплексных плоскостей $p \leftrightarrow s$, которые рассматриваются ниже, отвечают следующим необходимым требованиям:

- сохраняется дробно-рациональный вид функций $H(p)$ и $G(s)$;
- точки левой полуплоскости комплексной плоскости отображаются также в левую полуплоскость, т. е. устойчивые фильтры преобразуются в устойчивые.

Преобразование НЧ фильтра. Это самый простой тип преобразования частот ω синтезируемого фильтра в частоты Ω нормированного НЧ фильтра-прототипа. Переход от функции $H(p)$ к $G(s)$ осуществляется масштабированием переменной: $s = p/\omega_p$, $G(s) = H(s\omega_p)$, где ω_p — заданная граница полосы пропускания (рад/с). Формула пересчета частот представляет собой частный случай переменных $s = i\Omega$, $p = i\omega$ и потому имеет аналогичный вид:

$$\Omega = \omega/\omega_p. \quad (4.45)$$

Граничная частота полосы задерживания для фильтра-прототипа в нормированных частотах принимает значение $\Omega_s = \omega_s/\omega_p$. После решения задачи синтеза НЧ-прототипа и нахождения его передаточной функции $G(s)$ полюсы и нули $G(s)$ преобразуются в полюсы и нули $H(p)$ в результате обратного масштабирования (денормирования):

$$p = s\omega_p. \quad (4.46)$$

Преобразование ВЧ фильтра. Формула преобразования частоты имеет вид:

$$\Omega = \omega_p/\omega. \quad (4.47)$$

Параметр граничной частоты полосы задерживания для фильтра-прототипа принимает значение $\Omega_s = \omega_p/\omega_s$. Схема отображения

ЧХ ВЧ-фильтра в ЧХ нормированного НЧ-прототипа отражена на рис. 4.15.

Нули и полюсы функции $G(s)$ преобразуются в нули и полюсы $H(p)$ посредством формулы денормирования

$$p = \omega_p / s. \quad (4.48)$$

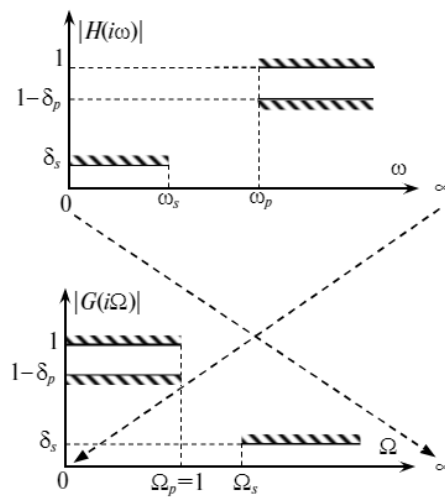


Рис. 4.15. Преобразование ЧХ ВЧ-фильтра в ЧХ нормированного НЧ-прототипа

Преобразование полосового фильтра. Переход от ПФ полосового фильтра $H(p)$ к ПФ нормированного НЧ прототипа $G(s)$ осуществляется заменой переменных

$$s = \frac{\omega_0}{\omega_{p2} - \omega_{p1}} \left(\frac{p}{\omega_0} + \frac{\omega_0}{p} \right), \quad (4.49)$$

где $\omega_0 = \sqrt{\omega_{p1}\omega_{p2}}$. Подставив в (4.49) значения переменных $s = i\Omega$ и $p = i\omega$, получим формулу пересчета частот для аналогового прототипа:

$$\Omega = \frac{\omega_0}{\omega_{p2} - \omega_{p1}} \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right). \quad (4.50)$$

Несложно убедиться, что при подстановке в (4.50) значений $\omega = \omega_{p2}$ и $\omega = \omega_{p1}$ получим частоты для нормированного НЧ-

прототипа $\Omega = 1$ и $\Omega = -1$ соответственно. Подстановка в формулу (4.50) значений границ полос подавления $\omega = \omega_{s2}$ и $\omega = \omega_{s1}$ в общем случае дает частоты нормированного прототипа, различающиеся по абсолютной величине: $-\Omega_{s1} \neq \Omega_{s2}$. Так как АЧХ фильтров с вещественными коэффициентами должна быть всегда симметрична относительно нулевой частоты, то для выполнения условия $\Omega_s = \Omega_{s2} = -\Omega_{s1}$ границу полосы подавления для нормированного НЧ-прототипа выбираем по наиболее жесткому требованию: $\Omega_s = \min(-\Omega_{s1}, \Omega_{s2})$. Связь АЧХ нормированного НЧ-прототипа и заданного полосового фильтра отражена на рис. 4.16.

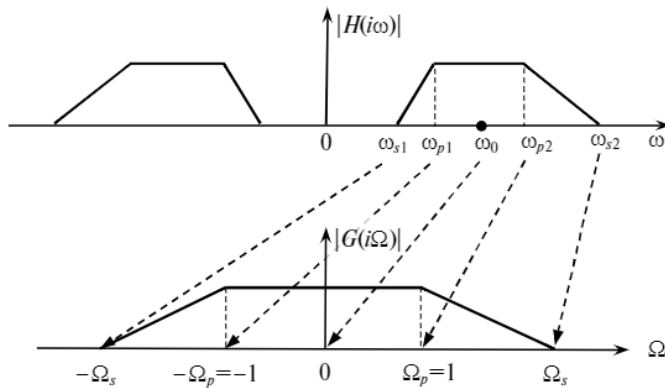


Рис. 4.16. Соответствие частот нормированного НЧ-прототипа и полосового фильтра, определяемое соотношением (4.50)

После определения нулей и полюсов $\{s_k\}$ ПФ $G(s)$ фильтра-прототипа нули и полюсы функции $H(p)$ находятся в результате решения уравнения (4.49) относительно переменной p . Так как уравнение (4.49) сводится к квадратному, то каждой точке s соответствует два значения p :

$$p_{1,2} = \frac{1}{2} \left((\omega_{p2} - \omega_{p1})s \pm \sqrt{(\omega_{p2} - \omega_{p1})^2 s^2 - 4\omega_{p2}\omega_{p1}} \right). \quad (4.51)$$

Упражнение. Докажите справедливость соотношения (4.51).

Преобразование режекторного фильтра. Переход к ПФ нормированного НЧ-прототипа осуществляется в результате замены переменных

$$s = \frac{(\omega_{p2} - \omega_{p1})p}{p^2 + \omega_0^2}, \tag{4.52}$$

где $\omega_0^2 = \omega_{p1}\omega_{p2}$. Отсюда, подставляя $s = i\Omega$ и $p = i\omega$, получаем:

$$\Omega = \frac{(\omega_{p2} - \omega_{p1})\omega}{\omega_0^2 - \omega^2}. \tag{4.53}$$

Упражнение. Убедитесь, что величины $s_1 = s(p)$ (4.49) и $s_2 = s(p)$ (4.52) взаимно обратны, т. е. связаны соотношением $s_1s_2 = 1$.

Соответствие осей частот режекторного фильтра и НЧ-прототипа отражено на рис. 4.17. Аналогично случаю полосового фильтра, рассчитав по (4.53) для значений $\omega = \omega_{s2}$ и $\omega = \omega_{s1}$ соответствующие величины Ω_{s2} и Ω_{s1} , далее в качестве параметра границы полосы подавления для нормированного НЧ-прототипа выбираем $\Omega_s = \min(-\Omega_{s2}, \Omega_{s1})$. Каждый ноль (полюс) ПФ $G(s)$ вновь порождает по два нуля (полюса) функции $H(p)$, которые находятся из решения уравнения (4.52) относительно p (убедитесь!):

$$p_{1,2} = \frac{\omega_{p2} - \omega_{p1}}{2s} \pm \frac{1}{2} \sqrt{\left(\frac{\omega_{p2} - \omega_{p1}}{s}\right)^2 - 4\omega_{p2}\omega_{p1}}. \tag{4.54}$$

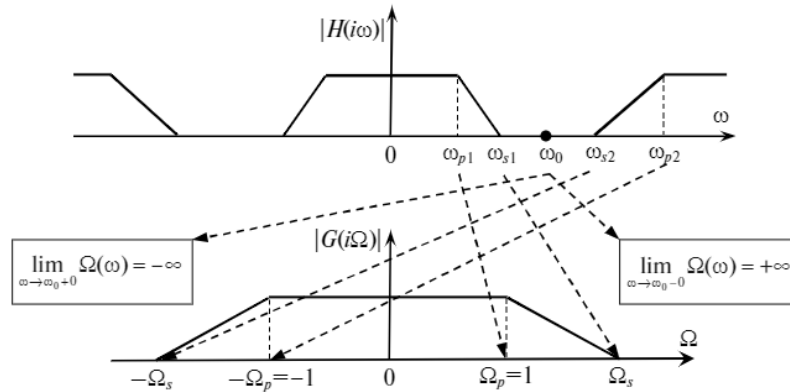


Рис. 4.17. Соответствие частот нормированного НЧ-прототипа и режекторного фильтра, определяемое соотношением (4.53)

После того как синтезируемый аналоговый фильтр приведен к нормированному низкочастотному фильтру-прототипу, имеющему те же самые ограничения на неравномерность АЧХ δ_p и δ_s , единичную граничную частоту $\Omega_p = 1$ полосы пропускания и пересчитанную (по одной из формул (4.45), (4.47), (4.50), (4.53)) частоту Ω_s , далее осуществляется аппроксимация идеальной АЧХ нормированного фильтра-прототипа.

Для нормированного НЧ-прототипа в качестве аппроксимирующих используются четыре стандартных вида функции АЧХ $|G(i\Omega)|$, которые определяют типы соответствующих аналоговых фильтров: Баттерворта, Чебышева I рода, Чебышева II рода, Золотарева — Кауэра (или *эллиптический*). Виды АЧХ этих фильтров приведены на рис. 4.18.

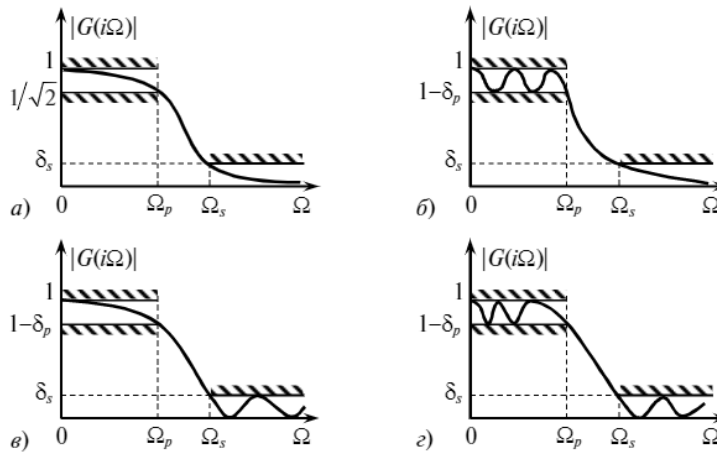


Рис. 4.18. Схематический вид АЧХ стандартных НЧ фильтров-прототипов ($\Omega_p = 1$): а) Баттерворта, б) Чебышева I рода, в) Чебышева II рода, з) эллиптического

Вместо АЧХ $|G(i\Omega)|$ часто оказывается удобнее рассматривать квадрат АЧХ $|G(i\Omega)|^2$. Для этого введем вспомогательную функцию $F(s) = G(s)G(-s)$. Так как ЧХ $G(i\Omega)$ есть преобразование Фурье вещественной импульсной характеристики, то в силу свойства сопряженной симметрии интеграла Фурье имеем $G(-i\Omega) = \overline{G(i\Omega)}$ и $F(i\Omega) = |G(i\Omega)|^2$.

Фильтры Баттерворта

Для фильтров Баттерворта N -го порядка для аппроксимации квадрата АЧХ нормированного НЧ-прототипа используется функция

$$F(s) = G(s)G(-s) = \frac{1}{1 + (-s^2)^N}, \quad (4.55)$$

откуда

$$F(i\Omega) = |G(i\Omega)|^2 = \frac{1}{1 + \Omega^{2N}}. \quad (4.56)$$

Из (4.56) видно, что у фильтра Баттерворта любого порядка N на граничной частоте полосы пропускания $\Omega_p = 1$ АЧХ $|G(i\Omega_p)| = 1/\sqrt{2}$. АЧХ данного фильтра монотонно убывает и имеет максимальное значение $|G(0)| = 1$.

Первые $2N - 1$ производных квадрата АЧХ (4.56) обращаются в ноль при $\Omega = 0$, что является максимально возможным количеством нулевых производных по сравнению с другими типами фильтров того же порядка N . Поэтому АЧХ фильтров Баттерворта называют *максимально плоской* в полосе пропускания. Другим достоинством этих фильтров является то, что в полосе пропускания их ФЧХ ближе к линейной по сравнению с ФЧХ фильтров Чебышева и Золотарева — Кауэра.

С увеличением порядка фильтра N крутизна спада АЧХ увеличивается, переходная полоса сужается, см. рис. 4.19.

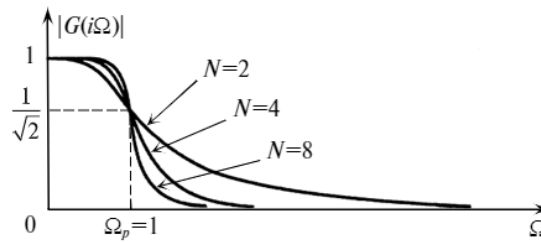


Рис. 4.19. АЧХ фильтров Баттерворта различных порядков N

Упражнение. Покажите, что для определяемой (4.56) АЧХ порядок фильтра N может быть найден по заданной границе полосы подавления Ω_s и неравномерности в полосе подавления δ_s из соотношения

$$N \geq \frac{\lg\left(\frac{1}{\delta_s^2} - 1\right)}{2\lg\Omega_s}. \quad (4.57)$$

Пример 4.13. Найти передаточную функцию фильтра Баттерворта.

◀ Найдем сначала нули и полюсы функции $F(s) = G(s)G(-s)$ (4.55). Для $F(s)$ точка комплексной плоскости $s = \infty$ является нулем кратности $2N$, поэтому $G(s)$ должна иметь в точке $s = \infty$ ноль кратности N .

Для нахождения полюсов функции $F(s)$ приравняем знаменатель (4.55) нулю, имеем: $1 + (is)^{2N} = 0$. Отсюда находим $2N$ полюсов:

$$s_k = \frac{1}{i} \sqrt[2N]{-1} = \frac{1}{i} \left(\cos \frac{\pi + 2\pi k}{2N} + i \sin \frac{\pi + 2\pi k}{2N} \right), \quad k = 0, 1, \dots, 2N - 1;$$

или окончательно

$$s_k = \sin \frac{\pi(2k+1)}{2N} - i \cos \frac{\pi(2k+1)}{2N}. \quad (4.58)$$

Несложно заметить, что $\forall k |s_k| = 1$ и $s_k = -s_{k+N}$, т.е. каждый полюс имеет пару с обратным знаком. Для задания передаточной функции $G(s)$ можно взять любые N полюсов, по одному из N пар, тогда в (4.55) войдут все полюсы (4.58). Однако из теории аналоговых фильтров известно, что для их устойчивости необходимо и достаточно, чтобы все полюсы ПФ лежали в левой полуплоскости комплексной плоскости. Потому ПФ фильтра Баттерворта будем строить, выбирая такие полюсы из каждой пары $s_k = -s_{k+N}$, для которых $\operatorname{Re} s_k = \sin \frac{\pi(2k+1)}{2N} < 0$, т.е. из набора (4.58) выбираем только полюсы с индексами $k = N, \dots, 2N - 1$.

По точкам нулей и полюсов передаточная функция определяется с точностью до вещественной нормировочной константы K и для фильтра Баттерворта принимает вид:

$$G(s) = K / \prod_{k=N}^{2N-1} (s - s_k), \quad (4.59)$$

где полюсы s_k находятся по (4.58). ▶

Упражнение. Покажите, что условие $G(0) = 1$ (см. (4.55) и рис. 4.18) определяет в (4.59) множитель $K = 1$. *Указание:* учтите, что в (4.59) полюсы $|s_k| = 1$ и образуют комплексно сопряженные пары $s_{N+m} = \bar{s}_{2N-1-m}$, $m = 0, 1, \dots, N - 1$.

Фильтры Чебышева

Мы наблюдали выше, что АЧХ фильтра Баттерворта нормированного НЧ-прототипа монотонно убывает с увеличением частоты, как в полосе пропускания, так и в полосе подавления. Если допустить наличие пульсаций АЧХ, то при тех же спецификациях порядок фильтра можно понизить. Это позволяют сделать фильтры Чебышева. В зависимости от того, в какой области частот минимизируется ошибка аппроксимации идеальной АЧХ, различают (см. рис. 4.18):

- фильтры первого рода, АЧХ которых равномерно пульсирует в области пропускания и монотонно убывает в области подавления;
- фильтры второго рода, АЧХ которых монотонно убывает в области пропускания и равномерно пульсирует в области подавления.

Фильтры Чебышева II рода называются также *инверсными* фильтрами I рода.

Для фильтров Чебышева I рода порядка N квадрат АЧХ нормированного НЧ-прототипа аппроксимируется при помощи функции

$$|G_1(i\Omega)|^2 = \frac{1}{1 + \varepsilon^2 T_N^2(\Omega)}, \quad (4.60)$$

где $T_N(\Omega)$ — полином Чебышева N -го порядка (см. (4.24)), а ε — параметр, определяющий величину равномерных пульсаций в полосе пропускания:

$$1 - \delta_p = 1/\sqrt{1 + \varepsilon^2}. \quad (4.61)$$

По заданным спецификациям порядок фильтра N для нормированного НЧ-прототипа находится из соотношения

$$N \geq \text{arch} \left(\frac{\sqrt{\frac{1}{\delta_s^2} - 1}}{\sqrt{(1 - \delta_p)^2 - 1}} \right) / \text{arch } \Omega_s. \quad (4.62)$$

Передаточная функция фильтра Чебышева I рода имеет ноль кратности N в точке $s = \infty$, а полюсы находятся по формуле ($k = 0, 1, \dots, N-1$):

$$s_k = -\text{sh} \frac{\text{arsh}(1/\varepsilon)}{N} \sin \frac{(2k+1)\pi}{2N} + i \text{ch} \frac{\text{arsh}(1/\varepsilon)}{N} \cos \frac{(2k+1)\pi}{2N}. \quad (4.63)$$

Функция квадрата АЧХ для фильтров Чебышева II рода (инверсных) порядка N находится по (4.60) как $|G_{II}(i\Omega)|^2 = 1 - |G_I(i\Omega_s/\Omega)|^2$ и принимает вид

$$|G_{II}(i\Omega)|^2 = \frac{\varepsilon^2 T_N^2(\Omega_s/\Omega)}{1 + \varepsilon^2 T_N^2(\Omega_s/\Omega)}, \quad (4.64)$$

где параметр ε определяет уже величину пульсаций в полосе подавления:

$$\delta_s = \varepsilon / \sqrt{1 + \varepsilon^2}. \quad (4.65)$$

По заданной спецификации АЧХ порядок фильтра второго рода для нормированного НЧ-прототипа находится из того же условия, что и для фильтра первого рода, — по формуле (4.62).

Полюсы ПФ фильтра Чебышева II рода $\{p_k\}_{k=1}^N$ определяются по полюсам передаточной функции фильтра первого рода (4.63) (где параметр ε выражается из (4.65)) как величины $p_k = \Omega_s/s_k$.

Для фильтра Чебышева II рода передаточная функция имеет также N нулей, которые находятся как величины $z_k = i\Omega_s/t_k$, где t_k — нули (все — вещественные!) многочлена Чебышева $T_N(t)$, $k = 1, \dots, N$.

Упражнение. Основываясь на свойствах полиномов Чебышева (см. раздел 4.4), обоснуйте справедливость соотношений (4.61) и (4.65).

Эллиптические фильтры Золотарева — Кауэра

Этот тип фильтров имеет равномерные пульсации АЧХ как в области пропускания, так и в области подавления, см. рис. 4.18г. Квадрат АЧХ нормированного НЧ-прототипа имеет в данном случае вид:

$$|G(i\Omega)|^2 = \frac{1}{1 + \varepsilon^2 U_N^2(\Omega, k)},$$

где параметр ε^2 определяется по (4.61), а $U_N(\Omega, k)$ представляет собой эллиптическую функцию Якоби, параметр k который определенным образом вычисляется по δ_p и δ_s . Эллиптические фильтры среди прочих рассмотренных нами типов обеспечивают минимальную ширину переходной полосы для заданного порядка фильтра N и фиксированных значений неравномерностей ЧХ δ_p и δ_s . Недостатком данного типа фильтров является существенная нелинейность ФЧХ. Поэтому эллиптические фильтры применяются,

когда необходимо получить высокую частотную избирательность системы (резкий спад АЧХ), а линейность ФЧХ не является важным требованием. Для нахождения нулей и полюсов передаточной функции фильтров Золотарева — Кауэра необходимо использовать численные методы, которые алгоритмически реализованы в различных компьютерных программах синтеза фильтров.

Рассмотрев основные идеи методов синтеза аналоговых частотно-избирательных фильтров, представим теперь общую формализованную схему синтеза цифровых БИХ-фильтров на основе аналоговых прототипов в виде следующей последовательности шагов.

1. По спецификации частотных характеристик цифрового фильтра формулируются требования к аналоговому фильтру-прототипу.
2. Требования к аналоговому фильтру-прототипу (НЧ, ВЧ, полосовому, режекторному) преобразуются в требования к нормированному аналоговому НЧ-прототипу с единичной граничной частотой полосы пропускания $\Omega_p = 1$.
3. В зависимости от требований к линейности фазы и сложности реализации фильтра, имеющихся возможностей использования средств поддержки автоматизированной разработки выбирается тип аппроксимации АЧХ.
4. Для выбранного типа аппроксимации осуществляется синтез аналогового НЧ-прототипа, обеспечивающего требования к ЧХ при минимальном порядке фильтра N . То есть находятся нули и полюсы передаточной функции нормированного аналогового НЧ-прототипа.
5. Найденные нули и полюсы денормируются, т.е. пересчитываются в нули и полюсы передаточной функции аналогового прототипа цифрового фильтра.
6. Производится отображение (пересчет) нулей и полюсов передаточной функции фильтра из аналоговой области в цифровую.
7. Проводится проверка выполнения заданных спецификаций цифрового фильтра.

Прокомментируем приведенную схему. Шаг 1 — это, фактически, перенос значений частотных границ полос пропускания и подавления, неравномерностей в этих полосах на требования, предъявляемые к аналоговому фильтру. Простой перенос спецификации,

однако, не учитывает последующие возможные эффекты искажения ЧХ из-за наложения частот, и может случиться, что итоговый фильтр придется синтезировать повторно, несколько ужесточив требования к аналоговому прототипу по сравнению с заданными исходными спецификациями цифрового фильтра (по неравномерности АЧХ, по ширине переходной полосы).

Шаги 2–5 — это собственно процедура синтеза аналогового фильтра, основные идеи которой кратко рассмотрены нами в данном разделе. На сегодняшний день эта процедура имеет мощную поддержку в виде различных специализированных программных средств автоматизации проектирования фильтров.

Конкретный метод синтеза цифрового фильтра по аналоговому прототипу определяет именно шаг 6. Это своего рода «проектирование» фильтра из аналоговой области в цифровую, которое неизбежно связано с некоторой ошибкой. Далее мы рассмотрим два таких наиболее распространенных метода: метод инвариантности импульсной характеристики и метод билинейного преобразования.

Шаг 7 представляет собой проверку того, удалось ли разработать в итоге фильтр, отвечающий исходной спецификации. Наличие упомянутых ошибок, также как и возможных ошибок применявшихся численных методов, может привести к тому, что итоговые характеристики синтезированного фильтра не отвечают требованиям исходной спецификации. В этом случае необходимо повторить процедуру синтеза фильтра, несколько ужесточив по сравнению с исходными начальными требованиями спецификации для аналогового прототипа.

Перейдем к рассмотрению конкретных методов проектирования БИХ-фильтров по аналоговым прототипам.

4.8. Синтез БИХ-фильтров методом инвариантности импульсной характеристики

Будем использовать далее обозначения $g(t)$ и $G(p)$ для ИХ и ПФ аналогового фильтра-прототипа, а для соответствующего ему

дискретного фильтра обозначим ИХ и ПФ соответственно $h(n)$ и $H(z)$.

Идея метода состоит в том, чтобы, синтезировав аналоговый фильтр-прототип с импульсной характеристикой $g(t)$, затем выполнить ее дискретизацию: $h(n) = Tg(nT)$. Полученная ИХ $\{h(n)\}_{n=0}^{\infty}$ определяет дискретный фильтр, передаточная функция которого находится как Z -преобразование ИХ: $H(z) = Z\{h(n)\}$. Равенство дискретной и аналоговой ИХ в точках дискретизации, $t_n = nT$, и понимается под термином *инвариантность* (фактически, относительно дискретизации) в названии метода.

Рассмотрим подробнее специфику такого отображения аналогового фильтра в дискретный. Пусть ПФ денормированного аналогового фильтра-прототипа $G(p)$ содержит N простых (т.е. единичной кратности) полюсов $\{p_k\}_{k=1}^N$ и менее чем N нулей. Тогда она может быть представлена в виде:

$$G(p) = \sum_{k=1}^N \frac{C_k}{p - p_k}, \quad (4.66)$$

где C_k — некоторые числа (возможно, комплексные), которые могут быть найдены по формуле $C_k = \lim_{p \rightarrow p_k} (p - p_k)G(p)$. Импульсная характеристика аналогового фильтра восстанавливается по ПФ при помощи обратного преобразования Лапласа и для (4.66) принимает вид (см. (4.41), (4.42)):

$$g(t) = \sum_{k=1}^N C_k e^{p_k t}.$$

Выполнив дискретизацию ИХ $g(t)$ с периодом T , получим выражение для ИХ дискретного фильтра:

$$h(n) = Tg(nT) = T \sum_{k=1}^N C_k (e^{p_k T})^n. \quad (4.67)$$

Имеющий смысл нормировочного, множитель T вводится в (4.67) для того, чтобы избежать возрастания коэффициента усиления фильтра при уменьшении периода дискретизации T . Тогда ЧХ дискретного фильтра будет примерно соответствовать ЧХ аналогового фильтра (см. также раздел 4.6):

$$G(i\omega) = \int_0^{\infty} g(t)e^{-i\omega t} dt \approx T \sum_{n=0}^{\infty} g(nT)e^{-i\omega nT} = \sum_{n=0}^{\infty} h(n)e^{-i\omega nT} = H(e^{i\omega T}) = K(\omega).$$

Приближенное равенство будет выполняться здесь тем точнее, чем меньше период дискретизации T .

Передаточную функцию синтезируемого дискретного фильтра найдем как Z-преобразование ИХ (4.67):

$$H(z) = T \sum_{n=0}^{\infty} \left(\sum_{k=1}^N C_k e^{p_k T n} \right) z^{-n} = T \sum_{k=1}^N C_k \sum_{n=0}^{\infty} \left(\frac{e^{p_k T}}{z} \right)^n = T \sum_{k=1}^N \frac{C_k}{1 - e^{p_k T} z^{-1}}. \quad (4.68)$$

Таким образом, каждое слагаемое из суммы (4.66) (p -плоскость, область ПФ аналогового фильтра) отображается в соответствующее слагаемое суммы (4.68) (z -плоскость, область ПФ дискретного фильтра) по следующему правилу:

$$\frac{C_k}{p - p_k} \rightarrow \frac{T \cdot C_k}{1 - e^{p_k T} z^{-1}}. \quad (4.69)$$

Из (4.69) видно также, что устойчивый аналоговый фильтр-прототип преобразуется в устойчивый дискретный фильтр, так как если для всех полюсов аналогового фильтра $\text{Re } p_k < 0$, то порождаемые полюсы ПФ дискретного фильтра $z_k = e^{p_k T}$ таковы, что $|z_k| < 1$, $k = 1, \dots, N$. Поэтому дискретный фильтр также будет устойчивым (см. теорему 3.4).

Так как полиномы числителя и знаменателя ПФ аналогового фильтра (4.40) имеют вещественные коэффициенты, то ее полюсы либо вещественные, либо образуют комплексно-сопряженные пары $p_{k_1} = \bar{p}_{k_2}$. Такие пары дают в разложении ПФ (4.66) слагаемые, коэффициенты которых также комплексно сопряжены: $C_{k_1} = \bar{C}_{k_2}$. Соответствующие слагаемые в ПФ дискретного фильтра (4.68) следует объединить приведением к общему знаменателю, для того чтобы получить звено второго порядка с вещественными коэффициентами передаточной функции:

$$\frac{C}{1 - e^{sT} z^{-1}} + \frac{\bar{C}}{1 - e^{\bar{s}T} z^{-1}} = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}}, \quad (4.70)$$

где (убедитесь!): $b_0 = 2\text{Re } C$, $b_1 = -2\text{Re}(\bar{C} e^{sT})$, $a_1 = -2\text{Re}(e^{sT})$, $a_2 = |e^{sT}|^2$. В результате из разложения (4.68) получим представление ПФ в виде (3.11), готовом для реализации параллельной структуры фильтра (см. рис. 3.14) из звеньев первого и второго порядков.

Пример 4.14. Синтезировать цифровой фильтр нижних частот, имеющий граничную частоту полосы пропускания $\omega_p = 20$ рад/с, частоту подавления $\omega_s = 50$ рад/с. Максимальное допустимое отклонение

в полосе пропускания $\delta_p = 0,1$; в полосе подавления $\delta_s = 0,2$. Период дискретизации $T = 0,01$ с. Найти ПФ соответствующего цифрового фильтра.

◀ В качестве прототипа возьмем фильтр Чебышева I рода. Порядок фильтра находим из условия (4.62), где для ФНЧ по (4.45) следует положить $\Omega_s = \omega_s / \omega_p = 2,5$. Выбрав порядок фильтра

$$N \geq \text{arch} \sqrt{\left(\frac{1}{0,04} - 1\right) / \left(\frac{1}{0,81} - 1\right)} / \text{arch}(2,5) \approx 1,92$$

как минимально возможное целое число, т.е. $N = 2$, по формуле (4.63) находим полюсы нормированного НЧ-прототипа $s_{1,2} = -0,5688052 \pm 0,9074907i$, которые по (4.46) дают такие полюсы денормированного аналогового прототипа: $p_{1,2} = -11,37610 \pm 18,14981i$.

Передаточная функция фильтра Чебышева I рода не имеет нулей (кроме точки $s = \infty$), поэтому найденные полюсы определяют ЧХ аналогового прототипа с точностью до вещественной константы-множителя K . Для представления ПФ в форме (4.66) тогда несложно получить выражение:

$$G(p) = K \left(\frac{i}{p + 11,37610 + i \cdot 18,14981} + \frac{-i}{p + 11,37610 - i \cdot 18,14981} \right).$$

Отсюда на основании (4.69) и (4.70) находим ПФ дискретного фильтра:

$$H(z) = \frac{-2KT e^{-0,113761} \sin(0,1814981)z^{-1}}{1 - 2e^{-0,113761} \cos(0,1814981)z^{-1} + e^{-2 \times 0,113761} z^{-2}}.$$

Объединяя вещественные множители в числителе в новую нормировочную константу $A = -2K \times 0,01 \times e^{-0,113761} \sin(0,1814981)$ и подчитывая численные значения коэффициентов в знаменателе, получаем:

$$H(z) = \frac{A \cdot z^{-1}}{1 - 1,7556z^{-1} + 0,7965z^{-2}}. \quad (4.71)$$

Если учесть, что при синтезе аналогового прототипа использовалась чебышевская модель для квадрата АЧХ (4.60), которая в условиях нашего примера для постоянной составляющей дает коэффициент усиления

$$\left| H(e^{i\omega T}) \right|_{\omega=0} = |G(i\omega)|_{\omega=0} = 1 / \sqrt{1 + \varepsilon^2 T_2^2(0)} = 1 - \delta_p = 0,9,$$

то масштабный множитель в (4.71) следует принять равным $A \approx 0,0368$. Тогда в полосе пропускания синтезированного дискретного фильтра с ПФ (4.71) ЧХ будет в точности находиться в диапазоне $0,9 \leq |H(e^{j\omega T})| \leq 1$.

Отметим, что стоящий в числителе (4.71) множитель z^{-1} соответствует элементу задержки на входе (выходе) фильтра и не изменяет его АЧХ. Поэтому множитель z^{-1} в данном случае можно исключить и принять окончательно следующую ПФ для разработанного фильтра:

$$H(z) = \frac{0,0368}{1 - 1,7556z^{-1} + 0,7965z^{-2}}.$$

Изменяющий только ФЧХ множитель z^{-1} нельзя было бы удалить без искажения АЧХ, если бы в структуре разработанного фильтра присутствовали другие параллельные звенья, не имеющие в явном виде множитель z^{-1} в ПФ звена. ►

Упражнение. Используя какой-либо из математических программных пакетов, постройте графики АЧХ и ФЧХ фильтра, разработанного в примере 4.14. Убедитесь, что полученная АЧХ удовлетворяет спецификации, заданной в условиях примера. Можно ли считать, что ФЧХ в полосе пропускания близка к линейной?

Так как ЧХ $K(\omega) = H(e^{j\omega T})$ представляет собой спектр дискретного сигнала $h(n) = Tg(nT)$, то ЧХ дискретного фильтра представляет собой функцию периода $2\pi/T$, полученную из сдвинутых копий ЧХ аналогового прототипа $G(j\omega)$ (спектра сигнала $g(t)$)¹:

$$H(e^{j\omega T}) = \sum_{m=-\infty}^{\infty} G\left(j\left(\omega + \frac{2\pi m}{T}\right)\right). \quad (4.72)$$

Поэтому при проектировании цифровых фильтров методом инвариантности импульсной характеристики необходимо учитывать эффект наложения частот. Данный метод вызывает сложности и малопригоден для проектирования ВЧ и режекторных фильтров. В большей степени он подходит для проектирования низкочастотных, а также полосовых фильтров, если на частотах, близких

¹ Соответствующая связь спектров аналогового и дискретного сигналов в переменных линейной частоты $\nu = \frac{\omega}{2\pi}$ была установлена ранее в формуле (2.10).

к половине частоты дискретизации (и больших ее), аналоговый прототип имеет высокое подавление АЧХ. Очевидно, что повышение частоты дискретизации может быть одним из способов снижения негативного влияния эффекта наложения частот.

Упражнение. Передаточная функция аналогового фильтра-прототипа задана формулой:

$$G(p) = \frac{100}{p^2 + 10\sqrt{2}p + 100}.$$

Методом инвариантного преобразования импульсной характеристики найдите передаточную функцию цифрового фильтра с частотой дискретизации 200π рад/с. Устойчив ли полученный фильтр?

4.9. Синтез БИХ-фильтров методом билинейного Z-преобразования

Отображение комплексной p -области, на которой определяется ПФ аналогового фильтра, в комплексную z -область цифрового фильтра задается формулой преобразования переменных $z = \exp(pT)$, где T — период дискретизации аналоговой ИХ (см. рис. 4.12). Поэтому для обратного отображения можно предложить соотношение¹ $p = \frac{1}{T} \ln z$. Функцию логарифма представим сначала в виде разложения в ряд, которое находится в следующем примере.

Пример 4.15. Показать, что

$$\ln z = 2 \sum_{k=0}^{\infty} \frac{1}{2k+1} \left(\frac{z-1}{z+1} \right)^{2k+1}. \quad (4.73)$$

◀ Используя известное разложение в ряд Тейлора $\ln(1+u) = \sum_{n=1}^{\infty} (-1)^{n-1} u^n/n$ (абсолютно сходящегося при $|u| < 1$), несложно получить представление для функции

¹ Строго говоря, $p = \frac{1}{T} \ln z = \frac{1}{T} (\ln|z| + i \arg z)$ — это лишь одна из счетного множества точек $p_k = \frac{1}{T} \text{Ln } z = \frac{1}{T} (\ln|z| + i \arg z + i2\pi k)$, $k \in \mathbb{Z}$, удовлетворяющих соотношению $z = \exp(pT)$.

$$\ln \frac{1+u}{1-u} = \ln(1+u) - \ln(1-u) = 2 \sum_{k=0}^{\infty} \frac{u^{2k+1}}{2k+1}.$$

Отсюда, обозначив $z = \frac{1+u}{1-u}$ и выразив $u = \frac{z-1}{z+1}$, получаем ряд (4.73), который абсолютно сходится (проверьте по признаку Даламбера или Коши) в области $\operatorname{Re} z > 0$. ▶

Возьмем первый член ряда (4.73) и вместо преобразования комплексных переменных $p = \frac{1}{T} \ln z$ рассмотрим так называемое *билинейное преобразование*:

$$p = \frac{2}{T} \left(\frac{z-1}{z+1} \right) = \frac{2}{T} \left(\frac{1-z^{-1}}{1+z^{-1}} \right). \quad (4.74)$$

Обозначая $\gamma = 2/T$, обратное выражение находим как

$$z = \frac{\gamma + p}{\gamma - p}. \quad (4.75)$$

Отметим, что пара преобразований (4.74) и (4.75) является взаимно однозначным отображением.

В методе билинейного Z-преобразования ПФ дискретного фильтра $H(z)$ находится по ПФ аналогового фильтра-прототипа $G(p)$ с использованием подстановки (4.74) следующим образом:

$$H(z) = G \left(\gamma \frac{1-z^{-1}}{1+z^{-1}} \right), \quad (4.76)$$

где $\gamma = 2/T$.

Рассмотрим, как отображается p -плоскость в z -плоскость в результате билинейного преобразования (4.74) $p \rightarrow z$. Для комплексных чисел p -плоскости, $p = \sigma + i\omega$, имеем на основании (4.75):

$$|z| = \left| \frac{\gamma + \sigma + i\omega}{\gamma - \sigma - i\omega} \right| = \sqrt{\frac{(\gamma + \sigma)^2 + \omega^2}{(\gamma - \sigma)^2 + \omega^2}}. \quad (4.77)$$

Для аналоговой частотной оси $p = i\omega$ ($\sigma = 0$) все отображенные точки лежат на единичной окружности $|z| = 1$. Поэтому $p = i\omega \rightarrow z = e^{i\hat{\omega}}$, где нормированная частота дискретного фильтра

$$\hat{\omega} = \arg \left(\frac{\gamma + i\omega}{\gamma - i\omega} \right) = \arg \left(\frac{(\gamma + i\omega)^2}{\gamma^2 + \omega^2} \right) = 2 \arg \left(\frac{2}{T} + i\omega \right) = 2 \operatorname{arctg} \frac{\omega T}{2}. \quad (4.78)$$

Несложно видеть, что изменению аналоговой частоты в пределах $-\infty < \omega < +\infty$ соответствует изменение нормированной частоты

ты дискретного фильтра $-\pi < \hat{\omega} < \pi$. Для некоторых характерных аналоговых частот имеем:

$$\omega = 0 \rightarrow \hat{\omega} = 0, \quad \omega = \pm 2/T \rightarrow \hat{\omega} = \pm \pi/2, \quad \omega = \pm \infty \rightarrow \hat{\omega} = \pm \pi.$$

Для всех комплексных чисел p -плоскости, $p = \sigma + i\omega$, имеем $|z| \leq 1$ при $\sigma \leq 0$ и $|z| \geq 1$ при $\sigma \geq 0$. Это следует из того, что числитель под корнем выражения (4.77) меньше знаменателя при $\sigma < 0$ и больше знаменателя при $\sigma > 0$. Таким образом, левая полуплоскость p -плоскости отображается внутрь, а правая полуплоскость — вовне единичной окружности $|z| = 1$ на z -плоскости. Поэтому если устойчив аналоговый фильтр с ПФ $G(p)$ (все ее полюсы на p -плоскости лежат в левой полуплоскости), то будет устойчив и цифровой фильтр, все полюсы ПФ (4.76) которого будут находиться на z -плоскости в области $|z| < 1$.

При проектировании фильтров методом билинейного преобразования следует учитывать его нелинейный, хотя и взаимно однозначный, характер преобразования частот аналогового прототипа и цифрового фильтра, который определяется соотношением (4.78) и обратным к нему $\omega = \frac{2}{T} \operatorname{tg} \frac{\hat{\omega}}{2}$. На рис. 4.20 отражено, как АЧХ $|G(i\omega)|$ аналогового фильтра преобразуется в ЧХ $|H(e^{i\hat{\omega}})|$ дискретного нормированного фильтра.

Поскольку наложения частот при билинейном преобразовании нет, то частотная характеристика дискретного фильтра $H(e^{i\hat{\omega}})$ сохраняет те же величины и характер пульсаций в заданных полосах частот, что и ЧХ аналогового прототипа $G(i\omega)$. Нужно лишь помнить, что границы полос частот необходимо пересчитать в соответствии с (4.78), см. также рис. 4.20. Можно сказать, что ЧХ $H(e^{i\hat{\omega}})$ представляет собой сжатую по оси частот в соответствии с (4.78) ЧХ $G(i\omega)$: $\omega \in (-\infty; +\infty) \rightarrow \hat{\omega} \in (-\pi; \pi)$. При этом на частотах, связанных соотношением (4.78), значения АЧХ фильтров совпадают: $|H(e^{i\hat{\omega}})| = |G(i\omega)|$.

Билинейное преобразование сохраняет количество нулей и полюсов передаточной функции аналогового прототипа. Нули и полюсы прототипа преобразуются в нули и полюсы дискретного фильтра по формуле (4.75).

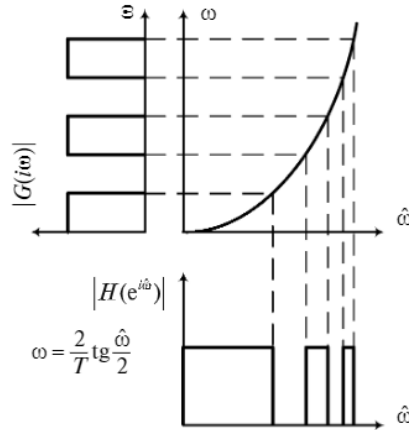


Рис. 4.20. Деформация оси частот при билинейном Z-преобразовании: $\hat{\omega} \in (-\pi; \pi) \leftrightarrow \omega \in (-\infty; \infty)$

Пример 4.16. Найти ПФ цифрового фильтра методом билинейного Z-преобразования, если ПФ аналогового фильтра-прототипа имеет вид:

$$G(p) = \frac{1}{p+a}.$$

◀ Передаточная функция аналогового прототипа имеет полюс $p_* = -a$ и ноль $p_o = \infty$. Запишем $H(z)$, произведя подстановку (4.76):

$$H(z) = \frac{1}{\frac{2}{T} \left(\frac{1-z^{-1}}{1+z^{-1}} \right) + a} = \frac{1+z^{-1}}{\left(\frac{2}{T} + a \right) - \left(\frac{2}{T} - a \right) z^{-1}}.$$

Видим, что $H(z)$ имеет полюс в точке $z_* = \frac{(2/T)+p}{(2/T)-p} \Big|_{p=-a} = \frac{(2/T)-a}{(2/T)+a}$,

а также ноль в точке $z_o = \lim_{p \rightarrow \infty} \left(\frac{2}{T} + p \right) / \left(\frac{2}{T} - p \right) = -1$, см. (4.75). ▶

Деформация оси частот делает затруднительным использование метода билинейного преобразования для синтеза фильтров с произвольными АЧХ. Однако данный метод хорошо подходит для проектирования частотно-избирательных фильтров всех типов (НЧ, ВЧ, полосовых, режекторных). При этом следует иметь в виду, что ни импульсная характеристика, ни фазочастотная характеристика

аналогового прототипа не сохраняются. Поэтому для проектирования фильтров с близкой к линейной ФЧХ метод билинейного преобразования также не подходит.

Рассмотрим пример проектирования БИХ-фильтра с использованием метода билинейного преобразования по шагам.

Пример 4.17. Синтезировать дискретный БИХ-фильтр верхних частот Чебышева II рода с помощью метода билинейного Z-преобразования. Требуемые спецификации нормированного дискретного фильтра:

$$\begin{cases} |H(e^{j\hat{\omega}})| \leq 0,178 & \text{при } \hat{\omega} \in [0; 0,2\pi] \\ 0,81 \leq |H(e^{j\hat{\omega}})| \leq 1 & \text{при } \hat{\omega} \in [0,3\pi; \pi] \end{cases}$$

Положить период дискретизации $T = 1$.

◀ 1. Учет деформации оси частот. Сначала преобразуем частоты цифрового фильтра в частоты аналогового фильтра-прототипа. Обращая (4.78), находим граничные частоты полос подавления и пропускания (для $T = 1$):

$$\omega_s = \frac{2}{T} \operatorname{tg} \frac{\hat{\omega}_s}{2} = 2 \operatorname{tg} \frac{0,2\pi}{2} \approx 0,649839; \quad \omega_p = \frac{2}{T} \operatorname{tg} \frac{\hat{\omega}_p}{2} = 2 \operatorname{tg} \frac{0,3\pi}{2} \approx 1,01905.$$

2. Определение спецификаций нормированного НЧ-прототипа.

Преобразуя аналоговый ВЧ-прототип к аналоговому нормированному НЧ-прототипу, по формуле (4.47) находим $\Omega_s = \omega_p / \omega_s \approx 1,56816$. Требования к нормированному НЧ-прототипу принимают следующий вид:

$$\begin{cases} 0,81 \leq |G(i\Omega)| \leq 1 & \text{при } \Omega \in [0; 1] \\ |G(i\Omega)| \leq 0,178 & \text{при } \Omega \in [1,56816; +\infty) \end{cases}$$

Параметры неравномерности АЧХ имеют здесь значения $\delta_p = 0,19$ и $\delta_s = 0,178$.

3. Синтез нормированного НЧ-прототипа. Порядок полинома для аппроксимации фильтра Чебышева II рода находим из (4.62):

$$N \geq \operatorname{arch} \left(\frac{\sqrt{\frac{1}{\delta_s^2} - 1}}{\sqrt{\frac{1}{(1 - \delta_p)^2} - 1}} \right) / \operatorname{arch} \Omega_s \approx \frac{\operatorname{arch}(7,636)}{\operatorname{arch}(1,568)} \approx 2,67.$$

Принимаем $N = 3$.

Как отмечалось в разделе 4.7, для фильтра Чебышева II рода полюсы ПФ s_{*k} определяются по полюсам передаточной функции

фильтра первого рода s_k (4.63) как величины $s_{*k} = \Omega_s / s_k$, $k = 1, 2, 3$. Нули s_{*k} ПФ фильтра Чебышева II рода определяются по нулям t_k многочлена Чебышева $T_3(t)$ как величины $s_{*k} = i\Omega_s / t_k$, $k = 1, 2, 3$.

Нули и полюсы определяют ПФ нормированного прототипа с точностью до некоторого вещественного множителя.

4. Пересчет нормированного НЧ-прототипа в денормированный аналоговый ВЧ-прототип. Переход от нормированного НЧ-прототипа к аналоговому ВЧ-прототипу осуществляется пересчетом нулей и полюсов по формуле (4.48): $p_k = \omega_p / s_k$. Получаем тогда для ПФ аналогового прототипа:

$$G(p) = K \prod_{k=1}^3 (p - p_{*k}) / \prod_{k=1}^3 (p - p_{*k}) = K \prod_{k=1}^3 \left(p - \frac{\omega_p}{i\Omega_s} t_k \right) / \prod_{k=1}^3 \left(p - \frac{\omega_p}{\Omega_s} s_k \right).$$

В этом выражении следует положить $K = 1$, так как АЧХ полученного аналогового ВЧ-прототипа $|G(i\omega)|$ при $\omega \rightarrow \infty$ должна принимать значение, равное АЧХ нормированного НЧ-прототипа (4.64) при $\Omega = 0$, т.е. $\lim_{\omega \rightarrow \infty} |G(i\omega)| = 1$ (см. также рис. 4.18б). Заметим, что $\omega_p / \Omega_s = \omega_s \approx 0,649839$. Поэтому ПФ аналогового ВЧ-прототипа принимает следующий вид:

$$G(p) = \prod_{k=1}^3 (p + i\omega_s t_k) / \prod_{k=1}^3 (p - \omega_s s_k).$$

Найдем корни многочлена Чебышева $T_3(t) = 4t^3 - 3t$: $t_1 = -\sqrt{3}/2$, $t_2 = 0$, $t_3 = \sqrt{3}/2$; определим полюсы ПФ фильтра Чебышева I рода из формул (4.63), (4.65): $s_1 = -0,4465308 + 1,1611069i$; $s_2 = -0,8930615$; $s_3 = \bar{s}_1 = -0,4465308 - 1,1611069i$. Тогда для ПФ аналогового ВЧ-прототипа получаем далее:

$$G(p) = \frac{p(p + i\omega_s \sqrt{3}/2)(p - i\omega_s \sqrt{3}/2)}{(p - \omega_s s_1)(p - \omega_s \bar{s}_1)(p - \omega_s s_2)} = \frac{p(p^2 + 3\omega_s^2/4)}{(p^2 - 2p\omega_s \operatorname{Re} s_1 + \omega_s^2 |s_1|^2)(p - \omega_s s_2)}.$$

Подставляя в это выражение найденные значения числовых параметров, имеем:

$$G(p) = \frac{p(p^2 + a)}{(p^2 + bp + c)(p + b)},$$

где $a = 0,316718$; $b = 0,580346$; $c = 0,653520$.

5. Билинейное преобразование. Учитывая, что $T = 1$, в результате подстановки (4.76) получаем для ПФ дискретного фильтра:

$$H(z) = G\left(2\frac{z-1}{z+1}\right) = \frac{2\frac{z-1}{z+1}\left(4\left(\frac{z-1}{z+1}\right)^2 + a\right)}{\left(4\left(\frac{z-1}{z+1}\right)^2 + 2b\frac{z-1}{z+1} + c\right)\left(2\frac{z-1}{z+1} + b\right)}$$

Умножив числитель и знаменатель на величину $(z + 1)^3$, после несложных преобразований имеем (убедитесь):

$$H(z) = \frac{8+2a}{(2+b)(4+2b+c)} \cdot \frac{\left(z^2 + \frac{2a-8}{a+4}z + 1\right)(z-1)}{\left(z^2 + \frac{2c-8}{4+2b+c}z + \frac{4-2b+c}{4+2b+c}\right)\left(z + \frac{b-2}{b+2}\right)}$$

Подставив численные значения коэффициентов a, b, c , окончательно получаем следующее выражение для ПФ дискретного фильтра:

$$H(z) = \frac{0,5755(z^2 - 1,7065z + 1)(z - 1)}{(z^2 - 1,1511z + 0,6007)(z - 0,5502)}$$

График АЧХ $|H(e^{i\hat{\omega}})|$ синтезированного фильтра приведен на рис. 4.21. Видно, что заданная спецификация АЧХ выполнена. ►

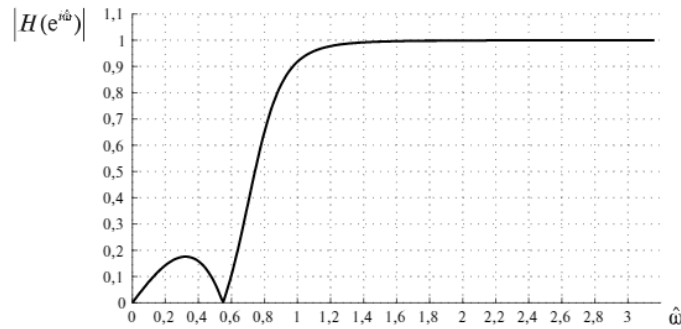


Рис. 4.21. АЧХ ФВЧ, синтезированного в примере 4.17, в основной полосе нормированных циклических частот $\hat{\omega} \in [0; \pi]$

Упражнение. Разработать цифровой фильтр с АЧХ, удовлетворяющей для нормированных циклических частот $\hat{\omega}$ требованиям:

$$\begin{cases} 0,7 \leq |H(e^{i\hat{\omega}})| \leq 1 & \text{при } \hat{\omega} \in [0; 0,2\pi] \\ |H(e^{i\hat{\omega}})| \leq 0,1 & \text{при } \hat{\omega} \in [0,3\pi; \pi] \end{cases}$$

В качестве аналогового фильтра-прототипа использовать фильтр Чебышева I рода, имеющего квадрат АЧХ вида (4.60). Используя какой-либо программный пакет, построить графики АЧХ и ФЧХ синтезированного фильтра, убедиться в выполнении заданной спецификации АЧХ.

Упражнение. ПФ аналогового фильтра-прототипа задана формулой:

$$G(p) = \frac{100}{p^2 + 10\sqrt{2}p + 100}.$$

Дискретный фильтр имеет частоту дискретизации 200π рад/с. Методом билинейного Z-преобразования найти ПФ цифрового фильтра. Используя какой-либо программный пакет, построить графики АЧХ и ФЧХ синтезированного фильтра.

Упражнение. С использованием метода билинейного преобразования по спецификации цифрового ФВЧ, имеющего частоту дискретизации $f_d = 20000$ Гц, была получена следующая спецификация аналогового ФВЧ-прототипа: $f_s = \frac{f_d}{\pi\sqrt{3}}$ (Гц), $f_p = \frac{f_d}{\pi}$ (Гц), $\delta_s = 0,001$ и $\delta_p = 0,01$. Покажите, что в исходной спецификации АЧХ цифрового ФВЧ граничные частоты полос подавления и пропускания соответствуют значениям $\hat{f}_s = f_d/6 \approx 3333$ Гц и $\hat{f}_p = f_d/4 = 5000$ Гц, и сформулируйте требования к нормированному аналоговому НЧ-прототипу.

Метод билинейного преобразования для расчета частотно-избирательных фильтров (НЧ, ВЧ, полосовой, режекторный) был обобщен Константиноидисом [19]. Если выделить пять шагов проектирования цифрового фильтра в соответствии с тем, как это было сделано при рассмотрении примера 4.17, то метод обобщенного билинейного преобразования фактически определяет способ замены шагов 1 и 2 одним шагом (со сходным, но все же отличающимся результатом). Шаги 4 и 5, которые задают обратное преобразование нормированного аналогового НЧ-прототипа в цифровой частотно-избирательный фильтр, в методе обобщенного билинейного преобразования также объединяются в один. Составляющие основу методики Константиноидиса формулы прямого и обратного отображения частот проектируемого цифрового полосового фильтра

в частоты нормированного аналогового НЧ-прототипа приведены в различных источниках [8, 9, 19, 43], где можно найти и описание самого метода обобщенного билинейного преобразования.

4.10. Выбор структуры для реализации фильтра

Получив представление для ПФ, обеспечивающее требуемую спецификацию фильтра, далее необходимо выбрать структурную схему для реализации фильтра (основные структуры фильтров рассмотрены в разделе 3.4). При использовании цифровой аппаратуры с большой разрядностью рассмотренными в разделе 3.11 эффектами квантования умножителей и переполнения сумматоров можно пренебречь, и тогда в качестве возможного варианта реализации фильтра может быть использована любая удобная структура.

Однако собственными шумами фильтра и опасностью переполнения сумматоров пренебречь можно далеко не всегда. Поэтому возникает необходимость анализа различных вариантов реализации фильтра с заданной ПФ с целью минимизации уровня собственных шумов системы и рисков переполнений.

Для КИХ-фильтров выбор структуры практически всегда предопределен трансверсальной схемой рис. 4.2 (ограничения аппаратуры могут повлечь необходимость реализации всей системы в виде последовательного соединения КИХ-звеньев меньшего порядка).

Реализация БИХ-фильтров оставляет разработчику достаточный простор для выбора структуры. В случае реализации фильтра на вычислительной платформе, имеющей высокую разрядность умножителей и сумматоров, выбор различных структур реализации оказывает малое влияние на итоговые характеристики фильтров. Однако из-за большей подверженности влиянию различных негативных эффектов от конечной разрядности представления данных в большинстве случаев использование прямых форм реализации БИХ-фильтров является нежелательным. Как отмечалось в разделе 4.1, основной формой реализации БИХ-фильтров является каскадная или параллельная структуры, которые состояются из звеньев второго или первого порядка (см. также раздел 3.4). При грамотном

подходе такие формы реализации дискретного фильтра позволяют обеспечить меньшую мощность (дисперсию) его собственных шумов по сравнению с прямыми формами, снижают риск возникновения переполнения в сумматорах и уменьшают искажения ЧХ от квантования коэффициентов фильтра. Рассмотрим специфику построения параллельных и каскадных структур БИХ-фильтров по заданной ПФ.

Для каскадной реализации передаточную функцию фильтра необходимо представить в виде

$$H(z) = \prod_{k=1}^K H_k(z), \quad (4.79)$$

после чего структура фильтра реализуется (см. раздел 3.4) как последовательное соединение звеньев второго порядка с передаточными функциями

$$H_k(z) = \frac{b_{k,0} + b_{k,1}z^{-1} + b_{k,2}z^{-2}}{1 + a_{k,1}z^{-1} + a_{k,2}z^{-2}}. \quad (4.80)$$

Фильтр второго порядка, имеющий ПФ вида (4.80), называют также *биквадратным звеном*. Каждое такое звено имеет в общем случае два нуля и два полюса передаточной функции $H_k(z)$.

По заданным нулям $\{z_{*,k}\}$ и полюсам $\{z_{*k}\}$ передаточной функции $H(z)$ реализовать соответствующий фильтр в виде каскадной структуры можно различными способами, которые определяются вариантами группировки нулей и полюсов, формирующих ПФ звеньев (4.80). Например, если ПФ фильтра содержит (всего лишь!) 4 нуля и 4 полюса, то для формирования первого звена фильтра существует 6 вариантов выбора пары нулей¹, каждому из которых соответствует такое же количество вариантов выбора полюсов, и всего имеем 36 вариантов построения фильтра из двух биквадратных звеньев. Чаще всего как нули, так и полюсы имеют комплексно-сопряженные пары, а такие пары всегда объединяются в одном звене, поэтому количество вариантов факторизации ПФ фильтра (4.79) становится меньшим. Тем не менее даже для пар связанных нулей и связанных полюсов количество вариантов группировки числителей и знаменателей при формировании ПФ биквадратных звеньев (4.80) во многих случаях оказывается столь велико, что не позво-

¹ Определяется количеством сочетаний из 4 по 2: $C_4^2 = \frac{4!}{2!2!} = 6$.

ляет провести анализ величины шумов системы для каждого варианта в отдельности и выбрать среди них оптимальный. Однако для каскадной реализации БИХ-фильтра при заданных наборах нулей $\{z_{\circ k}\}$ и полюсов $\{z_{*k}\}$ его ПФ необходимо знать ответы на следующие вопросы.

1. Как сгруппировать нули и полюсы фильтра для формирования биквадратных звеньев?
2. В каком порядке расположить полученные звенья?

Существуют некоторые простые эмпирические правила [29, 43], которые позволяют формализовать процесс построения каскадной структуры БИХ-фильтра по заданным наборам нулей и полюсов всей системы и дают ответы на поставленные выше вопросы. Конечно, речь идет только об устойчивых системах, все полюсы которых лежат на комплексной плоскости в области $|z| < 1$.

Правило 1. Группировать в звенья полюсы и нули, ближе всего расположенные друг к другу на комплексной z -плоскости.

Правило 2. Упорядочивать звенья в каскадной структуре в порядке удаления полюсов от нуля: первыми идут звенья, полюсы которых находятся ближе к точке $z = 0$.

Поясним, на чем основываются данные правила. Первое правило вызвано стремлением ограничения неравномерности коэффициента усиления каждого звена, который для гармонического колебания частоты ω определяется величиной $|H_k(e^{j\omega})|$. Звено с ПФ $H_k(z)$ (4.80) с высоким коэффициентом усиления пропорционально увеличивает и шумы квантования предшествующих звеньев, а также повышает опасность переполнения сумматоров. Если же в k -м биквадратном звене нули и полюсы имеют близкие значения, т. е. $z_{\circ k_1} \approx z_{*k_1}$ и $z_{\circ k_2} \approx z_{*k_2}$, то АЧХ k -го звена становится менее пульсирующей, без больших изменений коэффициента передачи звена C_k :

$$|H_k(e^{j\omega})| = C_k \left| \frac{(e^{j\omega} - z_{\circ k_1}) \cdot (e^{j\omega} - z_{\circ k_2})}{(e^{j\omega} - z_{*k_1}) \cdot (e^{j\omega} - z_{*k_2})} \right| \approx C_k.$$

Второе правило имеет следующее обоснование: чем ближе полюс звена (находящийся внутри единичного круга!) к окружности

$|z|=1$, тем ближе такое звено к состоянию неустойчивости, больше неравномерность коэффициента усиления звена на разных частотах и выше коэффициент усиления звена на частоте $\omega_k = \arg(z_k)$, соответствующей полюсу z_k . Поскольку любые ошибки, проявляющиеся в звеньях каскада, усиливаются всеми последующими звеньями, то «менее стабильные» звенья следует располагать в конце каскада.

Пример группировки нулей и полюсов в биквадратные звенья для фильтра пятого порядка, имеющего по две пары комплексно-сопряженных нулей и полюсов, а также один вещественный полюс и один вещественный ноль ПФ, приведен на рис. 4.22. В соответствии с правилом 2 последовательность звеньев в каскадной структуре должна быть выбрана такой: $H_3(z), H_1(z), H_2(z)$. При этом звено, соответствующее ПФ $H_3(z)$, будет звеном первого порядка, которое имеет один вещественный ноль z_{o3} и один вещественный полюс z_{*3} , и в выражении (4.80) для $H_3(z)$ имеем: $a_{3,2} = b_{3,2} = 0$.

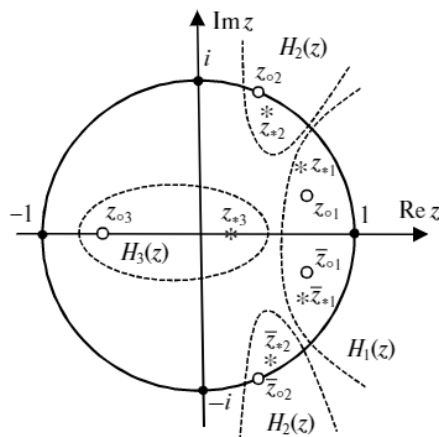


Рис. 4.22. Пример объединения нулей и полюсов для системы пятого порядка

Для параллельной реализации фильтра его ПФ необходимо представить в виде разложения (3.11). В параллельной структуре все звенья являются равноправными и вопрос о порядке следования звеньев не возникает. Однако параллельная реализация цифровых фильтров является менее популярной. Отчасти это связано с тем, что коэффициенты умножителей звеньев в каскадных

структурах (в отличие от параллельных структур) часто могут быть выбраны из набора $\{0, \pm 1, \pm 2\}$, что упрощает аппаратную реализацию фильтров [1].

Упражнение. Фильтр, имеющий ПФ

$$H(z) = \frac{z^{-2} - 1,5z^{-1} + 1,5}{(1 + 0,25z^{-2})(1 + 0,75z^{-1})}$$

представьте в виде параллельной и каскадной структур, определив порядок следования звеньев при каскадной реализации.

При проектировании фильтров часто возникает вопрос: какая структура, КИХ или БИХ, наилучшим образом подходит для заданного конкретного применения? Сравним основные характеристики нерекурсивных КИХ- и рекурсивных БИХ-фильтров, см. табл. 4.3.

Таблица 4.3. Некоторые сравнительные характеристики рекурсивных и нерекурсивных фильтров

| Характеристика | БИХ-фильтры | Нерекурсивные КИХ-фильтры |
|---|---------------------------------------|---------------------------|
| Количество умножителей и регистров памяти | Наименьшее | Наибольшее |
| Чувствительность к квантованию коэффициентов | Может оказаться критично высокой* | Очень низкая |
| Вероятность переполнений | Может оказаться критично высокой* | Очень низкая |
| Устойчивость | Должна обеспечиваться проектировщиком | Всегда гарантирована |
| Линейность ФЧХ | Невозможна | Может быть гарантирована |
| Возможность аппроксимации АЧХ аналоговых фильтров | Да | Нет |

* Снижается при использовании каскадных или параллельных реализаций (отказе от прямых форм структурных схем).

Выбор КИХ- или БИХ-структуры должен определяться наиболее важными для заданного применения характеристиками проектируемого фильтра. Кратко, главное отличие заключается в том, что БИХ-фильтры в целом обладают более высокой

производительностью и могут аппроксимировать заданные аналоговые фильтры, а КИХ-структуры обладают большими возможностями в аппроксимации произвольных (нестандартных) АЧХ и могут гарантировано обеспечить линейность ФЧХ.

Однако если при реализации заданной спецификации фильтра для КИХ-структуры получено, например, в два раза большее количество умножений на отсчет сигнала, чем для БИХ-структуры, то это еще не означает, что итоговое быстродействие КИХ-фильтра на имеющейся вычислительной платформе обязательно будет более низким. Так, многие программируемые процессоры цифровой обработки сигналов имеют архитектуру, оптимизированную для реализации КИХ-фильтров (например, за счет реализации параллельных вычислений) — это необходимо учитывать при оценке итогового быстродействия фильтра.

4.11. Адаптивная фильтрация. Фильтр Винера

Адаптивными системами ЦОС называются системы с самонастраивающимися характеристиками, способные приспособиваться к неизвестным заранее внешним условиям (случайные помехи, изменяющиеся параметры каналов связи и т. п.).

Основой наиболее распространенных адаптивных систем является трансверсальный фильтр с регулируемыми коэффициентами $\{h_n\}_{n=0}^{N-1}$, схема которого представлена на рис. 4.23. Такой фильтр называют также *адаптивным линейным сумматором*.

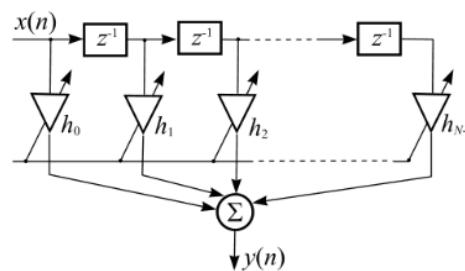


Рис. 4.23. Адаптивный КИХ-фильтр

Рассмотрим следующую задачу фильтрации. На вход фильтра поступает последовательность отсчетов $X(n)$, которая представляет собой полезный (рассматриваемый в общем случае как случайный) сигнал $Y(n)$, искаженный аддитивным шумом $V(n)$ с дисперсией σ^2 :

$$X(n) = Y(n) + V(n).$$

Требуется подобрать такие коэффициенты фильтра $\{h_n\}_{n=0}^{N-1}$, чтобы на его выходе получить оценку $\hat{Y}(n)$, максимально соответствующую в смысле некоторого выбранного критерия качества неискаженному сигналу $Y(n)$.

Для выходного сигнала фильтра имеем:

$$\hat{Y}(n) = \sum_{k=0}^{N-1} h_k X(n-k),$$

или в матричном виде:

$$\hat{Y}(n) = \mathbf{X}^T(n) \cdot \mathbf{H} = \mathbf{H}^T \mathbf{X}(n), \tag{4.81}$$

где $\mathbf{X}(n) = (X(n), X(n-1), \dots, X(n-N+1))^T$ — вектор-столбец входного сигнала, а $\mathbf{H} = (h_0, h_1, \dots, h_{N-1})^T$ — вектор-столбец коэффициентов фильтра. В качестве функции качества оценивания сигнала $Y(n)$ (целевой функции) $F(\mathbf{H}) = F(h_0, h_1, \dots, h_{N-1})$ чаще всего используют ошибку адаптации, равную математическому ожиданию квадрата ошибки оценивания $E(n) = Y(n) - \hat{Y}(n) = Y(n) - \mathbf{H}^T \mathbf{X}(n)$:

$$F(\mathbf{H}) = M(E^2(n)) = M\left\{ \left(Y(n) - \hat{Y}(n) \right)^2 \right\} = M\left\{ \left(Y(n) - \mathbf{H}^T \mathbf{X}(n) \right)^2 \right\}.$$

Таким образом, задача фильтрации сводится к следующей задаче безусловной оптимизации:

$$F(\mathbf{H}) = M\left\{ \left(Y(n) - \mathbf{H}^T \mathbf{X}(n) \right)^2 \right\} \rightarrow \min_{\mathbf{H} \in \mathbb{R}^N}. \tag{4.82}$$

Найдем $\text{grad } F(\mathbf{H})$. Так как $\mathbf{H}^T \mathbf{X}(n) = \sum_{j=0}^{N-1} h_j X(n-j)$ и $\partial(\mathbf{H}^T \mathbf{X}(n))/\partial h_j = X(n-j)$, то

$$\text{grad } F(\mathbf{H}) = \begin{pmatrix} \partial F / \partial h_0 \\ \partial F / \partial h_1 \\ \vdots \\ \partial F / \partial h_{N-1} \end{pmatrix} = -2M \left\{ \left(Y(n) - \mathbf{H}^T \mathbf{X}(n) \right) \begin{pmatrix} \partial(\mathbf{H}^T \mathbf{X}(n))/\partial h_0 \\ \partial(\mathbf{H}^T \mathbf{X}(n))/\partial h_1 \\ \vdots \\ \partial(\mathbf{H}^T \mathbf{X}(n))/\partial h_{N-1} \end{pmatrix} \right\} =$$

$$= -2\mathbf{M}\left\{\left(Y(n) - \mathbf{H}^T \mathbf{X}(n)\right) \begin{pmatrix} X(n) \\ X(n-1) \\ \vdots \\ X(n-N+1) \end{pmatrix}\right\} = -2\mathbf{M}\left\{\left(Y(n) - \mathbf{H}^T \mathbf{X}(n)\right) \mathbf{X}(n)\right\}.$$

Для оптимального в смысле (4.82) вектора коэффициентов $\mathbf{H} = \mathbf{H}_0$ должно выполняться необходимое условие минимума: $\text{grad } F(\mathbf{H}_0) = \mathbf{0}$, т.е.

$$\mathbf{M}\left\{\left(Y(n) - \mathbf{H}_0^T \mathbf{X}(n)\right) \mathbf{X}(n)\right\} = \mathbf{0}. \quad (4.83)$$

Далее мы увидим, что функция $F(\mathbf{H})$ в (4.82) является выпуклой, поэтому условие минимума (4.83) является необходимым и достаточным.

Перепишем соотношение (4.83) в терминах векторов-строк:

$$\mathbf{M}\left\{\underbrace{\left(Y(n) - \mathbf{H}_0^T \mathbf{X}(n)\right)}_{\text{скаляр}} \underbrace{\mathbf{X}^T(n)}_{\text{вектор-строка}}\right\} = \underbrace{\mathbf{0}^T}_{\text{вектор-строка}}.$$

Тогда, в силу линейности операции математического ожидания, имеем:

$$\mathbf{M}\left(Y(n) \mathbf{X}^T(n)\right) = \mathbf{H}_0^T \mathbf{M}\left(\mathbf{X}(n) \mathbf{X}^T(n)\right). \quad (4.84)$$

Вводя обозначения $\mathbf{R}(n) = \mathbf{M}\left(\mathbf{X}(n) \mathbf{X}^T(n)\right)$, $\mathbf{P}(n) = \mathbf{M}\left(Y(n) \mathbf{X}(n)\right)$, можно записать соотношение (4.84) как $\mathbf{P}^T(n) = \mathbf{H}_0^T \mathbf{R}(n)$, или, так как $\mathbf{R}(n) = \mathbf{R}^T(n)$:

$$\mathbf{P}(n) = \mathbf{R}(n) \mathbf{H}_0. \quad (4.85)$$

Равенство (4.85) называется матричным *уравнением Винера — Хопфа*; из него получаем оптимальное по критерию наименьших квадратов решение задачи (4.82) в виде следующего выражения для коэффициентов адаптивного КИХ-фильтра:

$$\mathbf{H}_0 = \mathbf{R}^{-1}(n) \mathbf{P}(n). \quad (4.86)$$

Отметим, что матрица $\mathbf{R}(n)$ размера $N \times N$ есть второй начальный момент случайного вектора входных отсчетов $\mathbf{X}(n)$. Если этот вектор центрирован, т.е. $\mathbf{m}_x(n) = \mathbf{M}(\mathbf{X}(n)) = \mathbf{0}$, то матрица $\mathbf{R}(n)$ совпадает с ковариационной матрицей этого вектора $\mathbf{K}(n) = \mathbf{M}\left[\left(\mathbf{X}(n) - \mathbf{m}_x(n)\right)\left(\mathbf{X}(n) - \mathbf{m}_x(n)\right)^T\right]$. Вектор-столбец $\mathbf{P}(n)$ размерности N при $\mathbf{M}(Y(n)) = 0$ и $\mathbf{m}_x(n) = \mathbf{0}$ описывает взаимные

ковариационные моменты между оцениваемым сигналом $Y(n)$ и отсчетами входной последовательности $X(n)$.

КИХ-фильтр, коэффициенты которого определяются соотношением (4.86), называется *фильтром Винера*. Оценка (4.86) является результатом блочной обработки, которая применяется для N -компонентного набора данных ограниченного размера — вектора $X(n)$.

Пересчет коэффициентов H_0 по (4.86) для каждого значения дискретного времени n обычно приводит к недопустимым объемам вычислений даже для невысоких порядков N фильтра Винера. Однако в случае нестационарного входного сигнала $X(n)$ матрицы $R(n)$ и $P(n)$ все же требуется периодически обновлять по выборочным оценкам сигнала для учета изменений его статистических свойств, а коэффициенты H_0 фильтра Винера должны соответствующим образом адаптироваться по (4.86). Для стационарных сигналов матрицы $R(n) = R$ и $P(n) = P$ полагают постоянными. Если априори они неизвестны, то их также находят в результате некоторого статистического анализа сигнала.

Пример 4.18. Пусть входное воздействие $X(n) = \sin(\pi n/4) + V(n)$ представляет собой детерминированный сигнал $y(n) = \sin(\pi n/4)$, искаженный аддитивным дискретным белым шумом $V(n)$ с нулевым математическим ожиданием и дисперсией σ^2 . Необходимо найти оптимальное по Винеру значение коэффициентов $H_0 = (h_0, h_1)^T$.

◀ Так как $M(V(n)) = 0$, то $\forall n$:

$$\begin{aligned} M(X^2(n)) &= M[\sin^2(\pi n/4) + 2V(n)\sin(\pi n/4) + V^2(n)] = \\ &= \sin^2 \frac{\pi n}{4} + 2M(V(n))\sin \frac{\pi n}{4} + M(V^2(n)) = \sin^2 \frac{\pi n}{4} + \sigma^2 = \frac{1 - \cos(\pi n/2)}{2} + \sigma^2. \end{aligned}$$

В силу некоррелированности белого шума имеем:

$$\begin{aligned} M(X(n)X(n-1)) &= \\ &= M\left(\sin \frac{\pi n}{4} \sin \frac{\pi(n-1)}{4} + V(n-1)\sin \frac{\pi n}{4} + V(n)\sin \frac{\pi(n-1)}{4} + V(n)V(n-1)\right) = \\ &= \sin \frac{\pi n}{4} \sin \frac{\pi(n-1)}{4} = \frac{1}{2} \left(\cos \frac{\pi}{4} - \cos \frac{\pi(n-0,5)}{2} \right) = \frac{\sqrt{2}}{4} - \frac{1}{2} \cos \frac{\pi(n-0,5)}{2}. \end{aligned}$$

Тогда для матрицы $R(n) = M \begin{pmatrix} X^2(n) & X(n)X(n-1) \\ X(n)X(n-1) & X^2(n-1) \end{pmatrix}$ получаем:

$$\mathbf{R}(n) = \begin{pmatrix} \frac{1 - \cos(\pi n/2)}{2} + \sigma^2 & \frac{1}{2\sqrt{2}} - \frac{1}{2} \cos \frac{\pi(n-0,5)}{2} \\ \frac{1}{2\sqrt{2}} - \frac{1}{2} \cos \frac{\pi(n-0,5)}{2} & \frac{1 - \cos \frac{\pi(n-1)}{2}}{2} + \sigma^2 \end{pmatrix}.$$

Приведенную выше матрицу $\mathbf{R}(n)$ оказалось возможно аналитически выразить, зная точный вид полезного сигнала $y(n) = \sin(\pi n/4)$. На практике точные значения сигнала $y(n)$ неизвестны (иначе не требовалось бы искать его оценку), и вместо матрицы $\mathbf{R}(n)$ для обработки стационарного сигнала можно использовать получаемую в результате статистического анализа сигнала $X(n)$ «усредненную» матрицу $\mathbf{R} = \frac{1}{M} \sum_{k=0}^{M-1} \mathbf{R}(n-k)$. Тогда при $M \rightarrow \infty$ периодические составляющие компонент $\mathbf{R}(n)$, которые имеют нулевое среднее, будут устранены, и в результате получим

$$\mathbf{R} = \begin{pmatrix} \frac{1}{2} + \sigma^2 & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2} + \sigma^2 \end{pmatrix},$$

откуда

$$\mathbf{R}^{-1} = \frac{1}{\sigma^4 + \sigma^2 + \frac{1}{8}} \begin{pmatrix} \frac{1}{2} + \sigma^2 & -\frac{1}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} & \frac{1}{2} + \sigma^2 \end{pmatrix}.$$

Аналогично,

$$\mathbf{P}(n) = \mathbf{M} \left[y(n) \begin{pmatrix} X(n) \\ X(n-1) \end{pmatrix} \right] = \mathbf{M} \left[y(n) \begin{pmatrix} y(n) + V(n) \\ y(n-1) + V(n-1) \end{pmatrix} \right] = \begin{pmatrix} y^2(n) \\ y(n)y(n-1) \end{pmatrix},$$

а усреднение по n дает матрицу-вектор $\mathbf{P} = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{n=1}^L \mathbf{P}(n) = \left(\frac{1}{2}; \frac{1}{2\sqrt{2}} \right)^T$ (убедитесь).

Для компонент вектора оптимальных коэффициентов $\mathbf{H}_0 = (h_0, h_1)^T$ получаем в результате выражение:

$$\mathbf{H}_0 = \begin{pmatrix} h_0 \\ h_1 \end{pmatrix} = \mathbf{R}^{-1} \mathbf{P} = \frac{1}{\sigma^4 + \sigma^2 + \frac{1}{8}} \begin{pmatrix} \frac{1}{2} \sigma^2 + \frac{1}{8} \\ \frac{\sigma^2}{2\sqrt{2}} \end{pmatrix}. \quad (4.87)$$

Несложно видеть, что при $\sigma^2 = 0$ (шум на входе отсутствует) фильтр не изменяет сигнал: ЧХ фильтра $K(\omega) = h_0 + h_1 e^{-i\omega} = 1$. ►

Итак, для реализации винеровской фильтрации необходимо знать матрицы $\mathbf{R}(n)$ и $\mathbf{P}(n)$ или иметь их оценки. Для матрицы $\mathbf{R}(n) = \mathbf{M}(\mathbf{X}(n)\mathbf{X}^T(n))$ эта оценка $\hat{\mathbf{R}}(n)$ может быть получена по выборочным значениям входного зашумленного сигнала $\mathbf{X}(n) = (x(n), x(n-1), \dots, x(n-N+1))^T$ как усреднение:

$$\hat{\mathbf{R}}(n) = \frac{1}{L} \sum_{k=n+1-L}^n \mathbf{X}(k)\mathbf{X}^T(k). \quad (4.88)$$

Если сигнал $X(n)$ является стационарным, то $\forall n \mathbf{R}(n) = \mathbf{R}$, и с увеличением числа L в (4.88) точность оценки $\mathbf{R}(n) \approx \hat{\mathbf{R}}(n)$ повышается, так как $\hat{\mathbf{R}}(n) \xrightarrow[L \rightarrow \infty]{\mathbb{P}} \mathbf{R}$ (в смысле сходимости по вероятности). Если же сигнал является нестационарным, то для сохранения способности адаптации оценки (4.88) к локально изменяющимся характеристикам сигнала число L должно быть ограничено некоторым значением, в том числе возможен вариант $L = 1$.

Для непосредственного нахождения сходной с (4.88) выборочной оценки $\hat{\mathbf{P}}(n)$ вектора $\mathbf{P}(n) = \mathbf{M}(Y(n)\mathbf{X}(n))$, помимо знания реализованных значений входного (искаженного шумом) сигнала $\mathbf{X}(n)$, необходима также информация о реализации полезного (неискаженного) сигнала $Y(n)$. Эта информация является недоступной, иначе задача поиска оценки $\hat{Y}(n)$ сигнала $Y(n)$ не возникла бы. Однако если некоррелированный с полезным сигналом $Y(n)$ дискретный шум $V(n)$ — белый, имеет нулевое математическое ожидание $m_v = 0$ и значение дисперсии $\mathbf{M}(V^2(n)) = \sigma^2$ известно, то получение выборочной оценки $\hat{\mathbf{P}}(n)$ для вектора $\mathbf{P}(n)$ возможно по доступным значениям зашумленного сигнала $X(n) = Y(n) + V(n)$. Действительно, обозначив первый столбец матрицы $\mathbf{R}(n) = \mathbf{M}(\mathbf{X}(n)\mathbf{X}^T(n))$ как $\mathbf{R}_1(n) = \mathbf{M}(X(n)\mathbf{X}(n))$, имеем:

$$\mathbf{R}_1(n) = \mathbf{M}[(Y(n) + V(n))\mathbf{X}(n)] = \mathbf{P}(n) + \mathbf{M}[V(n)(\mathbf{Y}(n) + \mathbf{V}(n))].$$

В силу некоррелированности полезного сигнала $Y(n)$ и шума $V(n) \forall k, m: \mathbf{M}(Y(k)V(m)) = \text{cov}(Y(k), V(m)) + m_y m_v = 0$ и $\mathbf{M}[V(n)\mathbf{Y}(n)] = \mathbf{0}$; кроме того, для белого шума $\mathbf{M}(V(n)\mathbf{V}(n)) = (\sigma^2, 0, \dots, 0)^T$. Поэтому

$$\mathbf{P}(n) = \mathbf{R}_1(n) - (\sigma^2, 0, \dots, 0)^T \quad (4.89)$$

и в качестве оценки для вектора $\mathbf{P}(n) = \mathbf{M}(Y(n)\mathbf{X}(n))$ можно использовать соотношение

$$\hat{\mathbf{P}}(n) = \hat{\mathbf{R}}_1(n) - (\sigma^2, 0, \dots, 0)^T, \quad (4.90)$$

где вектор $\hat{\mathbf{R}}_1(n)$ представляет собой первый столбец матрицы (4.88).

Заметим, что справедливость соотношения (4.89) была проиллюстрирована ранее в примере 4.18 и могла использоваться при его решении.

Найдем остаточную среднеквадратическую ошибку оценивания $F(\mathbf{H}_0) = \mathbf{M}[E_0^2(n)]$, соответствующую оптимальному вектору коэффициентов фильтра \mathbf{H}_0 . Учитывая (4.81) и (4.82), получаем:

$$\begin{aligned} \mathbf{M}[E_0^2(n)] &= \mathbf{M}[E_0(n)(Y(n) - \hat{Y}(n))] = \mathbf{M}[(Y(n) - \hat{Y}(n))Y(n) - E_0(n)\hat{Y}(n)] = \\ &= \mathbf{M}[Y^2(n)] - \mathbf{M}[Y(n)\mathbf{H}_0^T \mathbf{X}(n)] - \mathbf{M}[E_0(n)\mathbf{H}_0^T \mathbf{X}(n)] = \\ &= \mathbf{M}[Y^2(n)] - \mathbf{H}_0^T \mathbf{M}[Y(n)\mathbf{X}(n)] - \mathbf{H}_0^T \mathbf{M}[E_0(n)\mathbf{X}(n)]. \end{aligned}$$

Так как $\mathbf{M}[E_0(n)\mathbf{X}(n)] = \mathbf{0}$ (см. (4.83)), то с использованием принятого ранее обозначения $\mathbf{P}(n) = \mathbf{M}[Y(n)\mathbf{X}(n)]$ окончательно получаем:

$$\mathbf{M}[E_0^2(n)] = \mathbf{M}[Y^2(n)] - \mathbf{H}_0^T \mathbf{P}(n). \quad (4.91)$$

Подчеркнем, что величина (4.91) представляет собой характеристику ошибки прогноза (оценки) сигнала $Y(n)$ по выходному сигналу $\hat{Y}(n)$ винеровского фильтра, а не дисперсию выходного шума.

Пример 4.19. Для входного сигнала $X(n) = \sin(\pi n/4) + V(n)$ из примера 4.18 и полученного винеровского фильтра с ИХ (4.87) найти величину ошибки (4.91) и подавление шумового сигнала $V(n)$ при входной дисперсии шума $\sigma^2 = 1/2$.

◀ Вновь, как и в предыдущем примере 4.18, понимая математическое ожидание для детерминированного сигнала как усреднение по реализации, для полезного входного сигнала $y(n) = \sin(\pi n/4)$ получаем необходимое в (4.91) значение:

$$y^2 = \mathbf{M}[Y^2(n)] = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{n=1}^L y^2(n) = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{n=1}^L \frac{1 - \cos(\pi n/2)}{2} = \frac{1}{2}.$$

Видим, что на входе фильтра отношение сигнал/шум: $y^2/\sigma^2 = 1$, или, в логарифмическом масштабе, $10 \lg(y^2/\sigma^2) = 0$ (дБ), т. е. уровень шума равен уровню полезного сигнала.

Коэффициенты фильтра Винера при дисперсии входного шума $\sigma^2 = 1/2$ получаем по формуле (4.87): $\mathbf{H}_0 = \left(\frac{3}{7}; \frac{\sqrt{2}}{7} \right)$. Используя найденный в примере 4.18 вектор $\mathbf{P} = \left(\frac{1}{2}; \frac{1}{2\sqrt{2}} \right)^T$, для ошибки (4.91) имеем: $M[E_0^2(n)] = \frac{1}{2} - \left(\frac{3}{7}; \frac{\sqrt{2}}{7} \right) \begin{pmatrix} 1/2 \\ \sqrt{2}/4 \end{pmatrix} = \frac{3}{14}$.

Так как отсчеты входного шума $V(n)$ некоррелированы, то дисперсия выходного шума фильтра может быть найдена по формуле (3.22):

$$\sigma_y^2 = \sigma^2 (h_0^2 + h_1^2) = \frac{1}{2} \left(\frac{9}{49} + \frac{2}{49} \right) = \frac{11}{98}.$$

Мощность шума уменьшилась в $\sigma^2/\sigma_y^2 = 49/11 \approx 4,455$ раза, т. е. в логарифмических единицах коэффициент передачи шума составил величину

$$K_m = 10 \lg(\sigma_y^2/\sigma^2) \approx -6,49 \text{ дБ.}$$

Отметим, что полезный сигнал $y(n) = \sin(\pi n/4)$ также несколько ослабляется, однако меньше, чем шум. Уменьшение амплитуды отклика на гармонический сигнал частоты $\omega = \pi/4$ характеризует АЧХ фильтра $|H(e^{i\pi/4})|$ (см. раздел 3.6), для фильтра с ИХ $\mathbf{H}_0 = \left(\frac{3}{7}; \frac{\sqrt{2}}{7} \right)$ имеем:

$$|H(e^{i\pi/4})| = |h_0 + h_1 e^{-i\pi/4}| = \left| \frac{3}{7} + \frac{\sqrt{2}}{7} \left(\cos \frac{\pi}{4} - i \sin \frac{\pi}{4} \right) \right| = \frac{1}{7} |4 - i| = \sqrt{17}/7.$$

Квадрат амплитуды сигнала изменяется в $|H(e^{i\pi/4})|^2 = 17/49$ раза, усиление полезного сигнала является отрицательным и в логарифмических единицах дает

$$K_c = 20 \lg |H(e^{i\pi/4})| = 10 \lg |H(e^{i\pi/4})|^2 = 10 \lg(17/49) \approx -4,60 \text{ дБ.}$$

Как видим, даже простейший фильтр Винера второго порядка обеспечил в рассмотренном примере понижение уровня шума по отношению к сигналу. Если на входе фильтра уровень шума был равен уровню полезного сигнала, $y^2 = \sigma^2 = 1/2$, то на выходе отношение сигнал/шум повысилось на $K_c - K_m = 1,89$ дБ. ►

Упражнение. Входное воздействие $X(n) = y(n) + V(n)$ представляет собой детерминированный сигнал $y(n) = \sin(2\pi n/M)$, искаженный аддитивным дискретным белым шумом $V(n)$ с нулевым

математическим ожиданием и дисперсией σ^2 . Обобщая решение примера 4.18, покажите, что необходимые для определения винеровского фильтра второго порядка с ИХ $\mathbf{H}_0 = (h_0, h_1)^T$ матрицы $\mathbf{R}(n) = \mathbf{R}$ и $\mathbf{P}(n) = \mathbf{P}$ в (4.86) имеют вид:

$$\mathbf{R} = \frac{1}{2} \begin{pmatrix} 1 + 2\sigma^2 & \cos \frac{2\pi}{M} \\ \cos \frac{2\pi}{M} & 1 + 2\sigma^2 \end{pmatrix}, \quad \mathbf{P} = \frac{1}{2} \begin{pmatrix} 1 & \\ \cos \frac{2\pi}{M} & \end{pmatrix}.$$

Применение винеровских фильтров на практике в большинстве случаев ограничивается следующим фактором: вычисление оптимальных коэффициентов фильтра по (4.86) требует нахождения обратной матрицы $\mathbf{R}^{-1}(n)$. Это является непростой и вычислительно сложной задачей: используемые на практике адаптивные фильтры нередко содержат сотни коэффициентов \mathbf{H} . Данная проблема вызвала необходимость разработки альтернативных, вычислительно менее сложных алгоритмов адаптивной адаптации параметров фильтра — весовых коэффициентов \mathbf{H} . Такие алгоритмы основываются на теории методов оптимизации. Напомним некоторые необходимые термины и положения этой теории [24].

Функция вида

$$F(\mathbf{H}) = \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} \alpha_{k,j} h_k h_j + \sum_{j=0}^{N-1} b_j h_j + c$$

называется *квадратичной функцией* N переменных. Положив $a_{k,j} = \alpha_{k,j} + \alpha_{j,k}$, получим симметрическую матрицу $\mathbf{A} = (a_{k,j})_{k,j=0}^{N-1}$, с помощью которой $F(\mathbf{H})$ представляется в виде

$$F(\mathbf{H}) = \frac{1}{2} \mathbf{H}^T \mathbf{A} \mathbf{H} + \mathbf{H}^T \mathbf{V} + c, \quad (4.92)$$

где $\mathbf{V} = (b_0, \dots, b_{N-1})^T$ — вектор-столбец коэффициентов. Градиент квадратичной функции (4.92) равен

$$\text{grad } F(\mathbf{H}) = \mathbf{A} \mathbf{H} + \mathbf{V}, \quad (4.93)$$

а матрица вторых производных (гессиан) равна $F''(\mathbf{H}) = \mathbf{A}$. Если матрица Гессе квадратичной функции (4.92) положительно определена, то эта функция является *сильно выпуклой*. Это означает, что она достигает глобального минимума в единственной точке $\mathbf{H}^0 \in \mathbb{R}^N$, причем условие $\text{grad } F(\mathbf{H}_0) = \mathbf{A} \mathbf{H}_0 + \mathbf{V} = \mathbf{0}$ или

$$\mathbf{A}\mathbf{H}_0 = -\mathbf{B} \quad (4.94)$$

является необходимым и достаточным для того, чтобы \mathbf{H}_0 была точкой глобального минимума $F(\mathbf{H})$. Для проверки положительной определенности матрицы обычно используют критерий Сильвестра.

При решении задачи минимизации для дифференцируемой целевой функции $F(\mathbf{H})$ наиболее общим подходом является использование методов градиентного спуска, в которых строится итерационная последовательность приближений $\{\mathbf{H}^k\}$ к точке минимума \mathbf{H}_0 по общей формуле

$$\mathbf{H}^{k+1} = \mathbf{H}^k - \alpha_k \cdot \text{grad } F(\mathbf{H}^k), \quad (4.95)$$

$k = 0, 1, \dots$, где $\mathbf{H}^0 \in \mathbb{R}^N$ — начальное приближение, а величину k -го шага $\alpha_k > 0$ можно выбрать различными способами, которые и определяют конкретный метод из семейства градиентных.

Если $F(\mathbf{H})$ — квадратичная функция (4.92) с положительно определенной матрицей \mathbf{A} , а l и L — наименьшее и наибольшее собственные значения этой матрицы соответственно, то при постоянном шаге $\alpha_k = \alpha = \text{const} \in (0; 2/L)$ последовательность $\{\mathbf{H}^k\}$, найденная по формуле (4.95), сходится в евклидовой векторной норме к единственной точке минимума \mathbf{H}_0 этой функции со скоростью геометрической прогрессии: $\|\mathbf{H}^k - \mathbf{H}_0\|_E \leq q^k \|\mathbf{H}^0 - \mathbf{H}_0\|_E$, где $q = \max\{|1 - \alpha l|, |1 - \alpha L|\}$ [24].

В другом варианте метода градиентного спуска выбирают начальный шаг α и на каждой итерации (4.95) проверяют выполнение условия $F(\mathbf{H}^{k+1}) < F(\mathbf{H}^k)$. Если оно нарушается, то шаг уменьшают, например, полагая его равным $\alpha/2$, затем снова проверяют указанное условие и т.д. В любом варианте градиентного спуска итерации (4.95) выполняются до тех пор, пока не выполнится приближенное условие минимума $F(\mathbf{H})$, а именно: $\|\text{grad } F(\mathbf{H}^k)\| < \varepsilon$, где $\varepsilon > 0$ — заданное достаточно малое число.

Для квадратичной функции $F(\mathbf{H})$ с учетом равенства (4.93) формула метода градиентного спуска (4.95) примет вид

$$\mathbf{H}^{k+1} = \mathbf{H}^k - \alpha_k \cdot (\mathbf{A}\mathbf{H}^k + \mathbf{B}). \quad (4.96)$$

С использованием обозначений $\mathbf{R}(n) = \mathbf{M}[\mathbf{X}(n)\mathbf{X}^T(n)]$, $\mathbf{P}(n) = \mathbf{M}[Y(n)\mathbf{X}(n)]$, целевую функцию $F(\mathbf{H})$ из (4.82) можно представить в виде

$$F(\mathbf{H}) = \mathbf{H}^T \mathbf{R}(n) \mathbf{H} - 2\mathbf{H}^T \mathbf{P}(n) + M[Y^2(n)] \quad (4.97)$$

(убедитесь самостоятельно). Из сравнения (4.92) и (4.97) видно, что целевая функция $F(\mathbf{H})$ из (4.82) является квадратичной¹, и в данном случае

$$\mathbf{A} = 2\mathbf{R}(n), \quad \mathbf{B} = -2\mathbf{P}(n), \quad c = M[Y^2(n)]. \quad (4.98)$$

Упражнение. Убедитесь, что $F(\mathbf{H})$ (4.97) — выпуклая функция, поскольку матрица квадратичной формы $\mathbf{A} = 2\mathbf{R}(n)$ положительно полуопределена.

Указание: воспользуйтесь определением положительной полуопределённости, показав, что $\forall \mathbf{H}: \mathbf{H}^T \mathbf{A} \mathbf{H} \geq 0$.

Упражнение. Убедитесь, что с учетом введенных обозначений (4.98) уравнение Винера — Хопфа (4.85) эквивалентно уравнению (4.94).

Для случая целевой функции $F(\mathbf{H})$ из (4.82) формула (4.96) для итераций градиентного спуска с учетом (4.98) примет вид:

$$\mathbf{H}^{k+1} = \mathbf{H}^k - 2\alpha_k \cdot [\mathbf{R}(n)\mathbf{H}^k - \mathbf{P}(n)]. \quad (4.99)$$

Недостаток винеровской оценки (4.86) оптимальных коэффициентов адаптивного фильтра состоит в необходимости нахождения матрицы $\mathbf{R}(n)$ и вектора $\mathbf{P}(n)$, а также обращения матрицы $\mathbf{R}(n)$. При непосредственном применении метода градиентного спуска (4.99) также необходимо знать (или вычислять) матрицу $\mathbf{R}(n)$ и вектор $\mathbf{P}(n)$. Однако необходимая в (4.86) вычислительно наиболее сложная процедура обращения матрицы $\mathbf{R}^{-1}(n)$ в итерационном процессе (4.99) оказалась исключена. Использование итерационного поиска (4.99) позволяет повысить порядок фильтров N , для которых возможно (с заданной точностью) нахождение оптимальных коэффициентов \mathbf{H}_0 .

Для практической реализации адаптивной фильтрации к управляемому линейному сумматору, представленному на рис. 4.23, необходимо добавить блок, который реализует алгоритм адаптации коэффициентов фильтра. В зависимости от решаемых задач на практике могут использоваться разные способы включения

¹ Очевидно, что квадратичная функция (4.82) достигает глобального минимума в некоторой точке $\mathbf{H}^* \in \mathbb{R}^N$, поскольку ограничена снизу: $F(\mathbf{H}) \geq 0$.

адаптивного фильтра Винера. Часто структуру адаптивного цифрового фильтра-шумоподавителя можно представить в виде блок-схемы, приведенной на рис. 4.24. В этой схеме предполагается, что доступный алгоритму адаптации сигнал $S(n)$ коррелирован с неизвестным шумом $V(n)$ и потому может быть использован для нахождения его оценки $\hat{V}(n)$.

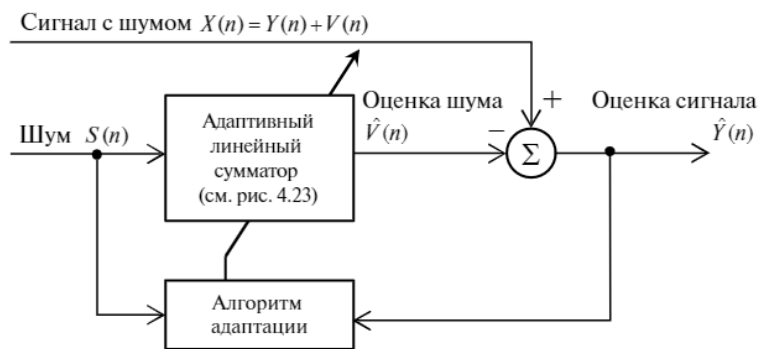


Рис. 4.24. Блок-схема адаптивного фильтра подавления шума

Упражнение. Покажите, что при использовании схемы рис. 4.24 для получения наилучшей в смысле $\min_{\mathbf{H} \in \mathbb{R}^N} M[(\hat{Y}(n) - Y(n))^2]$ оценки $\hat{Y}(n)$ полезного сигнала $Y(n)$ импульсная характеристика $\mathbf{H} = (h_0, h_1, \dots, h_{N-1})^T$ фильтра Винера находится по формуле (4.86), где нужно положить:

$$\mathbf{R}(n) = M[\mathbf{S}(n)\mathbf{S}^T(n)], \quad \mathbf{P}(n) = M[V(n)\mathbf{S}(n)].$$

Указание: используйте ту же схему рассуждений, что и ранее при выводе формулы (4.86), обозначив $\mathbf{S}(n) = (S(n), S(n-1), \dots, S(n-N+1))^T$.

Для реализации адаптивной фильтрации по схеме рис. 4.24 алгоритм адаптации должен формировать оценки $\hat{\mathbf{R}}(n)$ и $\hat{\mathbf{P}}(n)$ матриц $\mathbf{R}(n)$ и $\mathbf{P}(n)$. Оценка $\hat{\mathbf{R}}(n)$ формируется аналогично (4.88) по доступным значениям «опорного» шумового сигнала $S(n)$.

Для получения оценки $\hat{P}(n)$ алгоритм адаптации в схеме рис. 4.24 использует значения $\hat{Y}(n)$ и $S(n)$. Эта оценка основана на следующей связи:

$$\begin{aligned} M[\hat{Y}(n)S(n)] &= M[(Y(n) + V(n) - S(n)H^T)S(n)] = \\ &= M[Y(n)S(n)] + P(n) - M[S(n)H^T S(n)] = M[Y(n)S(n)] + P(n) - R(n)H, \end{aligned}$$

так как скаляр $H^T S(n) = S^T(n)H$. Для шумов $S(n)$ и $V(n)$, некоррелированных с полезным сигналом $Y(n)$, имеем $\forall k, j: \text{cov}(Y(k), S(j)) = 0$;

$$M[Y(n)S(n)] = \text{cov}(Y(n), S(n)) + M[Y(n)]M[S(n)] = M[Y(n)]M[S(n)].$$

Тогда

$$P(n) = M[\hat{Y}(n)S(n)] - M[Y(n)]M[S(n)] + R(n)H.$$

Полученное соотношение используется для нахождения необходимой алгоритму адаптации оценки вектора $P(n) = M[V(n)S(n)]$ по выборочным оценкам вектора $M[\hat{Y}(n)S(n)]$ и матрицы $R(n)$. Обычно $M[S(n)] = 0$.

Приведем пример практического использования адаптивного фильтра Винера, представленного на рис. 4.24, для подавления эха в устройствах, которые имеют акустическую обратную связь, см. рис. 4.25. Данной схеме соответствуют, например, фильтры эхоподавления в мобильных телефонах, системах селекторной связи и т. п.

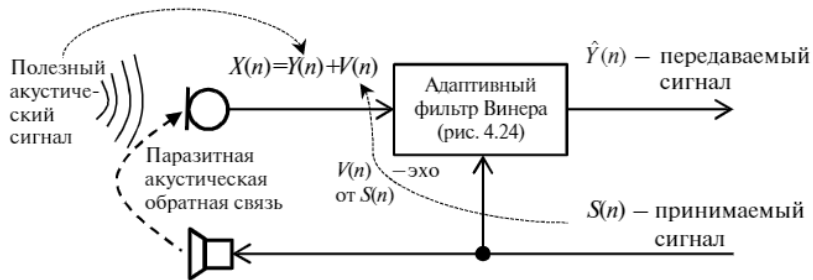


Рис. 4.25. Использование адаптивной винеровской фильтрации для подавления эха. Шумовой сигнал $V(n)$ есть результат прохождения принимаемого сигнала $S(n)$ по акустическому тракту динамик—микрофон

4.12. Построение фильтра Винера в частотной области

В данном разделе мы будем рассматривать модель искажения

$$X(n) = Y(n) + V(n), \tag{4.100}$$

где все дискретные сигналы являются стационарными случайными процессами, причем шумовая помеха $V(n)$ не коррелирована с полезным сигналом $Y(n)$ и $M[Y(n)] = M[V(n)] = 0$, т. е.

$$\text{cov}(Y(n), V(n)) = M[Y(n)V(n)] - \underbrace{M[Y(n)]}_0 \underbrace{M[V(n)]}_0 = 0. \tag{4.101}$$

Перед тем как перейти к формулировке требований к фильтру Винера в частотной области, рассмотрим следующий пример.

Пример 4.20. Покажем, что спектр

$$\tilde{F}_f(v) = \sum_{n=-\infty}^{\infty} f(n)e^{-2\pi inv}$$

любого дискретного сигнала $f(n)$ обладает следующим свойством:

$$\int_0^1 |\tilde{F}_f(v)|^2 dv = \sum_{n=-\infty}^{\infty} |f(n)|^2. \tag{4.102}$$

◀ Так как $\forall m \in \mathbb{Z}: \int_0^1 e^{2\pi imu} du = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$, то интегрированием по отрезку $v \in [0, 1]$ выражения

$$|\tilde{F}_f(v)|^2 = \tilde{F}_f(v) \overline{\tilde{F}_f(v)} = \sum_{n=-\infty}^{\infty} f(n)e^{-2\pi inv} \sum_{k=-\infty}^{\infty} \overline{f(k)}e^{2\pi ikv} = \sum_{n=-\infty}^{\infty} f(n) \sum_{k=-\infty}^{\infty} \overline{f(k)}e^{2\pi i(k-n)v}$$

получим соотношение (4.102). ▶

Выражение (4.102), по сути, представляет собой равенство Парсеваля, а величина $E = \int_0^1 |\tilde{F}_f(v)|^2 dv = \sum_{n=-\infty}^{\infty} |f(n)|^2$ имеет смысл энергии дискретного сигнала $f(n)$. Моделируя сигнал с вещественными отсчетами случайным процессом $X(n)$, его энергией называем тогда

$$E = \int_0^1 M(|\tilde{F}_X(v)|^2) dv = \sum_{n=-\infty}^{\infty} M(X^2(n)). \tag{4.103}$$

При условии $M[X(n)] = 0$ каждое слагаемое в сумме (4.103) для стационарного случайного процесса представляет собой его дисперсию $\sigma_X^2 = M(X^2(n))$. Тогда состоящий из одинаковых слагаемых ряд

в (4.103) расходитя, а энергия стационарного процесса обращается в бесконечность. В этом нет ничего удивительного, поскольку мы рассматриваем дискретный сигнал $X(n)$ неограниченной длительности, который имеет счетное количество отсчетов, $n \in \mathbb{Z}$.

Усечем сигнал $X(n)$ и возьмем только часть его отсчетов с номерами $n = 0, \pm 1, \dots, \pm L$, положив остальные равными нулю. Для такого фрагмента сигнала будем использовать обозначение

$$X_L(n) = \begin{cases} X(n), & -L \leq n \leq L, \\ 0, & |n| > L. \end{cases}$$

На основании соотношения (4.102) можем записать:

$$\int_0^1 |\tilde{F}_{X_L}(v)|^2 dv = \sum_{n=-L}^L X^2(n), \quad (4.104)$$

где спектр дискретного сигнала $X_L(n)$ является случайной функцией:

$$\tilde{F}_{X_L}(v) = \sum_{n=-L}^L X(n) e^{-2\pi i n v}. \quad (4.105)$$

Вместо энергии (4.103) для стационарного случайного процесса с нулевым математическим ожиданием $M[X(n)] = 0$ будем рассматривать его мощность P , т. е. энергию, приходящуюся в среднем на один отсчет. Используя представление (4.104), получаем:

$$P = \lim_{L \rightarrow \infty} \int_0^1 \frac{M(|\tilde{F}_{X_L}(v)|^2)}{2L+1} dv = \lim_{L \rightarrow \infty} \frac{\sum_{n=-L}^L M(X^2(n))}{2L+1} = M(X^2(n)) = \sigma_X^2. \quad (4.106)$$

Пример 4.21. Показать, что для спектра (4.105) выполняется соотношение

$$\tilde{s}_X(v) = \lim_{L \rightarrow \infty} \frac{1}{2L+1} M(|\tilde{F}_{X_L}(v)|^2), \quad (4.107)$$

где

$$\tilde{s}_X(v) = \sum_{k=-\infty}^{\infty} K_X(k) e^{-2\pi i k v} \quad (4.108)$$

— спектр мощности¹ дискретного сигнала $X(n)$, а

$$K_X(k) = \text{cov}(X(n), X(n-k)) = M[X(n)X(n-k)]$$

— его ковариационная функция (при условии, что $M[X(n)] = 0$).

¹ Напомним, что спектр мощности представляет собой преобразование Фурье ковариационной функции и рассматривался как для непрерывных (см. раздел 1.8), так и для дискретных сигналов (см. раздел 3.13, пример 3.36).

◀ Подставим выражение (4.105) в правую часть (4.107):

$$\begin{aligned} & \lim_{L \rightarrow \infty} \frac{1}{2L+1} \mathbb{M} \left(\underbrace{\sum_{n=-L}^L X(n) e^{-2\pi i n \nu}}_{\tilde{F}_{X_L}(\nu)} \underbrace{\sum_{k=-L}^L X(k) e^{2\pi i k \nu}}_{\tilde{F}_{X_L}(\nu)} \right) = \\ & = \lim_{L \rightarrow \infty} \frac{\sum_{n=-L}^L \sum_{k=-L}^L \mathbb{M}[X(n)X(k)] e^{2\pi i (k-n)\nu}}{2L+1} = \lim_{L \rightarrow \infty} \frac{\sum_{n=-L}^L \sum_{k=-L-n}^{L-n} \mathbb{M}[X(n)X(k+n)] e^{2\pi i \nu k}}{2L+1} = \\ & = \lim_{L \rightarrow \infty} \frac{\sum_{n=-L}^L \sum_{k=n-L}^{n+L} \mathbb{M}[X(n)X(n-k)] e^{-2\pi i \nu k}}{2L+1} = \lim_{L \rightarrow \infty} \frac{\sum_{n=-L}^L \sum_{k=n-L}^{n+L} K_X(k) e^{-2\pi i \nu k}}{2L+1} = \\ & = \sum_{k=-\infty}^{\infty} K_X(k) e^{-2\pi i \nu k} = \tilde{s}_X(\nu). \quad \blacktriangleright \end{aligned}$$

Установив справедливость выражения (4.107), мы показали, что дисперсия дискретного стационарного процесса $X(n)$ связана с его спектром мощности (4.108) следующим соотношением, вытекающим из (4.106) и (4.107):

$$\sigma_X^2 = \int_0^1 \tilde{s}_X(\nu) d\nu. \quad (4.109)$$

Спектр мощности дискретного сигнала $X(n)$, как несложно видеть из (4.107) и (4.108), является вещественной и неотрицательной функцией единичного периода: $\tilde{s}_X(\nu) = \tilde{s}_X(\nu+1) \geq 0$.

В этом разделе обозначение $w(n)$ будем использовать для ИХ фильтра Винера, имеющего ЧХ $W(\nu) = \sum_n w(n) e^{-2\pi i n \nu}$. Для сигнала (4.100) фильтр Винера должен обеспечивать формирование наилучшей оценки $\hat{Y}(n) = \sum_k X(n-k)h(k)$ в смысле минимума мощности сигнала-ошибки $E(n) = Y(n) - \hat{Y}(n)$ (см. также (4.82)):

$$\sigma_E^2 = \mathbb{M} \left\{ \left(Y(n) - \hat{Y}(n) \right)^2 \right\} \rightarrow \min. \quad (4.110)$$

Сформулируем требования к фильтру Винера в частотной области.

Для рассматриваемой модели искажения (4.100) обозначим:

$$\tilde{F}_{Y_L}(\nu) = \sum_{n=-\infty}^{\infty} Y_L(n) e^{-2\pi i n \nu} = \sum_{n=-L}^L Y(n) e^{-2\pi i n \nu}$$

– спектр усеченного полезного сигнала $Y_L(n)$, $Y_L(n) = 0$ при $|n| > L$,

$$\tilde{F}_{\hat{Y}_L}(\nu) = \sum_{n=-\infty}^{\infty} \hat{Y}_L(n) e^{-2\pi i n \nu}$$

– спектр выходного сигнала фильтра Винера (см. (4.100))

$$\hat{Y}_L(n) = \sum_{k=-L}^L X(k)w(n-k) = \sum_{k=-L}^L Y(k)w(n-k) + \sum_{k=-L}^L V(k)w(n-k), \quad (4.111)$$

отсчеты которого могут быть ненулевыми и для индексов n : $|n| > L$. Тогда на основании (4.106) для мощности сигнала-ошибки в (4.110) имеем:

$$\sigma_E^2 = \lim_{L \rightarrow \infty} \int_0^1 \frac{\mathbb{M}(|\tilde{F}_{E_L}(v)|^2)}{2L+1} dv = \lim_{L \rightarrow \infty} \int_0^1 \frac{\mathbb{M}(|\tilde{F}_{Y_L}(v) - \tilde{F}_{\hat{Y}_L}(v)|^2)}{2L+1} dv, \quad (4.112)$$

поскольку

$$\tilde{F}_{E_L}(v) = \sum_{n=-\infty}^{\infty} E_L(n)e^{-2\pi i n v} = \sum_{n=-\infty}^{\infty} (Y_L(n) - \hat{Y}_L(n))e^{-2\pi i n v} = \tilde{F}_{Y_L}(v) - \tilde{F}_{\hat{Y}_L}(v).$$

Пример 4.22. Показать, что для имеющих нулевое математическое ожидание взаимно некоррелированных сигналов $Y(n)$ и $V(n)$ спектр мощности $\tilde{s}_E(v)$ сигнала-ошибки $E(n) = Y(n) - \hat{Y}(n)$ можно представить в следующем виде:

$$\tilde{s}_E(v) = |W(v)|^2 (\tilde{s}_Y(v) + \tilde{s}_V(v)) + (1 - W(v) - \overline{W(v)}) \tilde{s}_Y(v), \quad (4.113)$$

где $\tilde{s}_Y(v)$ и $\tilde{s}_V(v)$ – спектры мощности полезного сигнала $Y(n)$ и шумовой помехи $V(n)$ соответственно.

◀ В частотной области уравнению (4.111) соответствует соотношение

$$\tilde{F}_{\hat{Y}_L}(v) = W(v) (\tilde{F}_{Y_L}(v) + \tilde{F}_{V_L}(v)),$$

где $\tilde{F}_{V_L}(v) = \sum_{n=-L}^L V(n)e^{-2\pi i n v}$ – спектр усеченного шумового сигнала (помехи) $V_L(n)$. Поэтому для спектра $\tilde{F}_{E_L}(v)$ ошибки $E_L(n) = Y_L(n) - \hat{Y}_L(n)$ имеем:

$$\begin{aligned} \tilde{F}_{E_L}(v) &= \tilde{F}_{Y_L}(v) - \tilde{F}_{\hat{Y}_L}(v) = \tilde{F}_{Y_L}(v) - W(v) (\tilde{F}_{Y_L}(v) + \tilde{F}_{V_L}(v)) = \\ &= (1 - W(v)) \tilde{F}_{Y_L}(v) - W(v) \tilde{F}_{V_L}(v). \end{aligned}$$

Тогда

$$\begin{aligned} \mathbb{M}(|\tilde{F}_{E_L}(v)|^2) &= \mathbb{M}(\tilde{F}_{E_L}(v) \overline{\tilde{F}_{E_L}(v)}) = \\ &= \mathbb{M} \left[\left((1 - W(v)) \tilde{F}_{Y_L}(v) - W(v) \tilde{F}_{V_L}(v) \right) \left((1 - \overline{W(v)}) \overline{\tilde{F}_{Y_L}(v)} - \overline{W(v)} \overline{\tilde{F}_{V_L}(v)} \right) \right] = \\ &= (1 - W(v)) (1 - \overline{W(v)}) \mathbb{M} \left[\tilde{F}_{Y_L}(v) \overline{\tilde{F}_{Y_L}(v)} \right] - \underbrace{W(v) (1 - \overline{W(v)}) \mathbb{M} \left[\tilde{F}_{V_L}(v) \overline{\tilde{F}_{Y_L}(v)} \right]}_0 - \end{aligned}$$

$$\begin{aligned}
 & - (1-W(v))\overline{W(v)}\underbrace{\mathbb{M}\left[\tilde{F}_{V_L}(v)\overline{\tilde{F}_{V_L}(v)}\right]}_0 + W(v)\overline{W(v)}\mathbb{M}\left[\tilde{F}_{V_L}(v)\overline{\tilde{F}_{V_L}(v)}\right] = \\
 & = (1-W(v))(1-\overline{W(v)})\mathbb{M}\left(|\tilde{F}_{V_L}(v)|^2\right) + |W(v)|^2\mathbb{M}\left(|\tilde{F}_{V_L}(v)|^2\right),
 \end{aligned}$$

поскольку в силу некоррелированности сигнала $Y(n)$ и шума $V(n)$

$$\begin{aligned}
 \mathbb{M}\left[\tilde{F}_{Y_L}(v)\overline{\tilde{F}_{V_L}(v)}\right] &= \overline{\mathbb{M}\left[\tilde{F}_{V_L}(v)\overline{\tilde{F}_{Y_L}(v)}\right]} = \mathbb{M}\left[\sum_{n=-L}^L Y(n)e^{-2\pi i n v} \sum_{k=-L}^L V(k)e^{2\pi i k v}\right] = \\
 &= \sum_{n=-L}^L \sum_{k=-L}^L \underbrace{\mathbb{M}[Y(n)V(k)]}_0 e^{2\pi i(k-n)v} = 0.
 \end{aligned}$$

Для спектра мощности $\tilde{s}_E(v)$ сигнала-ошибки $E(n) = Y(n) - \hat{Y}(n)$ получаем следующее выражение, см. (4.112):

$$\begin{aligned}
 \tilde{s}_E(v) &= \lim_{L \rightarrow \infty} \frac{1}{2L+1} \mathbb{M}\left(|\tilde{F}_{E_L}(v)|^2\right) = \\
 &= (1-W(v))(1-\overline{W(v)}) \lim_{L \rightarrow \infty} \frac{\mathbb{M}\left(|\tilde{F}_{V_L}(v)|^2\right)}{2L+1} + |W(v)|^2 \lim_{L \rightarrow \infty} \frac{\mathbb{M}\left(|\tilde{F}_{V_L}(v)|^2\right)}{2L+1} = \\
 &= (1-W(v))(1-\overline{W(v)})\tilde{s}_Y(v) + |W(v)|^2\tilde{s}_V(v).
 \end{aligned}$$

Раскрыв в последнем выражении скобки и перегруппировав слагаемые, получим соотношение (4.113). ►

Поскольку (см. 4.109) мощность сигнала-ошибки $\sigma_E^2 = \int_0^1 \tilde{s}_E(v)dv$, то условие оптимальной фильтрации (4.110) с учетом полученного выражения (4.113) определяет следующие требования к фильтру Винера в частотной области:

$$\sigma_E^2 = \int_0^1 |W(v)|^2 (\tilde{s}_Y(v) + \tilde{s}_V(v)) + (1-W(v)-\overline{W(v)})\tilde{s}_Y(v) dv \rightarrow \min_{W(v)}. \quad (4.114)$$

Необходимо найти такую ЧХ $W(v)$, которая обеспечивает минимальное значение σ_E^2 .

Пример 4.23. Покажем, что для имеющих нулевое математическое ожидание взаимно некоррелированных сигналов $Y(n)$ и $V(n)$ ЧХ оптимального фильтра Винера, обеспечивающего выполнение условия (4.114), определяется выражением

$$W(v) = \frac{\tilde{s}_Y(v)}{\tilde{s}_Y(v) + \tilde{s}_V(v)}, \quad (4.115)$$

где $\tilde{s}_Y(v)$ и $\tilde{s}_V(v)$ — спектры мощности полезного сигнала $Y(n)$ и шумовой помехи $V(n)$ соответственно.

◀ Очевидно, что для обеспечения минимума значения интеграла в выражении (4.114) нужно выбирать ЧХ $W(v)$ в каждой точке $v \in [0, 1]$ такой, чтобы подынтегральная функция принимала минимальное значение:

$$\tilde{s}_E(v) = |W(v)|^2 (\tilde{s}_Y(v) + \tilde{s}_V(v)) + (1 - W(v) - \overline{W(v)}) \tilde{s}_Y(v) \rightarrow \min_{W(v)}. \quad (4.116)$$

Обозначив $A = |W(v)|$ и $\varphi = \arg W(v)$, несложно найти минимум функции двух вещественных переменных, которая получается в результате преобразования соотношения (4.116):

$$g(A, \varphi) = A^2 (\tilde{s}_Y(v) + \tilde{s}_V(v)) + (1 - 2A \cos \varphi) \tilde{s}_Y(v) = \tilde{s}_E(v) \rightarrow \min_{A, \varphi}. \quad (4.117)$$

Из условий $\partial g / \partial A = 0$ и $\partial g / \partial \varphi = 0$ получаем:

$$A = \frac{\tilde{s}_Y(v) \cos \varphi}{\tilde{s}_Y(v) + \tilde{s}_V(v)}, \quad \sin \varphi = 0.$$

Значение $\varphi = \pi$ дало бы модуль $A = |W(v)| < 0$ (поскольку $\tilde{s}_Y(v) \geq 0$, $\tilde{s}_V(v) \geq 0$), поэтому полагаем $\varphi = 0$ и в результате получаем для ЧХ фильтра Винера выражение (4.115).

Покажите самостоятельно, что выбор значения $A = |W(v)| = W(v)$ по (4.115) и $\varphi = 0$ действительно определяет точку минимума функции (4.117). Для этого убедитесь в положительной определенности матрицы вторых производных функции $g(A, \varphi)$ при найденных значениях A и φ . ▶

Как и спектры мощности $\tilde{s}_Y(v)$ и $\tilde{s}_V(v)$, ЧХ оптимального фильтра Винера $W(v)$ (4.115) является вещественной, неотрицательной и четной функцией. Вещественный характер $W(v)$ означает симметричность ИХ: $w(n) = w(-n)$. Ее значения могут быть найдены по ЧХ $W(v)$ (как в примере 3.17):

$$w(\pm n) = \int_{-1/2}^{1/2} W(v) e^{\pm 2\pi i v n} dv = 2 \int_0^{1/2} W(v) \cos(2\pi v n) dv, \quad n = 0, 1, \dots \quad (4.118)$$

В общем случае формула (4.118) определяет физически нереализуемый БИХ-фильтр, поэтому его практическая реализация по ЧХ (4.115) или ИХ (4.118) осуществляется приближенно, например с использованием методов, которые были рассмотрены в разделах 4.3 и 4.4.

ГЛАВА 5

ОСНОВЫ ПРИКЛАДНОЙ ТЕОРИИ ИНФОРМАЦИИ

5.1. Мера количества информации для дискретного источника сообщений без памяти

Рассмотрим модель *дискретного источника сообщений* X , который изменяет свое состояние в заданные моменты времени, причем число

| | | | |
|-------------|-------|-----|-------|
| Состояние | x_1 | ... | x_N |
| Вероятность | p_1 | ... | p_N |

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_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 количество 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 году 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 переходе в очередное состояние создает источник сообщений без памяти, имеющий два равновероятных состояния.

Пример 5.1. Из колоды, содержащей 32 карты, 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») равновероятными 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 (5.1) не получила широкого распространения, поскольку применима лишь к весьма узкому классу источников сообщений, для которых состояния равновероятны.

Дискретный источник с неравновероятными состояниями

Рассмотрим следующие два источника X и Y , имеющие по два возможных состояния, т. е. алфа-

| | | | |
|-------|-------|-------|-------|
| x_1 | x_2 | y_1 | y_2 |
| 0,5 | 0,5 | 0,001 | 0,999 |

бит каждого из которых состоит из двух символов. 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 мы почти наверняка не ошибемся, сказав, что после очередного перехода 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 заданном N) количество информации принимает в случае равной вероятности состояний $p = 1/N$, т. е. $I_{\max} = \log N$. Если же какое-то состояние источник 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 Шенноном 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 | b | v | z |
| 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 $H(X)$ (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 виде математического ожидания $H(X) = \bar{I} = \sum_{k=1}^N p_k I_k$, 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 $\exists k: p_k = 1$.

◀ 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 $\exists k: p_k = 1$ (т. е. источник имеет единственное состояние). ▶

2°. Для независимых источников сообщений X и Y энтропия их ансамбля

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

◀ Пусть источник X имеет N состояний, а источник Y — M состояний. Тогда ансамбль источников (X, Y) описывается вероятностями совместной реализации состояний $p_{n,m} = P\{X = x_n, Y = y_m\}$, $n = 1, \dots, N$; $m = 1, \dots, M$. 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

$$\begin{aligned} H(X, Y) &= - \sum_{n=1}^N \sum_{m=1}^M p(x_n)p(y_m)(\log p(x_n) + \log p(y_m)) = \\ &= - \underbrace{\sum_{n=1}^N p(x_n) \log p(x_n)}_{H(X)} \underbrace{\sum_{m=1}^M p(y_m)}_1 + \underbrace{\sum_{n=1}^N p(x_n)}_1 \underbrace{\left(- \sum_{m=1}^M p(y_m) \log p(y_m) \right)}_{H(Y)} = \\ &= H(X) + H(Y). \end{aligned} \quad \blacktriangleright$$

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 $H(p_1, \dots, p_N) = - \sum_{k=1}^N p_k \log p_k$ (5.2) при условии $\sum_{k=1}^N p_k = 1$. Для решения задачи нахождения условного экстремума составим функцию Лагранжа:

$$L(p_1, \dots, p_N) = - \sum_{k=1}^N p_k \log p_k + \lambda \left(\sum_{k=1}^N p_k - 1 \right)$$

и найдем ее критические точки, приравняв нулю частные производные:

$$\partial L(p_1, \dots, p_N) / \partial p_j = - \log p_j - 1 + \lambda = 0, \quad j = 1, \dots, N,$$

и
$$\partial L(p_1, \dots, p_N) / \partial \lambda = \sum_{k=1}^N p_k - 1 = 0.$$

Отсюда получаем единственную критическую точку $p_1 = p_2 = \dots = p_N = 1/N$, при этом $\lambda = 1 - \log N$.

Достаточным условием безусловного максимума функции Лагранжа (соответствующего максимуму энтропии (5.2) при условии $\sum_{k=1}^N p_k = 1$) является отрицательный знак второго дифференциала в найденной критической точке, что соответствует отрицательной определенности матрицы вторых производных (матрицы Гессе):

$$\mathbf{H} = \left\{ h_{k,m} = \frac{\partial^2 L(p_1, \dots, p_N)}{\partial p_k \partial p_m} \Big|_{p_1 = \dots = p_N = 1/N} \right\}_{k,m=1}^N = \begin{pmatrix} -N & 0 & \dots & 0 \\ 0 & -N & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -N \end{pmatrix}.$$

В соответствии с критерием Сильвестра матрица размерности $N \times N$ отрицательно определена тогда и только тогда, когда угловые миноры $|\Delta_j|$ (определители подматриц, составленных из первых j строк и j столбцов исходной матрицы) являются знакопередающимися, а именно: $\text{sign}|\Delta_j| = (-1)^j, j = 1, \dots, N$. Видим, что данное условие для полученной матрицы \mathbf{H} выполняется, поскольку определитель диагональной матрицы равен произведению ее диагональных элементов. Таким образом, $H_{\max}(p_1, \dots, p_N) = H(1/N, \dots, 1/N) = \log N$. ►

- 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:

$$H(A) = H(B) = -0,1 \cdot \log_2 0,1 - 0,9 \cdot \log_2 0,9 \approx 0,47 \text{ (бита)}.$$

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, неэквивалентных понятий. В разделе 5.1 мы определили бит как *единицу измерения количества информации*. Другое значение термина «бит» — *двоичный символ*. Обычно, если нет указания на то, в каком смысле понимается термин «бит»,



имеется в виду второе значение. Двоичный символ несет количество информации, равное одному биту, тогда и только тогда, когда символ создан источником сообщений без памяти, 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: $N \leq 2^R$. То есть для представления алфавита $\{x_k\}_{k=1}^N$ исходного сообщения, порождаемого 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) мы также будем использовать далее двоичные единицы измерения (биты), выбрав основание логарифма равным 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 *кодовые слова* (конечные битовые последовательности) следующим образом: $a \leftrightarrow 0$, $b \leftrightarrow 10$, $v \leftrightarrow 110$, $z \leftrightarrow 111$. 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 бит однозначно декодируется, например: $\underbrace{11110100011}_{z \quad b \quad b \quad a \quad a \quad v \quad z} \dots$. Обозначая R_k — число битов в коде символа x_k , находим среднее количество битов, приходящееся на кодирование одного символа

$$R = \sum_{k=1}^4 p_k R_k = 0,5 \cdot 1 + 0,25 \cdot 2 + 0,125 \cdot 3 + 0,125 \cdot 3 = 1,75 \text{ (бита)}$$

— это 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

более 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 взаимнооднозначное соответствие граф, имеющий вид дерева. Например, префиксному коду из примера 5.3 соответствует граф, приведенный на рис. 5.1. Двоичные коды символов формируются последовательным приписыванием битов, соответствующих ребрам графа, которые нужно пройти от вершины дерева до окончательного 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 декодируемых кодов. Прежде чем доказывать соответствующую теорему, сформулируем ряд вспомогательных утверждений.

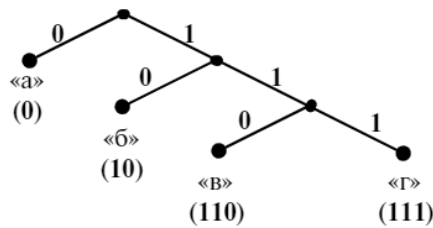


Рис. 5.1. Кодовое дерево, которое определяет приведенный в примере 5.3 префиксный код

Лемма 5.1. Если целые числа $I_j \geq 0$ ($j = 1, \dots, n$) таковы, что выполняется неравенство

$$\sum_{j=1}^n I_j 2^{-j} \leq 1, \quad (5.3)$$

то верно также и неравенство

$$\sum_{j=1}^{n-1} I'_j 2^{-j} \leq 1,$$



где для $j = 1, \dots, n-2$: $I'_j = I_j$,

$$а \quad I'_{n-1} = \begin{cases} I_{n-1} + \frac{I_n}{2}, & \text{если } I_n \text{ — четное} \\ I_{n-1} + \frac{I_n+1}{2}, & \text{если } I_n \text{ — нечетное} \end{cases}.$$

◀ 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 :

$$\underbrace{\sum_{j=1}^{n-1} I_j 2^{n-j}}_{\text{четное число}} + \underbrace{I_n}_{\text{нечетное число}} \leq \underbrace{2^n}_{\text{четное число}}.$$

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, содержащего N кодовых слов с битовыми 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 приписать бит кода (например, ноль для левого ребра и единицу для правого).

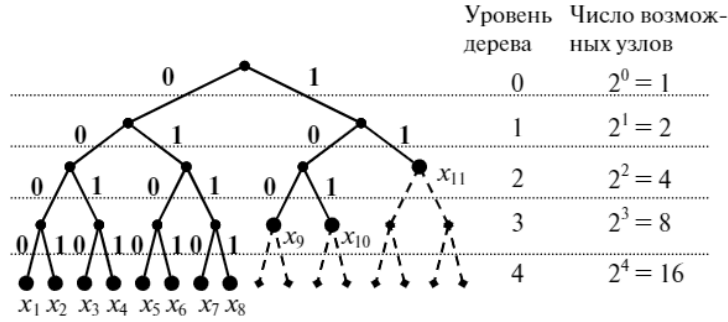


Рис. 5.2. К доказательству леммы 5.2: пример кодового графа для $n = 4$ и $N = 11$

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 (оконечные 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, допустить наличие узлов, имеющих одного потомка. Так, свойство декодируемости кода сохранится, если приведенный на рис. 5.2 граф преобразовать, например, к виду, изображенному на следующем 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).

Приведенные рассуждения можно провести аналогичным образом для любого числа n . Взяв произвольный префиксный код и построив порождающий его граф, мы сможем убедиться, что неравенство Крафта (5.4) выполняется.

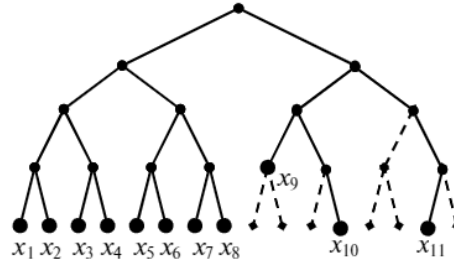


Рис. 5.3. Возможная модификация графа рис. 5.2. Свойство префиксности кода сохраняется

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.

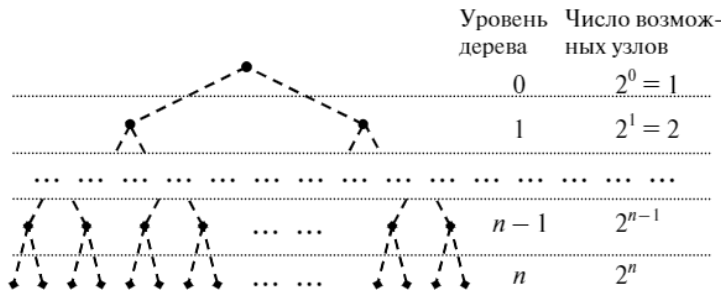


Рис. 5.4. Каркас (шаблон) графа для построения кодового дерева

Обозначим $I_j = |\{x_k \in X \mid R_k = j\}|$ — число элементов алфавита X , которым должны быть приписаны битовые коды длины $j = 1, \dots, n$. Тогда на кодовом дереве T , которое необходимо построить на j -м уровне, должно быть I_j узлов-листьев, соответствующих некоторым символам алфавита X , и неравенство Крафта (5.4) можно переписать в виде (5.3).

Из неравенства Крафта в форме (5.3) следует, что количество 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 $\{x_k \in X \mid R_k = n\}$ 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 $n-1$, см. 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.

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 количество формируемых узлов на каждом j -м слое не будет превышать максимально возможной величины 2^j . ▶

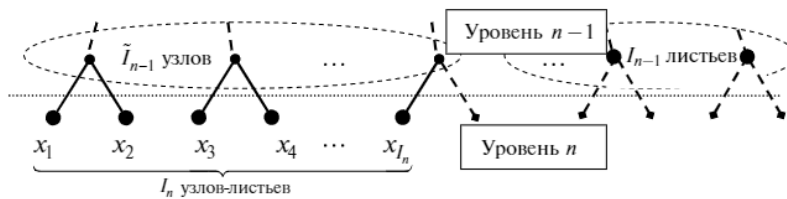


Рис. 5.5. Процедура, сводящая процесс построения n -уровневого кодового дерева к построению $(n-1)$ -уровневого дерева. Число I_n выбрано нечетным

Пример 5.4. Для дискретного источника сообщений без памяти, алфавит которого состоит из пяти символов, заданы длины

соответствующих кодовых слов (в битах): $R_5 = R_4 = 2, R_3 = R_2 = R_1 = 3$. Проверить возможность существования такого декодируемого кода и, если это возможно, то построить соответствующий префиксный код.

◀ Проверим выполнение неравенства Крафта:

$$\sum_{k=1}^5 2^{-R_k} = \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{7}{8} \leq 1.$$

Неравенство выполняется, декодируемый код существует. Следуя методике, описанной при доказательстве леммы 5.2, получаем граф (рис. 5.6), который определяет префиксный код.

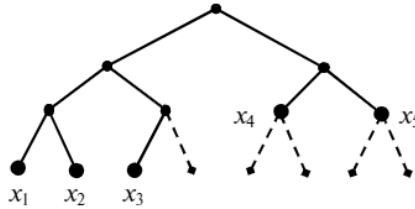


Рис. 5.6

Заметим, что символу x_3 можно поставить в соответствие код не из трех, а из двух битов, укоротив соответствующую ветвь дерева (см. рис. 5.7), тогда неравенство Крафта будет выполняться как равенство. ▶

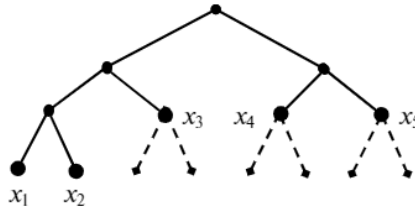


Рис. 5.7

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 $\{(x_k, R_k)\}_{k=1}^N$ справедливо 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). Лемма 5.3 определяет возможное правило выбора величин $\{R_k\}_{k=1}^N$ и верхнюю оценку для битовых затрат соответствующего кода.

Лемма 5.3. 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

$$\sum_{k=1}^N 2^{-R_k} \leq \sum_{k=1}^N p_k = 1.$$

Выбрав для каждого k наименьшее целое число R_k из условия $2^{-R_k} \leq p_k$, эквивалентного требованию $R_k \geq -\log_2 p_k$, мы сможем затем построить префиксный код по схеме, описанной при доказательстве леммы 5.2. Для так выбранных значений R_k имеем:

$$R_k - 1 < -\log_2 p_k, \quad k = 1, 2, \dots, N.$$

Умножив неравенства на p_k и просуммировав их 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 леммы. ►

Из теоремы 5.1 и леммы 5.3 следует, что, выбирая для заданного источника $X = \{(x_k, p_k)\}_{k=1}^N$ величины R_k как наименьшие целые числа, отвечающие 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

$$H(X) \leq R < H(X) + 1, \tag{5.5}$$

где $H(X)$ — двоичная энтропия (5.2) источника X .



Упражнение. Выбирая длины битовых кодов символов как наименьшие целые числа из условия $R_k \geq -\log_2 p_k$ и следуя методике, приведенной при доказательстве леммы 5.2 и проиллюстрированной в примере 5.4, постройте префиксные коды для источника сообщений Z , который имеет следующее распределение 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 |

Сформулируем следующее основополагающее утверждение теории информации, вытекающее из оценки (5.5).

Теорема 5.2. (Шеннон.) При бесконечной длительности сообщения 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 неравенством (5.5) можно закодировать со средними битовыми затратами R_M , приходящимися 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, называют *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 Шеннона — Фано и Хаффмана

Теорема Шеннона 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 V , имеющим два состояния (см. рис. 5.8). При этом, очевидно, информация, содержащаяся в битовой последовательности, в точности та же, которая была создана исходным источником X .

При таком рассмотрении процесса двоичного кодирования из теоремы Шеннона следует важное следствие: появление нулевых и единичных битов при эффективном двоичном кодировании



Рис. 5.8. Интерпретация битового кода как сообщения двоичного источника



является равновероятным, т. е. $p_0 = p_1 = 1/2$. Действительно, изменения состояний источника V кодируются с битовыми затратами $R = 1$ (бит). При этом количество информации $H(V) = -p_0 \log_2 p_0 - p_1 \log_2 p_1$, создаваемое источником, максимально и равно 1 биту, когда состояния источника V равновероятны. Если же $p_0 \neq p_1$, т. е. $H(V) < 1$ (бит), то в соответствии с теоремой Шеннона существует способ повышения эффективности кодирования, при котором среднее количество двоичных символов R , приходящееся на кодирование очередного состояния источника V , может быть уменьшено до величины $R < H(V) + \varepsilon < 1$ (бит), где число ε может быть выбрано сколь угодно малым.

Способы построения эффективных двоичных кодов основаны на необходимости получения равных вероятностей появления нулевого и единичного бита в двоичном представлении информации. Практически не отличающиеся друг от друга методы построения эффективных кодов первоначально были предложены независимо друг от друга Шенноном и Фано.

Метод построения эффективных кодов по Шеннону — Фано

1. Упорядочиваем таблицу символов (состояний источника) $\{x_1, \dots, x_N\}$ в порядке убывания 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 5.5. Используя методику Шеннона — Фано, построить эффективные коды для источника сообщений Z , который имеет следующее распределение 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 получаем $R = \sum_{k=1}^8 p_k R_k = 2,84$ бита, где R_k — 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 Шеннона — Фано обладает некоторой неоднозначностью построения кода: она не регламентирует, как следует производить разбиение блока на подблоки в том случае, когда возможны два варианта разбиения. 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 5.5 кодах все-таки осталась. Из теоремы Шеннона следует, что избыточность кодов, т.е. величину $\Delta = R - H$, можно сделать сколь угодно малой. Для этого символы сообщения нужно кодировать не по одному, а объединять их в блоки (по два, три и т.д.), следуя идее, приведенной в доказательстве теоремы 5.2.



Пример 5.6. Построить эффективные коды для источника сообщений Z , имеющего вероятности появления символов: $p(z_1) = 0,9$ и $p(z_2) = 0,1$.

◀ Энтропия источника $H(Z) = -0,1 \cdot \log_2 0,1 - 0,9 \cdot \log_2 0,9 \approx 0,47$ 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 100%-ю 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 Шеннона—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 $2R = 0,81 \cdot 1 + 0,09 \cdot 2 + 0,09 \cdot 3 + 0,01 \cdot 3 = 1,29$ 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%. Кодировав сообщение источника Z тройками символов, т. е. рассматривая при построении кодов Шеннона—Фано алфавит из восьми символов, можно дополнительно повысить эффективность кодирования. ▶

Упражнение. Используя новый алфавит, полученный для источника из примера 5.6 объединением символов в блоки по три, построить по методу Шеннона—Фано эффективные коды и найти их избыточность.

Следуя общей схеме объединения символов в пары, тройки и т. д., теоретически можно как угодно приблизить величину битовых 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 5.5, методика Шеннона — 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 Шеннона — Фано.

Метод построения эффективных кодов по Хаффману

1. Построение *дерева Хаффмана*.
 - 1.1. Упорядочиваем текущую таблицу символов в порядке убывания 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 5.7. Построить по методу Хаффмана эффективный двоичный код для источника, имеющего пять равновероятных состояний.

◀ Для определенности положим, что символами исходного алфавита являются буквы $\{A, B, C, D, E\}$. Процедуру построения дерева Хаффмана (шаг 1) тогда можно изобразить в виде схемы рис. 5.9. Построение соответствующего двоичного кода по дереву Хаффмана (выполнение шага 2) отражает рис. 5.10.

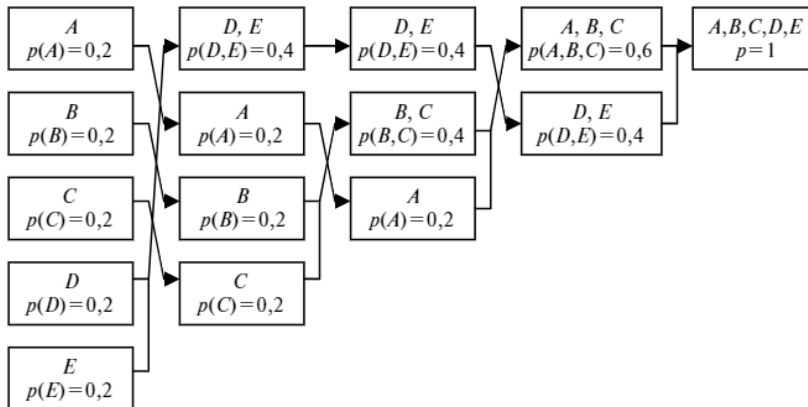


Рис. 5.9. Построение дерева Хаффмана для источника сообщений из примера 5.7

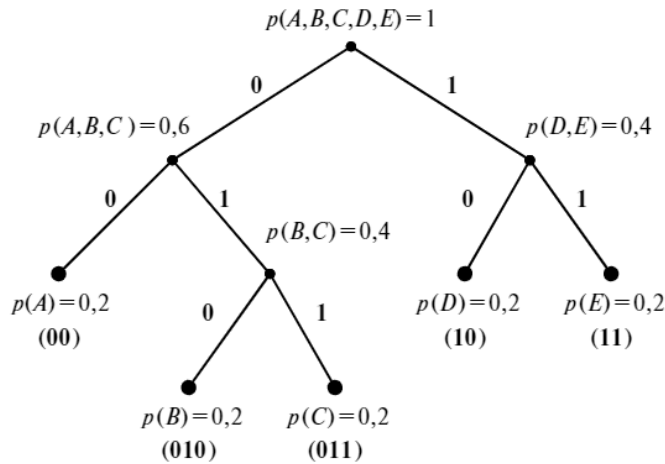


Рис. 5.10. Построение битового кода по дереву Хаффмана для источника сообщений из примера 5.7

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%). ►

Упражнение. Постройте самостоятельно коды Хаффмана для источника сообщений, рассмотренного в примере 5.5. Убедитесь в том, что битовые затраты составляют 2,80 бита на символ, т. е. соответствуют лучшему из рассмотренных в примере 5.5 вариантов построения кода Шеннона — Фано.

5.4. Кодирование длин серий

Если энтропия источника сообщений мала за счет наличия состояния, в котором источник находится с близкой к единице вероятностью, то вместо кодирования сообщений по Хаффману или Шеннону — Фано с объединением символов в блоки (см. пример 5.6) 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 количество символов, а кодируются кодами фиксированной битовой длины b .

Рассмотрим метод КДС на примере двоичного источника $Z = \{z_1, z_2\}$. Обозначим вероятности $P\{Z = z_1\} = p$, $P\{Z = z_2\} = 1 - p = q$ и положим для определенности, что $p > q$. Введем в рассмотрение следующую систему из $M = 2^b$ блоков, состоящих из разного числа символов сообщения.

| Блок символов | Кол-во символов в блоке | Вероятность появления блока | Двоичный код блока (b битов) |
|-------------------------------|-------------------------|-----------------------------|---------------------------------|
| $S_0 = z_2$ | 1 | $1 - p$ | 00...000 |
| $S_1 = z_1 z_2$ | 2 | $p(1 - p)$ | 00...001 |
| $S_2 = z_1 z_1 z_2$ | 3 | $p^2(1 - p)$ | 00...010 |
| ... | ... | ... | |
| $S_{M-2} = z_1 \dots z_1 z_2$ | $M - 1$ | $p^{M-2}(1 - p)$ | 11...110 |
| $S_{M-1} = z_1 \dots z_1 z_1$ | $M - 1$ | p^{M-1} | 11...111 |

Сообщение из последовательности символов, созданных источником Z , можно 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 бит) индексы соответствующих блоков. Фактически кодируется длина серии символов z_1 , разделяющих в сообщении появления 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. 5.11 приведены графики функций $R = (b, p)$ 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 рис. 5.11, параметр b следует выбирать в зависимости от значения вероятности p , которая определяет модель двоичного источника без памяти. При надлежащем выборе b по минимуму величины (5.6) избыточность кода, полученного по методу КДС, будет незначительной.

5.5. Арифметическое кодирование

Рассмотрим метод кодирования, который был предложен в 70-х годах прошлого века и представляет собой альтернативу методам Шеннона — 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 сначала на примере.

Пример 5.9. Рассмотрим источник, алфавит которого и соответствующие вероятности появления символов заданы следующей таблицей. Найдем число, соответствующее сообщению «ДАВАС».

| A | B | C | D |
|-----|-----|-----|-----|
| 0,5 | 0,2 | 0,2 | 0,1 |

◀ Сначала по таблице распределения вероятностей необходимо построить *интервалы вероятностей* $\Delta_k = [L_{x_k}; U_{x_k})$ 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 следующий текущий интервал вложен в предыдущий. Так, для нашего случая кодирование сообщения «ДАВАС» сводится к процедуре построения последовательности текущих интервалов, показанной на рис. 5.12. Получаем окончательный интервал $[0,9285; 0,9295)$.

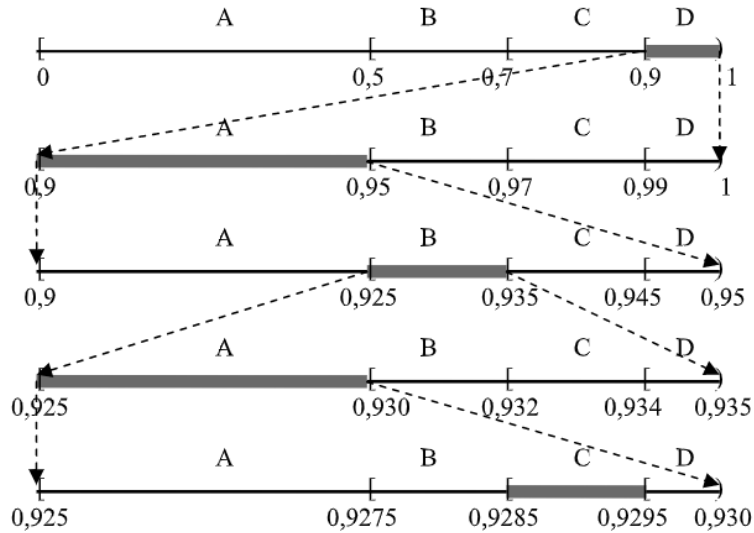


Рис. 5.12. Изменение текущего интервала при арифметическом кодировании последовательности символов из примера 5.9

Из него можно взять любое число, например $B = 0,929$ (как содержащее наименьшее число цифр), и восстановить по нему исходную последовательность из пяти символов «DABAC».

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 длинам интервалов 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: $L = 0$, $U = 1$.
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 границы текущего интервала: $L = 0$, $U = 1$.
Установить счетчик символов $j = 1$.
- Шаг 2.** Спроектировать текущий интервал $[L; U)$ на интервал $[0; 1)$, положив
- $$B = \frac{B - L}{U - L}.$$
- Шаг 3.** Найти интервал Δ_k , в который попадает число $B \in \Delta_k$.
Выдать символ $x(j) = x_k$. Положить

$$U = U_{x(j)}, \quad L = L_{x(j)}.$$
- Шаг 4.** Увеличить счетчик символов: $j = j + 1$.
Если $j \leq M$, то перейти на шаг 2, иначе — закончить работу.

Проектирование текущего интервала $[L; U)$ на интервал $[0; 1)$ на шаге 2 сводится к изменению значения кодового числа: $B \rightarrow B'$, см. рис. 5.13. Это выполняется для того, чтобы не пересчитывать границы всех интервалов символов перед последующим поиском интервала на шаге 3.

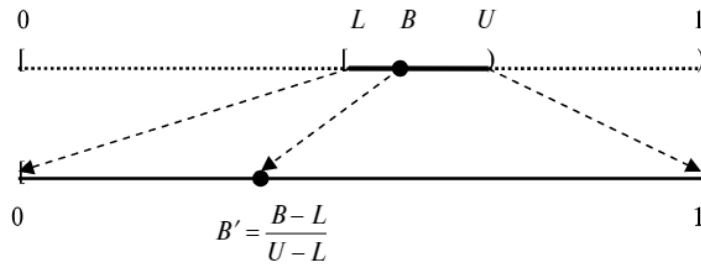


Рис. 5.13. Геометрическая интерпретация вычислений шага 2 алгоритма арифметического декодирования

Поскольку в компьютере используется двоичное представление чисел, выходным битовым кодом арифметического кодера является двоичная запись числа B . Поэтому на шаге 4 в процедуре кодирования следует выбирать число, которое можно представить наименьшим количеством двоичных цифр $\{0; 1\}$.

Пример 5.10. В условиях примера 5.9 определить кодовое число, содержащее наименьшее количество двоичных цифр.

◀ Границы текущего интервала, полученные после окончания рассмотренной в примере 5.9 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

$$t_m = [T \cdot 2^m] \bmod 2, \quad t_m \in \{0; 1\},$$

используя которую получим:

$$U = 0,9295_{(10)} = 0,111011011111\dots_{(2)},$$

$$L = 0,9285_{(10)} = 0,111011011011\dots_{(2)}.$$

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 совпавших 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:

$$\begin{aligned} U = 0, b_1 b_2 \dots b_K 1 u_{K+2} u_{K+3} \dots_{(2)} &= \sum_{k=1}^K b_k 2^{-k} + \frac{1}{2^{K+1}} + \sum_{k=K+2}^{\infty} u_k 2^{-k}, \\ L = 0, b_1 b_2 \dots b_K 0 l_{K+2} l_{K+3} \dots_{(2)} &= \sum_{k=1}^K b_k 2^{-k} + \sum_{k=K+2}^{\infty} l_k 2^{-k}. \end{aligned} \quad (5.7)$$

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: $\sum_{k=K+2}^{\infty} 2^{-k} = 1/2^{K+1}$.

Очевидно, что число совпавших 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)) равна вероятности реализации серии $\{x(1), \dots, x(M)\}$:

$$P(x(1), x(2), \dots, x(M)) = \prod_{m=1}^M P(x(m)) = U - L \leq \frac{1}{2^{K+1}} + \frac{1}{2^{K+1}} = \frac{1}{2^K},$$

откуда

$$K = K(x(1), \dots, x(M)) \leq -\log_2 P(x(1), \dots, x(M)).$$

Усредняя последнее неравенство по всем возможным реализациям серии $\{x(1), \dots, x(M)\}$, получаем для математического ожидания количества совпавших старших разрядов в (5.7) после кодирования M символов:

$$\begin{aligned} \bar{K} &= \sum_{k_1=1}^N \dots \sum_{k_M=1}^N P(x_{k_1}, \dots, x_{k_M}) K(x_{k_1}, \dots, x_{k_M}) \leq \\ &\leq -\sum_{k_1=1}^N \dots \sum_{k_M=1}^N P(x_{k_1}, \dots, x_{k_M}) \log_2 P(x_{k_1}, \dots, x_{k_M}) = M \cdot H(X), \end{aligned} \quad (5.8)$$

поскольку M символов, созданных источником без памяти, могут рассматриваться как символы, созданные M независимыми источ-

никами, 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, 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, который можно не передавать, а при декодировании всегда сначала добавлять единичный бит в конец кодовой последовательности. Тогда кодом сообщения будет 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): $R \leq H(X)$. Из теоремы 5.1 следует, 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 данное 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 сообщений, отличающихся количеством символов ($M = 1, 2, 3, \dots$), при этом короткие сообщения представляют собой начало более длинных.

Упражнение. Для источника, описанного в примере 5.9, определить кодовое число в двоичном виде, которое соответствует последовательности символов «ДАВАСА».

Так, для источника сообщений из примера 5.9 по последовательности бит «111011011», соответствующей кодовому числу $B = 0,111011011_{(2)} = 0,9287109375_{(10)}$ (см. также пример 5.10), можно восстановить сообщения: D, DA, DAB, DABA, DAVAC, DAVACA, ... Нельзя указать, какое именно сообщение было закодировано. Соответственно, описанный способ арифметического кодирования



является неоднозначным. Поэтому на практике необходимо дополнительно кодировать количество символов M в сообщении или дополнительно вводить в алфавит источника специальный символ признака конца сообщения. В результате битовые затраты на кодирование всего сообщения становятся большими, чем его энтропия, но все-таки близкими к ней.

Пример 5.11. Оценить битовые затраты на кодирование одного символа сообщения при использовании специального символа для обозначения признака конца сообщения.

◀ В алфавит источника сообщений $X = \{(x_k, p_k)\}_{k=1}^N$ введем дополнительный символ-признак конца сообщения x_{N+1} , приписав ему некоторую вероятность $\tilde{p}_{N+1} = \delta$. Остальные вероятности пересчитаем пропорционально их исходным значениям: $\tilde{p}_k = (1 - \delta)p_k$, $k = 1, \dots, N$.

При кодировании последовательности из M символов арифметический кодер будет использовать разбиение на интервалы вероятностей, где новые длины интервалов равны величинам $|\Delta_k| = \tilde{p}_k$, $k = 1, \dots, N+1$. После кодирования M символов $x(m) \in \{x_1, x_2, \dots, x_N\}$ ($m = 1, 2, \dots, M$) кодируется $(M+1)$ -й символ — признак конца сообщения x_{N+1} . Длина текущего интервала после обработки всех $M+1$ символов определяется пересчитанными вероятностями (см. (5.7)):

$$\prod_{m=1}^{M+1} \tilde{P}(x(m)) = U - L \leq \frac{1}{2^{K+1}} + \frac{1}{2^{K+1}} = \frac{1}{2^K},$$

где $\tilde{P}(x(m)) \in \{\tilde{p}_1, \dots, \tilde{p}_N\}$ для $m = 1, \dots, M$ и $\tilde{P}(x(M+1)) = \delta$. Отсюда число K — количество совпавших старших разрядов в двоичных представлениях (5.7), равное длине выходного битового кода:

$$K = K(x(1), \dots, x(M)) \leq -\log_2 \prod_{m=1}^{M+1} \tilde{P}(x(m)) = -\log_2 \delta - \sum_{m=1}^M \log_2 \tilde{P}(x(m)).$$

Усредняя последнее неравенство по всем возможным реализациям серии $\{x(1), \dots, x(M)\}$, получаем математическое ожидание длины выходного битового кода после кодирования M символов:

$$\begin{aligned} \bar{K} &= \sum_{k_1=1}^N \dots \sum_{k_M=1}^N P(x_{k_1}, \dots, x_{k_M}) K(x_{k_1}, \dots, x_{k_M}) \leq \\ &\leq -\log_2 \delta - \sum_{m=1}^M \sum_{k_m=1}^N \underbrace{P(x(m) = x_{k_m})}_{p_k} \log_2 \underbrace{\tilde{P}(x(m) = x_{k_m})}_{p_k(1-\delta)} = \\ &= -\log_2 \delta - M \sum_{k=1}^N p_k \log_2 (p_k(1-\delta)) = -M \sum_{k=1}^N p_k \log_2 p_k - \log_2 ((1-\delta)^M \delta) = \end{aligned}$$

$$= M \cdot H(X) - \log_2((1-\delta)^M \delta).$$

Тогда битовые затраты $R = \bar{K}/M$ на кодирование одного символа M -символьной последовательности с учетом дополнительных расходов на кодирование признака конца сообщения оцениваем как

$$R \leq H(X) - \frac{1}{M} \log_2((1-\delta)^M \delta).$$

Минимальное значение для величины $\tau(\delta) = -\frac{1}{M} \log_2((1-\delta)^M \delta)$ достигается при $\delta = \frac{1}{M+1}$ (убедитесь!) и равно

$$\tau_M = \frac{M+1}{M} \log_2(M+1) - \log_2 M,$$

поэтому с учетом утверждений теоремы 5.1 окончательная оценка для битовых затрат примет вид:

$$H(X) \leq R \leq H(X) + \tau_M. \quad (5.9)$$

Несложно убедиться, что $\lim_{M \rightarrow \infty} \tau_M = 0$, т.е. при увеличении длины M кодируемой последовательности битовые затраты R на кодирование одного символа приближаются к энтропии источника. ►

Упражнение. Найдите численные оценки битовых затрат по формуле (5.9) для случаев, когда длина кодируемой последовательности символов $M = 16, 128, 1024, 8096$.

Описанный в примере 5.11 способ арифметического кодирования и полученная верхняя оценка битовых затрат (5.9) носят теоретический характер и призваны продемонстрировать возможность построения арифметического кодера, эффективность кодирования которого асимптотически приближается к теоретическому пределу.

Приведенные выше рассуждения раскрывают общие идеи метода арифметического кодирования и предполагают использование гипотетического вычислительного устройства, которое оперирует бесконечной разрядностью двоичного представления чисел. Реализация арифметического кодирования на реальных вычислительных устройствах обладает спецификой, связанной с 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:

$$\begin{aligned} W_r &= U_r - L_r + 1, \\ U_r &= L_r + (W_r \times U_{x(j),r}) 2^{-r} - 1, \\ L_r &= L_r + (W_r \times L_{x(j),r}) 2^{-r}. \end{aligned}$$

Каждый раз после выполнения приведенных вычислений проверяются на совпадение старшие биты чисел U_r и L_r . Совпавшие 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 существует также опасность возникновения ситуации, когда верхняя и нижняя границы недопустимо сблизятся: $U_r = \underbrace{100\dots 00}_r$, $L_r = \underbrace{011\dots 11}_r$. В этом случае алгоритм заикнется, так как дальнейшее уменьшение текущего интервала (дальнейшее сближение границ) невозможно. Для избежания подобной ситуации существуют специальные меры технического характера, описание которых мы опустим.

В заключение отметим достоинства и недостатки метода арифметического кодирования в сравнении с рассмотренными ранее методами Шеннона—Фано, Хаффмана, кодирования длин серий. Важным преимуществом является то, что арифметическое кодирование не требует специальных ухищрений при кодировании источников с малой энтропией, при этом оно 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 Шеннона—Фано (см. пример 5.6). В сравнении с кодированием длин серий (когда предполагается наличие «доминирующего»

состояния, которое источник принимает с большой вероятностью), арифметическое кодирование представляет собой более универсальный метод. В итоге в большинстве случаев использование арифметического кодирования приводит к более эффективному сжатию данных по сравнению с другими рассмотренными нами методами статистического кодирования, но за это приходится расплачиваться существенно большей вычислительной сложностью реализации метода.

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; можно предположить, 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 i и j , т.е. пределы суммирования определяются способом нумерации и количеством возможных состояний источников сообщений). Справедливость неравенства (5.10) мы докажем несколько позже в теореме 5.3.

Ансамбль из 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 $\forall i, j$:

$$P(X = x_i, Y = y_j) = \underbrace{P(X = x_i)}_{p(x_i)} \underbrace{P(Y = y_j)}_{p(y_j)}.$$

В общем 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:

$$\begin{aligned} H(X, Y) &= -\sum_n \sum_m p(x_n, y_m) \log p(x_n, y_m) = \\ &= -\sum_n \sum_m p(x_n) p(y_m | x_n) \log(p(x_n) p(y_m | x_n)) = \\ &= \underbrace{-\sum_n p(x_n) \log p(x_n)}_{H(X)} \underbrace{\sum_m p(y_m | x_n)}_1 + \sum_n p(x_n) \underbrace{\left(-\sum_m p(y_m | x_n) \log p(y_m | x_n) \right)}_{H(Y|x_n)}. \end{aligned}$$

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 по отношению к состоянию 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 по отношению к источнику X :

$$H(Y | X) = \sum_n p(x_n) H(Y | x_n). \quad (5.12)$$

Аналогичное (5.12) выражение можно записать и для условной энтропии источника X по отношению к источнику Y .

С использованием понятия условной энтропии можно записать общее соотношение для энтропии ансамбля двух, вообще говоря, зависимых источников сообщений X и Y :

$$H(X, Y) = H(X) + H(Y | X) = H(Y) + H(X | Y). \quad (5.13)$$

Пример 5.12. Для ансамбля источников U и V найти по заданному закону совместного распределения вероятностей двоичную условную энтропию $H(V|U)$.

| | | | |
|------------------|-------|-------|-------|
| $U \backslash V$ | 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:

$$H(UV) = -\sum_i \sum_j p_{i,j} \log_2 p_{i,j} = -0,4 \cdot \log_2 0,4 - 0,1 \cdot \log_2 0,1 - 0,2 \cdot \log_2 0,2 - 0,3 \cdot \log_2 0,3 \approx 1,85 \text{ (бита)}.$$

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:

$$H(V) = -0,4 \cdot \log_2 0,4 - 0,3 \cdot \log_2 0,3 - 0,3 \cdot \log_2 0,3 \approx 1,57 \text{ (бита)}. \quad \blacktriangleright$$

Для обоснования неравенства (5.10) осталось доказать важное свойство условной энтропии (5.12), которое мы сформулируем в виде следующей теоремы.

Теорема 5.3. 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:

$$\begin{aligned} H(X|Y) - H(X) &= -\sum_m p(y_m) \sum_n p(x_n | y_m) \log p(x_n | y_m) + \\ &+ \sum_n p(x_n) \log p(x_n) \sum_m p(y_m | x_n) = \sum_n \sum_m p(x_n, y_m) \log \frac{p(x_n)}{p(x_n | y_m)} \leq \\ &\leq \sum_n \sum_m p(x_n, y_m) \left(\frac{p(x_n)}{p(x_n | y_m)} - 1 \right) = \underbrace{\sum_n p(x_n)}_1 \underbrace{\sum_m p(y_m)}_1 - \underbrace{\sum_n \sum_m p(x_n, y_m)}_1 = 0. \end{aligned}$$

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. Если между источниками существует статистическая зависимость, то знание о реализации состояния одного источника может только уменьшить неопределенность другого.

Пример 5.13. Для ансамбля источников X и Y найти по заданному закону совместного распределения вероятностей двоичные условную $H(X|Y)$ и безусловную $H(X)$ энтропии.

| $X \backslash Y$ | y_1 | y_2 | y_3 |
|------------------|-------|-------|-------|
| x_1 | 1/8 | 2/8 | 0 |
| x_2 | 1/8 | 0 | 2/8 |
| x_3 | 1/8 | 1/8 | 0 |

◀ Найдем частные условные энтропии:

$$H(X|y_1) = \log_2 3 \approx 1,585 \text{ (бита);}$$

$$H(X|y_2) = -\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3} \approx 0,9183 \text{ (бита);}$$

$$H(X|y_3) = 0 \text{ (битов).}$$

Отсюда двоичная условная энтропия (см. (5.12))

$$H(X|Y) = \sum_{n=1}^3 p(y_n) H(X|y_n) \approx \frac{3}{8} \cdot 1,585 + \frac{3}{8} \cdot 0,9183 \approx 0,9387 \text{ (бита).}$$

Безусловная двоичная энтропия источника

$$H(X) = -\frac{3}{8} \log_2 \frac{3}{8} - \frac{3}{8} \log_2 \frac{3}{8} - \frac{2}{8} \log_2 \frac{2}{8} \approx 1,561 \text{ (бита).}$$

Видим, что источники X и Y зависимы, поэтому $H(X|Y) < H(X)$. Обратите внимание на то, что при этом частная энтропия источника $H(X|y_1) > H(X)$, и здесь нет противоречия (объясните почему). ▶

Пример 5.14. 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 $p(y_k | x_m) = P\{Y = y_k | X = x_m\} = \begin{cases} 1, & y_k = \varphi(x_m) \\ 0, & y_k \neq \varphi(x_m) \end{cases}$, 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), $H(X, Y) = H(X)$. ▶

Упражнение. Покажите, что энтропию ансамбля из n дискретных источников сообщений (X_1, \dots, X_n) можно представить в следующем виде:

$H(X_1, X_2, \dots, X_n) = H(X_1) + H(X_2|X_1) + H(X_3|X_1X_2) + \dots + H(X_n|X_1 \dots X_{n-1})$,
 где (см. (5.12), (5.11))

$$H(X_j|X_1X_2 \dots X_{j-1}) = \sum_{k_1} \dots \sum_{k_{j-1}} p(x_{k_1}, \dots, x_{k_{j-1}}) H(X_j|x_{k_1}, \dots, x_{k_{j-1}}) \quad (5.14)$$

— условная энтропия источника X_j по отношению к ансамблю (X_1, \dots, X_{j-1}) , а

$$H(X_j|x_{k_1}, \dots, x_{k_{j-1}}) = - \sum_{k_j} p(x_{k_j}|x_{k_1}, \dots, x_{k_{j-1}}) \log p(x_{k_j}|x_{k_1}, \dots, x_{k_{j-1}}) \quad (5.15)$$

— частная условная энтропия источника X_j по отношению к реализации ансамбля $(X_1 = x_{k_1}, X_2 = x_{k_2}, \dots, X_{j-1} = x_{k_{j-1}})$. *Указание:* в первом равенстве соотношения (5.13) произведите подстановки: $X = (X_1, X_2, \dots, X_{n-1})$, $Y = X_n$, и примените далее полученную формулу рекуррентно.

Можно показать (доказательство аналогично приведенному ранее для теоремы 5.3), что для условной энтропии (5.14) также верна оценка

$$H(X_j|X_1X_2 \dots X_{j-1}) \leq H(X_j).$$

Отсюда

$$H(X_1, X_2, \dots, X_n) \leq \sum_{j=1}^n H(X_j)$$

(равенство выполняется только для ансамбля независимых источников).

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 очередным символом, создаваемым источником сообщения, и реализациями 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. То есть вероятности перехода задаются *условными* законами распределения. Обозначим m -е 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 , используя 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 длительности M представляет собой дискретный марковский процесс M -го порядка¹.

Для того чтобы полностью описать математическую модель источника с конечной памятью, необходимо задать закон *совместного* распределения вероятностей, т. е. $(M+1)$ -мерную таблицу

¹ С простой цепью Маркова с непрерывными состояниями мы уже сталкивались ранее в разделе 3.13.

$$p(x_{k_0}, x_{k_1}, \dots, x_{k_M}) = P(X(m) = x_{k_0}, X(m-1) = x_{k_1}, \dots, X(m-M) = x_{k_M}), \quad (5.17)$$

$$k_j = 1, \dots, N; \quad j = 0, \dots, M$$

(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

$$p(x_{k_1}, \dots, x_{k_M}) = P(X(m-1) = x_{k_1}, \dots, X(m-M) = x_{k_M}) = \sum_{k_0=1}^N p(x_{k_0}, x_{k_1}, \dots, x_{k_M}).$$

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 информации, связанное с очередным переходом источника в новое состояние и выдачей очередного символа сообщения, оценивается по условной энтропии (сравните с (5.14)):

$$H_{M+1}(X) = H(X(m) | X(m-1), X(m-2), \dots, X(m-M)) =$$

$$= \sum_{k_1=1}^N \dots \sum_{k_M=1}^N p(x_{k_1}, \dots, x_{k_M}) H(X | x_{k_1}, \dots, x_{k_M}). \quad (5.19)$$

Условная энтропия (5.19) характеризует неопределенность источника с памятью и носит также название «энтропия $(M+1)$ -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), т. к. для любой предыстории получаем в (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 энтропии первого порядка) $H_1(X) = H(X)$ из (5.2). Энтропией нулевого порядка назовем величину $H_0(X) = \log N$.

Для энтропии (5.19) верна теорема 5.2 (Шеннона), т. е. значение $H_{M+1}(X)$ (5.19) является точной оценкой минимально возможных битовых затрат, приходящихся в среднем на кодирование одного символа сообщения источника с памятью длительности M . Схема построения эффективных битовых кодов для источника с памятью основана на тех же общих идеях, которые изложены ранее в доказательствах леммы 5.2 и теорем 5.1 и 5.2.

Если алфавит источника состоит из N символов, то при кодировании очередного символа нужно 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 использования 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 учитывать 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 5.15. 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

| $X(m) \backslash X(m-1)$ | x_1 | x_2 | x_3 |
|--------------------------|-------|-------|-------|
| x_1 | 1/8 | 1/8 | 1/4 |
| x_2 | 1/8 | 1/8 | 0 |
| x_3 | 1/4 | 0 | 0 |

$$p_{m_0, m_1} = P(X(m) = x_{m_0}, X(m-1) = x_{m_1})$$

задан приведенной таблицей. Построить эффективный двоичный код, если предыдущее состояние источника:

- а) не учитывается, б) учитывается.

а) Если нет возможности учитывать предыдущее состояние источника, то эффективные коды строим по безусловному распределению вероятностей. Используя методику

| | | | |
|--------------|-------|-------|-------|
| X | x_1 | x_2 | x_3 |
| p | 1/2 | 1/4 | 1/4 |
| Двоичный код | 1 | 01 | 00 |

Шеннона — Фано или Хаффмана, получим среднюю длину битового кода $R = 1,5$ бита, что совпадает с двоичной энтропией первого порядка (5.2) источника X .

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 таблицей вероятностей. Частная условная энтропия (5.18) источника совпадает со средними битовыми затратами полученного кода:

| | | | |
|--------------|-------|-------|-------|
| $X(m)$ | x_1 | x_2 | x_3 |
| $p(x_k x_1)$ | 1/4 | 1/4 | 1/2 |
| двоичный код | 11 | 10 | 0 |

$$H(X|x_1) = R(X|x_1) = 1,5 \text{ (бита)}.$$

- $x(m-1) = x_2$. Частная условная энтропия (5.18) совпадает со средними битовыми затратами полученного кода и равна

| | | | |
|--------------|-------|-------|-------|
| $X(m)$ | x_1 | x_2 | x_3 |
| $p(x_k x_2)$ | 1/2 | 1/2 | 0 |
| Двоичный код | 1 | 0 | — |

$$H(X|x_2) = R(X|x_2) = 1 \text{ (бит)}.$$

- $x(m-1) = x_3$. В этом случае очередное состояние источника полностью определено, кодирование состояния $X(m) = x_1$ не требуется:

$$H(X|x_3) = R(X|x_3) = 0 \text{ (битов)}.$$

Декодер при появлении символа $x(m-1) = x_3$ будет сразу выдавать следующий символ $x(m) = x_1$.

В данном примере при использовании информации о предыдущем состоянии источника мы получили минимально возможную среднюю длину битового кода одного символа, равную энтропии второго порядка источника X :

$$H_2(X) = H(X(m)|X(m-1)) = \sum_{k=1}^3 p(x_k)H(X|x_k) =$$

$$\begin{aligned}
 &= R(X(m) | X(m-1)) = \sum_{k=1}^3 p(x_k) R(X | x_k) = \\
 &= 0,5 \times 1,5 + 0,25 \times 1 + 0,25 \times 0 = 1 \text{ (бит)}.
 \end{aligned}$$

Следующее упражнение также иллюстрирует общее утверждение о том, что $H_2(X) \leq H_1(X)$, а учет зависимостей между символами сообщения позволяет кодировать их более эффективно.

Упражнение. 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. Оценить минимальные

| $X(m) \backslash X(m-1)$ | -1 | 0 | 1 |
|--------------------------|------|------|------|
| -1 | 3/16 | 3/16 | 2/16 |
| 0 | 4/16 | 0 | 0 |
| 1 | 1/16 | 1/16 | 2/16 |

битовые затраты R , которые необходимы для эффективного кодирования символа сообщения, и построить эффективный двоичный код, если предыдущее состояние источника: а) не учитывается, б) учитывается.

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 $H_0(X) \geq H_1(X) \geq \dots \geq H_K(X) \geq \dots$, что вытекает из следующей теоремы.

Теорема 5.4. Для энтропии (5.19) любого стационарного источника сообщений X с алфавитом $\{x_1, \dots, x_N\} \forall K = 1, 2, \dots$ справедливо неравенство

$$H_K(X) \geq H_{K+1}(X).$$

◀ Схема доказательства та же, что и для теоремы 5.3. Имеем:

$$H_{K+1}(X) - H_K(X) =$$

$$\begin{aligned}
 &= - \sum_{k_1=1}^N \dots \sum_{k_K=1}^N p(x_{k_1}, \dots, x_{k_K}) \sum_{k_0=1}^N p(x_{k_0} | x_{k_1}, \dots, x_{k_K}) \log p(x_{k_0} | x_{k_1}, \dots, x_{k_K}) + \\
 &+ \sum_{k_1=1}^N \dots \sum_{k_{K-1}=1}^N p(x_{k_1}, \dots, x_{k_{K-1}}) \sum_{k_0=1}^N p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}}) \log p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}}) \times \\
 &\times \underbrace{\sum_{k_K=1}^N p(x_{k_K} | x_{k_0}, \dots, x_{k_{K-1}})}_1 = - \sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K}) \log p(x_{k_0} | x_{k_1}, \dots, x_{k_K}) + \\
 &+ \sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K}) \log p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}}) = \\
 &= \sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K}) \log \frac{p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}})}{p(x_{k_0} | x_{k_1}, \dots, x_{k_K})} \leq \\
 &\leq \sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K}) \left(\frac{p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}})}{p(x_{k_0} | x_{k_1}, \dots, x_{k_K})} - 1 \right) = \\
 &= \underbrace{\sum_{k_1=1}^N \dots \sum_{k_K=1}^N p(x_{k_1}, \dots, x_{k_K}) \sum_{k_0=1}^N p(x_{k_0} | x_{k_1}, \dots, x_{k_{K-1}})}_1 - \underbrace{\sum_{k_0=1}^N \dots \sum_{k_K=1}^N p(x_{k_0}, \dots, x_{k_K})}_1 = 0,
 \end{aligned}$$

откуда следует утверждение теоремы. ►

Следствие. Несложно показать, что $\forall K: H_K(X) \geq 0$ (см. доказательство свойства 1° энтропии (5.2)). 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 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 $K > M+1$, 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 $\{p(x_{k_0} | x_{k_1}, \dots, x_{k_M})\}_{k_0, \dots, k_M=1}^N$ (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 $x(1) = x_j$, кодеру необходимо 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 или 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 — длительность памяти источника, при этом порядок энтропии (5.19), который используется при описании источника, равен $M+1$. 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 Шеннона—Фано реально не позволяет производить изменение таблиц кодов после каждого обработанного символа, и в этом случае их обновление производится «время от времени», т. е. после обработки очередного блока из некоторого числа символов.

5.9. Непрерывный источник сообщений. Дифференциальная энтропия

Рассмотрим *непрерывный* источник сообщений (без памяти), состояния которого описываются не дискретным алфавитом, а непрерывным числовым множеством. Выбор очередного состояния таким источником отождествим с реализацией случайной величины непрерывного типа, тогда описать модель источника можно при помощи функции плотности распределения вероятностей.

Рассмотрим возможность применения понятия энтропии к случайной величине непрерывного типа X . Будем считать, что X принимает значения из полуинтервала $x \in [0, 1)$ и описывается

заданной непрерывной функцией плотности распределения вероятностей $f(x)$, $x \in [0, 1)$, $\int_0^1 f(x) dx = 1$. Разобьем 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:

$$\begin{aligned} H(\hat{X}) &= - \sum_{j=0}^{N-1} \int_{\Delta_j} f(x) dx \cdot \log \int_{\Delta_j} f(x) dx = - \sum_{j=0}^{N-1} f(\theta_j) q \cdot \log (f(\theta_j) q) = \\ &= -q \sum_{j=0}^{N-1} f(\theta_j) \log f(\theta_j) - \log q \sum_{j=0}^{N-1} \int_{\Delta_j} f(x) dx = \\ &= - \sum_{j=0}^{N-1} f(\theta_j) \log f(\theta_j) \cdot q - \log q, \end{aligned} \quad (5.20)$$

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 $\varepsilon^2 = M((X - \hat{X})^2)$:

$$\varepsilon^2 = \sum_{j=0}^{N-1} \int_{\Delta_j} \underbrace{(x - (j + \frac{1}{2})q)^2}_{\geq 0} \underbrace{f(x) dx}_{\text{непрерывна}} = \sum_{j=0}^{N-1} f(\xi_j) \int_{\Delta_j} (x - (j + \frac{1}{2})q)^2 dx,$$

где ξ_j — некоторая точка, которая в соответствии с одной из теорем о среднем для интегралов 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 дискрета (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, $H(\hat{X}) \sim -\log q$. 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 $q \rightarrow 0$ 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 соответствующие 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 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*.

Пример 5.16. Найти энтропию (5.24) случайной величины, имеющей стандартное нормальное распределение. Основание логарифма в формуле (5.24) положить равным e .

◀ Учитывая, что $\int_{-\infty}^{+\infty} e^{-x^2/2} dx = \int_{-\infty}^{+\infty} x^2 e^{-x^2/2} dx = \sqrt{2\pi}$, имеем:

$$H(X) = -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \ln\left(\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right) dx = -\ln \frac{1}{\sqrt{2\pi}} + \frac{1}{2} = \frac{1 + \ln(2\pi)}{2}. \quad \blacktriangleright$$

Для непрерывного источника информации (источника сообщений) X дифференциальная энтропия является мерой априорной неопределенности его состояния. Если положить в формуле (5.24) основание логарифма равным 2, то смысловое содержание понятия «дифференциальная энтропия» будет следующим. Пусть для некоторых непрерывных источников информации X , Y выполняется

неравенство $H(X) < H(Y)$, тогда при достаточно малом шаге квантования $q \rightarrow 0$ 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} потребуется 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. Отметим, что дифференциальная энтропия (5.24) может принимать и отрицательные значения.

Упражнение. Покажите, что если случайная величина непрерывного типа X имеет плотность распределения $f_X(t)$ и энтропию $H(X)$, то для случайной величины $Y = aX + b$ ($a \neq 0$) дифференциальная энтропия (5.24): $H(Y) = H(X) + \log|a|$. *Указание:* воспользуйтесь соотношением связи для функций плотности случайных величин Y и X : $f_Y(t) = \frac{1}{|a|} f_X\left(\frac{t-b}{a}\right)$.

Приведенное упражнение показывает, что смещение случайной величины на константу, $X' = X - b$, не влияет на значение дифференциальной энтропии. Масштабирование случайной величины, $X' = aX$, изменяет энтропию (5.24) следующим образом: $H(X') = H(X) + \log|a|$.

Пример 5.17. Пусть нормализованная¹ случайная величина непрерывного типа X принимает значения из некоторого интервала $X \in (-A; A)$. Необходимо найти плотность распределения $p(x)$ случайной величины, которая дает наибольшее значение дифференциальной энтропии (5.24):

$$H(X) = - \int_{-\infty}^{+\infty} p(x) \log p(x) dx = - \int_{-A}^A p(x) \log p(x) dx.$$

◀ Фактически необходимо найти максимум функционала

$$H(p(x)) = - \int_{-A}^A p(x) \log p(x) dx \quad (5.25)$$

при ограничениях:

¹ То есть имеющая нулевое математическое ожидание и единичную дисперсию.

$$\int_{-A}^A p(x)dx = 1 \tag{5.26}$$

(условие нормировки функции плотности вероятности) и

$$\int_{-A}^A x^2 p(x)dx = 1 \tag{5.27}$$

(требование единичной дисперсии). Математическое ожидание $m_x = 0$.

Решим сначала следующую вспомогательную задачу. Разобьем интервал $(-A; A)$ на N равных частей $\Delta_j = (-A + (j-1)q; -A + jq)$ длины $q = |\Delta_j| = 2A/N$ с серединами в точках $x_j = -A + (j - \frac{1}{2})q$, $j = 1, \dots, N$. Найдем такую функцию p дискретного аргумента, $p_j = p(x_j)$, чтобы функция от N переменных

$$H(p_1, p_2, \dots, p_N) = -q \sum_{j=1}^N p_j \log p_j \tag{5.28}$$

обращалась в максимум при условиях

$$q \sum_{j=1}^N p_j = 1, \tag{5.29}$$

$$q \sum_{j=1}^N x_j^2 p_j = 1. \tag{5.30}$$

Составив по (5.28)–(5.30) функцию Лагранжа

$$L(p_1, p_2, \dots, p_N) = -q \sum_{j=1}^N p_j \log p_j + \lambda_0 \left(q \sum_{j=1}^N p_j - 1 \right) + \lambda_1 \left(q \sum_{j=1}^N x_j^2 p_j - 1 \right)$$

и приравняв нулю ее частные производные

$$\frac{\partial L(p_1, \dots, p_N)}{\partial p_k} = q(-\log p_k - 1) + \lambda_0 q + \lambda_1 q x_k^2 = 0,$$

после сокращения множителя $q = 2A/N$ находим критические точки:

$$p_k = p(x_k) = e^{\lambda_0 - 1 + \lambda_1 x_k^2} = C e^{\lambda x_k^2}, \tag{5.31}$$

где обозначено $C = e^{\lambda_0 - 1}$, $\lambda = \lambda_1$. Несложно видеть, что соотношение (5.31) определяет максимум функции Лагранжа, т. к. ее матрица Гессе имеет диагональный вид и отрицательно определена:

$$\frac{\partial^2 L(p_1, \dots, p_N)}{\partial p_k \partial p_j} = \begin{cases} -q/p_k, & j = k, \\ 0, & j \neq k. \end{cases}$$

Теперь вернемся к основной задаче, определяемой соотношениями (5.25)–(5.27), и заметим, что выражения (5.28)–(5.30) представляют собой интегральные суммы для уравнений (5.25)–(5.27) соответственно. Поэтому решение задачи (5.25)–(5.27) получаем предельным переходом $N \rightarrow \infty$ в уравнениях (5.28)–(5.31), что дает из (5.31) выражение

$$p(x) = \begin{cases} Ce^{\lambda x^2}, & x \in (-A; A), \\ 0, & x \notin (-A; A). \end{cases} \quad (5.32)$$

Необходимые здесь параметры C и λ должны быть заданы исходя из требований (5.26) и (5.27), которые принимают вид:

$$C \int_{-A}^A e^{\lambda x^2} dx = 1, \quad C \int_{-A}^A x^2 e^{\lambda x^2} dx = 1.$$

Отсюда видим, что параметр λ должен удовлетворять уравнению

$$\int_{-A}^A e^{\lambda x^2} dx = \int_{-A}^A x^2 e^{\lambda x^2} dx, \quad (5.33)$$

решив которое для заданного значения A получим затем и нормировочный коэффициент $C = 1 / \int_{-A}^A e^{\lambda x^2} dx$.

Вместо (5.33) рассмотрим эквивалентное (в силу четности подынтегральных функций) уравнение

$$\int_0^A (1 - x^2) e^{\lambda x^2} dx = 0. \quad (5.34)$$

Очевидно, что для значений $0 < A \leq 1$ решение уравнения (5.34) относительно λ не существует, т. к. определенный интеграл от положительной функции не может быть равен нулю. Значению $A \rightarrow +\infty$, очевидно, соответствует $\lambda = -1/2$, когда (5.32) определяет стандартный нормальный закон. Другую очевидную пару (A, λ) можно найти, подставив в (5.34) значение $\lambda = 0$ — тогда получим (убедитесь!) $A = \sqrt{3}$, а плотность вероятности (5.32) будет определять равномерное распределение на отрезке $[-\sqrt{3}; \sqrt{3}]$.

Пары (A, λ) , удовлетворяющие (5.34), можно найти численно, определяя параметр A для различных значений $\lambda \in (-1/2; +\infty)$. Функция

$$F(A, \lambda) = \int_0^A (1 - x^2) e^{\lambda x^2} dx$$

для допустимого диапазона $1 < A < +\infty$ монотонна по переменной A , т. к. $\partial F(A, \lambda) / \partial A = (1 - A^2) e^{\lambda A^2} < 0$, поэтому для каждого значения

λ решение уравнения $F(A, \lambda) = 0$ (5.34) будет единственным. График функции $A(\lambda)$ приведен на рис. 5.14. Можно показать, что $\lim_{\lambda \rightarrow +\infty} A(\lambda) = 1$. ►

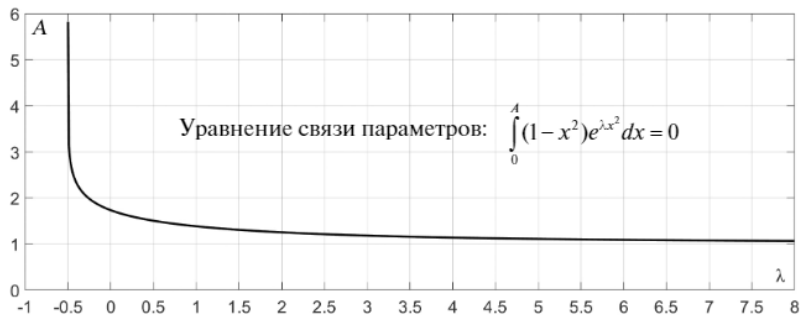


Рис. 5.14. Зависимость параметров (A, λ) функции плотности распределения вероятности (5.32)

Упражнение. Нарисуйте эскизы графика плотности вероятности (5.32) для значений параметра $A = 2$ и $A = 1,5$.

Результаты решения примера 5.17 позволяют сделать следующие важные выводы.

- При фиксированном значении дисперсии (например, не тая общности, единичном) максимальной дифференциальной энтропией обладает нормальное распределение.
- Если ограничить диапазон возможных значений случайной величины конечным отрезком длины $B = 2A$, то на этом отрезке равномерное распределение будет иметь наибольшую дифференциальную энтропию среди всех возможных распределений с той же дисперсией $D = B^2/12$.

Упражнение. Вычислите дифференциальную энтропию (с натуральным логарифмом) для случайной величины, имеющей равномерное распределение с единичной дисперсией, и сравните найденное значение с результатом, полученным в примере 5.16.

На непрерывный источник информации также можно распространить понятие условной энтропии. Рассмотрим ансамбль двух непрерывных источников (X, Y) , который удобно интерпретировать



как случайный двухкомпонентный вектор; обозначим его функцию плотности совместного распределения компонент $f_{X,Y}(x,y)$. Тогда безусловные распределения компонент описываются функциями плотности вероятностей

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dy, \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dx,$$

а условные распределения компонент — функциями плотности условного распределения вероятностей

$$f_{Y|x}(x,y) = f_{X,Y}(x,y)/f_X(x), \quad f_{X|y}(x,y) = f_{X,Y}(x,y)/f_Y(y).$$

Найдем дифференциальную энтропию ансамбля (X, Y) :

$$\begin{aligned} H(X,Y) &= - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y) \log f_{X,Y}(x,y) dx dy = \\ &= - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_X(x) f_{Y|x}(x,y) \log(f_X(x) f_{Y|x}(x,y)) dx dy = \\ &= - \int_{-\infty}^{+\infty} f_X(x) \log f_X(x) \underbrace{\int_{-\infty}^{+\infty} f_{Y|x}(x,y) dy}_{1} dx - \int_{-\infty}^{+\infty} f_X(x) \int_{-\infty}^{+\infty} f_{Y|x}(x,y) \log f_{Y|x}(x,y) dy dx = \\ &= H(X) + \int_{-\infty}^{+\infty} f_X(x) H(Y|x) dx = H(X) + H(Y|X), \end{aligned}$$

где

$$H(Y|X) = \int_{-\infty}^{+\infty} f_X(x) H(Y|x) dx \tag{5.35}$$

— условная энтропия источника Y по отношению к источнику X , а

$$H(Y|x) = - \int_{-\infty}^{+\infty} f_{Y|x}(x,y) \log f_{Y|x}(x,y) dy \tag{5.36}$$

— частная условная энтропия источника Y по отношению к источнику X , находящемуся в реализованном (частном) состоянии x .

Аналогично можно записать, что $H(X, Y) = H(Y) + H(X|Y)$. Таким образом, формула (5.13) верна как для дискретных, так и для непрерывных источников сообщений. В общем случае для ансамбля непрерывных источников (X_1, X_2, \dots, X_n) также справедливо соотношение:

$$H(X_1, X_2, \dots, X_n) = H(X_1) + H(X_2|X_1) + H(X_3|X_1 X_2) + \dots + H(X_n|X_1 \dots X_{n-1}).$$

При выводе формул (5.35) и (5.36) мы фактически использовали схему, аналогичную применявшейся ранее для получения формул

(5.11)–(5.13), заменяя суммирование на интегрирование, а дискретные вероятности на функции плотности вероятностей. Используя такую аналогию, несложно обобщить результат теоремы 5.3 на случай непрерывных источников сообщений.

Упражнение. Используя схему доказательства теоремы 5.3, покажите, что для ансамбля двух непрерывных источников информации (X, Y) сохраняются в силе соотношения: $H(X|Y) \leq H(X)$, $H(Y|X) \leq H(Y)$.

5.10. Передача дискретного сообщения по каналу с помехами

Из-за сбоев аппаратуры, влияния шумов и других возможных причин при хранении или передаче сообщения, созданного дискретным источником информации X , могут возникнуть искажения (ошибки), состоящие в подмене некоторых символов сообщения.

Для уменьшения влияния ошибок в дискретное сообщение добавляется некоторая дополнительная служебная информация, которая приводит к повышению избыточности данных, но снижает вероятность появления ошибок в декодированном сообщении. Для этих целей разработаны специальные методы так называемого *помехоустойчивого кодирования*.

Идея построения помехоустойчивых кодов состоит в том, чтобы двоичные кодовые комбинации $H_n = (h_1 \dots h_k)_n$, $h_k \in \{0; 1\}$, приписываемые каждому символу (блоку символов) x_n источника X , различались максимальным количеством битов; при этом задействованными оказываются не все возможные битовые комбинации, часть из них оказываются «запрещенными». Тогда, если вследствие имевших место искажений при декодировании очередного символа источника сообщений возникает запрещенная комбинация $\tilde{H} = (\tilde{h}_1 \dots \tilde{h}_k)$, она декодируется как символ x_m источника X , двоичный код которого $H_m = (h_1 \dots h_k)_m$ отличается от \tilde{H} минимальным количеством битовых разрядов. Количество битов, различающихся в кодовых словах H_m и \tilde{H} , называется *кодовым расстоянием* (при совпадении слов кодовое расстояние между ними равно нулю).

Теория методов построения кодов, исправляющих ошибки, достаточно обширна и выходит за рамки нашего рассмотрения. Приведем небольшой пример построения *кодов Хэмминга* для одного частного случая.

Пример 5.18. Код Хэмминга для источника сообщений $X = \{x_n\}_{n=0}^{15}$ с алфавитом из 16 символов, устойчивый к искажению одного бита.

◀ Припишем каждому символу источника X четырехбитовую кодовую комбинацию¹ B_n : $x_n \leftrightarrow B_n = (b_3 b_2 b_1 b_0)_n$, где $n = \sum_{k=0}^3 b_k 2^k$, $b_k \in \{0; 1\}$. Каждое кодовое слово $H_n = (h_1 \dots h_7)_n$, соответствующее символу x_j источника X , составим из 7 битов по правилу:

$$\begin{aligned} h_1 &= b_3 \oplus b_2 \oplus b_0, & h_2 &= b_3 \oplus b_1 \oplus b_0, & h_3 &= b_3, \\ h_4 &= b_2 \oplus b_1 \oplus b_0, & h_5 &= b_2, & h_6 &= b_1, & h_7 &= b_0, \end{aligned}$$

где обозначение \oplus использовано для битовой операции «исключающее или» (сложение по модулю 2). Тогда кодовое расстояние между любыми двумя словами из полученного набора $\{H_n\}_{n=0}^{15}$ будет не меньше 3. При этом возможную ошибку в одном из семи битов любого кодового слова можно не только обнаружить, но и устранить. Декодер, приняв некоторое кодовое слово, должен вычислить три контрольных бита:

$$c_0 = h_1 \oplus h_3 \oplus h_5 \oplus h_7, \quad c_1 = h_2 \oplus h_3 \oplus h_6 \oplus h_7, \quad c_2 = h_4 \oplus h_5 \oplus h_6 \oplus h_7.$$

Полученное трехбитовое число $C = 4c_2 + 2c_1 + c_0$ равно нулю, если кодовое слово $H = (h_1 \dots h_7)$ не содержит ошибок. В противном случае бит h_c , номер которого определяется числом $C \in \{0, 1, \dots, 7\}$, декодирован ошибочно, и его нужно инвертировать.

Если в кодовом слове $H = (h_1 \dots h_7)$ искаженными окажутся два бита или более, то устранить ошибки с использованием приведенного способа кодирования-декодирования будет невозможно. Однако вероятность повторных ошибок намного меньше вероятности единичной ошибки. Например, если вероятность искажения одного бита $q = 0,005$ и не зависит от того, искажены ли другие

¹ Для источника X с 16 равновероятными состояниями такой код является оптимальным. Можно рассматривать символы источника X как блок $B = (b_3 b_2 b_1 b_0)$ из четырех символов сообщения источника без памяти с двумя равновероятными состояниями, который моделирует выходной код оптимального статистического кодера (см. рис. 5.8).

биты в кодовом слове, то с использованием схемы испытаний Бернулли для вероятности ошибки в одном бите из семи получим значение $p_1 \approx 0,034$, а для вероятности искажения двух битов из семи: $p_2 \approx 0,00051$. ►

Для повышения помехоустойчивости (т. е. способности исправлять большее количество ошибок) необходимо использовать коды с большими расстояниями между кодовыми словами. Очевидно, что за это придется платить увеличением избыточности применяемых кодов.

Упражнение. Используя персональный компьютер, проверьте полным перебором, что между любыми двумя кодовыми словами из примера 5.18 кодовое расстояние не меньше 3. Убедитесь также, что искажение одного (любого) из семи битов в любом из 16 кодовых слов устранимо по приведенному в примере 5.18 алгоритму декодирования.

Допустив возможность искажения сообщения при передаче и убедившись на примере, что в определенных пределах устранение ошибок возможно, рассмотрим соответствующую модель дискретного канала передачи информации с потерями.

Пусть алфавит источника без памяти X имеет N символов с заданным распределением вероятностей их появления, $X = \{(x_n, p(x_n))\}_{n=1}^N$. Символы сообщения источника X поступают на вход дискретного канала передачи информации, в результате чего на выходе канала появляются символы $\{y_k\}_{k=1}^M$ с вероятностями появления $\{p(y_k)\}_{k=1}^M$ (в общем случае алфавиты X и Y , а также количество символов в алфавитах N и M различны). То есть сообщение на выходе канала передачи информации можно рассматривать как результат работы источника $Y = \{(y_k, p(y_k))\}_{k=1}^M$.

Считаем, что канал не имеет памяти, т. е. вероятности появления символов на его выходе не зависят от предыстории. Тогда вероятности появления символов $\{y_k\}_{k=1}^M$ можно представить по формуле полной вероятности следующим образом:

$$p(y_k) = \sum_{j=1}^N p(x_j) p(y_k | x_j), \quad (5.37)$$

$k = 1, 2, \dots, M$.

Сформируем векторы-столбцы $\mathbf{P}(X) = (p(x_1), p(x_2), \dots, p(x_N))^T$ и $\mathbf{P}(Y) = (p(y_1), p(y_2), \dots, p(y_M))^T$. Вводя матрицу условных вероятностей

$$\mathbf{Q} = \begin{pmatrix} p(y_1|x_1) & p(y_1|x_2) & \dots & p(y_1|x_N) \\ p(y_2|x_1) & p(y_2|x_2) & \dots & p(y_2|x_N) \\ \dots & \dots & \dots & \dots \\ p(y_M|x_1) & p(y_M|x_2) & \dots & p(y_M|x_N) \end{pmatrix}, \quad (5.38)$$

уравнения (5.37) можно кратко записать тогда в матричном виде как

$$\mathbf{P}(Y) = \mathbf{Q}\mathbf{P}(X). \quad (5.39)$$

Матрицу $\mathbf{Q} = (q_{k,j} = p(y_k|x_j))$ (5.38) назовем *матрицей переходных вероятностей канала* или, кратко, *матрицей канала*.

Так как канал может исказить передаваемую информацию, появление на выходе некоторого символа y_k не несет полной информации о том, какой из символов $\{x_j\}_{j=1}^N$ был подан на вход канала передачи. Соответствующая неопределенность входного состояния выражается как частная условная энтропия дискретного источника X по отношению к наблюдаемому выходному состоянию канала y_k :

$$H(X|y_k) = -\sum_{j=1}^N p(x_j|y_k) \log p(x_j|y_k). \quad (5.40)$$

Усредняя данную величину по всем возможным выходным состояниям канала, получаем неопределенность входного состояния канала по отношению к выходному как условную энтропию

$$H(X|Y) = \sum_{k=1}^M p(y_k) H(X|y_k) = -\sum_{j=1}^N \sum_{k=1}^M p(x_j, y_k) \log p(x_j|y_k). \quad (5.41)$$

Отметим, что в случае полного отсутствия ошибок в канале матрица (5.38) является единичной, а значения неопределенностей (5.40) и (5.41) равны нулю, т. е. информация на выходе канала полностью описывает состояние источника X на входе.

Неотрицательную функцию

$$I(X, Y) = H(X) - H(X|Y) \quad (5.42)$$

назовем *взаимной информацией* источника X и канала Y . Так как $H(X) \geq H(X|Y)$ (см. теорему 5.3), то взаимная информация (5.42) не может принимать отрицательные значения. При отсутствии ошибок в канале величина (5.42) в точности равна $H(X)$, т. е. количеству

информации, создаваемой источником X . При наличии ошибок канала передачи информация на его выходе о состоянии источника X является неполной, уменьшаясь на величину (5.41). Таким образом, потери информации в канале передачи определяются величиной (5.41), а среднее количество информации, прошедшей через канал — величиной (5.42).

Воспользовавшись соотношением (5.41), преобразуем (5.42):

$$I(X, Y) = - \sum_{j=1}^N \sum_{k=1}^M \overbrace{p(x_j, y_k) \log p(x_j)}^{H(X)} + \sum_{j=1}^N \sum_{k=1}^M \overbrace{p(x_j, y_k) \log \frac{p(x_j, y_k)}{p(y_k)}}^{-H(X|Y)},$$

$$I(X, Y) = \sum_{j=1}^N \sum_{k=1}^M p(x_j, y_k) \log \frac{p(x_j, y_k)}{p(x_j)p(y_k)}. \quad (5.43)$$

Соотношение (5.43) наглядно показывает, что при независимости появления символов на входе и выходе канала, т. е. когда $\forall k, j: p(x_j, y_k) = p(x_j)p(y_k)$, взаимная информация принимает минимальное значение $I(X, Y) = 0$. Действительно, если появление того или иного символа на выходе канала связи совершенно не зависит от того, какой символ был подан на вход канала (создан источником X), то никакой передачи информации по каналу не происходит.

Учитывая (см. (5.38)), что $p(x_j, y_k) = p(x_j)p(y_k | x_j) = p(x_j)q_{k,j}$, $p(y_k) = \sum_{m=1}^N p(x_m, y_k) = \sum_{m=1}^N p(x_m)q_{k,m}$, и подставляя данные выражения в (5.33), получим:

$$I(X, Y) = \sum_{j=1}^N p(x_j) \sum_{k=1}^M q_{k,j} \log \frac{q_{k,j}}{\sum_{m=1}^N p(x_m)q_{k,m}}. \quad (5.44)$$

Как видим, среднее количество информации, получаемое на выходе канала передачи, полностью определяется свойствами источника X (распределением вероятностей $\{p(x_j)\}_{j=1}^N$) и матрицей канала (5.38).

Упражнение. Начав с выражения (5.44), покажите, что для взаимной информации верно симметричное к (5.42) выражение: $I(X, Y) = H(Y) - H(Y | X)$.

Связь между значениями взаимной информации и неопределенностями источника и канала можно представить в виде схемы рис. 5.15.

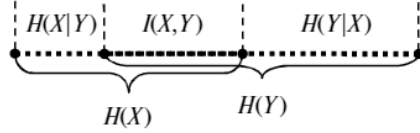


Рис. 5.15. Взаимосвязь информационных характеристик источника данных и канала передачи

Упражнение. Источник X имеет два состояния: $p(x_1) = 1/8, p(x_2) = 7/8$. Матрица канала имеет вид: $\mathbf{Q} = \begin{pmatrix} 3/4 & 1/8 \\ 1/4 & 7/8 \end{pmatrix}$. Используя при вычислениях двоичное основание логарифма, найдите взаимную информацию (5.42) и сравните значения двоичной энтропии $H(X)$ и $H(Y)$.

Определение. *Пропускной способностью* дискретного канала передачи информации с заданной матрицей канала (5.38) называется величина

$$C = \max_{P(X)} I(X, Y), \tag{5.45}$$

где максимум взаимной информации (5.42) (или (5.43), (5.44)) ищется по всем возможным распределениям вероятностей $\mathbf{P}(X) = (p(x_1), p(x_2), \dots, p(x_N))^T$ источника сообщений X .

Пропускная способность канала характеризует, какое максимальное количество информации может в среднем нести один выходной символ канала. Из (5.44) следует, что эта величина определяется только матрицей канала \mathbf{Q} (5.38). Очевидно, что для канала без потерь информации $C = \log N$.

Пример 5.19. Пусть алфавит источника сообщений состоит из двух символов, а матрица переходных вероятностей канала имеет вид:

$$\mathbf{Q} = \begin{pmatrix} 1-q & q \\ q & 1-q \end{pmatrix},$$

где q — вероятность ошибочной передачи символа. Такой канал передачи информации называется *двоичным симметричным*. Найдем его пропускную способность (5.45).

◀ Преобразуем формулу (5.44), для случая $N = M = 2$ получаем:

$$I(X, Y) = - \underbrace{\sum_{k=1}^2 \left(\sum_{j=1}^2 \overbrace{p(x_j)q_{k,j}}^{p(y_k)} \right) \log \left(\sum_{m=1}^2 \overbrace{p(x_m)q_{k,m}}^{p(y_k)} \right)}_{H(Y)} + \underbrace{\sum_{j=1}^2 p(x_j) \sum_{k=1}^2 q_{k,j} \log q_{k,j}}_{-H(Y|X)}.$$

Обозначим $p(x_1) = p$, $p(x_2) = 1 - p$. Несложно убедиться, что в полученном выражении для взаимной информации слагаемое $-H(Y|X)$ не зависит от распределения вероятностей источника X , т. е. от величины p :

$$-H(X|Y) = \sum_{j=1}^2 p(x_j) \sum_{k=1}^2 q_{k,j} \log q_{k,j} = p((1-q)\log(1-q) + q\log q) + (1-p)(q\log q + (1-q)\log(1-q)) = q\log q + (1-q)\log(1-q).$$

Поэтому для определения пропускной способности канала (5.45) необходимо найти максимально возможное значение

$$H_{\max}(Y) = \max_p H(Y) = \max_p (-p(y_1)\log p(y_1) - p(y_2)\log p(y_2)),$$

где (см. (5.39)):

$$\begin{pmatrix} p(y_1) \\ p(y_2) \end{pmatrix} = \mathbf{Q} \begin{pmatrix} p(x_1) \\ p(x_2) \end{pmatrix} = \begin{pmatrix} 1-q & q \\ q & 1-q \end{pmatrix} \begin{pmatrix} p \\ 1-p \end{pmatrix}.$$

Для имеющего два выходных состояния канала передачи значение $H_{\max}(Y) = \log 2$ достигается при $p(y_1) = p(y_2) = 1/2$, что для любого $q \neq 1/2$ определяет равновероятное распределение источника X : $p(x_1) = p(x_2) = 1/2$ (убедитесь!). Окончательно получаем для пропускной способности двоично-симметричного канала выражение:

$$C = \log 2 + q\log q + (1-q)\log(1-q).$$

Видим, что максимальная пропускная способность канала имеет место при отсутствии ошибок, когда матрица канала \mathbf{Q} — единичная ($q = 0$). ►

Упражнение. Покажите, что для вероятности ошибки двоичного симметричного канала $q = 1/2$ при любом распределении $p(x_1) = p$, $p(x_2) = 1 - p$ двоичного источника X пропускная способность канала $C = 0$, т. е. передача какой-либо информации невозможна.

Применение специальных кодов, исправляющих ошибки, совместно с использованием методов статистического кодирования источников сообщений позволяет получить эффективную и достоверную передачу информации даже при использовании каналов с потерями данных. Теоретический предел эффективности передачи информации при этом определяет следующая теорема Шеннона, которую мы приведем без доказательства.



Теорема 5.5. Если дискретный источник сообщений X имеет энтропию $H(X)$, то для канала с пропускной способностью $C > H(X)$ при кодировании сообщения бесконечной длительности $\forall \varepsilon > 0$ возможно построение таких кодов, для которых вероятность q ошибки декодирования очередного символа: $q < \varepsilon$. Если же $C < H(X)$, то существует число $q_0 > 0$, такое, что получение вероятности ошибки декодирования $q < q_0$ невозможно.

Таким образом, пропускная способность дискретного канала определяет максимально возможную энтропию источника, для которого при передаче информации по данному каналу может быть получена сколь угодно малая вероятность q ошибки декодирования символа (при достаточно большой длине сообщения).

5.11. 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 году Лемпелом (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 этом методе по последовательности кодируемых символов перемещается окно, состоящее из двух частей: словаря, в который

помещаются уже закодированные символы сообщения, и опережающего буфера, см. рис. 5.16. Обычно словарь содержит несколько тысяч символов, а буфер — несколько десятков.

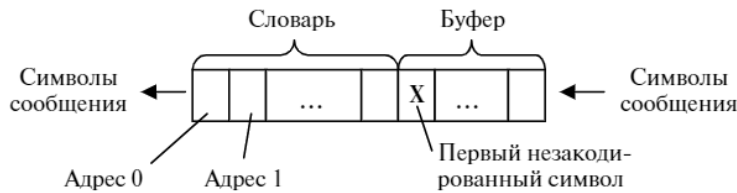


Рис. 5.16. Скользящее окно кодера LZ77

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, смещается на одну позицию. Например, если окно кодера находится в положении, приведенном на рис. 5.17а, то кодер обнаружит в словаре по адресу 5 совпадающую с началом буфера строку из двух символов «ес» и выдаст запись $(5, 2, \text{”т”})$. После этого скользящее окно переместится по сообщению на три символа, см. рис. 5.17б. Следующей записью, которую выдаст кодер, будет тройка $(0, 0, \text{”ь”})$, а окно переместится затем на одну позицию.



Рис. 5.17. Иллюстрация одного шага кодирования по методу LZ77



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-битовая 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 году, авторы — Сторер (Storer) и Шиманский (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 алгоритмах 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 году.

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 закодированный 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 выполняет следующие действия.

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 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 году.

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: $S_I = S$.
3. Присваивает текущей строке значение $S = c$.

Строка S_I 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 $P(k)$) надо приписать к строке $P(I)$ для формирования фразы S_I , 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 является 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 представить 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 литературе (например, [5]).



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 (zip, rar и др.).

ГЛАВА 6

ПРИМЕНЕНИЕ ДИСКРЕТНЫХ ОРТОГОНАЛЬНЫХ ПРЕОБРАЗОВАНИЙ ДЛЯ КОМПРЕССИИ И СПЕКТРАЛЬНОГО АНАЛИЗА СИГНАЛОВ

6.1. Корреляция как мера статистической зависимости данных. Преобразование Карунена—Лозва

Пусть $\mathbf{X} = (X_0, \dots, X_{N-1})^T$ — случайный вектор, состоящий из временных отсчетов дискретного сигнала. Для реальных физических процессов соседние отсчеты обычно имеют близкие значения, т. е. между компонентами вектора имеется статистическая зависимость. В терминах пятой главы такой 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 *ковариационным моментом* (ковариацией):

$$\text{cov}(X, Y) = \mathbf{M}[(X - m_X)(Y - m_Y)] = \mathbf{M}(XY) - m_X m_Y,$$



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 в общем случае не верно.

Упражнение. Пусть случайная величина имеет равномерное распределение $X \sim R(-1; 1)$, а случайная величина $Y = X^2$. Убедитесь, что $\text{cov}(X, Y) = 0$, т. е., несмотря на наличие функциональной зависимости, случайные величины X и Y некоррелированы.

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 $\rho(X, Y) = \text{cov}(X, Y) / \sqrt{\sigma_X^2 \sigma_Y^2}$, где σ_X^2, σ_Y^2 — их дисперсии. Коэффициент корреляции может принимать значения из диапазона $\rho \in [-1; 1]$ и характеризует наличие линейной зависимости случайных величин, что иллюстрирует следующий пример.

Пример 6.1. Пусть X, Y — случайные величины, имеющие математические ожидания m_X, m_Y и среднеквадратичные отклонения σ_X, σ_Y соответственно, с известным коэффициентом корреляции $\rho_{X,Y}$. Найти выражение для линейного прогноза \tilde{Y} случайной величины Y по значению случайной величины X , обеспечивающее минимальное значение среднего квадрата ошибки предсказания $\varepsilon^2 = M[(Y - \tilde{Y})^2]$.

◀ Будем искать линейную функцию прогноза в виде $\tilde{Y} = \alpha + \beta(X - m_X)$. Рассмотрим квадратичное отклонение

$$\begin{aligned} \varepsilon^2(\alpha, \beta) &= M[(Y - \tilde{Y})^2] = M[(Y - \alpha - \beta(X - m_X))^2] = \\ &= M[(Y - \alpha)^2] - 2\beta M[(Y - \alpha)(X - m_X)] + \beta^2 \sigma_X^2. \end{aligned}$$

Так как $M[(Y - \alpha)(X - m_X)] = M(XY - \alpha X + \alpha m_X - Y m_X) =$
 $= M(XY) - m_X m_Y = \text{cov}(X, Y),$

то $\varepsilon^2(\alpha, \beta) = M(Y^2) - 2\alpha m_Y + \alpha^2 - 2\beta \text{cov}(X, Y) + \beta^2 \sigma_X^2.$

Находя минимум полученного выражения по параметрам α и β , получаем: $\alpha = m_Y, \beta = \text{cov}(X, Y) / \sigma_X^2 = \rho_{X,Y} \sigma_Y / \sigma_X$. Тогда окончательно функция линейного прогноза принимает вид:

$$\tilde{Y} = m_Y + \rho_{X,Y} \frac{\sigma_Y}{\sigma_X} (X - m_X), \quad (6.1)$$

при этом для квадратичной ошибки прогноза имеем:

$$\varepsilon^2 = M\left[(Y - \tilde{Y})^2\right] = \sigma_Y^2(1 - \rho_{X,Y}^2) \quad (6.2)$$

(убедитесь!).

Чем ближе к единице абсолютная величина коэффициента корреляции, тем сильнее зависимость между случайными величинами. Как следует из формулы (6.2), при $|\rho_{X,Y}| = 1$ случайные величины жестко связаны линейной зависимостью, и с использованием формулы (6.1) реализация одной случайной величины позволяет указать точное значение другой. Таким образом, коэффициент корреляции является мерой линейной зависимости случайных величин: чем ближе модуль коэффициента корреляции к единице, тем сильнее эта зависимость. ►

Упражнение. Покажите, что минимум значения $M[(Y - \alpha)^2]$, где Y — некоторая случайная величина, α — константа, достигается, если выбрать $\alpha = M(Y)$.

В теории вероятностей вводится понятие *функции регрессии* случайной величины Y на случайную величину X . Данная функция $\varphi(x)$ представляет собой условное математическое ожидание Y при известном реализованном значении $X = x$: $m_{Y|x} = M(Y|X = x) = \varphi(x)$. Если функция регрессии является линейной, то (см. (6.1))

$$m_{Y|x} = m_Y + \rho_{X,Y} \frac{\sigma_Y}{\sigma_X} (x - m_X), \quad (6.3)$$

при этом величина $\sigma_Y^2(1 - \rho_{X,Y}^2)$ (см. (6.2)) представляет собой дисперсию условного распределения Y . В общем случае функция регрессии $m_{Y|x} = \varphi(x)$ может быть нелинейной, тогда выражение (6.3) следует рассматривать как ее линейное приближение.

Таким образом, наличие или отсутствие зависимости между случайными величинами в первом приближении можно характеризовать их коррелированностью, т.е. зависимостью линейного характера. Отметим, что для одного из наиболее важных распределений — двумерного (в общем случае, многомерного) нормального закона — корреляционные связи полностью описывают статистические зависимости между компонентами вектора, понятия



коррелированности и зависимости являются тождественными, а функция регрессии имеет линейный вид.

В данной главе понятия коррелированности и зависимости случайных величин мы также отождествим, т. е. ограничим изучение статистических зависимостей случайных величин рамками корреляционных связей. В соответствующей *корреляционной теории* закон распределения исследуемых случайных величин (векторов) обычно считается неизвестным, либо подразумевается многомерное нормальное распределение. 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, неотрицательно определена, а ее диагональные элементы представляют собой дисперсии компонент случайного вектора \mathbf{X} . Считая \mathbf{X} и \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 расположенными 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 5.15).

Как отмечалось в пятой главе, 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 сигнала* 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, функции регрессии 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 обратимого преобразования F , переводящего вектор X в некоторый вектор $Y = F\{X\}$, в котором зависимость между компонентами ослаблена (в рамках корреляционного подхода это означает, что компоненты менее коррелированы). Тогда независимое покомпонентное кодирование вектора Y , а не вектора X , становится более обоснованным. Вектор Y будем называть вектором трансформант (от англ. *transform* — преобразование).

Положим, что для вектора данных X известны ковариационная матрица K_X и вектор математических ожиданий компонент $\mathbf{m}_X = M(X) = (m_{x_0}, \dots, m_{x_{N-1}})^T$. Пусть вектор Y получен из вектора X в результате некоторого линейного преобразования, которое определяется матрицей W : $Y = WX$. Наша задача состоит в том, чтобы ослабить (в идеале — исключить) межкомпонентную зависимость данных в векторе Y . Поскольку статистическую зависимость мы отождествили с коррелированностью, это означает, что необходимо добиться того, чтобы ковариационная матрица вектора Y приняла диагональный вид. Возможно ли это? Для ответа на данный вопрос найдем сначала выражение для ковариационной матрицы вектора Y .

В силу линейности операции нахождения математического ожидания $\mathbf{m}_Y = (m_{y_0}, \dots, m_{y_{N-1}})^T = M(Y) = M(WX) = W \cdot M(X) = W \cdot \mathbf{m}_X$. Поэтому для ковариационной матрицы имеем:

$$\begin{aligned} K_Y &= M[(Y - \mathbf{m}_Y)(Y - \mathbf{m}_Y)^T] = M[(WX - W\mathbf{m}_X)(WX - W\mathbf{m}_X)^T] = \\ &= M[W(X - \mathbf{m}_X)(X - \mathbf{m}_X)^T W^T] = W \cdot M[(X - \mathbf{m}_X)(X - \mathbf{m}_X)^T] W^T, \end{aligned}$$



откуда

$$\mathbf{K}_Y = \mathbf{W}\mathbf{K}_X\mathbf{W}^T. \quad (6.4)$$

Напомним два утверждения, известные 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 $\mathbf{B} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ 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} , составленная 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 (6.4) следует, что ковариационная матрица \mathbf{K}_Y вектора $\mathbf{Y} = \mathbf{W}\mathbf{X}$ 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 — Лозва*.

Таким образом, по заданной ковариационной матрице вектора \mathbf{X} с помощью ортогонального преобразования Карунена — Лозва можно получить вектор $\mathbf{Y} = \mathbf{W}\mathbf{X}$, 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 — Лозва существенно ограничиваются следующими факторами. Прежде всего, это преобразование не является универсальным: 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 зависимости 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 неизвестны точные значения элементов

$\mathbf{K}_X = \{\text{cov}(X_k, X_m)\}_{k,m=0}^{N-1}$. Использование же некоторой «средней» ковариационной матрицы \mathbf{K}_X для построения преобразования Карунена — Лозва не позволяет говорить о его оптимальности для каждого отдельного вектора \mathbf{X} — очередной выборки из дискретного сигнала.

Другая проблема, связанная с использованием преобразования Карунена — Лозва, заключается в отсутствии быстрых алгоритмов его вычисления.

Упражнение. Для двухкомпонентного вектора \mathbf{X} , имеющего ковариационную матрицу $\mathbf{K}_X = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, где $\rho \in [-1; 1]$, построить матрицу \mathbf{W} преобразования Карунена — Лозва и найти ковариационную матрицу \mathbf{K}_Y вектора $\mathbf{Y} = \mathbf{W}\mathbf{X}$.

6.2. Эффективность использования дискретных ортогональных преобразований для кодирования коррелированных данных

Отсутствие быстрых алгоритмов вычисления и зависимость параметров оптимального преобразования Карунена — Лозва 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).

Наличие этапа квантования компонент вектора трансформант \mathbf{Y} (будем также называть этот вектор *спектром*, понимая данный термин обобщенно) вносит определенную ошибку $\|\mathbf{X} - \hat{\mathbf{X}}\|_E^2 = \|\mathbf{Y} - \hat{\mathbf{Y}}\|_E^2$ в восстановленные данные — вектор $\hat{\mathbf{X}}$ (см. также пример 2.16), которая может быть снижена до пренебрежимо малого уровня за счет квантования спектральных компонент с большим числом уровней. 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 декоррелирующая

эффективность применения вещественных ортогональных преобразований для целей кодирования, основан на следующих рассуждениях.



Рис. 6.1. Схема сжатия (а) и восстановления (б) коррелированных данных с использованием ортогонального преобразования

Мерой априорной неопределенности для непрерывной случайной величины является дифференциальная энтропия (5.24). Рассматривая N -компонентный вектор как ансамбль из N непрерывных источников информации, дифференциальную энтропию случайного вектора $\mathbf{Y} = (Y_0, Y_1, \dots, Y_{N-1})^T$ можем записать в виде:

$$H(\mathbf{Y}) = - \int_{\mathbb{R}^N} f_Y(y_0, \dots, y_{N-1}) \log f_Y(y_0, \dots, y_{N-1}) dy_0 \dots dy_{N-1} =$$

$$= - \int_{\mathbb{R}^N} f_Y(\mathbf{y}) \log f_Y(\mathbf{y}) d\mathbf{y},$$

где $f_Y(\mathbf{y}) = f_Y(y_0, \dots, y_{N-1})$ — функция плотности совместного распределения компонент вектора \mathbf{Y} . Если $\mathbf{Y} = \mathbf{W}\mathbf{X}$, где матрица \mathbf{W} — ортогональная (т. е. $\mathbf{W}^{-1} = \mathbf{W}^T$ и $\det \mathbf{W} = \det \mathbf{W}^{-1} = 1$), то, как можно показать¹:

$$f_Y(\mathbf{y}) = f_X(\mathbf{W}^{-1}\mathbf{y}) |\det \mathbf{W}^{-1}| = f_X(\mathbf{W}^T \mathbf{y}),$$

где $f_X(\mathbf{x}) = f_X(x_0, \dots, x_{N-1})$ — функция плотности совместного распределения компонент вектора \mathbf{X} . Поэтому

$$H(\mathbf{Y}) = - \int_{\mathbb{R}^N} f_Y(\mathbf{y}) \log f_Y(\mathbf{y}) d\mathbf{y} = - \int_{\mathbb{R}^N} f_X(\mathbf{W}^T \mathbf{y}) \log f_X(\mathbf{W}^T \mathbf{y}) d\mathbf{y}.$$

¹ См., например, вывод соотношения (5.31) в учебнике [34].

Выполнив в последнем интеграле замену переменных $\mathbf{x} = \mathbf{W}^T \mathbf{y}$ с якобианом $J = \det \mathbf{W} = 1$, получаем окончательно:

$$H(\mathbf{Y}) = - \int_{\mathbb{R}^N} f_X(\mathbf{x}) \log f_X(\mathbf{x}) d\mathbf{x} = H(\mathbf{X}).$$

Таким образом, после применения ортогонального преобразования дифференциальная энтропия вектора не изменяется.

С использованием понятия условной энтропии (5.35) дифференциальную энтропию вектора $\mathbf{Y} = \mathbf{W}\mathbf{X}$ можем представить в виде

$$H(\mathbf{Y}) = H(Y_0) + H(Y_1 | Y_0) + H(Y_2 | Y_0 Y_1) + \dots + H(Y_{N-1} | Y_0 Y_1 \dots Y_{N-2}).$$

Поскольку $H(Y_j | Y_0 \dots Y_{j-1}) \leq H(Y_j)$, $j = 1, \dots, N-1$ (см. раздел 5.9), то

$$H(\mathbf{Y}) \leq \sum_{j=0}^{N-1} H(Y_j).$$

Данное неравенство выполняется как равенство, если компоненты вектора $\mathbf{Y} = (Y_0, Y_1, \dots, Y_{N-1})^T$ являются независимыми случайными величинами. В общем случае можем записать $H(\mathbf{Y}) = \sum_{j=0}^{N-1} H(Y_j) - I_{\text{св}}(\mathbf{Y})$, где величина $I_{\text{св}}(\mathbf{Y}) \geq 0$ характеризует информацию, заключенную в статистических связях между компонентами вектора \mathbf{Y} : декоррелирующая эффективность преобразования $\mathbf{Y} = \mathbf{W}\mathbf{X}$ тем выше, чем ближе к нулю значение $I_{\text{св}}(\mathbf{Y})$. Учитывая, что $H(\mathbf{Y}) = H(\mathbf{X})$, получаем окончательно следующее соотношение:

$$\sum_{j=0}^{N-1} H(Y_j) = H(\mathbf{X}) + I_{\text{св}}(\mathbf{Y}).$$

Из данного выражения следует важный вывод: применение дискретного ортогонального преобразования с матрицей \mathbf{W} к вектору $\mathbf{X} = (X_0, X_1, \dots, X_{N-1})^T$ для устранения межкомпонентных статистических связей будет тем эффективнее, чем меньше будет суммарная безусловная энтропия $H_{\Sigma}(\mathbf{Y}) = \sum_{j=0}^{N-1} H(Y_j)$ компонент вектора $\mathbf{Y} = \mathbf{W}\mathbf{X}$.

Для возможности сравнения преобразований различной размерности N вместо величины $H_{\Sigma}(\mathbf{Y})$ будем рассматривать далее среднюю безусловную энтропию одной компоненты вектора трансформант: $H_{\text{ср}} = H_{\text{ср}}(\mathbf{Y}) = H_{\Sigma}(\mathbf{Y})/N$. 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 $f_k^0(x) = f_k(0, 1, x)$ и учитывая,

что $f_k(m_k, \sigma_k, x) = \frac{1}{\sigma_k} f_k^0\left(\frac{x-m_k}{\sigma_k}\right)$, $\int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) dx = 1$, среднюю безусловную энтропию можем записать следующим образом:

$$\begin{aligned} H_{\text{cp}} &= \frac{1}{N} \sum_{k=0}^{N-1} H(Y_k) = -\frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) \log f_k(m_k, \sigma_k, x) dx = \\ &= -\frac{1}{N} \sum_{k=0}^{N-1} \log \frac{1}{\sigma_k} \int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) dx - \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0\left(\frac{x-m_k}{\sigma_k}\right) \log f_k^0\left(\frac{x-m_k}{\sigma_k}\right) d\left(\frac{x}{\sigma_k}\right) = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \log \sigma_k - \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx. \end{aligned} \quad (6.5)$$

Чем меньше средняя энтропия (6.5), тем эффективнее будет последующее независимое кодирование компонент вектора \mathbf{Y} . Отметим, что требование ортогональности исследуемого преобразования является принципиальным. Действительно, анализируя среднюю неопределенность одной компоненты вектора $\mathbf{Y} = \mathbf{W}\mathbf{X}$ по величине средней дифференциальной энтропии (6.5), мы фактически предполагаем (см. раздел 5.9), 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 $\|\mathbf{Y} - \hat{\mathbf{Y}}\|_E^2 \sim Nq^2/12$. Эта ошибка всегда равна ошибке, внесенной в восстановленный вектор $\hat{\mathbf{X}}$, только для ортогональных преобразований (см. лемму 2.1). Если анализируемые преобразования не являются ортогональными, то для них нельзя говорить об эквивалентности ошибки $\|\mathbf{X} - \hat{\mathbf{X}}\|_E^2 \sim Nq^2/12$, вносимой скалярным квантованием компонент вектора \mathbf{Y} в восстановленный вектор $\hat{\mathbf{X}}$, и сравнение декоррелирующей эффективности преобразований по критерию (6.5), 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 $\mathbf{W} = \{w_{k,j}\}_{k,j=0}^{N-1}$. Можно допустить возможность применения центральной предельной теоремы и положить, что распределение каждой из компонент вектора \mathbf{Y} подчиняется нормальному закону, а отличия имеют место только в параметрах распределения. Такое допущение не является строгим (прежде всего потому, что в общем случае X_j представляют собой зависимые величины), однако часто применяется на практике.

Действительно, зависимость между отсчетами X_m и X_{m+n} для реальных дискретных сигналов ослабевает с увеличением $|n|$, и для некоторого достаточно большого значения $M \gg 1$ случайные величины X_m и X_{m+M} можно считать независимыми. Положим, что размерность преобразования представима в виде $N = ML$, где $L \gg 1$, тогда для компонент вектора \mathbf{Y} получаем выражение:

$$Y_k = \sum_{n=0}^{ML-1} w_{k,n} X_n = \sum_{m=0}^{M-1} \left(\sum_{l=0}^{L-1} \overbrace{w_{k,Ml+m} X_{Ml+m}}^{\substack{\text{При } L \gg 1 \text{ стремится} \\ \text{к нормальному закону}}} \right),$$

Независимы при $M \gg 1$

т.е. внутренние суммы $S_k(m) = \sum_{l=0}^{L-1} w_{k,Ml+m} X_{Ml+m}$ есть случайные величины, которые в силу центральной предельной теоремы имеют при $L \gg 1$ асимптотически нормальное распределение. Поэтому вследствие композиционной устойчивости нормального закона вид распределения для случайных величин $Y_k = \sum_{m=0}^{M-1} S_k(m)$ также можно считать нормальным и положить, что $\forall k = 0, 1, \dots, N-1$:

$$-\int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx \approx C = -\int_{-\infty}^{+\infty} p(x) \log p(x) dx,$$

где $p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Мы будем использовать менее жесткое ограничение, а именно:

$$-\frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx \approx C = \text{const}$$

(для любых преобразований при $N \gg 1$). Тогда среднюю энтропию (6.5) одной компоненты вектора трансформант можно записать в виде:

$$H_{\text{cp}} = \frac{1}{2N} \sum_{k=0}^{N-1} \log \sigma_k^2 + C = \frac{1}{2N} \log \prod_{k=0}^{N-1} \sigma_k^2 + C, \quad (6.6)$$



где C — некоторая константа, не зависящая от вида используемого преобразования и его размерности.

Необходимые в выражении (6.6) 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 (6.6) упрощается, так как матрица \mathbf{K}_Y имеет диагональный вид и

$$\prod_{k=0}^{N-1} \sigma_k^2 = \det \mathbf{K}_Y = \det \left(\mathbf{W}\mathbf{K}_X \frac{\mathbf{W}^T}{\mathbf{W}^{-1}} \right) = \det \mathbf{W} \det \mathbf{W}^{-1} \det \mathbf{K}_X = \det \mathbf{K}_X.$$

Поэтому для данного преобразования

$$H_{\text{cp}} = \frac{1}{2N} \log \prod_{k=0}^{N-1} \sigma_k^2 + C = \frac{1}{2N} \log \det \mathbf{K}_X + C. \quad (6.7)$$

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). Рассматривая среднюю безусловную энтропию одного элемента вектора трансформант (6.5) как характеристику декоррелирующих свойств ортогональных преобразований, естественно в качестве «точки отсчета» принять значение (6.5) 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 (6.6), (6.7):

$$\Delta H(\mathbf{W}, \mathbf{K}_X) = \frac{1}{2N} \log \left(\prod_{k=0}^{N-1} \sigma_k^2 / \det \mathbf{K}_X \right). \quad (6.8)$$

Назовем данную величину *средней избыточной энтропией* [49]. Величина (6.8) характеризует избыточность данных в векторе \mathbf{Y} , связанную с неполной декорреляцией компонент (чем больше значение (6.8), тем меньше эффективность декоррелирующего преобразования с матрицей \mathbf{W}).

Можно показать, что величина (6.8) всегда является неотрицательной и принимает нулевое значение для преобразования Карунена—Лоэва. Если использовать в (6.8) логарифм с основанием 2, то физический смысл величины (6.8) 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 анализируемое преобразование к оптимальному.

Лемма 6.1. Для всех ортогональных преобразований след ковариационной матрицы инвариантен:

$$\begin{aligned} \text{trace } \mathbf{K}_Y &= \sum_{k=0}^{N-1} \sigma_k^2 = \text{trace } \mathbf{K}_X = \sum_{k=0}^{N-1} \sigma_{X_k}^2. \\ \leftarrow \sum_{k=0}^{N-1} \sigma_k^2 &= \sum_{k=0}^{N-1} \mathbb{M}((Y_k - m_{Y_k})^2) = \mathbb{M} \left(\underbrace{(\mathbf{Y} - \mathbf{m}_Y)^T}_{\text{строка}} \underbrace{(\mathbf{Y} - \mathbf{m}_Y)}_{\text{столбец}} \right) = \\ &= \mathbb{M} \left((\mathbf{X} - \mathbf{m}_X)^T \underbrace{\mathbf{W}^T \mathbf{W}}_{\mathbf{E}} (\mathbf{X} - \mathbf{m}_X) \right) = \mathbb{M}((\mathbf{X} - \mathbf{m}_X)^T (\mathbf{X} - \mathbf{m}_X)) = \sum_{k=0}^{N-1} \sigma_{X_k}^2. \rightarrow \end{aligned}$$

Аналогично показывается, что для ортогональных преобразований в векторе трансформант сохраняется «энергия»:

$$\mathbb{M} \left(\sum_{k=0}^{N-1} Y_k^2 \right) = \mathbb{M} \left(\sum_{k=0}^{N-1} X_k^2 \right),$$

которая равна следу ковариационной матрицы в случае $\mathbf{m}_X = 0$.

Из формулы (6.8) следует, что при фиксированной размерности N наилучший вариант использования ортогонального преобразования дает максимальное значение параметра

$$P = \prod_{k=0}^{N-1} \sigma_k^2. \quad (6.9)$$

Упражнение. 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 убедитесь в том, что максимум выражения (6.9) достигается при $\sigma_0^2 = \sigma_1^2 = \dots = \sigma_{N-1}^2$ (удобнее искать максимум от логарифма выражения (6.9)).

Таким образом, наилучший случай использования преобразований имеет место тогда, когда суммарная дисперсия $\sum_{k=0}^{N-1} \sigma_k^2 = \text{const}$ (фактически энергия сигнала) равномерно распределена по всем коэффициентам-трансформантам: $\sigma_0^2 = \sigma_1^2 = \dots = \sigma_{N-1}^2$.

Тогда оптимальность 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 эффективности 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 [59].

6.3. ДПФ в вещественной форме. Дискретное преобразование Хартли

При выводе критерия средней избыточной энтропии (6.8) мы предполагали, что ортогональное преобразование, определяемое матрицей \mathbf{W} , является вещественным. Вместе с тем при использовании ДПФ (2.20), 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 устранить, если воспользоваться свойством 1° ДПФ (см. раздел 2.6). Для коэффициента ДПФ y_k из (2.20) 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 $y_{N/2} = \bar{y}_{N-(N/2)} = \bar{y}_{N/2} = a_{N/2}$. Остальные спектральные коэффициенты в общем случае являются комплексными числами, и в силу свойства 1° ДПФ получаем для них:

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i}{N}kj} = \underbrace{\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \cos \frac{2\pi kj}{N}}_{a_k} + i \underbrace{\frac{-1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \sin \frac{2\pi kj}{N}}_{b_k} = a_k + ib_k,$$

$$y_{N-k} = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{\frac{2\pi i k j}{N}} = \underbrace{\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \cos \frac{2\pi k j}{N}}_{a_{N-k}} + i \underbrace{\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \sin \frac{2\pi k j}{N}}_{b_{N-k}} = a_k - i b_k.$$

Из вектора-спектра \mathbf{Y} , полученного в результате ДПФ вещественного вектора \mathbf{X} , сформируем вектор $\tilde{\mathbf{Y}}$ по следующему правилу:

$$\tilde{y}_k = \begin{cases} a_k = y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \cos \frac{2\pi k j}{N}, & \text{для } k = 0 \text{ или } k = N/2 \\ a_k \sqrt{2} = \sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} x_j \cos \frac{2\pi k j}{N}, & \text{для } 1 \leq k < N/2 \\ -b_k \sqrt{2} = \sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} x_j \sin \frac{2\pi k j}{N}, & \text{для } N/2 < k \leq N-1 \end{cases} \quad (6.10)$$

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

Лемма 6.2. Дискретное преобразование вещественного вектора \mathbf{X} в вектор $\tilde{\mathbf{Y}} = \mathbf{W}\mathbf{X}$, определяемое правилом (6.10), 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 2.6) для любого вектора \mathbf{X} имеем $\|\mathbf{Y}\|_E = \|\mathbf{X}\|_E$, поэтому $\|\tilde{\mathbf{Y}}\|_E = \|\mathbf{X}\|_E$ и на основании леммы 2.1 получим, что определяемое формулой (6.10) преобразование является ортогональным, т. е. $\tilde{\mathbf{W}}^T = \tilde{\mathbf{W}}^{-1}$.

Для нечетного числа N докажите лемму самостоятельно. ▶

Преобразование (6.10) назовем ДПФ в вещественной форме или вещественным ДПФ (ВДПФ).

Другим родственным ДПФ (2.20) 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:

$$h_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} f_j \left(\cos \frac{2\pi k j}{N} + \sin \frac{2\pi k j}{N} \right), \quad (6.11)$$

$k = 0, 1, \dots, N-1$. В матричном виде: $\mathbf{h} = \mathbf{C} \cdot \mathbf{f}$, где

$$\mathbf{C} = \left\{ c_{k,j} = \frac{1}{\sqrt{N}} \left(\cos \frac{2\pi k j}{N} + \sin \frac{2\pi k j}{N} \right) \right\}_{k,j=0}^{N-1}.$$

Очевидно, что $\mathbf{C} = \mathbf{C}^T$.

Лемма 6.3. Дискретное преобразование Хартли (6.11) — ортогонально.

◀ Необходимо показать (см. лемму 2.1), что преобразование любого вектора \mathbf{f} по формуле (6.11) сохраняет евклидову норму: $\|\mathbf{f}\|_E = \|\mathbf{h}\|_E$.

Пусть $\mathbf{Y} = \mathbf{W}\mathbf{f}$ — ДПФ (2.20) вещественного вектора \mathbf{f} . Из формул (6.11) и (2.20) видно, что $\mathbf{h} = \mathbf{C} \cdot \mathbf{f} = \text{Re}\mathbf{Y} - \text{Im}\mathbf{Y}$ (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, $a_k = \text{Re } y_k$, $b_k = \text{Im } y_k$. Для нормы вектора \mathbf{h} имеем:

$$\|\mathbf{h}\|_E^2 = \sum_{k=0}^{N-1} h_k^2 = \sum_{k=0}^{N-1} (a_k - b_k)^2 = \sum_{k=0}^{N-1} \underbrace{(a_k^2 + b_k^2)}_{|y_k|^2} - 2 \sum_{k=0}^{N-1} a_k b_k = \|\mathbf{Y}\|_E^2 - 2 \sum_{k=0}^{N-1} a_k b_k.$$

Положим, что N — четное число. Так как, с учетом свойства 1° ДПФ (см. раздел 2.6), $b_k = -b_{N-k}$, $a_k = -a_{N-k}$, 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 $\|\mathbf{f}\|_E = \|\mathbf{Y}\|_E$ и потому $\|\mathbf{f}\|_E = \|\mathbf{h}\|_E$. ▶

Следствие. Матрица обратного преобразования Хартли совпадает с матрицей прямого преобразования: $\mathbf{C} = \mathbf{C}^{-1}$.

Для вычисления преобразования (6.11) удобным может оказаться использование следующего тригонометрического тождества:

$$\cos t + \sin t = \sqrt{2} \sin\left(t + \frac{\pi}{4}\right) = \sqrt{2} \cos\left(t - \frac{\pi}{4}\right).$$

Упражнение. Найдите ВДПФ и дискретное преобразование Хартли вектора $\mathbf{X} = (1, 0, 0, 1)^T$.

Упражнение. Найдите ВДПФ и дискретное преобразование Хартли каждой из приведенных ниже конечных последовательностей длины N .

$$\text{а) } x(n) = \tilde{\delta}(n) = \begin{cases} 1, & n = 0; \\ 0, & n \neq 0; \end{cases} \quad \text{б) } x(n) = \tilde{\delta}(n - n_0), \quad 0 \leq n_0 \leq N - 1.$$

6.4. Дискретное косинусное преобразование

Одной из важнейших корреляционных моделей для дискретных сигналов является модель дискретного стационарного марковского процесса первого порядка¹. 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 $X(k)$, $X(k+m)$ зависит только от расстояния между моментами наблюдения дискретного процесса, т. е. от величины $|m|$.

Пример 6.2. Для дискретного стационарного марковского процесса первого порядка, имеющего коэффициент корреляции соседних отсчетов $r(X(k), X(k+1)) = \rho$, найти коэффициент корреляции произвольных отсчетов $X(k)$, $X(k+m)$.

◀ Приведем решение для случая, когда состояния дискретного процесса — также дискретные, т. е. отсчеты сигнала представляют собой случайные величины дискретного типа, $X(k) \in \{x_1, x_2, \dots, x_N\}$. Так как рассматривается стационарный марковский процесс

¹ Определение дискретного марковского процесса первого порядка см. в разделе 3.13.



первого порядка, то вероятность $P\{X(k) = x_n | X(k-1) = x_j\} = p(x_n | x_j)$ перехода процесса в очередное состояние x_n зависит только от текущего состояния x_j и не зависит от момента наблюдения (индекса k). Имеем:

$$\begin{aligned} P\{X(k) = x_{j_0}, X(k+m) = x_{j_m}\} &= P\{X(0) = x_{j_0}, X(m) = x_{j_m}\} = \\ &= \sum_{j_1=1}^N \sum_{j_2=1}^N \dots \sum_{j_{m-1}=1}^N P\{X(0) = x_{j_0}, X(1) = x_{j_1}, \dots, X(m-1) = x_{j_{m-1}}, X(m) = x_{j_m}\} = \\ &= \sum_{j_1=1}^N \sum_{j_2=1}^N \dots \sum_{j_{m-1}=1}^N P\{X(0) = x_{j_0}\} P\{X(1) = x_{j_1} | X(0) = x_{j_0}\} P\{X(2) = x_{j_2} | X(1) = x_{j_1}\} \dots \\ &\quad \dots P\{X(m-1) = x_{j_{m-1}} | X(m-2) = x_{j_{m-2}}\} P\{X(m) = x_{j_m} | X(m-1) = x_{j_{m-1}}\} = \\ &= p(x_{j_0}) \sum_{j_1=1}^N p(x_{j_1} | x_{j_0}) \sum_{j_2=1}^N p(x_{j_2} | x_{j_1}) \dots \sum_{j_{m-1}=1}^N p(x_{j_{m-1}} | x_{j_{m-2}}) p(x_{j_m} | x_{j_{m-1}}). \end{aligned}$$

Значения дисперсии и математического ожидания не влияют на величину коэффициента корреляции, поэтому положим $m_X = 0$, $\sigma_X^2 = 1$. Тогда с учетом последнего полученного соотношения искомый коэффициент корреляции

$$\begin{aligned} r(X(k), X(k+m)) &= M[X(k)X(k+m)] = \\ &= \sum_{j_0=1}^N \sum_{j_m=1}^N x_{j_0} x_{j_m} P\{X(k) = x_{j_0}, X(k+m) = x_{j_m}\} = \\ &= \sum_{j_0=1}^N p(x_{j_0}) x_{j_0} \underbrace{\sum_{j_1=1}^N p(x_{j_1} | x_{j_0}) \dots \sum_{j_{m-1}=1}^N p(x_{j_{m-1}} | x_{j_{m-2}})}_{\rho^m x_{j_0}} \overbrace{\sum_{j_m=1}^N p(x_{j_m} | x_{j_{m-1}}) x_{j_m}}^{M[X(m) | X(m-1) = x_{j_{m-1}}] = \rho x_{j_{m-1}}}. \end{aligned}$$

Последняя внутренняя сумма представляет собой математическое ожидание значения случайного процесса при условии, что его предыдущее состояние $x_{j_{m-1}}$. Поскольку в рамках рассматриваемой в данной главе корреляционной модели данных мы полагаем, что вся информация о статистических связях заключена в ковариационных моментах (напомним, это в полной мере справедливо для многомерного нормального закона), то, воспользовавшись результатами решения примера 6.1 и формулой (6.3), имеем:

$$M[X(m) | x(m-1) = x_j] = \rho x_j.$$

Полученное соотношение верно для любых значений m ; применяя его далее к следующим вложенным суммам, окончательно получаем:

$$r(X(k), X(k+m)) = \sum_{j_0=1}^N p(x_{j_0}) x_{j_0} \rho^m x_{j_0} = \rho^m \sigma_x^2 = \rho^m.$$

Таким образом, для марковского процесса первого порядка коэффициент корреляции $r(X(k), X(j)) = \rho^{|k-j|}$, где ρ — коэффициент корреляции соседних отсчетов сигнала. ►

Из рассмотренного примера следует, что 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:

$$\mathbf{K}_X(\rho) = \sigma_x^2 \begin{pmatrix} 1 & \rho & \dots & \rho^{N-1} \\ \rho & 1 & \dots & \rho^{N-2} \\ \dots & \dots & \dots & \dots \\ \rho^{N-1} & \rho^{N-2} & \dots & 1 \end{pmatrix}. \quad (6.12)$$

Модель (6.12) представляет особый интерес, так как часто используется на практике для описания реальных дискретных сигналов, причем обычно параметр $\rho > 0$.

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 переводит вектор $\mathbf{X} = (x_0, x_1, \dots, x_{N-1})^T$ в вектор $\mathbf{Y} = (y_0, y_1, \dots, y_{N-1})^T$ по формуле

$$y_k = \sqrt{\frac{2}{N}} \cdot c(k) \sum_{j=0}^{N-1} x_j \cos\left(\frac{\pi k}{N}(j+0,5)\right), \quad (6.13)$$

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

где
$$c(k) = \begin{cases} 1/\sqrt{2} & \text{при } k = 0 \\ 1 & \text{при } k \neq 0 \end{cases}$$

Преобразование (6.13) можно представить в матричном виде как $Y = WX$, где

$$W = \left\{ w_{k,j} = \sqrt{\frac{2}{N}} c(k) \cos\left(\frac{\pi k}{N}(j+0,5)\right) \right\}_{k,j=0}^{N-1}.$$

Выясним «происхождение» формулы (6.13), для этого 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 (1.15) с коэффициентами, найденными по формуле (1.16), то свойства сходимости полученного ряда будут определяться характером функции $\check{f}(t)$ — продолжения с периодом T на всю числовую ось функции $f(t)$, см. рис. 6.2.

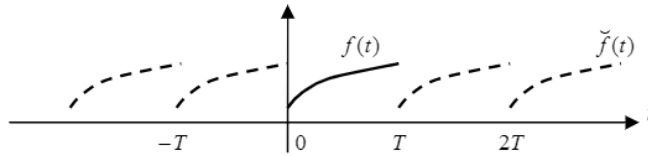


Рис. 6.2. Периодическое продолжение функции

Как отмечалось в разделе 2.6, с точностью до нормировочного множителя ДПФ представляет собой приближенное вычисление первых коэффициентов (1.16) ряда Фурье (1.15) с использованием квадратурной формулы прямоугольников:

$$c_k = \frac{1}{T} \int_0^T f(t) \exp\left(-i \frac{2\pi k}{T} t\right) dt \approx \frac{1}{T} \sum_{j=0}^{N-1} f(j\Delta t) \exp\left(-i \frac{2\pi k}{T} j\Delta t\right) \Delta t = \frac{1}{N} \sum_{j=0}^{N-1} x_j e^{-i \frac{2\pi k j}{N}},$$

где $N = T/\Delta t$, $x_j = f(j\Delta t)$. Это позволяет предположить, что чем выше скорость сходимости ряда (1.15) (т. е. чем быстрее убывают абсолютные величины коэффициентов Фурье), тем сильнее выражена локализация энергии сигнала вблизи первых коэффициентов вектора трансформант ДПФ.

Если функция $\check{f}(t)$ — непрерывная, то соответствующий ряд (1.15) будет сходиться быстрее по сравнению со случаем разрывной функции $\check{f}(t)$. Чтобы гарантировать непрерывность $\check{f}(t)$, сначала дополним функцию $f(t)$ четным образом на отрезке $t \in [-T, 0]$, а за-

тем полученную функцию (заданную уже на области $t \in [-T, T]$) продолжим на всю числовую ось с периодом $2T$, см. рис. 6.3.

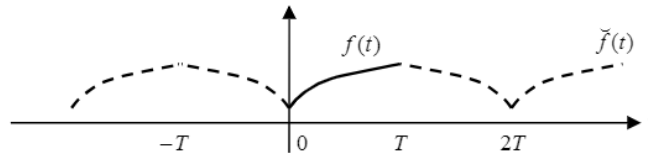


Рис. 6.3. Четное дополнение и периодическое продолжение функции

Найдем коэффициенты (1.16) ряда Фурье (1.15) функции $\tilde{f}(t)$ периода $2T$. В силу четности $\tilde{f}(t)$ имеем:

$$c_k = \frac{1}{2T} \int_{-T}^T \tilde{f}(t) e^{-i\frac{2\pi k}{2T}t} dt = \frac{1}{2T} \int_{-T}^T \tilde{f}(t) \cos\left(\frac{\pi k}{T}t\right) dt.$$

Окончательно получаем:

$$c_k = \frac{1}{T} \int_0^T f(t) \cos\left(\frac{\pi k}{T}t\right) dt. \tag{6.14}$$

Таким образом, коэффициенты Фурье являются вещественными и $c_k = c_{-k}$.

Подсчитаем приближенно значение интеграла (6.14) по квадратурной формуле прямоугольников, выбрав узлы в точках $t_j = (j+0,5)\Delta t$, $j = 0, 1, \dots, N-1$, где $N = T/\Delta t$. Обозначая $x_j = f((j+0,5)\Delta t)$, имеем:

$$c_k = \frac{1}{T} \int_0^T f(t) \cos\left(\frac{\pi k}{T}t\right) dt \approx \frac{1}{N} \sum_{j=0}^{N-1} x_j \cos\left(\frac{\pi k}{N}(j+0,5)\right).$$

Полученное выражение с точностью до нормировочного множителя совпадает с формулой ДКП (6.13). Оказывается, выбор узлов квадратурной формулы прямоугольников в точках $t_j = (j+0,5)\Delta t$ обеспечивает ортогональность ДКП.

Лемма 6.4. Матрица W ДКП (6.13) ортогональна.

◀ Необходимо показать, что матрица $U = WW^T$ является единичной. В соответствии с определением (6.13) получаем для произвольного элемента матрицы U :

$$\begin{aligned}
 u_{k,m} &= \sum_{j=0}^{N-1} w_{k,j} w_{m,j} = \frac{2}{N} c(k)c(m) \sum_{j=0}^{N-1} \cos\left(\frac{\pi k}{N}(j+0,5)\right) \cos\left(\frac{\pi m}{N}(j+0,5)\right) = \\
 &= \frac{1}{N} c(k)c(m) \left(\underbrace{\sum_{j=0}^{N-1} \cos\left(\frac{\pi(k+m)}{N}(j+0,5)\right)}_{s(k+m)} + \underbrace{\sum_{j=0}^{N-1} \cos\left(\frac{\pi(k-m)}{N}(j+0,5)\right)}_{s(k-m)} \right), \\
 & \quad k = 0, \dots, N-1; \quad m = 0, \dots, N-1.
 \end{aligned}$$

Найдем значение суммы $s(l) = \sum_{j=0}^{N-1} \cos(\pi l(j+0,5)/N) = s(-l)$ для $l = 0, \dots, 2N-2$. Очевидно, что $s(0) = N$, а при $l = 1, \dots, 2N-2$ находим по формуле для суммы геометрической прогрессии:

$$\begin{aligned}
 s(l) &= \operatorname{Re} \left(\sum_{j=0}^{N-1} e^{i \frac{\pi l}{N} (j+0,5)} \right) = \operatorname{Re} \left(e^{i \frac{\pi l}{2N}} \frac{e^{i \pi l} - 1}{e^{i \frac{\pi l}{N}} - 1} \right) = \operatorname{Re} \left(e^{i \frac{\pi l}{2}} \frac{e^{i \frac{\pi l}{2}} - e^{-i \frac{\pi l}{2}}}{e^{i \frac{\pi l}{2N}} - e^{-i \frac{\pi l}{2N}}} \right) = \\
 &= \operatorname{Re} \left(e^{i \frac{\pi l}{2}} \frac{\sin \frac{\pi l}{2}}{\sin \frac{\pi l}{2N}} \right) = \frac{\cos \frac{\pi l}{2} \sin \frac{\pi l}{2}}{\sin \frac{\pi l}{2N}} = \frac{\sin \pi l}{2 \sin \frac{\pi l}{2N}} = 0.
 \end{aligned}$$

Поэтому для $k = 0, \dots, N-1$, $m = 0, \dots, N-1$ получаем:

$$u_{k,m} = \frac{c(k)c(m)}{N} (s(k+m) + s(k-m)) = \begin{cases} 0, & \text{при } k \neq m \\ 2c(k)c(m), & \text{при } k = m = 0, \\ c(k)c(m), & \text{при } k = m \neq 0 \end{cases}$$

и, так как $c(j) = \begin{cases} 1/\sqrt{2} & \text{при } j = 0 \\ 1 & \text{при } j \neq 0 \end{cases}$, имеем $u_{k,m} = \begin{cases} 0, & \text{при } k \neq m \\ 1, & \text{при } k = m \end{cases}$ ▶

Следствие. Обратное для ДКП (6.13) преобразование (обратное ДКП, или ОДКП) определяется формулой

$$x_j = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} c(k) y_k \cos\left(\frac{\pi k}{N}(j+0,5)\right), \quad (6.15)$$

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

где $c(k) = \begin{cases} 1/\sqrt{2} & \text{при } k = 0 \\ 1 & \text{при } k \neq 0 \end{cases}$.

Упражнение. Вектор $Y = (1, 0, 1, 0)^T$ есть результат ДКП некоторого вектора X . Найдите вектор X .

Упражнение. Найдите ДКП каждой из приведенных ниже конечных последовательностей длины N .

$$\text{а) } x(n) = \tilde{\delta}(n) = \begin{cases} 1, & n=0 \\ 0, & n \neq 0 \end{cases}; \quad \text{б) } x(n) = \tilde{\delta}(n - n_0), \quad 0 \leq n_0 \leq N-1.$$

Формула (6.15) наглядно показывает, что при $k \neq 0$ коэффициент y_k ДКП (6.13) характеризует вклад базисного вектора — функции дискретного аргумента $w_k(j) = \sqrt{2/N} \cos\left(\frac{\pi k}{N}(j+0,5)\right)$, представляющей собой отсчеты гармонического колебания, в разложение дискретного сигнала — вектора X . Чем больше значение индекса k , тем выше частота гармонического колебания, а при $k = 0$ базисный вектор $w_0(j) = \sqrt{1/N}$ представляет собой константу, т.е. определяет постоянную составляющую разложения дискретного сигнала.

Положительный знак коэффициента корреляции соседних отсчетов ρ в модели (6.12) определяет тот факт, что соседние отсчеты вектора-сигнала X имеют статистически близкие значения. Чем ближе величина ρ к единице, тем существеннее зависимость соседних отсчетов и тем меньше вероятность резких изменений сигнала. Поэтому для вектора X , описываемого моделью (6.12) с $\rho > 0$, вклад высокочастотных составляющих ДКП-разложения (6.13) тем меньше, чем ближе к единице значение ρ . Расчеты по формуле (6.4) показывают, что при использовании для сигнала X модели (6.12) с $\rho \geq 0$ для дисперсий коэффициентов Y ДКП (6.13) имеют место неравенства: $\sigma_0^2 \geq \sigma_1^2 \geq \dots \geq \sigma_{N-1}^2$. Равенства дисперсий коэффициентов ДКП наблюдаются при $\rho = 0$, а чем ближе значение ρ к единице, тем быстрее убывание σ_k^2 с ростом k ; при $\rho = 1$ имеем $\sigma_0^2 = N\sigma_X^2$ и $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_{N-1}^2 = 0$.

Приведем расчетные значения дисперсий $\sigma_k^2 = D(Y_k)$ компонент вектора ДКП $Y = WX$ размерности $N = 32$, полученного для модели дискретного сигнала X , которая описывается марковским процессом первого порядка. На рис. 6.4 для удобства визуального восприятия дискретные значения $(x, y) = (k, \sigma_k^2)$ соединены линиями и отображены в виде непрерывных графиков, построенных для

ряда значений параметра ρ модели (6.12). При расчетах принималось, что $\sigma_x^2 = 1$. Как можно видеть из рис. 6.4, наибольшей дисперсией обладают те коэффициенты ДКП (6.13), индексы k которых близки к нулю, т.е. соответствуют весам низкочастотных компонент в представлении сигнала (6.15). Неравномерность распределения дисперсий по компонентам ДКП с преимущественным вкладом нижних частот выражена тем сильнее, чем ближе к единице коэффициент корреляции соседних отсчетов ρ в дискретном сигнале X .

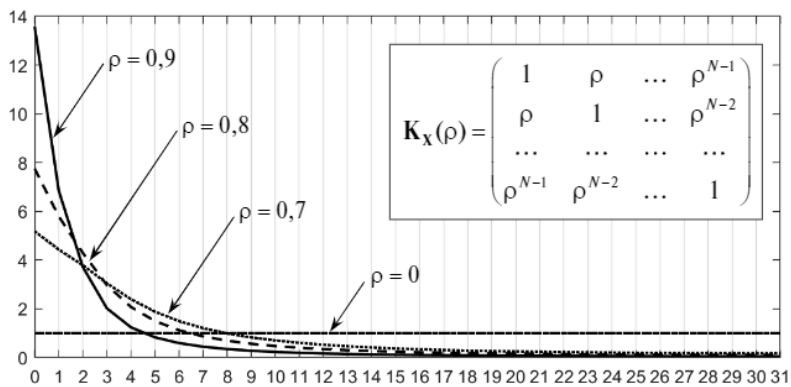


Рис. 6.4. Дисперсии компонент вектора Y , полученного в результате ДКП (6.13) размерности $N = 32$ из вектора X , моделируемого дискретным марковским процессом первого порядка с ковариационной матрицей вида (6.12)

Для иллюстрации декоррелирующих свойств ДКП приведем расчетные значения средней избыточной энтропии (6.8), соответствующие некоторым дискретным преобразованиям при обработке вектора X , который состоит из 16 компонент и описывается моделью марковского процесса первого порядка с ковариационной матрицей (6.12). На рис. 6.5 отражены результаты сравнения следующих ортогональных преобразований, имеющих быстрые алгоритмы вычислений: ВДПФ, ДКП, дискретные преобразования Уолша (ДПУ) и Хаара (ДПХ) (см. разделы 2.9, 2.10). Дискретное преобразование Хартли не отражено на графиках, поскольку оно показывает

практически те же характеристики, что и ВДПФ (значение средней избыточной энтропии (6.8) для преобразования Хартли по сравнению с ВДПФ меньше в среднем на 2%). Как видно из рис. 6.5, наиболее близкими характеристиками к оптимальному преобразованию Карунена — Лозва 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.

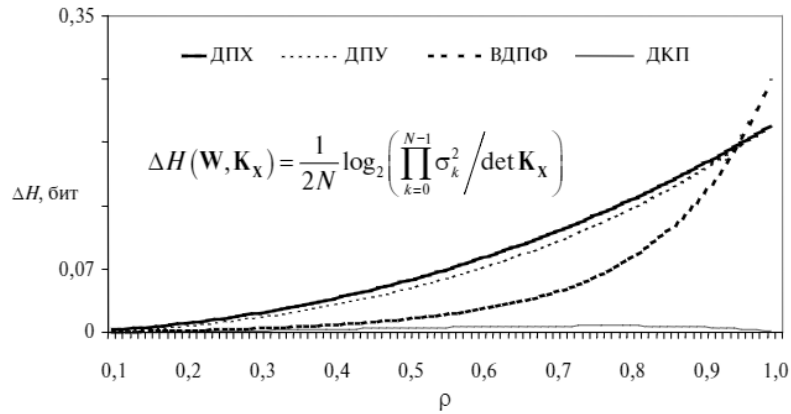


Рис. 6.5. Графики средней избыточной энтропии (6.8) при использовании некоторых ортогональных преобразований для обработки 16 компонентного вектора данных с ковариационной матрицей (6.12)

Упражнение. Для двухкомпонентного вектора \mathbf{X} , имеющего ковариационную матрицу $\mathbf{K}_x = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, где $\rho \in [-1; 1]$, найти ковариационную матрицу \mathbf{K}_y вектора $\mathbf{Y} = \mathbf{W}\mathbf{X}$, полученного из \mathbf{X} в результате ДКП.

Для ДКП разработаны эффективные быстрые алгоритмы вычислений, рассмотрение которых мы опустим.

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 $X = \{x_{j,m}\}$, ($j = 0, \dots, N-1$; $m = 0, \dots, M-1$) в матрицу Y той же размерности задано формулой

$$y_{k,l} = \sum_{j=0}^{N-1} w_{k,j}^{(N)} \sum_{m=0}^{M-1} w_{l,m}^{(M)} x_{j,m}, \quad (6.16)$$

$$k = 0, \dots, N-1, \quad l = 0, \dots, M-1,$$

где $W_K = \{w_{j,n}^{(K)}\}_{j,n=0}^{K-1}$ — матрица одномерного преобразования вида $U = W_K V$: $u_j = \sum_{n=0}^{K-1} w_{j,n}^{(K)} v_n$, $j = 0, \dots, K-1$. Тогда двумерное преобразование (6.16) называется дискретным *сепарабельным* преобразованием.

Из формулы (6.16) видно, что вычисление сепарабельного преобразования сводится к выполнению одномерных дискретных преобразований размерности M вдоль строк $j = 0, \dots, N-1$ матрицы X :

$$z_{j,l} = \sum_{m=0}^{M-1} w_{l,m}^{(M)} x_{j,m}, \quad l = 0, 1, \dots, M-1$$

(или $Z = XW_M^T$) и к последующим одномерным преобразованиям размерности N вдоль столбцов $l = 0, \dots, M-1$ полученной матрицы $Z = \{z_{j,l}\}$:

$$y_{k,l} = \sum_{j=0}^{N-1} w_{k,j}^{(N)} z_{j,l}, \quad k = 0, 1, \dots, N-1$$

(или $Y = W_N Z$). Причем сначала можно выполнить одномерные преобразования по столбцам матрицы, а потом по строкам, так как порядок суммирования в выражении (6.16) можно поменять. Из сказанного следует, что обратить двумерное сепарабельное преобразование можно, выполнив одномерные обратные преобразования сначала вдоль столбцов, а потом вдоль строк матрицы Y (или наоборот).

Запишем сепарабельное преобразование (6.16) в эквивалентном матричном виде: $\mathbf{Y} = \mathbf{W}_N \mathbf{X} \mathbf{W}_M^T$. Тогда если соответствующее одномерное преобразование является ортогональным, то обратное к (6.16) преобразование записывается как $\mathbf{X} = \mathbf{W}_N^T \mathbf{Y} \mathbf{W}_M$, а обращение формулы (6.16) принимает вид (убедитесь!):

$$x_{j,m} = \sum_{k=0}^{N-1} w_{k,j}^{(N)} \sum_{l=0}^{M-1} w_{l,m}^{(M)} y_{k,l}, \quad (6.17)$$

$$j = 0, \dots, N-1, \quad m = 0, \dots, M-1.$$

В соответствии с формулами (6.16), (6.17) и данными ранее определениями одномерных преобразований вводятся двумерные дискретные преобразования Фурье, Уолша, Хаара, Хартли. В частности, прямое и обратное двумерные ДКП (см. (6.13), (6.15)) задаются соответственно следующими формулами:

$$y_{k,l} = \frac{2}{\sqrt{MN}} c(k)c(l) \sum_{j=0}^{N-1} \cos\left(\frac{\pi k}{N}\left(j + \frac{1}{2}\right)\right) \sum_{m=0}^{M-1} x_{j,m} \cos\left(\frac{\pi l}{M}\left(m + \frac{1}{2}\right)\right), \quad (6.18)$$

$$k = 0, 1, \dots, N-1, \quad l = 0, 1, \dots, M-1;$$

$$x_{j,m} = \frac{2}{\sqrt{MN}} \sum_{k=0}^{N-1} c(k) \cos\left(\frac{\pi k}{N}\left(j + \frac{1}{2}\right)\right) \sum_{l=0}^{M-1} c(l) y_{k,l} \cos\left(\frac{\pi l}{M}\left(m + \frac{1}{2}\right)\right), \quad (6.19)$$

$$j = 0, 1, \dots, N-1, \quad m = 0, 1, \dots, M-1,$$

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 предполагает 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 *полутонного* цифрового изображения, которое можно представить в виде матрицы, элементами кото-

рой являются значения яркости точек растра. 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, декоррелирующие 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. (6.12)). Фрагменты обрабатываются последовательно (очередной обрабатываемый блок является соседним к предыдущему), общая схема обработки (рис. 6.1) принимает вид, приведенный на рис. 6.6, и включает в себя четыре основных шага.



Рис. 6.6. Схема сжатия изображений по стандарту JPEG

- Шаг 1.** Вычисление двумерного ДКП (6.18) блока размерности 8×8 пикселей.
- 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 $Q = \{q_{k,l}\}_{k,l=0}^7$. 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.



появления ненулевого элемента¹, а в полубайте L_2 записывается число B , которое равно числу двоичных разрядов, необходимых для представления элемента S , см. таблицу.

| 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 |
| ... | ... |
| $\{-2^{15} + 1, \dots, -2^{14}\} \cup \{2^{14}, \dots, 2^{15} - 1\}$ | 15 |

Во второй поток заносится B бит двоичного кода элемента S . Таким образом, первый поток данных — байтовый, второй — битовый. После записи последнего ненулевого элемента из «змейки» в байтовый 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.

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 таблицы кодов).

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 дискретов равномерного квантования $q_{k,l}$), тем большая ошибка возникнет в изображении после восстановления, но тем выше степень сжатия данных.

¹ Фактически, используется вариант метода КДС, описанного в разделе 5.4.

При квантовании коэффициенты ДКП, соответствующие высокочастотным компонентам разложения, квантуются более грубо, поскольку человеческий глаз менее чувствителен к ним при визуальном восприятии. Это наглядно иллюстрирует одна из рекомендуемых стандартных матриц квантования — матрица Пеннебакера:

$$Q = \begin{pmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{pmatrix}.$$

Каждому шагу в схеме сжатия (рис. 6.6) соответствует шаг в схеме восстановления изображения (рис. 6.8), 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 $Q = \{q_{k,l}\}_{k,l=0}^7$.



Рис. 6.8. 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. В приведенной матричной формуле предполагается, что каждая исходная компонента в представлениях (R, G, B) и (Y, Cr, Cb) является целым числом из диапазона $[0...255]$, поэтому после выполнения умножений и сложений результаты округляются до ближайшего целого.

По сравнению с яркостной составляющей Y , чувствительность человеческого глаза к компонентам Cr , Cb значительно ниже, что учитывается при построении методов компрессии цветных изображений. Так, для цветоразностных компонент в схеме JPEG применяется дополнительное прореживание массивов пикселей и затем используется обработка по той же общей схеме, которая была рассмотрена выше для яркостной компоненты, но с более грубым квантованием компонент ДКП. При восстановлении изображений используется обратное преобразование цветового пространства (с округлением результатов до целых чисел):

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1,4020 & 0 \\ 1 & -0,7141 & -0,3441 \\ 1 & 0 & 1,772 \end{pmatrix} \begin{pmatrix} Y \\ Cr - 128 \\ Cb - 128 \end{pmatrix}.$$

6.6. Дискретное псевдокосинусное преобразование

Рассмотренное выше ДКП нашло широкое применение на практике, главным образом для целей сжатия дискретных сигналов, имеющих статистику марковского процесса первого порядка. Однако ДКП, несмотря на наличие специальных быстрых алгоритмов, требует выполнения относительно большого объема вычислений, в том числе операций умножения. Рассмотрим возможность уменьшения необходимых вычислений.

Запишем ДКП (6.13) и ОДКП (6.15) в матричном виде соответственно:

$$\mathbf{Y} = \mathbf{W}_N \mathbf{X}, \quad \mathbf{X} = \mathbf{W}_N^T \mathbf{Y}$$

(здесь нижний индекс матриц указывает на размерность преобразований).

Для случаев $N = 2, 3$ матрицы ДКП имеют вид:

$$\mathbf{W}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad (6.20)$$

$$\begin{aligned} \mathbf{W}_3 &= \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -2 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}. \quad (6.21) \end{aligned}$$

Матрица 4-точечного ДКП ($N = 4$)

$$\mathbf{W}_4 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} \cos \frac{\pi}{8} & \frac{1}{\sqrt{2}} \sin \frac{\pi}{8} & -\frac{1}{\sqrt{2}} \sin \frac{\pi}{8} & -\frac{1}{\sqrt{2}} \cos \frac{\pi}{8} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} \sin \frac{\pi}{8} & -\frac{1}{\sqrt{2}} \cos \frac{\pi}{8} & \frac{1}{\sqrt{2}} \cos \frac{\pi}{8} & -\frac{1}{\sqrt{2}} \sin \frac{\pi}{8} \end{bmatrix} \quad (6.22)$$

с использованием идеи быстрого алгоритма ДКП (см., например, [59]) может быть представлена в виде произведения матриц:

$$\mathbf{W}_4 = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & \sin \frac{\pi}{8} & \cos \frac{\pi}{8} \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -\cos \frac{\pi}{8} & \sin \frac{\pi}{8} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}. \quad (6.23)$$

В последней записи заменим $\sin(\pi/8) \approx 0,38$ на $1/\sqrt{5} \approx 0,45$, а $\cos(\pi/8) \approx 0,92$ — на $2/\sqrt{5} \approx 0,89$. Тогда из \mathbf{W}_4 (6.23) получим матрицу

$$\tilde{\mathbf{W}}_4 = \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/\sqrt{10} & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/\sqrt{10} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}, \quad (6.24)$$

т. е.

$$\tilde{\mathbf{W}}_4 = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{2}{\sqrt{10}} & \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & -\frac{2}{\sqrt{10}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{10}} & -\frac{2}{\sqrt{10}} & \frac{2}{\sqrt{10}} & -\frac{1}{\sqrt{10}} \end{bmatrix}. \quad (6.25)$$

Матрицу $\tilde{\mathbf{W}}_4$ назовем матрицей дискретного псевдокосинусного преобразования (ДПКП) для $N=4$ отсчетов. Для $N=2, 3$ матрицы ДПКП определим так: $\tilde{\mathbf{W}}_2 = \mathbf{W}_2$ (6.20), $\tilde{\mathbf{W}}_3 = \mathbf{W}_3$ (6.21).

Заметим, что матрицы ДПКП ($N=2, 3, 4$) можно представить в виде (см. (6.20), (6.21), (6.25)):

$$\tilde{\mathbf{W}}_N = \mathbf{D}_N \mathbf{C}_N,$$

где \mathbf{D}_N — диагональная матрица нормирующих множителей, \mathbf{C}_N — матрица, состоящая только из чисел $\{-2, -1, 0, 1, 2\}$. Поэтому для произведения матрицы $\tilde{\mathbf{W}}_N$ ($N=2, 3, 4$) на вектор \mathbf{X} имеем:

$$\tilde{\mathbf{W}}_N \mathbf{X} = \mathbf{D}_N \mathbf{C}_N \mathbf{X}.$$

Отметим, что реализация умножения матрицы \mathbf{C}_N на вектор \mathbf{X} сводится только к операциям сложения и вычитания чисел (умножение на 2 эквивалентно сложению: $2x = x + x$), а необходимость в операциях умножения чисел возникает только при окончательной нормировке (умножении на матрицу \mathbf{D}_N). Так, умножение на матрицу $\tilde{\mathbf{W}}_2 = \mathbf{W}_2$ (6.20) требует две операции сложения, умножение на матрицу $\tilde{\mathbf{W}}_3 = \mathbf{W}_3$ (6.21) — пять операций сложения, на $\tilde{\mathbf{W}}_4$ (6.25) — 10 сложений (вычитаний) (см. (6.24)).

Легко убедиться, что матрицы ДПКП $\tilde{\mathbf{W}}_2, \tilde{\mathbf{W}}_3, \tilde{\mathbf{W}}_4$ — ортогональны, поэтому $\tilde{\mathbf{W}}_2^{-1} = \tilde{\mathbf{W}}_2^T, \tilde{\mathbf{W}}_3^{-1} = \tilde{\mathbf{W}}_3^T, \tilde{\mathbf{W}}_4^{-1} = \tilde{\mathbf{W}}_4^T$. Реализация обратного ДПКП, $\mathbf{X} = \tilde{\mathbf{W}}_N^{-1} \mathbf{Y} = \mathbf{C}_N^T \mathbf{D}_N \mathbf{Y}$ ($N = 2, 3, 4$), вновь не требует операций умножения чисел, за исключением первоначального масштабирования компонент вектора \mathbf{Y} (умножения на диагональную матрицу \mathbf{D}_N). Матрицы $\tilde{\mathbf{W}}_2, \tilde{\mathbf{W}}_3, \tilde{\mathbf{W}}_4$ ($\tilde{\mathbf{W}}_2^{-1}, \tilde{\mathbf{W}}_3^{-1}, \tilde{\mathbf{W}}_4^{-1}$) будем называть *элементарными матрицами ДПКП* (обратного ДПКП).

Определение. Тензорным (прямым) произведением матриц $\mathbf{D} = (d_{l,m})$ ($l = 0, \dots, L-1; m = 0, \dots, M-1$) и $\mathbf{G} = (g_{k,j})$ ($k = 0, \dots, K-1; j = 0, \dots, J-1$) назовем матрицу

$$\mathbf{S} = (s_{\alpha,\beta}) = \mathbf{D} \otimes \mathbf{G} = \begin{pmatrix} d_{0,0} \mathbf{G} & \dots & d_{0,M-1} \mathbf{G} \\ \dots & \dots & \dots \\ d_{L-1,0} \mathbf{G} & \dots & d_{L-1,M-1} \mathbf{G} \end{pmatrix},$$

$$\alpha = 0, 1, \dots, LK-1; \beta = 0, 1, \dots, MJ-1.$$

Упражнение. Пусть квадратная матрица $\mathbf{G} = \{g_{k,m}\}_{k,m=0}^{M-1}$ состоит из ненулевых элементов, а \mathbf{E}_N и \mathbf{E}_L — единичные матрицы размерностей $N \times N$ и $L \times L$ соответственно. Убедитесь, что матрица $\mathbf{A} = \mathbf{E}_N \otimes \mathbf{G} \otimes \mathbf{E}_L$ имеет размерность $(NML) \times (NML)$, причем в каждой строке (столбце) матрицы \mathbf{A} содержится ровно M ненулевых элементов, соответствующих некоторой строке (столбцу) матрицы \mathbf{G} .

Теорема 6.1. Пусть матрица \mathbf{K} размерности $N \times N$, где $N = N_1 \cdot N_2 \cdot \dots \cdot N_n$, представима в виде: $\mathbf{K} = \mathbf{K}_{N_1} \otimes \mathbf{K}_{N_2} \otimes \dots \otimes \mathbf{K}_{N_{n-1}} \otimes \mathbf{K}_{N_n}$, где \mathbf{K}_{N_j} — квадратная матрица размерности $N_j \times N_j$ ($j = 1, \dots, n$). Тогда справедливо представление: $\mathbf{K} = \mathbf{P}_N^{(1)} \mathbf{P}_N^{(2)} \dots \mathbf{P}_N^{(n)}$, где для $j = 1, \dots, n$

квадратные матрицы $\mathbf{P}_N^{(j)}$ размерности $N \times N$ имеют в каждой строке и каждом столбце по N_j ненулевых элементов, причем

$$\mathbf{P}_N^{(j)} = \mathbf{E}_{N_1} \otimes \mathbf{E}_{N_2} \otimes \dots \otimes \mathbf{E}_{N_{j-1}} \otimes \mathbf{K}_{N_j} \otimes \mathbf{E}_{N_{j+1}} \otimes \dots \otimes \mathbf{E}_{N_n} = \mathbf{E}_{N_1 N_2 \dots N_{j-1}} \otimes \mathbf{K}_{N_j} \otimes \mathbf{E}_{N_{j+1} \dots N_n}$$

(здесь \mathbf{E}_K — единичная матрица размерности $K \times K$).

Опустим несложное доказательство теоремы, которое носит технический характер [4, с. 84]. Из теоремы следует, что если некоторая матрица \mathbf{K} получена как тензорное произведение n матриц размерностей $N_1 \times N_1, N_2 \times N_2, \dots, N_n \times N_n$, то ее можно представить в виде произведения слабозаполненных матриц. Это произведение определяет быстрый алгоритм матричного умножения $\mathbf{Y} = \mathbf{KX}$, где \mathbf{X}, \mathbf{Y} — векторы размерности $N = N_1 \cdot N_2 \cdot \dots \cdot N_n$. В результате вместо $N^2 = (N_1 N_2 \dots N_n)^2$ умножений со сложениями, которые необходимы при непосредственной реализации операции $\mathbf{Y} = \mathbf{KX}$, при последовательном умножении вектора \mathbf{X} на матрицы-сомножители $\mathbf{P}_N^{(j)}$ ($j = n, \dots, 1$), $\mathbf{Y} = \mathbf{P}_N^{(1)} \mathbf{P}_N^{(2)} \dots \mathbf{P}_N^{(n)} \mathbf{X}$, необходимо выполнить $N(N_1 + N_2 + \dots + N_n)$ умножений со сложениями (убедитесь).

Для того чтобы получить дискретное преобразование, гарантированно имеющее на основании теоремы 6.1 быстрый алгоритм вычисления, матрицы ДПКП большей размерности ($N > 4$) будем строить как тензорные произведения элементарных матриц ДПКП следующим образом [48]:

$$\begin{aligned} \tilde{\mathbf{W}}_6 &= \tilde{\mathbf{W}}_2 \otimes \tilde{\mathbf{W}}_3, & \tilde{\mathbf{W}}_8 &= \tilde{\mathbf{W}}_2 \otimes \tilde{\mathbf{W}}_4, & \tilde{\mathbf{W}}_9 &= \tilde{\mathbf{W}}_3 \otimes \tilde{\mathbf{W}}_3, & \tilde{\mathbf{W}}_{12} &= \tilde{\mathbf{W}}_3 \otimes \tilde{\mathbf{W}}_4, \\ \tilde{\mathbf{W}}_{16} &= \tilde{\mathbf{W}}_4 \otimes \tilde{\mathbf{W}}_4, & \tilde{\mathbf{W}}_{18} &= \tilde{\mathbf{W}}_2 \otimes \tilde{\mathbf{W}}_3 \otimes \tilde{\mathbf{W}}_3 & \text{и т. д.} \end{aligned}$$

Так определенное ДПКП может иметь только такую размерность N , которая представима в виде произведения чисел 2, 3, 4. При этом располагать матрицы $\tilde{\mathbf{W}}_2, \tilde{\mathbf{W}}_3, \tilde{\mathbf{W}}_4$ при составлении тензорного произведения следует в порядке увеличения нижнего индекса, стремясь использовать меньшее число элементарных матриц большей размерности [48]. Матрицы ДПКП размерности $N > 4$ также ортогональны и $\tilde{\mathbf{W}}_N^{-1} = \tilde{\mathbf{W}}_N^T$.

Построим, например, матрицу ДПКП для $N = 8$. Имеем: $\tilde{\mathbf{W}}_8 = \tilde{\mathbf{W}}_2 \otimes \tilde{\mathbf{W}}_4$ (не $\tilde{\mathbf{W}}_4 \otimes \tilde{\mathbf{W}}_2$ и не $\tilde{\mathbf{W}}_2 \otimes \tilde{\mathbf{W}}_2 \otimes \tilde{\mathbf{W}}_2$):

$$\tilde{\mathbf{W}}_8 = \begin{bmatrix} \frac{1}{\sqrt{2}} \tilde{\mathbf{W}}_4 & \frac{1}{\sqrt{2}} \tilde{\mathbf{W}}_4 \\ \frac{1}{\sqrt{2}} \tilde{\mathbf{W}}_4 & -\frac{1}{\sqrt{2}} \tilde{\mathbf{W}}_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \frac{2}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{-2}{\sqrt{20}} & \frac{2}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{-2}{\sqrt{20}} \\ \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \frac{1}{\sqrt{20}} & \frac{-2}{\sqrt{20}} & \frac{2}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{-2}{\sqrt{20}} & \frac{2}{\sqrt{20}} & \frac{-1}{\sqrt{20}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} \\ \frac{2}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{-2}{\sqrt{20}} & \frac{-2}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{2}{\sqrt{20}} \\ \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} \\ \frac{1}{\sqrt{20}} & \frac{-2}{\sqrt{20}} & \frac{2}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{-1}{\sqrt{20}} & \frac{2}{\sqrt{20}} & \frac{-2}{\sqrt{20}} & \frac{1}{\sqrt{20}} \end{bmatrix}$$

На основании свойства тензорного произведения

$$(\mathbf{A}_1 \mathbf{B}_1) \otimes (\mathbf{A}_2 \mathbf{B}_2) = (\mathbf{A}_1 \otimes \mathbf{A}_2)(\mathbf{B}_1 \otimes \mathbf{B}_2)$$

несложно доказать по индукции, что для числа N , представимого в виде произведения $N = N_1 \cdot N_2 \cdot \dots \cdot N_n$, где $N_i \in \{2, 3, 4\}$, справедливо:

$$\tilde{\mathbf{W}}_N = \mathbf{D}_N \mathbf{C}_N,$$

где диагональная матрица $\mathbf{D}_N = \mathbf{D}_{N_1} \otimes \mathbf{D}_{N_2} \otimes \dots \otimes \mathbf{D}_{N_n}$, а матрица $\mathbf{C}_N = \mathbf{C}_{N_1} \otimes \mathbf{C}_{N_2} \otimes \dots \otimes \mathbf{C}_{N_n}$. Структуры матриц $\mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4, \mathbf{D}_2, \mathbf{D}_3, \mathbf{D}_4$ отражены в разложениях (6.20), (6.21), (6.25). Таким образом, реализация ДПКП $\mathbf{Y} = \tilde{\mathbf{W}}_N \mathbf{X} = \mathbf{D}_N \mathbf{C}_N \mathbf{X}$ сводится к умножению матрицы \mathbf{C}_N на вектор \mathbf{X} , $\tilde{\mathbf{Y}} = \mathbf{C}_N \mathbf{X}$, и последующей нормировке полученного вектора $\tilde{\mathbf{Y}}$: $\mathbf{Y} = \mathbf{D}_N \tilde{\mathbf{Y}}$.

Для вычисления ДПКП удобно использовать быстрые алгоритмы, основанные на факторизованном представлении для матрицы \mathbf{C}_N в виде произведения слабозаполненных матриц (см. теорему 6.1):

$$\mathbf{C}_N = \mathbf{T}_N^{(1)} \mathbf{T}_N^{(2)} \dots \mathbf{T}_N^{(n)},$$



где $\mathbf{T}_N^{(i)} = \mathbf{E}_{N_1 N_2 \dots N_{i-1}} \otimes \mathbf{C}_{N_i} \otimes \mathbf{E}_{N_{i+1} \dots N_n}$, \mathbf{E}_K — единичная матрица размерности $K \times K$. Быстрые алгоритмы обратного ДПКП строятся аналогично, так как

$$\tilde{\mathbf{W}}_N^{-1} = \tilde{\mathbf{W}}_N^T = \mathbf{C}_N^T \mathbf{D}_N = (\mathbf{T}_N^{(n)})^T (\mathbf{T}_N^{(n-1)})^T \dots (\mathbf{T}_N^{(1)})^T \mathbf{D}_N.$$

Пример 6.3. Найти факторизованное представление для матрицы ДПКП $\tilde{\mathbf{W}}_8$.

◀ Имеем $N_1 = 2$, $N_2 = 4$, $\tilde{\mathbf{W}}_8 = \tilde{\mathbf{W}}_2 \otimes \tilde{\mathbf{W}}_4 = \mathbf{D}_8 \mathbf{C}_8$; из (6.20) и (6.25):

$$\mathbf{D}_8 = \mathbf{D}_2 \otimes \mathbf{D}_4 = \begin{bmatrix} \frac{1}{\sqrt{2}} \mathbf{D}_4 & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{2}} \mathbf{D}_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2\sqrt{2}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2\sqrt{5}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2\sqrt{5}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2\sqrt{5}} \end{bmatrix},$$

$$\mathbf{C}_8 = \mathbf{C}_2 \otimes \mathbf{C}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 & 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 & 1 & -2 & 2 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 2 & 1 & -1 & -2 & -2 & -1 & 1 & 2 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -2 & 2 & -1 & -1 & 2 & -2 & 1 \end{bmatrix} = \mathbf{T}_8^{(1)} \mathbf{T}_8^{(2)},$$

$$\text{где } \mathbf{T}_8^{(1)} = \mathbf{C}_2 \otimes \mathbf{E}_4 = \begin{bmatrix} \mathbf{E}_4 & \mathbf{E}_4 \\ \mathbf{E}_4 & -\mathbf{E}_4 \end{bmatrix} = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{array} \right],$$

$$\mathbf{T}_8^{(2)} = \mathbf{E}_2 \otimes \mathbf{C}_4 = \left[\begin{array}{cccc|cccc} \mathbf{C}_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_4 \end{array} \right] = \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & -1 & -2 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 2 & -1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 & 2 & -1 \end{array} \right].$$

Умножение матрицы $\mathbf{T}_8^{(1)}$ на вектор требует восемь сложений (вычитаний) чисел. Умножение матрицы $\mathbf{T}_8^{(2)}$ на вектор эквивалентно двум матричным умножениям на \mathbf{C}_4 , каждое из которых сводится к 10 сложениям (вычитаниям) чисел, так как (см. (6.24), (6.25))

$$\mathbf{C}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}.$$

Таким образом, вычисление 8-точечного ДПКП по быстрому алгоритму: $\mathbf{Y} = \tilde{\mathbf{W}}_8 \mathbf{X} = \mathbf{D}_8 \mathbf{T}_8^{(1)} \mathbf{T}_8^{(2)} \mathbf{X}$ требует 28 операций типа сложения, а умножения чисел (не считая нормировки) не нужны. Аналогичный объем вычислений необходим и для обратного 8-точечного ДПКП. ►

Рассмотрим теперь некоторые характеристики ДПКП, связанные с его применением для сжатия сигналов, описываемых моделью марковского процесса первого порядка. Положив для ковариационной матрицы (6.12) $\rho = 0,9$, для вектора трансформант некоторых

ортогональных преобразований получим расчетные значения средней избыточной энтропии (6.8) с логарифмом по основанию 2, представленные в табл. 6.1.

Таблица 6.1. Средняя избыточная энтропия (в битах) для ряда ортогональных преобразований и модели сигнала с ковариационной матрицей (6.12) при $\rho = 0,9$

| Размерность | ДПХ | ДПУ | ДПКП | ВДПФ | ДКП |
|-------------|-------|-------|-------|-------|-------|
| $N = 8$ | 0,151 | 0,134 | 0,078 | 0,158 | 0,006 |
| $N = 16$ | 0,192 | 0,187 | 0,113 | 0,159 | 0,006 |

Все преобразования, приведенные в таблице, имеют быстрые алгоритмы вычислений. Первая группа — это дискретное преобразование Хаара (ДПХ), дискретное преобразование Уолша (ДПУ), ДПКП. Вторая группа — ВДПФ и ДКП. Преобразования первой группы не требуют для своего вычисления операций умножения чисел (за исключением, возможно, окончательной нормировки). Например, при размерности $N = 8$ отсчетов вычисление ДПХ сводится к 14 операциям типа сложения, ДПУ — к 24 сложениям, ДПКП — к 28 сложениям (см. пример 6.3). Преобразования второй группы, помимо сложений и вычитаний, принципиально требуют выполнения операций умножения на постоянные (причем, вообще говоря, иррациональные) множители. Таким образом, среди «сверхбыстрых» преобразований первой группы дискретное псевдокосинусное преобразование потенциально наиболее эффективно для декорреляции данных, описываемых моделью дискретного марковского процесса.

Отметим, что необходимая при вычислении ДПКП и обратного ДПКП нормировка, т.е. умножение на матрицу \mathbf{D}_N , для часто используемой на практике схемы сжатия данных (рис. 6.1) не влечет никакого усложнения вычислений. Использование для приведенной схемы сжатия-восстановления ДПКП, матрица которого представима в виде $\tilde{\mathbf{W}}_N = \mathbf{D}_N \mathbf{C}_N$, позволяет не учитывать операции умножения на нормирующие множители — элементы диагональной матрицы \mathbf{D}_N , так как нормировка может быть объединена с этапом скалярного квантования (при сжатии данных) $\tilde{y} = \text{round}(y/q)$ и этапом деквантования (при восстановлении данных) $\hat{y} = \tilde{y} \cdot q$.

Для этого шаг квантователя должен быть модифицирован и выбран для каждой компоненты y_j вектора трансформант $Y = C_N X$ по правилу: $q_j = q/d_{jj}$, где d_{jj} — j -й диагональный элемент нормировочной матрицы D_N , $j = 0, 1, \dots, N-1$.

Таким образом, дискретное псевдокосинусное преобразование может быть реализовано в практической схеме сжатия данных только при помощи операций типа сложения и вычитания; имеются быстрые алгоритмы реализации ДПКП. Для данных, представляющих собой реализацию марковского процесса первого порядка, ДПКП обладает большей эффективностью декорреляции данных по сравнению с другими быстрыми преобразованиями, реализация которых также сводится только к операциям сложения и вычитания. Поэтому для приложений, которые предполагают использование ограниченных вычислительных ресурсов, обоснованной является постановка вопроса об использовании ДПКП для сжатия дискретных сигналов. В частности, вариант двумерного ДПКП используется в стандарте сжатия видеозображений H.264 и H.265.

6.7. Оптимизация алгоритмов сжатия данных с потерями

В разделе 6.5 мы рассмотрели пример алгоритма сжатия данных с потерями, когда восстановленный дискретный сигнал \hat{X} (представленный вектором или матрицей) отличается от исходного X . Алгоритмы сжатия данных с потерями характеризуются прежде всего двумя следующими параметрами:

- вносимыми при кодировании искажениями D (т.е. ошибкой, описываемой, например, квадратом евклидовой нормы $D = \|X - \hat{X}\|_E^2$),
- битовыми затратами R на кодирование сигнала.

Для некоторого заданного алгоритма компрессии-декомпрессии F , результатом работы которого является восстановленный сигнал $\hat{X} = F(X)$, могут быть использованы различные параметры обработки \mathbf{u} , определяющие вносимые в сигнал искажения $D = D(X, \mathbf{u}) \geq 0$ и битовые затраты $R = R(X, \mathbf{u}) \geq 0$. Такими регулируемые параметрами \mathbf{u} в случае рассмотренного в разделе 6.5

алгоритма JPEG могут быть, например, значения компонент матрицы квантования \mathbf{Q} . Рассмотрим задачу выбора оптимальных параметров алгоритмов сжатия данных с потерями, которую сформулируем следующим образом.

Пусть для некоторого алгоритма сжатия F , варьируемые параметры которого составляют N -компонентный вектор $\mathbf{u} = (u_1, \dots, u_N)$, и заданного дискретного сигнала (вектора или матрицы) \mathbf{X} определены функции искажений $D = D(\mathbf{u})$ и битовых затрат $R = R(\mathbf{u})$. Необходимо найти такие параметры алгоритма кодирования $\mathbf{u}^* = (u_1^*, \dots, u_N^*)$, которые дают минимальные искажения сигнала при заданном *битовом бюджете* на кодирование R_b , т. е.

$$\begin{cases} D(\mathbf{u}^*) = \min_{\mathbf{u}} D(\mathbf{u}) \\ R(\mathbf{u}^*) = R_b \end{cases} \quad (6.26)$$

Поиск точки \mathbf{u}^* представляет собой классическую задачу условной минимизации. Если задача (6.26) имеет решение, то найдется такое число λ , что \mathbf{u}^* будет являться безусловным минимумом соответствующей функции Лагранжа:

$$J(\mathbf{u}^*) = D(\mathbf{u}^*) + \lambda R(\mathbf{u}^*) = \min_{\mathbf{u}} [J(\mathbf{u}) = D(\mathbf{u}) + \lambda R(\mathbf{u})]. \quad (6.27)$$

Если функции $D = D(\mathbf{u})$ и $R = R(\mathbf{u})$ заданы аналитически и дифференцируемы, то для поиска числа λ и точки \mathbf{u}^* может быть использован метод неопределенных множителей Лагранжа.

Если задача (6.26) имеет единственное решение $\mathbf{u}^* = \mathbf{u}^*(R_b)$ для каждого значения $R = R_b$ из некоторого отрезка $R \in [R_{\min}; R_{\max}]$, то в области определения $[R_{\min}; R_{\max}]$ можно задать функцию $D = D(\mathbf{u}^*(R)) = D(R)$. Каждая точка графика $D(R)$ соответствует минимуму функции Лагранжа (6.27) для какого-то числа λ ; для дифференцируемых функций $D(\mathbf{u})$ и $R(\mathbf{u})$ имеем $dJ(\mathbf{u}^*) = dD(\mathbf{u}^*) + \lambda dR(\mathbf{u}^*) = 0$, откуда $dD/dR = -\lambda$. Типичный вид зависимости $D(R)$ и геометрический смысл параметра λ функции Лагранжа (6.27) представлены на рис. 6.9.

Из рис. 6.9 видно, что если функция $D(R)$ является дифференцируемой и выпуклой, то величина λ может использоваться для параметрического задания этой функции: непрерывному уменьшению от значения $\lambda_{\max} = \lambda(R_{\min})$ до $\lambda_{\min} = \lambda(R_{\max})$ соответствует перемещение по графику от точки $D(R_{\min})$ до $D(R_{\max})$. При этом $[\lambda_{\min}; \lambda_{\max}] \subset [0; +\infty)$.

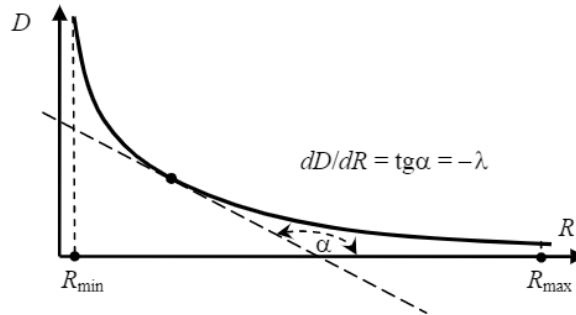


Рис. 6.9. Геометрический смысл параметра λ в функции Лагранжа (6.27)

Таким образом, при обработке дискретного сигнала с оптимальными параметрами алгоритма сжатия увеличению значения λ в (6.27) соответствует уменьшение битовых затрат R и увеличение искажений D . Действительно, параметр λ задает баланс между искажениями и битовыми затратами в функции Лагранжа (6.27): увеличение λ повышает вес слагаемого $R(\mathbf{u})$ и влечет более существенную минимизацию битовых затрат (за счет неизбежного увеличения искажений). Поэтому для решения задачи (6.26) на практике могут использоваться различные методы направленного поиска. Например, выбрав каким-то образом значение λ и найдя соответствующее решение (6.27) \mathbf{u}^* , определяем соответствующие битовые затраты $R(\mathbf{u}^*)$ и на следующей итерации увеличиваем λ , если $R(\mathbf{u}^*) > R_b$, или уменьшаем λ , если $R(\mathbf{u}^*) < R_b$. Однако если $D(R)$ не является

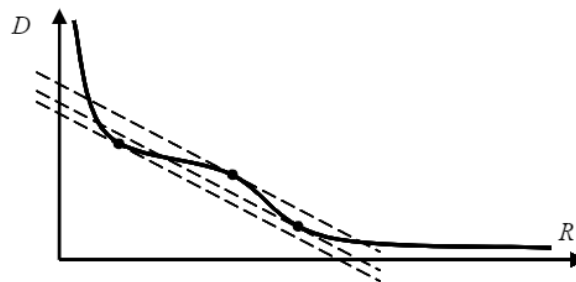


Рис. 6.10. Пример графика $D(R)$ для функции Лагранжа $J(\mathbf{u}) = D(\mathbf{u}) + \lambda R(\mathbf{u})$ с локальными экстремумами

выпуклой функцией, то такой итерационный поиск решения задачи (6.26) может не дать результат, так как одному значению параметра λ могут соответствовать различные точки графика $D(R)$ (см. рис. 6.10), определяющие локальные экстремумы функции Лагранжа $J(\mathbf{u}) = D(\mathbf{u}) + \lambda R(\mathbf{u})$.

Пример 6.4. Пусть имеющая стандартное нормальное распределение случайная величина X обрабатывается при помощи алгоритма F , который состоит в квантовании X по правилу:

$$\hat{X} = \begin{cases} -m, & \text{если } X < -T \\ 0, & \text{если } -T \leq X \leq T, \text{ где } 0 < T < m, \\ m, & \text{если } T < X \end{cases}$$

и последующем статистическом кодировании дискретной величины \hat{X} . Характеризуя ошибку кодирования по величине $D = M(|X - \hat{X}|^2)$, а битовые затраты — по значению двоичной энтропии $R = H(\hat{X})$, построить график зависимости $D(R)$, соответствующий оптимальному выбору параметров (m, T) алгоритма F .

◀ Случайная величина \hat{X} имеет три возможных значения. Обозначим:

$$p_0 = P\{\hat{X} = 0\} = \frac{1}{\sqrt{2\pi}} \int_{-T}^T \exp(-t^2/2) dt = \frac{2}{\sqrt{2\pi}} \int_0^T \exp(-t^2/2) dt = 2\Phi(T),$$

$$p_1 = P\{\hat{X} = -m\} = P\{\hat{X} = m\} = \frac{1}{\sqrt{2\pi}} \int_T^{\infty} \exp(-t^2/2) dt = \frac{1}{2} - \Phi(T),$$

где $\Phi(T) = \frac{1}{\sqrt{2\pi}} \int_0^T \exp(-t^2/2) dt$.

Оценивая битовые затраты на кодирование по величине двоичной энтропии, имеем:

$$R(T) = -p_0 \log_2 p_0 - 2p_1 \log_2 p_1.$$

Для ошибки (искажений) кодирования $D = M(|X - \hat{X}|^2)$ получаем:

$$\begin{aligned} D(m, T) &= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{-T} (t+m)^2 e^{-t^2/2} dt + \int_{-T}^T t^2 e^{-t^2/2} dt + \int_T^{+\infty} (t-m)^2 e^{-t^2/2} dt \right) = \\ &= \sqrt{\frac{2}{\pi}} \left(\int_0^T t^2 e^{-t^2/2} dt + \int_T^{+\infty} (t-m)^2 e^{-t^2/2} dt \right) = \end{aligned}$$

$$\begin{aligned}
 &= \sqrt{\frac{2}{\pi}} \left(\underbrace{\int_0^{+\infty} t^2 e^{-t^2/2} dt}_{\sqrt{\pi/2}} - 2m \underbrace{\int_T^{+\infty} t e^{-t^2/2} dt}_{\exp(-T^2/2)} + m^2 \int_T^{+\infty} e^{-t^2/2} dt \right) = \\
 &= 1 + \sqrt{\frac{2}{\pi}} \left(m^2 \int_T^{+\infty} e^{-t^2/2} dt - 2m e^{-T^2/2} \right). \quad (6.28)
 \end{aligned}$$

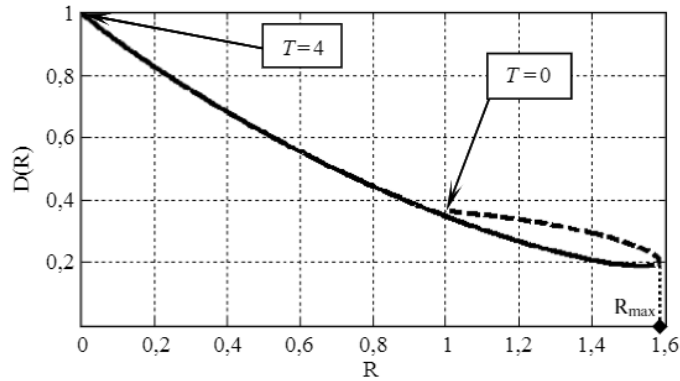
Составим функцию Лагранжа $L(m, T) = D(m, T) + \lambda R(T)$ и приравняем нулю ее частную производную по параметру m (необходимое условие безусловного минимума):

$$\frac{\partial L(m, T)}{\partial m} = \frac{\partial D(m, T)}{\partial m} = \sqrt{\frac{2}{\pi}} \left(2m \int_T^{+\infty} e^{-t^2/2} dt - 2e^{-T^2/2} \right) = 0.$$

Отсюда $m = e^{-T^2/2} / \int_T^{+\infty} e^{-t^2/2} dt$ — заметим, что это выражение представляет собой условное математическое ожидание стандартной нормальной случайной величины X , попавшей в интервал $X \in (T; +\infty)$, см. общую формулу (2.37). Подставляя последнее полученное выражение для m в соотношение (6.28), получаем:

$$D(T) = 1 - \frac{e^{-T^2}}{\sqrt{\pi/2} \int_T^{+\infty} e^{-t^2/2} dt} = 1 - \frac{2e^{-T^2}}{\pi(1 - 2\Phi(T))}. \quad (6.29)$$

Таким образом, рассматриваемый алгоритм кодирования F задается единственным параметром T , поскольку и битовые затраты, и искажение есть функции, зависящие только от T : $R = R(T)$, $D = D(T)$. Для определения оптимального значения T^* для каждого битового бюджета R_b необходимо решить одномерную задачу условной минимизации (6.26) относительно параметра T . Мы поступим иначе, выполнив полный перебор значений $0 \leq T \leq 4$ с достаточно малым шагом и построив соответствующую кривую $D = D(R)$ параметрически по точкам $(R(T); D(T))$, найденным в результате численных расчетов, см. рис. 6.11. Пунктиром обозначена та часть кривой, которая соответствует неоптимальным значениям параметра T . Сплошной линией показана зависимость $D = D(R)$ в области значений T , определяющих функцию оптимального кодирования $D = D(R)$. ►

Рис. 6.11. Функция $D = D(R)$ алгоритма F из примера 6.4

Упражнение. Поясните (см. рис. 6.11), почему для алгоритма кодирования F из примера 6.4:

- значению параметра $T = 0$ должна соответствовать величина $R = 1$ (бит);
- $\lim_{T \rightarrow \infty} R(T) = 0$ (битов);
- $R_{\max} = \max_{0 \leq T < \infty} R(T) = \log_2 3 \approx 1,585$ (бита).

Упражнение. В условиях примера 6.4 рассмотрим случай отсутствия ограничений на битовый бюджет, что соответствует значению $\lambda = 0$ в функции Лагранжа (6.27) и минимально возможной ошибке кодирования.

А) Покажите, что для оптимального параметра $T = T^*$ верно:

$$T = \frac{e^{-T^2/2}}{2 \int_T^{\infty} e^{-t^2/2} dt} = \frac{e^{-T^2/2}}{\sqrt{2\pi}(1-2\Phi(T))} = \frac{m}{2},$$

т. е. порог квантования T расположен в точности между соседними уровнями 0 и m , и алгоритм F определяет оптимальное квантование Ллойда — Макса с тремя уровнями (см. раздел 2.12 и формулы (2.36), (2.37)). *Указание:* учтите, что на графике $D(R)$ в соответствующей точке справедливо равенство $dD/dR = -\lambda = 0$, а $dD/dR = D'(T)/R'(T)$.

Б) Используя какой-либо программный пакет, позволяющий вычислять функцию $\Phi(T) = \frac{1}{\sqrt{2\pi}} \int_0^T \exp(-t^2/2) dt$, определите методом

простых итераций, $T_{j+1} = \frac{e^{-T_j^2/2}}{\sqrt{2\pi(1-2\Phi(T_j))}}$, численное значение оптимального параметра $T = T^*$ и соответствующее значение двоичной энтропии $R = R_b$.

Результаты решения примера 6.4 показывают, что задача оптимизации параметров алгоритма кодирования, представленная в формулировке (6.26), не всегда разрешима: если в условиях примера 6.4 задать слишком большой битовый бюджет $R_b > R_{\max}$, то решения задачи (6.26) для алгоритма F из примера 6.4 не существует. Другие причины неразрешимости (6.26) могут быть вызваны тем, что реальные функции $R(\mathbf{u})$ могут носить дискретный характер, изменяя значения возможно и небольшими, но «скачками». Заметим, что дискретными могут быть и компоненты набора \mathbf{u} , среди которых возможны также и логические переменные. Кроме того, аналитические выражения для $D = D(\mathbf{u})$ и $R = R(\mathbf{u})$ чаще всего отсутствуют и используют зависимости, получаемые экспериментально. Поэтому на практике, решая задачу (6.26), добиться точного равенства $R(\mathbf{u}^*) = R_b$ в большинстве случаев невозможно.

Окончательно задачу оптимизации параметров алгоритма сжатия данных с потерями в общем виде сформулируем следующим образом: для некоторого алгоритма сжатия F , имеющего N варьируемых параметров (возможно, различной природы) $\mathbf{u} = (u_1, \dots, u_N)$, и заданного дискретного сигнала (вектора или матрицы) \mathbf{X} найти такие значения параметров $\mathbf{u}^* = (u_1^*, \dots, u_N^*)$ алгоритма кодирования F , что

$$\begin{cases} D(\mathbf{u}^*) = \min_{\mathbf{u}} D(\mathbf{u}) \\ R(\mathbf{u}^*) \leq R_b \end{cases} \quad (6.30)$$

Подходы к решению задачи оптимизации в формулировке (6.30) на практике остаются теми же: итерационное уточнение параметра $\lambda \geq 0$ с минимизацией соответствующей функции Лагранжа (6.27).

Упражнение. Для алгоритма F из примера 6.4 построить график $D(R_b)$, соответствующий формулировке задачи оптимизации (6.30). Как получить соответствующую кривую из графика $D(R)$ рис. 6.11?



Решение возникающих на практике задач оптимизации параметров кодирования для реальных алгоритмов компрессии данных обычно сопряжено со значительными сложностями вычислительного характера.

6.8. Аппроксимационный подход к выбору преобразований для представления дискретных сигналов. Частотная трактовка

Ранее отмечалось (см. раздел 6.2), что целью использования дискретных ортогональных 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 важность неравномерного распределения энергии по компонентам преобразованного вектора \mathbf{Y} , когда основная часть энергии дискретного сигнала концентрируется в малом количестве компонент-трансформант — именно на этом принципе мы 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 $\{\varphi_k(t)\}_{k=0}^{\infty}$ (для упрощения изложения считаем, что система ортонормирована, $\|\varphi_k(t)\| = \sqrt{\langle \varphi_k(t), \varphi_k(t) \rangle} = 1$). Например, исходный аналоговый сигнал в случае использования систем Уолша и Хаара был представлен в виде $x(t) = \sum_{k=0}^{\infty} c_k \varphi_k(t)$, а дискретный — в виде $x_j = x(j\Delta t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} y_k \varphi_k(j\Delta t)$, где $c_k = \langle x(t), \varphi_k(t) \rangle \approx y_k / \sqrt{N}$ (см. замечания к формулам (2.30), (2.32)). Для энергии сигналов в силу ортонормированности систем и ортогональности преобразований имеем в непрерывном случае $\|x(t)\|_E^2 = \langle x(t), x(t) \rangle = \sum_{k=0}^{\infty} |c_k|^2$

(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 2.1).

Для концентрации основной энергии сигнала в малом числе коэффициентов ряда Фурье необходимо построить такой ортогональный базис $\{\varphi_k(t)\}_{k=0}^{\infty}$, который позволяет достаточно точно представить непрерывный сигнал $x(t) = \sum_{k=0}^{\infty} c_k \varphi_k(t)$ в виде конечной суммы с малым числом слагаемых: $x(t) \approx x_n(t) = \sum_{k=0}^n c_k \varphi_k(t)$. При этом квадрат ошибки, связанной с отбрасыванием членов ряда, на основании равенства Парсевала равен $\|x(t) - x_n(t)\|_E^2 = \sum_{k=n+1}^{\infty} |c_k|^2$. То есть искомый базис $\{\varphi_k(t)\}_{k=0}^{\infty}$ 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 $\{\varphi_k(t)\}_{k=0}^{\infty}$ в терминах его частотных свойств.

Частотный подход к исследованию реальных сигналов основан на том, что для реальных сигналов $f(t)$, представимых в виде интеграла Фурье (1.17), спектральная плотность убывает с увеличением частоты¹: $\lim_{\nu \rightarrow \infty} S(\nu) = 0$. Основная энергия сигнала во многих случаях локализована в области нижних 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 степенью точности считать ограниченным некоторой полосой $\nu \in (-F; F)$.

Из приведенных рассуждений следует, что обычно, выполнив частотное разложение дискретного сигнала, можно без особого ущерба для точности последующего восстановления отбросить (положить равными нулю) высокочастотные коэффициенты, например для спектра ДПФ (2.20) это те компоненты из $\{y_k\}_{k=0}^{N-1}$, номера которых близки к $N/2$ (см. также пример 2.16). Для ДКП (6.13) высокочастотными являются компоненты с большими индексами, т. е. близкими к N .

Пример 6.5. Аналогом производной для дискретного сигнала $x_j = x(j\Delta t)$ являются разности значений соседних отсчетов

¹ При разложении функций в тригонометрический ряд (1.13) или (1.15) соответствующее свойство коэффициентов Фурье $\{c_k\}$ принимает вид: $\lim_{k \rightarrow \infty} c_k = 0$.

$\Delta_j = x_j - x_{j-1}$, которые характеризуют «скорость» изменения сигнала. Получим оценки для модулей $|y_k|$ коэффициентов частотных разложений, определяемых ДКП (6.13) и ДПФ (2.20), в терминах разностей $\Delta_j = x_j - x_{j-1}$ отсчетов конечного дискретного сигнала (x_0, \dots, x_{N-1}) .

◀ Назовем ядром ДКП (6.13) для $k=1, \dots, N-1$ следующее выражение:

$$C_j(k) = \sum_{l=0}^j \cos\left(\frac{\pi k}{N}(l+0,5)\right).$$

Используя формулу для суммы геометрической прогрессии, при $k \neq 0$ получаем:

$$\begin{aligned} C_j(k) &= \operatorname{Re} \left(\sum_{l=0}^j e^{i \frac{\pi k}{N} (l + \frac{1}{2})} \right) = \operatorname{Re} \left(e^{i \frac{\pi k}{2N}} \frac{1 - e^{i \frac{\pi k}{N} (j+1)}}{1 - e^{i \frac{\pi k}{N}}} \right) = \\ &= \operatorname{Re} \left(e^{i \frac{\pi k}{2N}} \frac{e^{-i \frac{\pi k}{2N}} - e^{i \frac{\pi k}{2N} (j+1)}}{e^{-i \frac{\pi k}{2N}} - e^{i \frac{\pi k}{2N}}} \right) = \operatorname{Re} \left(e^{i \frac{\pi k}{2N} (j+1)} \frac{\sin \frac{\pi k (j+1)}{2N}}{\sin \frac{\pi k}{2N}} \right) = \\ &= \frac{\cos \frac{\pi k (j+1)}{2N} \sin \frac{\pi k (j+1)}{2N}}{\sin \frac{\pi k}{2N}} = \frac{\sin \frac{\pi k (j+1)}{N}}{2 \sin \frac{\pi k}{2N}}. \end{aligned}$$

Положим $C_{-1}(k) = 0$ и заметим также, что $C_{N-1}(k) = 0$.

Значение коэффициента ДКП $y_0 = (1/\sqrt{N}) \sum_{j=0}^{N-1} x_j$ с точностью до нормировочного множителя представляет собой среднее значение («постоянную» составляющую) дискретного сигнала. Для прочих компонент ДКП с номерами $k=1, \dots, N-1$ («переменных» составляющих) с использованием полученного выражения для ядра $C_j(k)$ имеем:

$$\begin{aligned} y_k &= \sqrt{\frac{2}{N}} \sum_{j=0}^{N-1} x_j (C_j(k) - C_{j-1}(k)) = \\ &= \sqrt{\frac{2}{N}} \left(\sum_{j=0}^{N-2} x_j C_j(k) + x_{N-1} \underbrace{C_{N-1}(k)}_0 - x_0 \underbrace{C_{-1}(k)}_0 - \sum_{j=1}^{N-1} x_j C_{j-1}(k) \right) = \end{aligned}$$

$$= \sqrt{\frac{2}{N}} \sum_{j=1}^{N-1} \underbrace{(x_{j-1} - x_j)}_{-\Delta_j} C_{j-1}(k) = \frac{-1}{\sqrt{2N} \sin \frac{\pi k}{2N}} \sum_{j=1}^{N-1} \Delta_j \sin \left(\frac{\pi k}{N} j \right).$$

Отсюда получаем следующую верхнюю оценку для модулей коэффициентов ДКП-разложения дискретного сигнала при $k = 1, \dots, N-1$:

$$|y_k| \leq \frac{1}{\sqrt{2N} \sin \frac{\pi k}{2N}} \sum_{j=1}^{N-1} |\Delta_j| = \frac{1}{\sqrt{2} \sin \frac{\pi k}{2N}} V_x, \quad (6.31)$$

где величина

$$V_x = \frac{1}{\sqrt{N}} \sum_{j=1}^{N-1} |\Delta_j| = \frac{1}{\sqrt{N}} \sum_{j=1}^{N-1} |x_j - x_{j-1}| \quad (6.32)$$

характеризует то, насколько сильно различается в среднем сигнал $\mathbf{X} = (x_0, \dots, x_{N-1})^T$ в соседних отсчетах. Быстро изменяющемуся дискретному сигналу \mathbf{X} , очевидно, будет соответствовать большее значение V_x .

Аналогичным образом поступим с ДПФ (2.20): введем в рассмотрение для $k = 1, \dots, N-1$ ядро

$$F_j(k) = \sum_{l=0}^j e^{-i\frac{2\pi}{N}kl} = \frac{1 - e^{-i\frac{2\pi}{N}k(j+1)}}{1 - e^{-i\frac{2\pi}{N}k}} = \frac{\sin \frac{\pi k(j+1)}{N}}{\sin \frac{\pi k}{N}} e^{-i\frac{\pi}{N}kj},$$

полагая $F_{-1}(k) = 0$ и замечая, что $F_{N-1}(k) = 0$. Тогда при $k \neq 0$ имеем:

$$\begin{aligned} y_k &= \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j (F_j(k) - F_{j-1}(k)) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N-1} (x_{j-1} - x_j) F_{j-1}(k) = \\ &= \frac{-1}{\sqrt{N} \sin \frac{\pi k}{N}} \sum_{j=1}^{N-1} \Delta_j \sin \left(\frac{\pi k}{N} j \right) e^{-i\frac{\pi}{N}k(j-1)}. \end{aligned}$$

Отсюда получаем для модулей коэффициентов ДПФ с номерами $k = 1, \dots, N-1$ следующую оценку:

$$|y_k| \leq \frac{1}{\sqrt{N} \sin \frac{\pi k}{N}} \sum_{j=1}^{N-1} |\Delta_j| = \frac{1}{\sin \frac{\pi k}{N}} V_x, \quad (6.33)$$

где характеристика скорости изменения сигнала V_x определена в (6.32).

Для размерности $N = 32$ верхние оценки (6.31) и (6.33) модулей компонент ДКП (вещественное преобразование) и ДПФ (комплексное преобразование) приведены в виде графиков их огибающих на рис. 6.12, при построении которых для величины (6.32) было принято значение $V_x = 1$ (оно определяет только масштаб отображения по оси ординат).

Приведенные на рис. 6.12 детерминированные оценки коэффициентов ДКП (6.13) согласуются с полученными ранее оценками для их дисперсий, которые были найдены для дискретного сигнала, моделируемого марковским процессом первого порядка (см. рис. 6.4). ▶

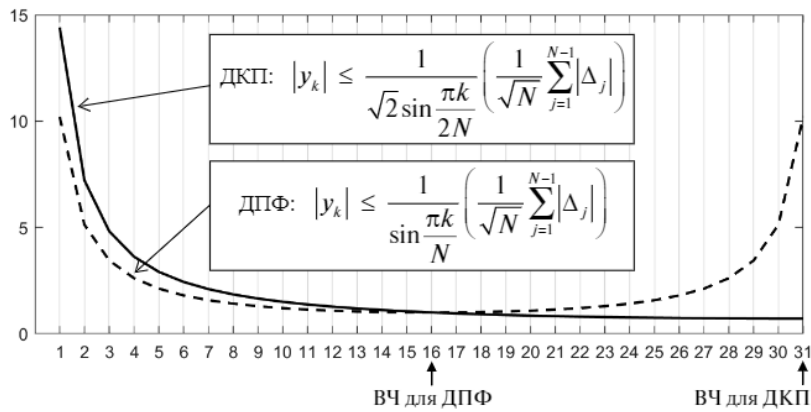


Рис. 6.12. Иллюстрация к примеру 6.5: верхние оценки (6.31) и (6.33) модулей коэффициентов ДКП (6.13) и ДПФ (2.20) для размерности преобразований $N = 32$

Помимо тригонометрической системы, для представления сигналов в виде ряда $x(t) = \sum_{k=0}^{\infty} c_k \varphi_k(t)$ можно использовать и другие функциональные базисы $\{\varphi_k(t)\}_{k=0}^{\infty}$. Основной вопрос при этом заключается в следующем: какую базисную систему функций предпочтительно использовать? Приведенные выше рассуждения о природе спектров реальных сигналов позволяют предположить, что частотные спектры используемых базисных функций $\{\varphi_k(t)\}_{k=0}^{\infty}$ должны различаться, образуя выраженные

низко- и высокочастотные области локализации энергии в спектральной области. Тогда заранее можно предположить, что коэффициенты Фурье, соответствующие «высокочастотным» базисным функциям, будут вносить малый вклад в разложение $x(t) = \sum_{k=0}^{\infty} c_k \varphi_k(t)$. Рассмотренные нами в первой главе системы Уолша и Хаара обладают такими общими свойствами: с увеличением номера базисной функции локализация ее энергии в частотной области имеет тенденцию смещаться в сторону высоких частот.

Исторически для представления сигнала на конечном интервале наблюдения длительности T все начиналось с тригонометрической системы (или базиса комплексных экспонент), см. (1.15):

$$\left\{ \varphi_k(t) = \exp\left(i \frac{2\pi}{T} kt\right) \right\}_{k=-\infty}^{\infty}. \quad (6.34)$$

Тригонометрический базис является наилучшим выбором для представления данных (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, которое мы хотим получить при использовании корреляционной модели дискретного сигнала.

Однако, как отмечалось в разделе 6.1, в силу возможного непостоянства статистических свойств дискретного сигнала его описание через ковариационную матрицу не всегда адекватно: реальные, интересующие нас сигналы, во многих случаях являются нестационарными случайными процессами, имеют локальные особенности во временной области («срывы» стационарности). Например, для цифровых изображений (двумерных массивов яркостей точек-пикселей) такие особенности, связанные с резким изменением яркости, проявляются вблизи контуров объектов, а стационарный характер данных наблюдается в фоновых областях и областях текстур. Для представления подобных особенностей тригонометрический базис неприемлем и вызывает хорошо известный эффект Гиббса (см. раздел 1.4), связанный с нелокальным распространением ошибки при отбрасывании части спектральных

коэффициентов. Приведем иллюстрацию подобного эффекта для дискретного случая.

Пример 6.6. 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.

◀ Заданный дискретный сигнал является наиболее ярким примером «особенности» — дискретным аналогом дельта-функции. Для его спектра, который находится с помощью ДПФ (2.20), 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 дискретное гармоническое колебание, т.е. влияние ошибки распространяется на все компоненты дискретного сигнала. ▶

Совершенно аналогичные проблемы с распространением ошибки возникают также, если вместо тригонометрического использовать другие базисы, имеющие неограниченный носитель во временной области. Действительно, представив сигнал в виде ряда Фурье (не обязательно тригонометрического), $x(t) = \sum_{k=0}^{\infty} c_k \varphi_k(t)$, и отбросив коэффициент ряда с номером m , мы получим ошибку $\varepsilon(t) = c_m \varphi_m(t)$, локализация которой во временной области полностью определяется локализацией базисной функции $\varphi_m(t)$. 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, обычно вносят малый вклад в разложение большинства реальных сигналов, то 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 высоко-частотные составляющие разложения значимы и должны быть сохранены.

Как отмечалось в разделе 1.6, 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 время×частота, приведенные требования к искомому базису находятся в противоречии. В следующем разделе мы обсудим возможный путь компромиссного разрешения данного противоречия.

6.9. Время–частотный анализ. Оконное преобразование Фурье

Частотный анализ сигналов нестационарной природы с изменяющимися во времени характеристиками сопряжен с определенными трудностями, так как для таких сигналов спектр (1.18) определяет некие «усредненные» частотные свойства. Поэтому для частотного анализа нестационарных сигналов интегральное преобразование (1.18) малоприспособно. Например, если анализируемый сигнал представляет собой звукозапись нескольких последовательно звучащих голосов разных людей во время беседы, то частотное разложение такого сигнала не будет характеризовать свойства голоса конкретно кого-то из говоривших.

Для того чтобы охарактеризовать локально изменяющиеся во времени частотные свойства сигнала, используется *оконное* (или *взвешенное*) преобразование Фурье (ОПФ):

$$S(\nu, \tau) = \int_{-\infty}^{+\infty} p(t - \tau) f(t) e^{-2\pi i \nu t} dt, \quad (6.35)$$

где $p(t)$ — некоторая вещественная и четная весовая функция окна, осуществляющая временной выбор интервала для частотного

анализа сигнала. В простейшем случае весовая функция представляет собой прямоугольный импульс

$$p(t) = \Pi_{T/2}(t) = \begin{cases} 1, & \text{при } t \in [-T/2; T/2] \\ 0, & \text{при } t \notin [-T/2; T/2] \end{cases} \quad (6.36)$$

хотя можно предположить, что в качестве окна лучше выбирать более гладкие функции — мы затронем этот вопрос в примере 6.7. Таким образом, вместо частотного преобразования (1.18) мы ввели в рассмотрение время-частотное (или частотно-временное) преобразование (6.35), которое характеризует частотные свойства сигнала в момент времени τ .

Ширина окна T (не обязательно прямоугольного) на практике должна выбираться так, чтобы на любом интервале длины T сигнал был *квазистационарным*, т. е. с некоторым допущением его можно было бы считать стационарным. Это часто вызывает трудности, поскольку априорные знания о длине интервала стационарности сигнала могут отсутствовать. Кроме того, нужно помнить о следующем существенном ограничении применения формулы (6.35) для частотного анализа: очевидно, что нельзя достоверно говорить о вкладе гармонических составляющих частоты ν , период колебаний которых $1/\nu$ сравним или больше длины интервала наблюдения (частотного анализа) — ширины окна T . Для таких значений частот ν использование формулы (6.35) является некорректным. Поэтому ширина окна T должна выбираться по возможности большей.

Не только длительность, но и форма весовой функции окна $p(t)$ влияет на характеристики время-частотного анализа. Это можно проиллюстрировать на следующем примере.

Пример 6.7. Сравнить частотные искажения время-частотного анализа сигнала $f(t) = 1$ при использовании весовых функций прямоугольного окна $p_1(t)$ и гауссова импульса $p_2(t)$.

◀ «Традиционный» спектр (1.18) сигнала $f(t) = 1$, как было установлено в разделе 1.7, представляет собой дельта-функцию $\delta(\nu)$ (см. (1.22)).

Для сигнала $f(t) = 1$ рассмотрим ОПФ (6.35) $S(\nu, \tau) = \int_{-\infty}^{+\infty} p(t - \tau) e^{-2\pi i \nu t} dt$ с заданными оконными функциями толь-

ко для $\tau = 0$. (При других значениях τ отличия будут наблюдаться лишь в фазовом спектре.)

Для прямоугольного окна (6.36) $p_1(t) = \Pi_{1/2}(t)$, имеющего единичную длительность $T = 1$ и единичную энергию, имеем:

$$S_1(\nu, 0) = \mathcal{F}\{p_1(t)\} = \frac{\sin \pi \nu}{\pi \nu}.$$

При решении примера 1.15 было установлено, что гауссову импульсу единичной энергии $p_0(t) = \frac{1}{\sqrt[4]{2\pi}} e^{-t^2/4}$ соответствует спектр $S_0(\nu) = \sqrt[4]{8\pi} e^{-4\pi^2\nu^2}$, который также представляет собой гауссов импульс. Так как функция $(p_0(t))^2$ совпадает с функцией плотности стандартного нормального закона, на основании правила «трех сигма» можно утверждать, что 99,7% энергии функции $p_0(t)$ локализовано в отрезке $t \in [-3; 3]$. Тогда для гауссова импульса $p_2(t) = \sqrt{6} p_0(6t) = \sqrt[4]{18/\pi} e^{-9t^2}$ энергия также будет единичной, а 99,7% этой энергии локализовано в отрезке $t \in [-0,5; 0,5]$. Тем самым мы практически уравнили во временной области длительность и энергию гауссова $p_2(t)$ и прямоугольного $p_1(t)$ импульсов.

На основании свойства 3° интеграла Фурье (см. раздел 1.5) находим, что

$$S_2(\nu, 0) = \mathcal{F}\{p_2(t)\} = \mathcal{F}\{\sqrt{6} p_0(6t)\} = \frac{1}{\sqrt{6}} S_0(\nu/6) = \sqrt[4]{\frac{2\pi}{9}} e^{-\pi^2\nu^2/9}.$$

Графики найденных спектральных функций $S_{p_1}(\nu) = \mathcal{F}\{p_1(t)\}$ и $S_{p_2}(\nu) = \mathcal{F}\{p_2(t)\}$ приведены на рис. 6.13. Каждый график представляет собой искаженное в результате влияния функции окна представление оригинального спектра — дельта-функции $\delta(\nu)$. Такого рода искажения нередко называют *утечкой спектра*, так как спектральная плотность, исходно представленная дельта-функцией, «растекается» по частотной оси.

Видно, что искажения, вызванные весовой функцией гауссова импульса, привели к меньшей утечке частотного спектра. Численные расчеты показывают, что примерно 95% энергии спектральных плотностей сосредоточены в интервалах $\nu \in [-2; 2]$ для $S_{p_1}(\nu) = \mathcal{F}\{p_1(t)\}$ и $\nu \in [-1; 1]$ для $S_{p_2}(\nu) = \mathcal{F}\{p_2(t)\}$. ►

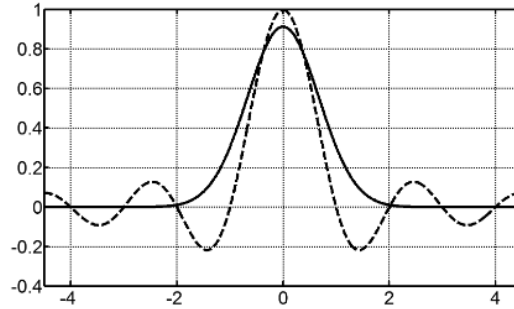


Рис. 6.13. Спектры прямоугольного окна $S_{p_1}(\nu) = \mathcal{F}\{p_1(t)\}$ (пунктирная линия) и гауссова окна $S_{p_2}(\nu) = \mathcal{F}\{p_2(t)\}$ (сплошная линия)

Если характеризовать длительность сигнала во временной области Δt и ширину спектра $\Delta \nu$ по эффективным значениям (см. (1.19) и (1.20)), то использование гауссова импульса в качестве весовой функции дает минимально возможную ширину спектра при фиксированной длительности окна, как это следует из теоремы 1.16 и примера 1.15. ОПФ с весовой функцией $p_\alpha(t) = \frac{1}{2\sqrt{\pi\alpha}} \exp(-t^2/4\alpha)$ носит название *преобразования Габора*, который впервые предложил его в 40-х годах прошлого века. Изменяя параметр α , можно менять соотношение между параметрами Δt и $\Delta \nu$. Впоследствии преобразование Габора было обобщено в виде (6.35) на другие весовые функции.

Упражнение. Найти преобразование Габора сигнала $f(t) = 1$ для $\alpha = 1/2$. *Указание:* воспользуйтесь результатами решения примера 1.15 и свойствами 2°, 3° преобразования Фурье.

Идею, заложенную в ОПФ (6.35), можно развить далее для построения функционального базиса, обладающего локализацией как во временной, так и в частотной областях.

Считая анализируемый сигнал в пределах любого временного интервала длительности T стационарным, выберем некоторую вещественную четную весовую функцию окна $p(t)$ длительности T и перепишем ОПФ (6.35) в несколько ином виде:

$$S(\nu, n) = \int_{-\infty}^{+\infty} p(t - n\tau_0) f(t) e^{-2\pi\nu t} dt, \quad (6.37)$$

где n — номер интервала наблюдения (частотного анализа), τ_0 — параметр сдвига окна, $\tau_0 \leq T$.

В простейшем случае, когда $p(t) = \Pi_{T/2}(t)$ (6.36) и $\tau_0 = T$ (окно — прямоугольное, интервалы наблюдения не перекрываются), формула (6.37) примет вид:

$$S(\nu, n) = \int_{(n-1/2)T}^{(n+1/2)T} f(t) e^{-2\pi\nu t} dt.$$

Возьмем отсчеты данной спектральной функции в точках $\nu_k = k/T$ и обозначим $c_{k,n} = S(\nu_k, n)/T$. Тогда функцию $f(t)$ можно представить в виде (см. формулы (1.15) и (1.16)):

$$f(t) = \sum_{n=-\infty}^{\infty} \Pi_{T/2}(t - nT) \sum_{k=-\infty}^{\infty} c_{k,n} e^{i\frac{2\pi}{T}kt}.$$

В общем случае ОПФ (6.37) интервалы наблюдения сигнала могут перекрываться, а весовая функция $p(t)$ может иметь, как мы уже видели в примере 6.7, и неограниченный носитель. Общей идеей остается использование вместо базиса (6.34), который нелокализован во временной и абсолютно локализован в частотной области, нового (возможно, не ортогонального) базиса

$$\left\{ \varphi_{k,n}(t) = p(t - n\tau_0) e^{i\frac{2\pi}{T}kt} \right\}_{n,k=-\infty}^{\infty} \quad (6.38)$$

для представления сигнала в виде

$$f(t) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{k,n} \varphi_{k,n}(t) = \sum_{n=-\infty}^{\infty} p(t - n\tau_0) \sum_{k=-\infty}^{\infty} c_{k,n} e^{i\frac{2\pi}{T}kt}. \quad (6.39)$$

Очевидно, что свойства локализации базиса (6.38) во временной области полностью определяются тем, какая весовая функция $p(t)$ была выбрана для его построения. Результаты решения примера 6.7 позволяют предположить, что и в частотной области свойства локализации базиса (6.38) также определяются выбором весовой функции. Покажем это.

Пример 6.8. Найти выражение для спектральной плотности функции $\varphi_{k,n}(t) = p(t - n\tau_0) e^{i\frac{2\pi}{T}kt}$, если известна спектральная плотность весовой функции $S_p(\nu) = \mathcal{F}\{p(t)\}$.



◀ На основании свойства 4° преобразования Фурье (см. раздел 1.5) для спектра сдвинутой оконной функции справедливо равенство

$$S_{p_n}(v) = \mathcal{F}\{\varphi_{0,n}(t)\} = \mathcal{F}\{p(t - n\tau_0)\} = S_p(v)e^{-2\pi i n\tau_0 v}.$$

Используя свойство 5° преобразования Фурье (см. раздел 1.5), окончательно получим:

$$S_{\varphi_{k,n}}(v) = \mathcal{F}\{\varphi_{k,n}(t)\} = S_{p_n}\left(v - \frac{k}{T}\right) = S_p\left(v - \frac{k}{T}\right)e^{-2\pi i n\tau_0\left(v - \frac{k}{T}\right)}. \quad \blacktriangleright$$

Умножению оконной функции $p(t - n\tau_0)$ на комплексное гармоническое колебание $e^{i\frac{2\pi}{T}kt}$ в частотной области соответствует сдвиг спектра, при котором нулевое значение частоты $v = 0$ переходит в точку $v = k/T$, а форма графика амплитудного спектра сохраняется: $|S_{\varphi_{k,n}}(v)| = \left|S_p\left(v - \frac{k}{T}\right)\right|$. Обозначив длины интервалов, характеризующих локализацию весовой функции $p(t)$ во временной и в частотной областях соответственно¹ Δt и Δv , можно схематично изобразить следующее (см. рис. 6.14) покрытие плоскости время×частота базисными функциями $\{\varphi_{k,n}(t)\}_{k,n=-\infty}^{\infty}$ (в общем случае соседние ячейки могут частично перекрываться).

Таким образом, свойства локализованности базисных функций (6.38) полностью определяются выбором весовой функции $p(t)$ не только для временной, но и для частотной области. Чем меньше значения величин Δt , Δv , тем выше локализация базисных функций

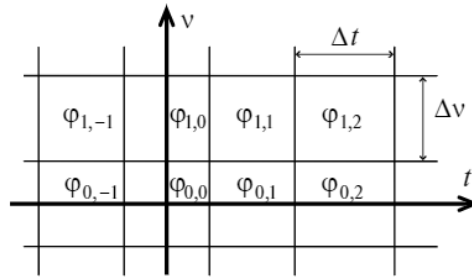


Рис. 6.14. Локализация на плоскости время×частота базисных функций (6.38), используемых для разложения (6.39)

¹ То, каким образом могут выбираться данные величины, обсуждалось в разделе 1.6.

во временной и частотной областях соответственно, которая ограничена принципом неопределенности (теорема 1.16).

Отметим, что рассмотренный в разделе 6.5 метод сжатия цифровых изображений JPEG на основе ДКП также использует идеи ОПФ, адаптированные для двумерного дискретного случая: окно (весовая функция) принимает единичное значение в некотором квадрате 8×8 пикселей и нулевое — вне этого квадрата, а вместо преобразования Фурье используется его модификация — косинусное преобразование.

Несмотря на использование базиса, имеющего как временную, так и частотную локализацию базисных функций, представление (6.39) обладает недостатком, вытекающим из общих свойств широкого класса сигналов, представляющих практический интерес. Природа многих реальных сигналов (в частности, мультимедийных) такова, что длина T временного интервала квазистационарности (т.е. относительного постоянства характеристик амплитуды и фазы) для низкочастотных составляющих больше, чем для высокочастотных составляющих. То есть если разложить такой сигнал $f(t)$ на низкочастотную и высокочастотную составляющие, $f(t) = f_{\text{нч}}(t) + f_{\text{вч}}(t)$, то оказывается, что $T_{\text{нч}} > T_{\text{вч}}$.

Для эффективного представления подобного рода сигналов высокочастотные базисные функции, которые характерно проявляются при спектральном разложении в областях резких изменений сигнала (т.е. в выраженных «особенностях»), должны обладать более высокой временной локализацией по сравнению

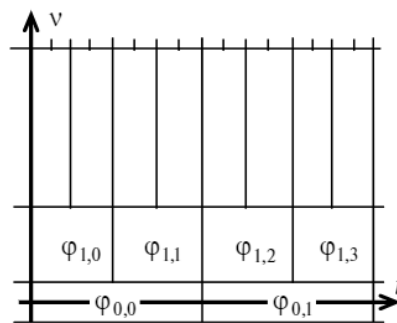


Рис. 6.15. Локализация базисных функций с переменным время-частотным разрешением



с низкочастотными базисными функциями. Поскольку повышение разрешения базисных функций во временной области возможно лишь за счет ухудшения разрешения в частотной области, то разбиение плоскости время×частота для конструируемого базиса должно соответствовать общим идеям, схематично отображенным на рис. 6.15. Такими свойствами переменного время-частотного разрешения обладают функциональные базисы *вейвлетов*, к рассмотрению которых мы перейдем в следующей главе.

6.10. Использование ДПФ для спектрального анализа

Практически сразу после появления алгоритмов быстрого преобразования Фурье в 60-е годы прошлого века было предложено использовать ДПФ (2.20) для сжатия сигналов (в простейшем случае, см. пример 2.16, высокочастотные компоненты дискретного спектра отбрасываются и при восстановлении сигнала заменяются нулями). Это послужило толчком к изучению возможностей и эффективности применения для сжатия сигналов других преобразований и появлению ДКП (6.13), в результате чего ДПФ (2.20) (или (6.10)) потеряло актуальность как способ эффективного представления сигналов для сжатия данных. Вместе с тем, помимо упомянутых аспектов кодирования, на практике часто возникает потребность в спектральном разложении дискретных сигналов (например, для выделения сигналов нужных частот или анализа доминирующих частот обрабатываемых сигналов). ДПФ является определенным приближением, которое может использоваться (и используется) для вычисления интегрального преобразования Фурье на конечных интервалах. Рассмотрим специфику применения ДПФ для спектрального анализа.

Вводя в предыдущем разделе оконное преобразование Фурье (6.35), мы преследовали цель построения способа частотного анализа сигнала на конечных временных интервалах — локальных областях, на которых частотные характеристики сигнала можно считать неизменными. При использовании для анализируемого сигнала модели случайного процесса в этом случае говорят также

о квазистационарности сигнала на соответствующих интервалах анализа. Вместе с тем на практике для частотного анализа любого (в том числе стационарного) сигнала может быть доступен лишь его фрагмент на конечном интервале наблюдения некоторой длины T .

Упражнение. Покажите, что для оценки $\hat{S}(v) = \int_0^T f(t)e^{-2\pi i vt} dt$ спектра $S(v) = \int_{-\infty}^{+\infty} f(t)e^{-2\pi i vt} dt$ верно соотношение:

$$\hat{S}(v) = \frac{e^{-i\pi T v}}{\pi} \int_{-\infty}^{+\infty} S(u) \frac{\sin(\pi T(v-u))}{v-u} e^{i\pi T u} du. \quad (6.40)$$

Указание: воспользуйтесь свойством 7° преобразования Фурье (см. раздел 1.5) и решением примера 1.14.

Формула (6.40) представляет собой интегральную свертку спектров анализируемой функции $f(t)$ и прямоугольного импульса длительности T (см. пример 1.14). Рассмотрим подробнее эффект искажений спектра (6.40), когда сигнал представляет собой гармоническое колебание $f(t) = \cos(2\pi\mu_0 t)$. В этом случае (убедитесь)

$$S(v) = \mathcal{F}\{\cos(2\pi\mu_0 t)\} = (\delta(v - \mu_0) + \delta(v + \mu_0))/2;$$

$$\hat{S}(v) = \frac{e^{-i\pi T v}}{2\pi} \left(\frac{\sin(\pi T(v - \mu_0))}{v - \mu_0} e^{i\pi T \mu_0} + \frac{\sin(\pi T(v + \mu_0))}{v + \mu_0} e^{-i\pi T \mu_0} \right) \quad (6.41)$$

есть спектр функции

$$\hat{f}(t) = \begin{cases} \cos(2\pi\mu_0 t), & \text{при } t \in [0; T] \\ 0, & \text{при } t \notin [0; T] \end{cases}. \quad (6.42)$$

На рис. 6.16 приведены два примера амплитудных спектров функций (6.42). Сигнал, амплитудный спектр которого изображен сверху, на интервале наблюдения длины T имеет ровно три периода колебания, а сигнал с амплитудным спектром, отображенным снизу — три с половиной периода.

Таким образом, «вырезание» фрагмента длительности T из гармонического колебания $f(t) = \cos(2\pi\mu_0 t)$ приводит к эффекту растекания спектральной плотности, превращая спектр из дельта-функций на частотах $\pm\mu_0$ в функцию вида (6.41). Подобный эффект утечки спектра непрерывного сигнала мы уже наблюдали ранее в примере 6.7 для постоянного сигнала $f(t) = 1$ (имеющего нулевую частоту $\mu_0 = 0$).

Выполним теперь дискретизацию сигнала (6.42), выбирая период дискретизации Δt достаточно малым для того, чтобы наложе-

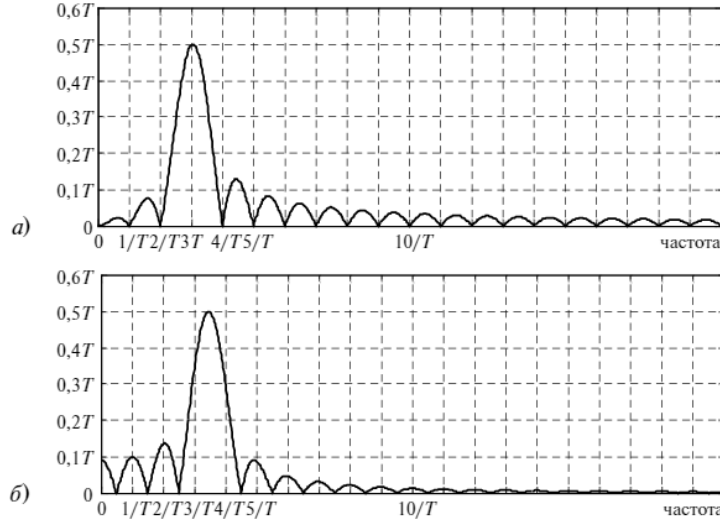


Рис. 6.16. Примеры амплитудных спектров $|\hat{S}(v)|$ сигналов $\hat{f}(t)$ (6.42): а) частота колебания $\mu_0 = 3/T$, б) частота колебания $\mu_0 = 3,5/T$

нием копий спектра (6.41) можно было пренебречь, и так, чтобы число $T/\Delta t = N$ было целым. Полученный дискретный сигнал

$$\hat{f}(t) = \Delta t \sum_{k=0}^{N-1} \cos(2\pi\mu_0 k \Delta t) \delta(t - k \Delta t) \quad (6.43)$$

имеет периодический спектр (см. (2.9))

$$\hat{S}(v) = \sum_{k=0}^{N-1} \cos(2\pi\mu_0 k \Delta t) e^{-2\pi i v k \Delta t} \Delta t, \quad (6.44)$$

который находится по (6.41) как $\hat{S}(v) = \sum_{k=-\infty}^{\infty} \hat{S}\left(v - \frac{k}{\Delta t}\right)$. На рис. 6.17 пунктиром изображены графики одного периода амплитудных спектров сигналов (6.43) (выбрано значение $N = T/\Delta t = 20$) для частот колебаний $\mu_0 = 3/T$, $\mu_0 = 3,5/T$. Сравните графики рис. 6.16 и 6.17.

Как установлено в разделе 2.6, для дискретного сигнала конечной длительности вида $\tilde{f}(t) = \Delta t \sum_{k=0}^{N-1} f(k \Delta t) \delta(t - k \Delta t)$ отсчеты спектральной функции (2.9) $\tilde{S}(v) = \Delta t \sum_{k=0}^{N-1} f(k \Delta t) e^{-2\pi i v k \Delta t}$ в точках $v_j = \frac{j}{N \Delta t} = j/T$, $j \in \mathbb{Z}$, могут быть найдены при помощи ДПФ

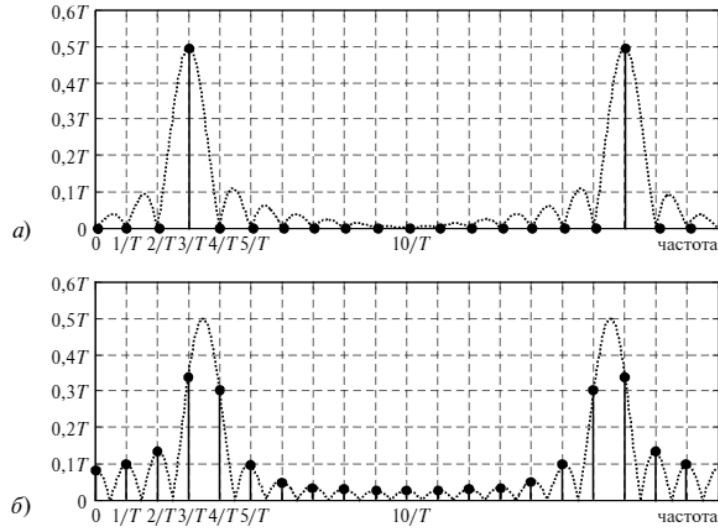


Рис. 6.17. Примеры непрерывных амплитудных спектров $|\tilde{S}(v)|$ (пунктирные кривые, отображен один период) сигналов (6.43) и их дискретные амплитудные спектры (обозначены точками) для $N = T/\Delta t = 20$: а) частота колебания $\mu_0 = 3/T$, б) частота колебания $\mu_0 = 3,5/T$

(2.18) (см. также пример 2.7). Последовательность $\tilde{s}_j = \tilde{S}(v_j)$ является периодической, $\tilde{s}_j = \tilde{s}_{j+N}$, и полностью описывается вектором $(\tilde{s}_0, \dots, \tilde{s}_{N-1})$, который мы ранее (в разделе 2.6) назвали дискретным спектром. Примеры амплитудных дискретных спектров $(|\tilde{s}_0|, \dots, |\tilde{s}_{N-1}|)$ сигналов (6.43) для частот колебаний $\mu_0 = 3/T$, $\mu_0 = 3,5/T$ отображены на рис. 6.17 точками на графиках непрерывных амплитудных спектров. Длительность $T = N\Delta t$ интервала наблюдения дискретного сигнала определяет шаг частоты отсчетов дискретного спектра $\Delta v = 1/N\Delta t$: чем больше значение T , тем меньше Δv и тем точнее отсчеты дискретного спектра $(\tilde{s}_0, \dots, \tilde{s}_{N-1})$ описывают непрерывный спектр дискретного сигнала $\tilde{S}(v)$.

Компоненты дискретного спектра $\tilde{s}_j = \tilde{S}(v_j)$, полученные при помощи ДПФ, соответствуют частотам гармонических колебаний $v_j = j/T$, которые имеют целое число периодов на интервале анализа сигнала длительности T . Такие колебания имеют в дискретном

спектре в основной полосе частот¹ единственную ненулевую компоненту, см. пример на рис. 6.17а.

В силу искажений спектра, вызванных конечностью интервала наблюдения сигнала (см. рис. 6.16), спектральные составляющие аналогового сигнала с частотами колебаний $\mu \neq \nu_j = j/T$ приводят к дискретному спектру, в котором уже все компоненты отличны от нуля — этот эффект размывания единственного пика и для случая дискретного спектра также называется утечкой. Тем не менее при достаточно малом значении $\Delta\nu$ максимум амплитуды дискретного спектра гармонического колебания частоты $\mu_0 \neq j/T$ также будет приходиться на близкую к ней частоту $\nu_j \approx \mu_0$, см. пример на рис. 6.17б.

Если значение длительности T интервала наблюдения сигнала $f(t)$ фиксировано, то увеличение числа временных отсчетов сигнала N на этом интервале (т. е. уменьшение периода дискретизации $\Delta t = T/N$) в спектре дискретного сигнала $\tilde{S}(\nu) = \sum_{k=-\infty}^{\infty} S(\nu - (k/\Delta t))$ повлечет раздвижение копий спектра непрерывного сигнала $S(\nu) = \mathcal{F}\{f(t)\}$. Но при этом не произойдет увеличение разрешения дискретного спектра по частоте, т. е. шаг частотной выборки $\Delta\nu = 1/T$ не уменьшится. Если бы перекрытия сдвинутых копий спектра $S(\nu)$ вообще не было (эффект наложения частот отсутствовал бы полностью), то, при увеличении плотности отсчетов сигнала во временной области, в дискретном спектре $(\tilde{s}_0, \dots, \tilde{s}_{N-1})$ (2.18) просто добавились бы новые нулевые компоненты, соответствующие верхним частотам (эти компоненты приходятся на середину отсчетов дискретного спектра и имеют индексы, близкие к $N/2$). Реально же ограниченность частотной полосы спектра непрерывного сигнала $S(\nu) = \mathcal{F}\{f(t)\}$ является некоторым допущением и определенное перекрытие частот, которым мы пренебрегаем, все же имеет место. Поэтому в вектор дискретного спектра на верхних частотах будут добавляться в общем случае не строго нулевые компоненты, а компоненты пренебрежимо малой абсолютной величины.

¹ Напомним, основной полосой частот для дискретного сигнала называют диапазон $\nu \in [0; f_d/2]$, где $f_d = 1/\Delta t$ — частота дискретизации сигнала. Основная полоса частот охватывает половину периода спектра дискретного сигнала.

Для повышения частотного разрешения $\Delta\nu = 1/T$ и, соответственно, более точного описания непрерывного спектра сигнала $\tilde{S}(\nu)$ дискретным $(\tilde{s}_0, \dots, \tilde{s}_{N-1})$, необходимо увеличение длительности интервала наблюдения $T = N\Delta t$. Но, поскольку интервал наблюдения сигнала в любом случае является конечным, при частотном анализе принципиально невозможно полностью устранить эффект утечки спектра. Для ослабления этого отрицательного эффекта, помимо увеличения длительности интервала наблюдения, применяются также весовые (оконные) функции, на которые умножается сигнал перед вычислением непрерывного или дискретного преобразования Фурье. Мы уже сталкивались с весовыми функциями при рассмотрении оконных методов синтеза КИХ-фильтров (раздел 4.3) и при изучении оконного преобразования Фурье в разделе 6.9. Как следует из рассмотренных нами ранее примеров 6.7 и 6.8, спектр применяемой оконной функции в значительной степени определяет искажения при частотном анализе сигналов, и прямоугольная форма окна в этом смысле может оказаться не самым лучшим вариантом.

Упражнение. Пусть оконная функция $w(t)$ длительности T такова, что $w(t) \geq 0$ при $-T/2 \leq t \leq T/2$ и $w(t) = 0$ для $|t| > T/2$, $f(t)$ — некоторый сигнал и $S(\nu) = \mathcal{F}\{f(t)\}$ — его спектр. Покажите, что для функции $\hat{f}(t) = w\left(t - \frac{T}{2}\right)f(t)$, представляющей собой взятую с весом $w(t)$ часть сигнала $f(t)$ на интервале $t \in [0; T]$, спектральная плотность $\hat{S}(\nu) = \mathcal{F}\{\hat{f}(t)\}$ удовлетворяет соотношению:

$$\hat{S}(\nu) = e^{-i\pi T\nu} \int_{-\infty}^{+\infty} S(u)S_w(\nu - u)e^{i\pi Tu} du = \int_{-\infty}^{+\infty} S(\nu - u)S_w(u)e^{-i\pi Tu} du, \quad (6.45)$$

где $S_w(\nu) = \mathcal{F}\{w(t)\}$ — спектр оконной функции $w(t)$. *Указание:* воспользуйтесь свойствами преобразования Фурье, приведенными в разделе 1.5.

Обобщая выражение (6.41) на случай произвольного (не обязательно прямоугольного) окна $w(t)$, спектр взвешенного гармонического колебания $\hat{f}(t) = w\left(t - \frac{T}{2}\right)\cos(2\pi\mu_0 t)$ можно представить тогда с использованием соотношения (6.45) в виде:

$$\hat{S}(v) = 0,5e^{-i\pi Tv} (S_w(v - \mu_0)e^{i\pi T\mu_0} + S_w(v + \mu_0)e^{-i\pi T\mu_0}), \quad (6.46)$$

где $S_w(v) = \mathcal{F}\{w(t)\}$ — спектр оконной функции $w(t)$.

Пример 6.9. Найти спектр весовой функции окна Ханна (или Хэннинга):

$$w(t) = \begin{cases} \frac{1}{2} \left(1 + \cos \frac{2\pi t}{T} \right), & |t| \leq T/2 \\ 0, & |t| > T/2 \end{cases}. \quad (6.47)$$

◀ Функцию (6.47) можно представить в виде $w(t) = \Pi(t)c(t)$, где

$$\Pi(t) = \begin{cases} 1, & |t| < T/2 \\ 0, & |t| \geq T/2 \end{cases}, \quad c(t) = \frac{1}{2} \left(1 + \cos \frac{2\pi t}{T} \right).$$

Поэтому спектр $S_w(v) = \mathcal{F}\{w(t)\}$ представляет собой интегральную свертку спектров $S_\Pi(v) = \mathcal{F}\{\Pi(t)\} = \frac{\sin(\pi Tv)}{\pi v}$ и $S_c(u) = \mathcal{F}\{c(t)\}$:

$$S_w(v) = \mathcal{F}\{w(t)\} = \frac{1}{\pi} \int_{-\infty}^{\infty} S_c(u) \frac{\sin(\pi T(v-u))}{v-u} du.$$

Так как (см. (1.22))

$$S_c(v) = \frac{1}{2} \mathcal{F} \left(1 + \frac{1}{2} e^{i\frac{2\pi}{T}t} + \frac{1}{2} e^{-i\frac{2\pi}{T}t} \right) = \frac{1}{2} \delta(v) + \frac{1}{4} \delta\left(v - \frac{1}{T}\right) + \frac{1}{4} \delta\left(v + \frac{1}{T}\right),$$

то окончательно имеем:

$$S_w(v) = \frac{1}{2\pi} \left[\frac{\sin(\pi Tv)}{v} + \frac{T}{2} \left(\frac{\sin \pi(Tv-1)}{Tv-1} + \frac{\sin \pi(Tv+1)}{Tv+1} \right) \right]. \quad \blacktriangleright$$

Обобщим выражение (6.43) и рассмотрим дискретные сигналы

$$\hat{f}(t) = \Delta t \sum_{k=-\infty}^{\infty} \hat{f}(k\Delta t) \delta(t - k\Delta t), \quad (6.48)$$

отсчеты которых получены из непрерывной функции

$\hat{f}(t) = w\left(t - \frac{T}{2}\right) \cos(2\pi\mu_0 t)$, где $w(t) = 0$ для $|t| > T/2$. Спектры таких

дискретных сигналов определяются спектрами $\hat{S}(v) = \mathcal{F}\{\hat{f}(t)\}$ не-

прерывных сигналов как $\hat{S}(v) = \sum_{k=-\infty}^{\infty} \hat{S}\left(v - \frac{k}{\Delta t}\right)$.

Воспользовавшись результатом решения примера 6.9 и соотношением (6.46), построим графики амплитудных спектров $|\hat{S}(v)|$ для сигналов (6.48), полученных с окном Ханна (6.47) при $\mu_0 = 3/T$, $\mu_0 = 3,5/T$ и периоде дискретизации $\Delta t = T/20$, см. рис. 6.18. Сравнение

этих графиков с графиками рис. 6.17, которые соответствуют сигналам (6.48) с прямоугольным окном и теми же параметрами Δt и μ_0 , позволяет говорить об уменьшении утечки спектра при использовании весовой функции (6.47) вместо окна прямоугольной формы. Обратим внимание на то, что в результате уменьшения средней амплитуды гармонического колебания после умножения на весовую функцию (6.47) произошло и уменьшение максимумов амплитудного спектра, приходящихся на частоты μ_0 .

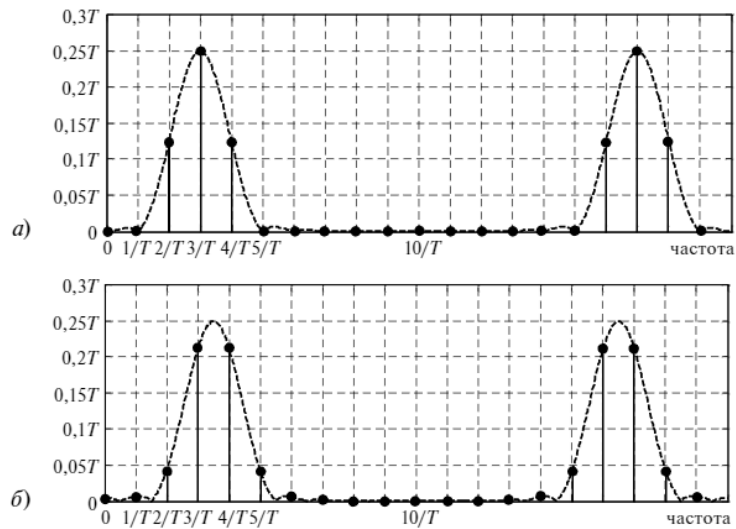


Рис. 6.18. Примеры непрерывных амплитудных спектров $|\tilde{S}(\nu)|$ (пунктирные кривые, отображен один период) сигналов (6.48) с окном (6.47) и их дискретные амплитудные спектры (обозначены точками) для $N = T/\Delta t = 20$: а) частота колебания $\mu_0 = 3/T$, б) частота колебания $\mu_0 = 3,5/T$

Таким образом, для спектрального анализа сигналов при помощи ДПФ предварительно может выполняться определенное взвешивание отсчетов дискретного сигнала, которое позволяет уменьшить утечку спектра. Общим свойством используемых весовых функций является то, что отсчеты, близкие к краям окна наблюдения сигнала, участвуют в частотном анализе с меньшими весами. Однако применение окон несколько снижает частотное разрешение

спектрального анализа, т. е. возможность отличить гармонические колебания близких частот, вследствие того, что после оконного взвешивания пики АЧХ гармонических колебаний становятся более размытыми. Расширение на частотной оси *главного лепестка* (части графика с глобальным максимумом между двумя ближайшими к нему нулевыми значениями АЧХ) — это плата за снижение уровня *боковых лепестков* (т. е. значения локального максимума АЧХ, ближайшего по величине к глобальному). Это можно наглядно наблюдать при сравнении графиков на рис. 6.17 и 6.18. С некоторыми примерами других ограниченных весовых окон, которые обеспечивают различный баланс между шириной основного лепестка и уровнем боковых лепестков, мы уже сталкивались ранее в разделе 4.3.

Упражнение. Рассмотрим сигнал $f(t) = \cos(8\pi t) + \cos(11\pi t)$, состоящий из двух гармонических колебаний с линейными частотами $\mu_1 = 4$ и $\mu_2 = 5,5$. Сигнал наблюдается на интервале $t \in [0; 1]$. Используя какой-либо программный пакет, постройте амплитудный спектр функции $\hat{f}(t) = w(t - 0,5)f(t)$, где

$$\text{а) } w(t) = \begin{cases} 1, & |t| \leq 0,5 \\ 0, & |t| > 0,5 \end{cases}, \quad \text{б) } w(t) = \begin{cases} 0,5(1 + \cos 2\pi t), & |t| \leq 0,5 \\ 0, & |t| > 0,5 \end{cases}.$$

Какой из амплитудных спектров точнее характеризует сигнал как сумму двух гармонических колебаний?

Указание: при нахождении спектра $\hat{S}(v) = \{\hat{f}(t)\}$ учтите, что

$$\hat{S}(v) = \mathcal{F}\{w(t - 0,5)\cos(8\pi t)\} + \mathcal{F}\{w(t - 0,5)\cos(11\pi t)\},$$

и воспользуйтесь соотношением (6.46).

Итак, характер вносимых в оцениваемый спектр искажений определяют форма и длительность весового окна $w(t)$, которое используется при спектральном анализе сигнала $f(t)$. Для приближенного описания спектральной плотности $S(v) = \mathcal{F}\{f(t)\}$ по ее оценке $\hat{S}(v) = \mathcal{F}\{w(t - \frac{T}{2})f(t)\}$ можно взять значения $\hat{S}(v)$ в точках $v_j = j/T$, где T — длительность интервала наблюдения сигнала (ширина окна $w(t)$), $j = 0, 1, \dots, N-1$. Выбирая число $N = T/\Delta t$ достаточно большим (чтобы период дискретизации $\Delta t = T/N$ был достаточно малым для устранения влияния эффекта наложения частот), значения отсчетов $\hat{S}(v_j)$ тогда можем найти с использованием формулы ДПФ (2.18)

по отсчетам сигнала во временной области $\hat{f}(k\Delta t) = w(k\Delta t - \frac{T}{2})f(k\Delta t)$, $k = 0, 1, \dots, N-1$. Однако при этом расстояние $\Delta\nu = 1/T$ между соседними частотными отсчетами $\hat{S}(v_j)$ не зависит от размерности N используемого ДПФ. Величина $\Delta\nu = 1/T$ может оказаться слишком большой, а полученное дискретное описание $\{\hat{S}(v_j)\}_{j=0}^{N-1}$ спектральной плотности $\hat{S}(v)$ — слишком грубым (см., например, рис. 6.18).

Для повышения точности дискретного описания спектра $\hat{S}(v)$ конечного сигнала $\{\hat{f}(k\Delta t)\}_{k=0}^{N-1}$ расположим отсчеты $\hat{s}_j = \hat{S}(v_j)$ на периоде спектральной плотности в K раз более плотно: $v_j = \frac{j}{KN\Delta t}$, $j = 0, 1, \dots, KN-1$, где $K = 1, 2, \dots$. Тогда имеем:

$$\hat{s}_j = \Delta t \sum_{k=0}^{N-1} \hat{f}(k\Delta t) e^{-2\pi i v_j k\Delta t} = \Delta t \sum_{k=0}^{N-1} \hat{f}(k\Delta t) e^{\frac{-2\pi i}{KN} jk} = \Delta t \sum_{k=0}^{KN-1} \hat{f}_k e^{\frac{-2\pi i}{KN} jk}, \quad (6.49)$$

где $\hat{f}_k = \begin{cases} \hat{f}(k\Delta t), & k = 0, 1, \dots, N-1 \\ 0, & k = N, \dots, KN-1 \end{cases}$. Видим, что выражение (6.49) представляет собой ДПФ (2.18) размерности KN , выполненное над вектором $(\hat{f}_0, \hat{f}_1, \dots, \hat{f}_{N-1}, 0, 0, \dots, 0)$, который получен дополнением $(K-1)N$ нулей в вектор исходного дискретного сигнала $(\hat{f}_0, \hat{f}_1, \dots, \hat{f}_{N-1})$.

Дополнение нулями вектора-сигнала с увеличением размерности используемого ДПФ позволяет более точно описать *искаженный* спектр $\hat{S}(v) = \mathcal{F}\{w(t - \frac{T}{2})f(t)\}$ сигнала $f(t)$ за счет более плотного расположения частотных отсчетов $\hat{S}(v_j)$ при дискретном спектральном анализе. Однако при этом не повышается реальное разрешение спектрального анализа, т.е. способность к разделению пиков в амплитудном спектре $|\hat{S}(v)|$ для сигнала $f(t)$, состоящего из двух гармонических колебаний с близкими частотами v и $v+\alpha$. Для повышения спектрального разрешения (уменьшения величины $|\alpha|$) необходимо прежде всего увеличение длительности T временного интервала наблюдения сигнала.

ГЛАВА 7

ВЕЙВЛЕТ- ПРЕОБРАЗОВАНИЯ И ИХ ПРИЛОЖЕНИЯ ДЛЯ ОБРАБОТКИ ДИСКРЕТНЫХ СИГНАЛОВ

7.1. Кратно-масштабный анализ

В данной главе будет рассматриваться гильбертово пространство $L_2(\mathbb{R})$ вещественных функций (см. пример 1.7) с нормой (1.2), индуцированной скалярным произведением

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

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: $\bigcap_{m \in \mathbb{Z}} V_m = \{\theta\}$.
- 5°. Замыкание множества всех 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 свойства 5° в виде $\lim_{m \rightarrow +\infty} V_m = L_2(\mathbb{R})$, а 4° — в виде $\lim_{m \rightarrow -\infty} V_m = \{\theta\}$, используя некоторое понятие предела последовательности подпространств. Однако такое общепринятое понятие отсутствует, и мы используем ту традиционную формулировку свойств 4° и 5°, которая приведена выше.

Лемма 7.1. Для КМА 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: $f(2^{-m}x) = \sum_{n \in \mathbb{Z}} c_n \phi(x - n)$, т.е. $f(x) = \sum_{n \in \mathbb{Z}} c_n \phi(2^m x - n)$ и система $\{\phi(2^m x - n)\}_{n \in \mathbb{Z}}$ образует базис в подпространстве V_m . В силу ортонормированности системы $\{\phi(x - n)\}_{n \in \mathbb{Z}} \subset V_0$ можем записать:

$$\int_{-\infty}^{\infty} \phi(2^m x - n) \phi(2^m x - k) dx = \frac{1}{2^m} \int_{-\infty}^{\infty} \phi(x - n) \phi(x - k) dx = \begin{cases} 1/2^m, & \text{при } k = n \\ 0, & \text{при } k \neq n \end{cases}$$

Тогда $\{\phi_{m,n}(x) = 2^{m/2} \phi(2^m x - n)\}_{n \in \mathbb{Z}}$ — ортонормированный базис в подпространстве 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:

$$\begin{aligned} h_n &= \langle \phi_{m,0}, \phi_{m+1,n} \rangle = 2^{m+1/2} \int_{-\infty}^{\infty} \phi(2^m x) \phi(2^{m+1} x - n) dx = \\ &= \sqrt{2} \int_{-\infty}^{\infty} \phi(x) \phi(2x - n) dx = \langle \phi, \phi_{1,n} \rangle. \end{aligned}$$

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 заданной 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° КМА 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 $f(x)$ не может понижаться, а $\lim_{m \rightarrow +\infty} f_m(x) = f(x)$. Сказанное иллюстрирует рис. 7.1.



Рис. 7.1. Структура вложенности подпространств КМА в пространстве $L_2(\mathbb{R})$

Обозначим 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 $a_{m,n} = a_{m,n}(f) = \langle f, \phi_{m,n} \rangle$.

Подпространство, представляющее собой ортогональное дополнение подпространства V_m до V_{m+1} (см. определение ортогонального дополнения в разделе 1.2), 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 для $m \rightarrow -\infty$ получим, что $V_N = \bigoplus_{m=-\infty}^{N-1} W_m$, откуда на основании свойства 5° КМА следует:

$$L_2(\mathbb{R}) = \overline{\bigoplus_{m \in \mathbb{Z}} W_m}. \tag{7.2}$$

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

$$\left\{ \psi_{m,n}(x) = 2^{m/2} \psi(2^m x - n) \right\}_{n \in \mathbb{Z}} \subset W_m$$

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:

$$f(x) = \lim_{M \rightarrow +\infty} \left(f_M(x) = \sum_{m=-\infty}^{M-1} y_m(x) \right) = \sum_{m \in \mathbb{Z}} y_m(x) = \sum_{m \in \mathbb{Z}} \underbrace{\sum_{n \in \mathbb{Z}} c_{m,n} \psi_{m,n}(x)}_{y_m(x) \in W_m},$$

где $f_M = A_M(f)$, $c_{m,n}(f) = \langle f, \psi_{m,n} \rangle$. 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:

$$\psi_{m,0}(x) = \sum_{n \in \mathbb{Z}} g_n \phi_{m+1,n}(x),$$

где $g_n = \langle \psi_{m,0}, \phi_{m+1,n} \rangle = \langle \psi, \phi_{1,n} \rangle$.

Позднее (в разделе 7.5) мы покажем, что коэффициенты Фурье в масштабирующих уравнениях (7.1) и (7.3) удовлетворяют одному из соотношений, $\forall n \in \mathbb{Z}: g_n = (-1)^{1-n} h_{1-n-2K}$ или $g_n = (-1)^{-n} h_{1-n-2K}$, где K — целое число, которое можно выбрать произвольно (одинаковым для всех n). 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 в разделе 7.6, 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.

Пример 7.1. Вейвлет Хаара. Масштабирующей в этом случае является функция $\phi(x) = \begin{cases} 1, & \text{при } x \in [0, 1) \\ 0, & \text{при } x \notin [0, 1) \end{cases}$. Несложно найти, что для коэффициентов уравнения (7.1) $h_n = \langle \phi, \phi_{1,n} \rangle$: $h_0 = h_1 = 1/\sqrt{2}$; $\forall k \notin \{0, 1\}$: $h_k = 0$. Поэтому масштабирующее уравнение (7.1) имеет вид:

$\phi(x) = \phi(2x) + \phi(2x - 1)$. Тогда коэффициенты для уравнения (7.3) можно выбрать так: $g_0 = h_1 = 1/\sqrt{2}$, $g_1 = -h_0 = -1/\sqrt{2}$; $\forall k \notin \{0, 1\}$: $g_k = 0$, а масштабирующее уравнение для вейвлетов (7.3) 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:

$$\phi_{m,n}(x) = \begin{cases} 2^{m/2}, & \text{при } x \in \Delta_n^m \\ 0, & \text{при } x \notin \Delta_n^m \end{cases}, \text{ где } \Delta_n^m = \left[\frac{n}{2^m}, \frac{n+1}{2^m} \right).$$

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 $S_{0,n}(v) = \mathcal{F}\{\phi_{0,n}(x)\}$ (см. раздел 1.6), 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 ширины частотной полосы (см. (1.19)) в силу свойств 2°, 3° интеграла Фурье (см. раздел 1.5) имеем:

$$\begin{aligned} \Delta v_j &= 2 \sqrt{\int_{-\infty}^{\infty} v^2 |S_{j,0}(v)|^2 dv} = 2 \sqrt{\frac{1}{2^j} \int_{-\infty}^{\infty} v^2 |S_{0,0}(v/2^j)|^2 dv} = \\ &= 2 \sqrt{2^{2j} \int_{-\infty}^{\infty} v^2 |S_{0,0}(v)|^2 dv} = 2^j \Delta v_0, \end{aligned}$$

поскольку $S_{j,0}(v) = \mathcal{F}\{2^{j/2} \phi_{0,0}(2^j x)\} = 2^{-j/2} S_{0,0}(v/2^j)$. При этом (убедитесь!) значение эффективной длительности сигнала (см. (1.20)) для базисных функций $\{\phi_{j,n}(x) = 2^{j/2} \phi_{0,0}(2^j x - n)\}$ пропорционально уменьшается: $\Delta t_j = \Delta t_0 / 2^j$.

С частотной точки зрения понижение точности проектирования функции $f(x) \in L_2(\mathbb{R})$ 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 $S(v) = \mathcal{F}\{f(x)\}$. 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 $f_{m+1} \in V_{m+1}$. Для КМА из примера 7.1 амплитудные спектры базисных функций $\{\phi_{1,k}(x)\} \in V_1$, $\{\phi_{0,k}(x)\} \in V_0$, $\{\psi_{0,k}(x)\} \in W_0$ отображены на рис. 7.2.

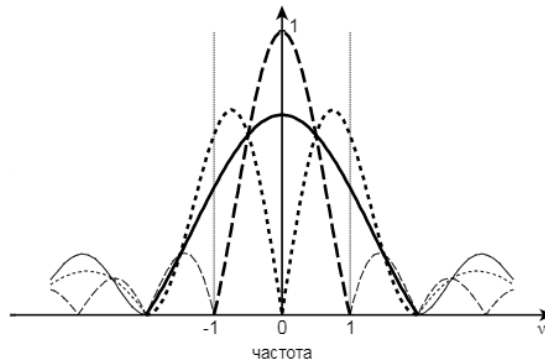


Рис. 7.2. Амплитудные спектры базисных функций КМА из примера 7.1 для подпространств V_1 (сплошная кривая), V_0 (пунктирная кривая), W_0 (точечная кривая). Утолщенными линиями выделены области основной локализации энергии функций

Рисунок 7.2 иллюстрирует общее свойство базисных функций и образуемых ими подпространств V_m , W_m , $m \in \mathbb{Z}$. 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

$$\int_{-\infty}^{+\infty} \psi(x) dx = 0.$$

Хотя для построения КМА компактность носителей масштабирующей функции $\phi(x)$ и материнского вейвлета $\psi(x)$ не является необходимым требованием, именно такие масштабирующие функции и порождаемые ими вейвелеты используются на практике в цифровой обработке сигналов. В дальнейшем мы будем полагать, что рассматриваемые нами масштабирующие функции и вейвелеты обладают ограниченными носителями.

Упражнение. Покажите, что для КМА из примера 7.1 амплитудные спектры базисных функций $\{\phi_{m,k}(x)\}_{k \in \mathbb{Z}} \subset V_m$, $\{\psi_{m,k}(x)\}_{k \in \mathbb{Z}} \subset W_m$:

$$|S_{\phi_{m,k}}(v)| = 2^{m/2} \left| \sin\left(\frac{\pi v}{2^m}\right) / (\pi v) \right|, \quad |S_{\psi_{m,k}}(v)| = 2^{m/2} \left| 2 \sin^2\left(\frac{\pi v}{2^{m+1}}\right) / (\pi v) \right|.$$

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 сколь угодно высокую заданную точность приближения исходной функции $f(x)$ ее проекцией $f_M(x)$.

¹ См., например, доказательство утверждения 5.8 в монографии [3].

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 подпространстве 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 $\phi_{N,n}(x) = \begin{cases} 2^{N/2}, & \text{при } x \in \Delta_n^N \\ 0, & \text{при } x \notin \Delta_n^N \end{cases}$, $\Delta_n^N = \left[\frac{n}{2^N}, \frac{n+1}{2^N} \right)$, и поэтому

$$a_{N,n} = \langle f, \phi_{N,n} \rangle = \int_{\Delta_n^N} f(x) \phi_{N,n}(x) dx = 2^{N/2} \int_{n/2^N}^{(n+1)/2^N} f(x) dx = 2^{-N/2} \tilde{f}_n,$$

где $\tilde{f}_n = 2^N \int_{\Delta_n^N} f(x) dx$ — среднее значение функции $f(x)$ на интервале $x \in \Delta_n^N$. Заметим, что в силу дифференцируемости $f(x)$ — непрерывна, поэтому $\forall n = 0, \dots, 2^N - 1 \exists x_n \in \Delta_n^N: \tilde{f}_n = f(x_n)$; 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 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 подынтегральной 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:

$$\varepsilon^2 = \sum_{n=0}^{2^N-1} \varepsilon_n^2 \leq h^3 \sum_{n=0}^{2^N-1} \max_{x \in \Delta_n^N} (f'(x))^2 \leq h^2 \max_{x \in [0;1]} (f'(x))^2 = h^2 \left(\max_{x \in [0;1]} |f'(x)| \right)^2,$$

$$\varepsilon = \|f - f_N\| \leq h \max_{x \in [0;1]} |f'(x)|. \quad (7.6)$$

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), 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 в общем случае 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.

Рассматриваемые нами масштабирующие функции, используемые в цифровой обработке сигналов, имеют ограниченные носители. В силу этого интеграл (7.7) можно представить в виде

$$a_{M,n} = \langle f, \phi_{M,n} \rangle = \int_{\Delta_{M,n}} f(x) \phi_{M,n}(x) dx, \quad (7.8)$$

где конечный отрезок $\Delta_{M,n} = \text{supp } \phi_{M,n}^{\Delta_{M,n}}$ — носитель функции $\phi_{M,n}(x)$.

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 сколь 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 $f(x) \forall x \in \Delta_{M,n} \exists \gamma(x) \in \Delta_{M,n}$:

$$f(x) = f(x_{M,n}) + \underbrace{f'(\gamma)(x - x_{M,n})}_{O(|\Delta_{M,n}|)},$$

т.е. $|f(x) - f(x_{M,n})| = O(|\Delta_{M,n}|) = O(2^{-M})$ и $f(x) = f(x_{M,n}) + O(2^{-M})$. 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 $f(x_{M,n})$ из-под знака интеграла в (7.8), получим:

$$\begin{aligned} a_{M,n} &= \int_{\Delta_{M,n}} f(x) \phi_{M,n}(x) dx \approx \tilde{a}_{M,n} = f(x_{M,n}) \int_{\Delta_{M,n}} \phi_{M,n}(x) dx = \\ &= f(x_{M,n}) 2^{M/2} \int_{\Delta_{M,n}} \phi(2^M x - n) dx = \frac{f(x_{M,n})}{2^{M/2}} \int_{\Delta_{M,n}} \phi(2^M x - n) d(2^M x - n) = \\ &= \frac{f(x_{M,n})}{2^{M/2}} \int_{\Delta_{M,0}} \phi(2^M x) d(2^M x) = \frac{f(x_{M,n})}{2^{M/2}} \int_{\Delta_{0,0}} \phi(x) dx = f(x_{M,n}) 2^{-M/2}, \end{aligned}$$

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 $f(x) \approx f(x+h)$:

$$a_{M,n} \approx \tilde{a}_{M,n} = f(x_{M,n}) 2^{-M/2}. \quad (7.9)$$

Оценим 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:

$$|a_{M,n} - \tilde{a}_{M,n}| \leq \frac{H^{3/2}}{2} \max_{x \in \Delta_{M,n}} |f'(x)|, \quad (7.10)$$

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 $\delta_n \leq \sqrt{H} |f(\xi_n) - f(x_{M,n})| \leq \frac{H^{3/2}}{2} \max_{x \in \Delta_{M,n}^*} |f'(x)|$. ►

Следствие. Если носитель $\Delta_{0,0}$ масштабирующей функции $\phi(x)$ конечен, то в пространстве $L_2[a; b]$ 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.

◀ Для упрощения технических выкладок будем рассматривать пространство $L_2[0; 1]$ и положим, что ограниченный носитель $\Delta_{0,0} = \text{supp} \phi_{0,0}(x) \subset [0; K]$, 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

$$f_M - \tilde{f}_M = \sum_{n=-K}^{2^M} a_{M,n} \phi_{M,n}(x) - \sum_{n=-K}^{2^M} \tilde{a}_{M,n} \phi_{M,n}(x) = \sum_{n=-K}^{2^M} (a_{M,n} - \tilde{a}_{M,n}) \phi_{M,n}(x),$$

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, при достаточно малом шаге дискретизации $h = x_{M,n+1} - x_{M,n} = 2^{-M}$ коэффициенты разложения (7.5)

дифференцируемой функции $f(x)$ можно трактовать как масштабированные отсчеты дискретного сигнала: $a_{M,n} \approx \sqrt{h}f(x_{M,n})$, где $x_{M,n} \in \Delta_{M,n} = \text{supp}\phi_{M,n}(x)$, 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)) на конечном отрезке в норме $L_2[a; b]$ погрешность $\delta = \|f_M - \tilde{f}_M\| = O(h)$.

Отметим, что общие оценки (7.10) и (7.11) в ряде случаев можно улучшить, повысив 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:

$$\delta_n = \left| \underbrace{\frac{1}{\sqrt{h}} \int_{nh}^{(n+1)h} f(x) dx}_{a_{N,n}} - \underbrace{\frac{f((n+0,5)h)}{\sqrt{h}} \int_{nh}^{(n+1)h} dx}_{\tilde{a}_{N,n}} \right| = \frac{1}{\sqrt{h}} \left| \int_{nh}^{(n+1)h} (f(x) - f(x_n)) dx \right|,$$

где $x_n = (n + 0,5)h$ — середина интервала $\Delta_{N,n} = \Delta_n^N = [nh, (n+1)h)$. 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:

$$\delta = \|f_N - \tilde{f}_N\| = O(h^2).$$

Таким образом, замена точного проектирования приближенным вносит дополнительную погрешность $\delta = O(h^2)$ более высокого порядка малости, чем порядок малости погрешности точного проектирования (7.6): $\varepsilon = O(h)$. 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. ►

Упражнение. Для функции

$$f(t) = \begin{cases} \sin \frac{\pi t}{5}, & t \in [0; 5] \\ 0, & t \notin [0; 5] \end{cases}$$

и КМА из примера 7.1 найдите:

- точные значения коэффициентов (7.8) для проекций (7.5) $f_0(t) \in V_0$, $f_1(t) \in V_1$;
- приближенные (с использованием формулы (7.9)) значения коэффициентов для проекций $\tilde{f}_0(t) \in V_0$, $\tilde{f}_1(t) \in V_1$;
- ошибки приближенного проектирования $\delta^{(0)} = \|f_0 - \tilde{f}_0\|$, $\delta^{(1)} = \|f_1 - \tilde{f}_1\|$. Сравните полученные значения ошибок с нормой $\|f\| = \sqrt{2,5}$ и найденными с использованием равенства Парсеваля — Стеклова (см. теорему 1.8) нормами $\|f_0\|$, $\|\tilde{f}_0\|$, $\|f_1\|$, $\|\tilde{f}_1\|$.

7.3. Вычисление дискретных 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 $f_m(x) \in V_m$ 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 $f_m(x)$, $f_{m-1}(x)$, $y_{m-1}(x)$ 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),

$$\begin{aligned} \phi_{m-1,k}(x) &= 2^{(m-1)/2} \phi(2^{m-1}x - k) = 2^{(m-1)/2} \sqrt{2} \sum_{j \in \mathbb{Z}} h_j \phi(2(2^{m-1}x - k) - j) = \\ &= \sum_{j \in \mathbb{Z}} h_j 2^{m/2} \phi(2^m x - j - 2k) = \sum_{j \in \mathbb{Z}} h_j \phi_{m,j+2k}(x) = \sum_{j \in \mathbb{Z}} h_{j-2k} \phi_{m,j}(x). \end{aligned}$$

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 $\phi_{m-1,k}(x)$, с учетом $\sum_{n \in \mathbb{Z}} a_{m-1,n} \langle \phi_{m-1,n}(x), \phi_{m-1,k}(x) \rangle = a_{m-1,k}$, $\sum_{n \in \mathbb{Z}} c_{m-1,n} \langle \psi_{m-1,n}(x), \phi_{m-1,k}(x) \rangle = 0$, подставим в полученное уравнение представление (7.17) для функции $\phi_{m-1,k}(x)$, тогда имеем

$$\begin{aligned} a_{m-1,k} &= \sum_{n \in \mathbb{Z}} a_{m,n} \langle \phi_{m,n}(x), \sum_{j \in \mathbb{Z}} h_{j-2k} \phi_{m,j}(x) \rangle = \\ &= \sum_{n \in \mathbb{Z}} a_{m,n} \sum_{j \in \mathbb{Z}} h_{j-2k} \langle \phi_{m,n}(x), \phi_{m,j}(x) \rangle. \end{aligned}$$

В силу ортонормированности системы $\{\phi_{m,k}(x)\}_{k \in \mathbb{Z}} \subset V_m$ получим окончательно $\forall k$

$$a_{m-1,k} = \sum_{n \in \mathbb{Z}} a_{m,n} h_{n-2k}. \tag{7.19}$$

Аналогично, умножив уравнение (7.16) скалярно на функцию $\psi_{m-1,k}(x)$ и затем представив последнюю при помощи выражения (7.18), получаем $\forall k$

$$c_{m-1,k} = \sum_{n \in \mathbb{Z}} a_{m,n} g_{n-2k}. \tag{7.20}$$

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 коэффициенты 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 (7.20) также конечны. Алгоритм соответствующих вычислений можно изобразить в виде схемы, приведенной на рис. 7.3.

На практике число шагов алгоритма вейвлет-преобразования (число повторений процедур (7.19), (7.20)) должно быть конечно.

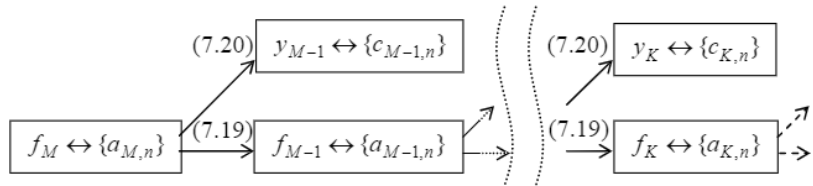


Рис. 7.3. Схема вычисления дискретного вейвлет-преобразования

Для некоторого числа $m = K$ вычисления прекращаются, т.е. вместо разложения (7.15) мы ограничиваемся представлением:

$$\begin{aligned} f_M(x) &= \underbrace{\sum_{n=-\infty}^{\infty} a_{K,n} \phi_{K,n}(x)}_{f_K(x) \in V_K} + \underbrace{\sum_{m=K}^{M-1} \sum_{n=-\infty}^{\infty} c_{m,n} \psi_{m,n}(x)}_{y_m(x) \in W_m} = \\ &= f_K(x) + y_K(x) + y_{K+1}(x) + \dots + y_{M-1}(x). \end{aligned} \quad (7.21)$$

Поскольку из представления (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 $\{a_{K,n}\}$ в (7.21) не отбрасывают. Число K обычно выбирается так, чтобы 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 считать некоррелированными. Так, если L — длина интервала корреляции для сигнала $f(x)$ (т.е. $\forall x_0, \forall \tau > L$ значения функции $f(x_0)$ и $f(x_0 + \tau)$ считаются некоррелированными), то выбирать 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 декомпозиции, 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 своим коэффициентам Фурье разложениям

$$f_{m-1}(x) = \sum_{n \in \mathbb{Z}} a_{m-1,n} \phi_{m-1,n}(x), \quad y_{m-1}(x) = \sum_{n \in \mathbb{Z}} c_{m-1,n} \psi_{m-1,n}(x)$$

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 $\phi_{m,k}(x)$:

$$\sum_{n \in \mathbb{Z}} a_{m,n} \langle \phi_{m,n}(x), \phi_{m,k}(x) \rangle =$$

$$= \sum_{n \in \mathbb{Z}} a_{m-1,n} \left\langle \underbrace{\sum_{j \in \mathbb{Z}} h_{j-2n} \phi_{m,j}(x)}_{\phi_{m-1,n}(x)}, \phi_{m,k}(x) \right\rangle + \sum_{n \in \mathbb{Z}} c_{m-1,n} \left\langle \underbrace{\sum_{j \in \mathbb{Z}} g_{j-2n} \phi_{m,j}(x)}_{\psi_{m-1,n}(x)}, \phi_{m,k}(x) \right\rangle.$$

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. 7.4).

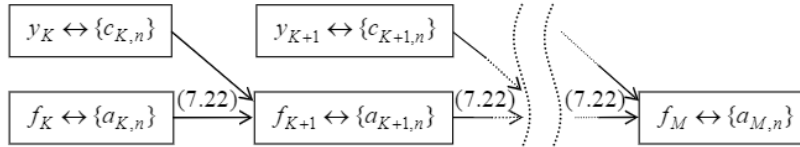


Рис. 7.4. Схема вычисления обратного дискретного вейвлет-преобразования

7.4. Квадратурно-зеркальные фильтры

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 (3.9), 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. 7.5.

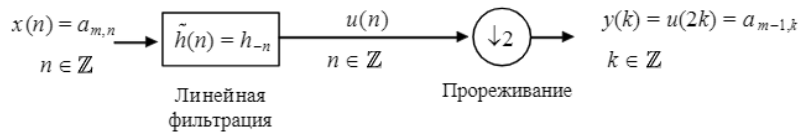


Рис. 7.5. Схема реализации процедуры (7.19) при помощи дискретного фильтра и последующего прореживания последовательности

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 индексы элементов обрабатываемых последовательностей не всегда обязательно отождествляются 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 $n < 0$ 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. 7.3) можно представить как параллельную обработку фильтрами с импульсными характеристиками $\{\tilde{h}(n) = h_{-n}\}$, $\{\tilde{g}(n) = g_{-n}\}$ и последующее прореживание выходных последовательностей, см. рис. 7.6.

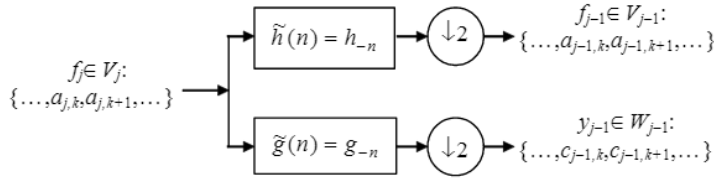


Рис. 7.6. Реализация одного шага вейвлет-преобразования (7.19), (7.20)

Пример 7.4. Для КМА из примера 7.1 найти выходные последовательности в схеме обработки рис. 7.6 при входном воздействии $\{x(0), x(1), \dots, x(2L-1)\}$.

◀ Воспользовавшись результатами примера 7.1, получаем следующие

$$\text{ИХ: } \tilde{h}(n) = \left\{ \dots, 0, 0, \underbrace{1/\sqrt{2}}_{\tilde{h}(-1)}, \underbrace{1/\sqrt{2}}_{\tilde{h}(0)}, 0, 0, \dots \right\}, \quad \tilde{g}(n) = \left\{ \dots, 0, 0, \underbrace{-1/\sqrt{2}}_{\tilde{g}(-1)}, \underbrace{1/\sqrt{2}}_{\tilde{g}(0)}, 0, 0, \dots \right\}.$$

Отклики фильтров находим по формулам свертки, которые примут вид:

$$u(k) = \sum_{n=-1}^0 \tilde{h}(n)x(k-n) = \frac{x(k) + x(k+1)}{\sqrt{2}},$$

$$v(k) = \sum_{n=-1}^0 \tilde{g}(n)x(k-n) = \frac{x(k) - x(k+1)}{\sqrt{2}}.$$

После прореживания, сохраняющего лишь четные номера выходных последовательностей, $\{u(2k)\}$ и $\{v(2k)\}$, окончательно получим из одной входной последовательности $\{x(0), x(1), \dots, x(2L-1)\}$ две вы-

ходные: $\left\{ \frac{x(2k) + x(2k+1)}{\sqrt{2}} \right\}_{k=0}^{L-1}, \quad \left\{ \frac{x(2k) - x(2k+1)}{\sqrt{2}} \right\}_{k=0}^{L-1}.$ ▶

Повторив последовательно преобразования по схеме, приведенной на рис. 7.6, 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 $\{a_{k,k}\}$, см. рис. 7.7.

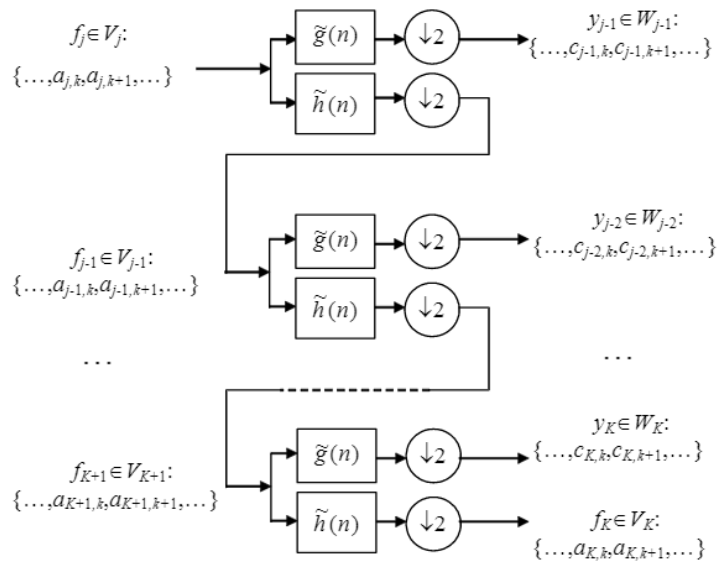


Рис. 7.7. Вычисление вейвлет-преобразования при помощи блоков фильтров

Упражнение. Пусть $\mathbf{F}_M = (f_M(0), f_M(1), \dots, f_M(7))^T$ — некоторый вектор-столбец, а $\mathbf{F}_{M-1} = (f_{M-1}(0), f_{M-1}(1), f_{M-1}(2), f_{M-1}(3))^T$ и $\mathbf{Y}_{M-1} = (y_{M-1}(0), y_{M-1}(1), y_{M-1}(2), y_{M-1}(3))^T$ — векторы, которые получены из \mathbf{F}_M по схеме обработки из примера 7.4, т. е.

$$\left\{ f_{M-1}(k) = \frac{x(2k) + x(2k+1)}{\sqrt{2}} \right\}_{k=0}^3, \quad \left\{ y_{M-1}(k) = \frac{x(2k) - x(2k+1)}{\sqrt{2}} \right\}_{k=0}^3.$$

А) Запишите матрицу \mathbf{H}_1 (размерности 8×8), определяющую преобразование

$$\begin{pmatrix} \mathbf{F}_{M-1} \\ \mathbf{Y}_{M-1} \end{pmatrix} = \begin{pmatrix} f_{M-1}(0) \\ f_{M-1}(1) \\ f_{M-1}(2) \\ f_{M-1}(3) \\ y_{M-1}(0) \\ y_{M-1}(1) \\ y_{M-1}(2) \\ y_{M-1}(3) \end{pmatrix} = \mathbf{H}_1 \begin{pmatrix} f_M(0) \\ f_M(1) \\ f_M(2) \\ f_M(3) \\ f_M(4) \\ f_M(5) \\ f_M(6) \\ f_M(7) \end{pmatrix} = \mathbf{H}_1 \mathbf{F}_M.$$

Б) Используя результаты решения пункта А, покажите, что трех-уровневое преобразование вектора $\mathbf{F}_M = (f_M(0), f_M(1), \dots, f_M(7))^T$ в вектор

$$\mathbf{U} = (f_{M-3}(0), y_{M-3}(0), y_{M-2}(0), y_{M-2}(1), y_{M-1}(0), y_{M-1}(1), y_{M-1}(2), y_{M-1}(3))^T$$

по схеме рис. 7.7 с фильтрами из примера 7.4 можно представить в матричном виде как $\mathbf{U} = \mathbf{H}\mathbf{F}_M$, где $\mathbf{H} = \mathbf{H}_3\mathbf{H}_2\mathbf{H}_1$, и найдите матрицы \mathbf{H}_3 , \mathbf{H}_2 . Убедитесь, что матрица \mathbf{H} определяет 8-точечное дискретное преобразование Хаара (ДПХ), а представление $\mathbf{H} = \mathbf{H}_3\mathbf{H}_2\mathbf{H}_1$ — быстрый алгоритм ДПХ (см. раздел 2.10).

Рассмотрим теперь формулу (7.22), которая обращает процедуру, определяемую формулами (7.19), (7.20) и схемой рис. 7.6. 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:

$$S_2(k) = \sum_{n=-\infty}^{\infty} c_{m-1,n} g_{k-2n} = \sum_{j=-\infty}^{\infty} d(j) g(k-j),$$

$$\text{где } g_m = g(m), \quad d(j) = \begin{cases} c_{m-1,j/2}, & \text{если } j \text{ — четное число} \\ 0, & \text{если } j \text{ — нечетное число} \end{cases}.$$

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 $g(n)$, см. рис. 7.8.

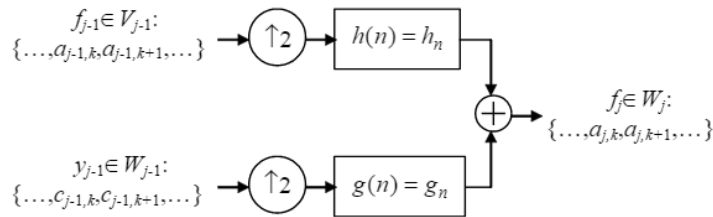


Рис. 7.8. Реализация процедуры (7.22) при помощи дискретных фильтров

Схема обработки, представленная на рис. 7.8, реализует обратную процедуру по сравнению со схемой обработки последовательности, приведенной ранее на рис. 7.6. Поэтому обращение многоуровневого преобразования, схема которого отражена на рис. 7.7, также достигается последовательным соединением соответствующих структурных блоков фильтров, см. рис. 7.9.

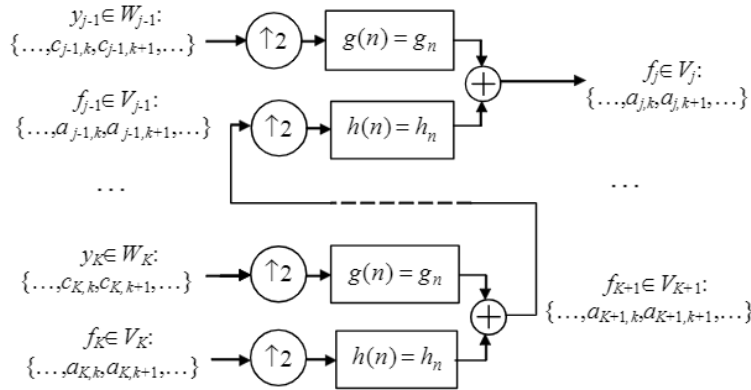


Рис. 7.9. Вычисление обратного вейвлет-преобразования при помощи блоков зеркальных фильтров

Для импульсных характеристик фильтров на рис. 7.6 и 7.8 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 $f_m \in V_m$, $y_m \in W_m$, можно трактовать как 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. 7.10.

Прежде всего, заметим, что если $K_H(\omega) = \sum_{k=-\infty}^{\infty} h(k)e^{-i\omega k}$ — ЧХ фильтра (см. раздел 3.6) с вещественной ИХ $h(n) = h_n$, то для зеркального фильтра, очевидно, ЧХ комплексно сопряжена:

$$K_{\tilde{H}}(\omega) = \sum_{k=-\infty}^{\infty} h(-k)e^{-i\omega k} = \sum_{k=-\infty}^{\infty} h(k)e^{i\omega k} = \overline{K_H(\omega)} = K_H(-\omega).$$

Аналогично, $K_G(\omega) = \sum_{k=-\infty}^{\infty} g(k)e^{-i\omega k}$, $K_{\tilde{G}}(\omega) = \overline{K_G(\omega)} = K_G(-\omega)$. Таким образом, АЧХ зеркальных фильтров одинаковы, а ФЧХ отличаются знаком.

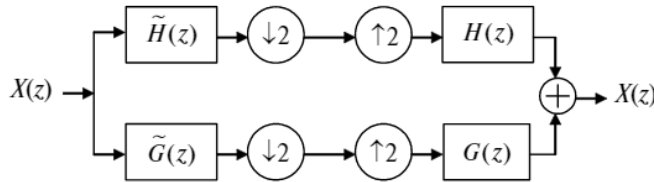


Рис. 7.10. Общая схема системы обработки сигналов при помощи КЗФ

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 1.5) его применение для масштабирующего уравнения (7.1) дает (убедитесь!):

$$S_{\phi}(\omega) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} h_n e^{-i\omega n/2} S_{\phi}(\omega/2) = \frac{1}{\sqrt{2}} K_H(\omega/2) S_{\phi}(\omega/2), \quad (7.25)$$

где спектр $S_{\phi}(\omega) = \int_{-\infty}^{\infty} \phi(t) e^{-i\omega t} dt$ рассматривается для циклической частоты ω .

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 $I(u)$.

$$\begin{aligned} S_I(\omega) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(t) \phi(t-u) dt e^{-i\omega u} du = \int_{-\infty}^{\infty} \phi(t) \int_{-\infty}^{\infty} \phi(t-u) e^{-i\omega u} du dt = \\ &= \int_{-\infty}^{\infty} \phi(t) e^{-i\omega t} \int_{-\infty}^{\infty} \phi(t-u) e^{-i\omega(u-t)} du dt = \int_{-\infty}^{\infty} \phi(t) e^{-i\omega t} \int_{-\infty}^{\infty} \phi(x) e^{i\omega x} dx dt = \\ &= \int_{-\infty}^{\infty} \phi(t) e^{-i\omega t} dt \int_{-\infty}^{\infty} \phi(x) e^{i\omega x} dx = S_{\phi}(\omega) \overline{S_{\phi}(\omega)} = |S_{\phi}(\omega)|^2. \end{aligned}$$

Выполняя затем обратное преобразование Фурье, получаем $\forall n \in \mathbb{Z}$:

$$\begin{aligned} I(n) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |S_{\phi}(\omega)|^2 e^{i\omega n} d\omega = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \int_{2\pi j}^{2\pi(j+1)} |S_{\phi}(\omega)|^2 e^{i\omega n} d\omega = \\ &= \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \int_0^{2\pi} |S_{\phi}(\omega + 2\pi j)|^2 e^{i(\omega+2\pi j)n} d\omega = \frac{1}{2\pi} \int_0^{2\pi} \sum_{j=-\infty}^{\infty} |S_{\phi}(\omega + 2\pi j)|^2 e^{i\omega n} d\omega. \end{aligned}$$

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 (1.16) равны

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} F(\omega) e^{-i\omega k} d\omega = I(-k) = \begin{cases} 1, k=0 \\ 0, k \neq 0 \end{cases}.$$

Следовательно,

$$F(\omega) = \sum_{j \in \mathbb{Z}} |S_{\phi}(\omega + 2\pi j)|^2 = 1 \quad (7.26)$$

(возможно, за исключением отдельных точек $\omega \in \mathbb{R}$, т.е. почти всюду).

Запишем формулу (7.26) для функции $F(2\omega) = 1$ и воспользуемся соотношением (7.25):

$$\begin{aligned}
 1 &= \sum_{j \in \mathbb{Z}} |S_{\phi}(2\omega + 2\pi j)|^2 = \frac{1}{2} \sum_{j \in \mathbb{Z}} |K_H(\omega + \pi j) S_{\phi}(\omega + \pi j)|^2 = \\
 &= \frac{1}{2} \left(\sum_{j \in \mathbb{Z}} |K_H(\omega + 2\pi j) S_{\phi}(\omega + 2\pi j)|^2 + \sum_{j \in \mathbb{Z}} |K_H(\omega + (2j+1)\pi) S_{\phi}(\omega + (2j+1)\pi)|^2 \right).
 \end{aligned}$$

С учетом 2π -периодичности ЧХ $K_H(\omega)$ получим:

$$2 = |K_H(\omega)|^2 \underbrace{\sum_{j \in \mathbb{Z}} |S_{\phi}(\omega + 2\pi j)|^2}_{F(\omega)=1} + |K_H(\omega + \pi)|^2 \underbrace{\sum_{j \in \mathbb{Z}} |S_{\phi}(\omega + \pi + 2\pi j)|^2}_{F(\omega+\pi)=1},$$

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} , причем $0 \leq |K_H(\omega)| = |K_{\tilde{H}}(\omega)| \leq \sqrt{2}$.

Проинтегрируем уравнение (7.1):

$$\int_{-\infty}^{\infty} \phi(x) dx = \sqrt{2} \sum_{n \in \mathbb{Z}} h_n \int_{-\infty}^{\infty} \phi(2x - n) dx = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} h_n \int_{-\infty}^{\infty} \phi(x - n) dx = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} \phi(x) dx \sum_{n \in \mathbb{Z}} h_n.$$

Отсюда, в силу (7.4), получаем $\forall j \in \mathbb{Z}$:

$$K_H(2\pi j) = \sum_{k=-\infty}^{\infty} h(k) e^{-i2\pi j k} = \sum_{k=-\infty}^{\infty} h(-k) e^{-i2\pi j k} = K_{\tilde{H}}(2\pi j) = \sum_{k \in \mathbb{Z}} h_k = \sqrt{2}. \quad (7.28)$$

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. раздел 3.5). 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. 7.10. Предварительно установим выражение для Z -образа сигнала, у которого отсчеты с нечетными номерами принудительно обнулены, т.е. сигнал подвергнут обработке по следующей схеме (рис. 7.11).

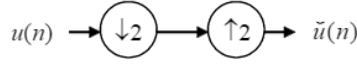


Рис. 7.11. Схема обработки, при которой обнуляются отсчеты сигнала с нечетными номерами

Очевидно, что (см. рис. 7.11)

$$\tilde{U}(z) = \sum_{n \in \mathbb{Z}} \tilde{u}(n)z^{-n} = \sum_{n \in \mathbb{Z}} u(2n)z^{-2n} = \sum_{n \in \mathbb{Z}} u(n) \frac{z^{-n} + (-z)^{-n}}{2} = \frac{U(z) + U(-z)}{2}.$$

Поэтому для точного восстановления сигнала при обработке по схеме рис. 7.10 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:

$$X(z)(H(z^{-1})H(z) + G(z^{-1})G(z)) + X(-z)(H(-z^{-1})H(z) + G(-z^{-1})G(z)) = 2X(z).$$

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

$$H(e^{i\omega}) = K_H(\omega), \quad H(e^{-i\omega}) = K_H(-\omega) = \sum_{n \in \mathbb{Z}} h(n)e^{i\omega n} = \overline{K_H(\omega)},$$

$$H(-e^{-i\omega}) = H(e^{-i(\omega+\pi)}) = \overline{K_H(\omega + \pi)},$$

и проводя аналогичную замену переменной $z = e^{i\omega}$ для G -фильтра, получим из (7.30) условия точного восстановления сигнала в терминах ЧХ:

$$\begin{cases} |K_H(\omega)|^2 + |K_G(\omega)|^2 = 2 \\ \overline{K_H(\omega + \pi)K_H(\omega)} + \overline{K_G(\omega + \pi)K_G(\omega)} = 0 \end{cases} \quad (7.31)$$

Равенства (7.31) должны выполняться для любых значений частоты ω . Из (7.27) и первого уравнения системы (7.31) следует, что

$$|K_G(\omega)| = |K_H(\omega + \pi)|.$$

Отсюда с учетом (7.29) 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 (7.28) имеем $\forall j \in \mathbb{Z}$:

$$K_{\tilde{G}}((2j+1)\pi) = K_G((2j+1)\pi) = \sum_{k \in \mathbb{Z}} (-1)^k g_k = \pm\sqrt{2}, \quad (7.33)$$

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), (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 $\forall \omega$: $b(\omega + \pi) - b(\omega) = (2k - 1)\pi$, $k \in \mathbb{Z}$. Полученному уравнению отвечает линейная функция $b(\omega) = (2k - 1)\omega + C$, где k — любое целое число, а C — произвольная вещественная константа, поэтому положим

$$K_G(\omega) = e^{i((2k-1)\omega+C)} \overline{K_H(\omega + \pi)}.$$

Отсюда

$$\begin{aligned} K_G(\omega) &= e^{i((2k-1)\omega+C)} \sum_{n \in \mathbb{Z}} h_n e^{in(\omega+\pi)} = e^{iC} \sum_{n \in \mathbb{Z}} h_n (-1)^n e^{i(2k+n-1)\omega} = \\ &= e^{iC} \sum_{n \in \mathbb{Z}} h_{n-2k+1} (-1)^{n-2k+1} e^{i\omega n} = e^{iC} \sum_{n \in \mathbb{Z}} h_{n-2k+1} (-1)^{n+1} e^{i\omega n} = \\ &= e^{iC} \sum_{n \in \mathbb{Z}} h_{1-n-2k} (-1)^{1-n} e^{-i\omega n} = \sum_{n \in \mathbb{Z}} g_n e^{-i\omega n}. \end{aligned}$$

Для того чтобы последнее равенство выполнялось всегда для любой частоты ω , коэффициенты ИХ G -фильтра должны

удовлетворять соотношению связи: $g_n = e^{iC} h_{1-n-2k} (-1)^{1-n}$, $n \in \mathbb{Z}$. 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 получаем, что условие точного восстановления сигнала после фильтрации по приведенной схеме рис. 7.10 устанавливает следующую 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:

$$G(z) = \alpha z^{2k-1} H(-z^{-1}). \quad (7.36)$$

Упражнение. Выразите передаточную функцию $\tilde{G}(z)$ через $H(z)$.

Соотношение (7.34) уже использовалось ранее при построении вейвлета Хаара в примере 7.1. Отметим, что выбор различных значений k в соотношениях (7.34)–(7.36) соответствует целочисленному смещению по оси абсцисс функции материнского вейвлета $\psi(x)$ и влечет лишь сдвиг нумерации системы функций $\{\psi_{0,n}(x) = \psi(x-n)\}_{n \in \mathbb{Z}} \subset W_0$.

Упражнение. Покажите, что если материнский вейвлет $\psi(x) = \psi^{(0)}(x)$ задан набором коэффициентов уравнения (7.3) $\{g_n = \alpha h_{1-n} (-1)^{1-n}\}$, то коэффициенты $\{g_n = \alpha h_{1-n-2k} (-1)^{1-n}\}$, $K \in \mathbb{Z}$, определяют материнский вейвлет $\psi(x) = \psi^{(0)}(x + K)$.

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

$$|K_{\tilde{H}}(\omega)| = |K_H(\omega)| = \left| \frac{1}{\sqrt{2}} + \frac{e^{-i\omega}}{\sqrt{2}} \right| = \sqrt{2} \left| e^{-\frac{i\omega}{2}} \right| \left| \frac{e^{i\omega/2} + e^{-i\omega/2}}{2} \right| = \sqrt{2} \left| \cos \frac{\omega}{2} \right| = \sqrt{1 + \cos \omega},$$

$|K_{\tilde{G}}(\omega)| = |K_G(\omega)| = |K_H(\omega + \pi)| = \sqrt{1 + \cos(\omega + \pi)} = \sqrt{1 - \cos\omega} = \sqrt{2} \left| \sin \frac{\omega}{2} \right|$.
 Графики АЧХ приведены на рис. 7.12. ►

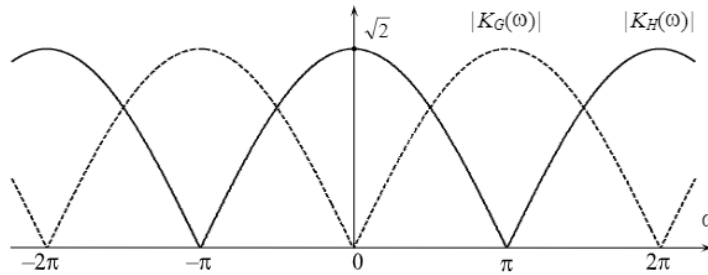


Рис. 7.12. Графики АЧХ для КЗФ, порожденных вейвлетом Хаара

Данный пример наглядно иллюстрирует отмеченное ранее общее свойство КЗФ: \tilde{H} и H являются НЧ-фильтрами, а \tilde{G} и G — фильтрами ВЧ. Процедуру фильтрации дискретного сигнала парой фильтров рис. 7.6 с частотной точки зрения можно пояснить как «расщепление» частотного спектра дискретного сигнала — последовательности $\{..., a_{j,k}, a_{j,k+1}, \dots\}$ — на два частотных диапазона равной ширины: низкочастотного (отклик \tilde{H} -фильтра) и высокочастотного (отклик \tilde{G} -фильтра).

Упражнение. Покажите, что для ФЧХ КЗФ, определяемых вейвлетом Хаара (см. пример 7.1), на периоде $\omega \in (-\pi; \pi]$ справедливы утверждения:

- а) $\varphi_H(\omega) = \arg K_H(\omega) = -\omega/2$ для $\omega \in (-\pi; \pi)$, а при $\omega = \pi$ значение $K_H(\omega)$ не определено;
- б) полагая в (7.35) $k = 0, \alpha = 1$, имеем для ФЧХ

$$\varphi_G(\omega) = \arg K_G(\omega) = \begin{cases} -\frac{\omega - \pi}{2}, & \text{при } \omega \in (-\pi; 0) \\ \text{не определена} & \text{при } \omega = 0. \\ -\frac{\omega + \pi}{2}, & \text{при } \omega \in (0; \pi] \end{cases}$$

Постройте графики $\varphi_H(\omega), \varphi_G(\omega)$.

7.6. Построение масштабирующих функций и вейвлетов по масштабирующим уравнениям

Как отмечено 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 $\{h_n\}$ при помощи соотношения (7.34), то задание набора $\{h_n\}$ оказывается достаточным также и для построения функции вейвлета.

Наше рассмотрение ограничено КМА, порождаемым масштабирующей функцией $\phi(x)$ с компактным носителем $\Delta_{0,0} = \text{supp } \phi(x) = [a; b]$. В этом случае количество слагаемых в масштабирующем уравнении (7.1) должно быть конечно:

$$\phi(x) = \sqrt{2} \sum_{n=N_1}^{N_2} h_n \phi(2x - n),$$

так как с ненулевыми коэффициентами h_n в сумму должны входить только те функции $\phi_{1,n}(x) = \sqrt{2} \phi(2x - n)$, носители которых $\Delta_{1,n} = \text{supp } \phi_{1,n}(x) = \left[\frac{a+n}{2}; \frac{b+n}{2} \right]$ образуют покрытие отрезка $\Delta_{0,0} = \text{supp } \phi(x) = [a; b]$, т. е.

$$\Delta_{0,0} = [a; b] = \bigcup_{n=N_1}^{N_2} \left[\frac{a+n}{2}; \frac{b+n}{2} \right] = \bigcup_{n=N_1}^{N_2} \Delta_{1,n}.$$

Отсюда $a = \frac{a+N_1}{2}$, $b = \frac{b+N_2}{2}$. Следовательно, если $\phi(x)$ имеет ограниченный носитель $\Delta_{0,0} = [a; b]$, то его границы должны быть целыми числами: $a = N_1$, $b = N_2$.

Занумеруем элементы системы функций $\{\phi(x - n)\}_{n=-\infty}^{+\infty} \in V_0$ так, чтобы для носителя $\Delta_{0,0} = \text{supp } \phi(x)$ масштабирующей функции $\phi(x) = \phi_{0,0}(x)$ получить $N_1 = 0$. Обозначив $N_2 = M - 1$, где целое число $M \geq 2$, для масштабирующей функции $\phi(x)$ с носителем $\Delta_{0,0} = [0; M - 1]$ окончательно получаем масштабирующее уравнение (7.1) в виде:

$$\phi(x) = \sqrt{2} \sum_{n=0}^{M-1} h_n \phi(2x - n). \tag{7.37}$$

Далее будем считать, что число M в уравнении (7.37) — четное. (При необходимости в сумму (7.37) можно ввести дополнительное слагаемое с коэффициентом $h_{M-1} = 0$.) Тогда, положив в (7.34) $2k = 2 - M$, 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, запись (7.38) означает также, что $\text{supp } \psi(t) \subset [0; M-1]$.

Рассмотрим два метода нахождения функции $\phi(x)$.

Первый метод. Из формул (7.9) и (7.12) следует, что для дифференцируемой функции верно:

$$f(x_{N,n}) 2^{-N/2} = a_{N,n} + O(2^{-3N/2}),$$

где $a_{N,n} = \langle f, \phi_{N,n} \rangle$, $x_{N,n}$ — любая точка из отрезка $\Delta_{N,n} = \text{supp } \phi_{N,n}(x) = \left[\frac{n}{2^N}; \frac{M-1+n}{2^N} \right]$. Поэтому, положив $x_{N,n} = n/2^N$, для отсчетов функции $f(x_{N,n}) = \phi(x_{N,n})$, взятых с шагом $h = 1/2^N$, можно записать:

$$\phi(n/2^N) = 2^{N/2} a_{N,n} + O(2^{-N}). \quad (7.39)$$

Отсюда видно, что, имея набор коэффициентов $\{a_{N,n}\}_{n \in \mathbb{Z}}$, который определяет проекцию $f_N = A_N(\phi) \in V_N$, можно с погрешностью порядка $O(h)$ найти значения функции $\phi(x)$ в точках $x_{N,n} = n/2^N$. То есть, увеличивая N , можно сколь угодно точно описать функцию $\phi(x)$: погрешность определения значения функции в отсчетах $\phi(x_{N,n})$ по формуле (7.39) будет уменьшаться, а сами отсчеты будут располагаться на числовой оси все более часто.

С учетом соотношения (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 $\phi(x) = \phi_{0,0}(x)$ — функция ортогонального базиса подпространства V_0 , то для разложения (7.5) функции $\phi(x)$ в подпространстве V_0 получим $a_{0,0} = 1$, а $\forall k \neq 0: a_{0,k} = 0$. При этом $\forall k: c_{0,k} = \langle \phi(x), \psi_{0,k}(x) \rangle = 0$ (так как $\phi(x) \perp W_0$). Поэтому с использованием формулы (7.22) можно найти коэффициенты $\{a_{1,k}\}_{k \in \mathbb{Z}}$ разложения масштабирующей функции $\phi(x) \in V_0$ по базису подпространства V_1 , получим: $a_{1,k} = h_k, k \in \mathbb{Z}$.

Рекуррентно применяя далее для $m = 2, 3, \dots, N$ полученную из (7.22) формулу

$$a_{m,k} = \sum_{n \in \mathbb{Z}} a_{m-1,n} h_{k-2n},$$

можно построить сколь угодно точное описание масштабирующей функции $\phi(x)$ через ее отсчеты в точках $x_{N,n} = n/2^N$. Для этого необходимо воспользоваться выражением (7.39).

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

$$\mathbf{B} = \sqrt{2} \begin{pmatrix} h_0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ h_2 & h_1 & h_0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ h_4 & h_3 & h_2 & h_1 & h_0 & \dots & 0 & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ h_{M-4} & h_{M-5} & h_{M-6} & h_{M-7} & h_{M-8} & \dots & h_1 & h_0 & 0 & 0 & 0 \\ h_{M-2} & h_{M-3} & h_{M-4} & h_{M-5} & h_{M-6} & \dots & h_3 & h_2 & h_1 & h_0 & 0 \\ 0 & h_{M-1} & h_{M-2} & h_{M-3} & h_{M-4} & \dots & h_5 & h_4 & h_3 & h_2 & h_1 \\ 0 & 0 & 0 & h_{M-1} & h_{M-2} & \dots & h_7 & h_6 & h_5 & h_4 & h_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & h_{M-1} & h_{M-2} & h_{M-3} & h_{M-4} & h_{M-5} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & h_{M-1} & h_{M-2} & h_{M-3} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & h_{M-1} \end{pmatrix}.$$

Равенство (7.40) означает, что матрица \mathbf{B} имеет собственное число $\lambda = 1$, которому соответствует собственный вектор $\Phi_0 = (\phi(0), \phi(1), \dots, \phi(M-1))^T$.

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 собствен-

ному числу $\lambda = 1$, введем дополнительное требование, которое вытекает из следующей леммы.

Лемма 7.3. Масштабирующая функция $\phi(x)$ удовлетворяет 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, $u(x) = u(x + 1)$, представим ее в виде ряда Фурье:

$$u(x) = \sum_{m \in \mathbb{Z}} c_m e^{i2\pi mx}.$$

Здесь

$$\begin{aligned} c_m &= \int_0^1 u(x) e^{-i2\pi mx} dx = \sum_{n \in \mathbb{Z}} \int_0^1 \phi(x - n) e^{-i2\pi mx} dx = \sum_{n \in \mathbb{Z}} \int_{-n}^{1-n} \phi(t) e^{-i2\pi m(t+n)} dt = \\ &= \sum_{n \in \mathbb{Z}} e^{-i2\pi mn} \int_{-n}^{1-n} \phi(t) e^{-i2\pi mt} dt = \int_{-\infty}^{+\infty} \phi(t) e^{-i2\pi mt} dt = S_\phi(2\pi m), \end{aligned}$$

$$\text{где } S_\phi(\omega) = \int_{-\infty}^{+\infty} \phi(t) e^{-i\omega t} dt.$$

Так как $S_\phi(0) = 1$ (см. (7.4)), то из соотношения (7.26) для $\omega = 0$ следует, что $\forall m \in \mathbb{Z}$

$$S_\phi(2\pi m) = \begin{cases} 1, & \text{при } m = 0 \\ 0, & \text{при } m \neq 0. \end{cases}$$

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 $\Phi_0 = (\phi(0), \phi(1), \dots, \phi(M-1))^T$ матрицы \mathbf{B} из уравнения (7.40) должен удовлетворять условию $\sum_{n=0}^{M-1} \phi(n) = 1$. После определения собственного вектора Φ_0 матрицы \mathbf{B} , отвечающего данному требованию нормировки, 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 $\Delta t = 2^{-K}$: $t_j = j\Delta t$, $j = 0, 1, \dots, 2^K(M-1)$.

В случае, если для конкретной матрицы \mathbf{B} поиск собственного вектора $\Phi_0 = (\phi(0), \phi(1), \dots, \phi(M-1))^T$ из уравнения (7.40) не влечет вычислительных проблем, второй из рассмотренных методов нахождения масштабирующей функции $\phi(x)$ по заданному набору коэффициентов $\{h_n\}$ является более предпочтительным, так как в узлах сетки аргумента $\{x_j = j2^{-m}\}_{j \in \mathbb{Z}}$ позволяет *точно* найти значения функции $\{\phi(x_j = j2^{-m})\}_{j \in \mathbb{Z}}$.

Упражнение. По аналогии с описанными процедурами итерационного построения масштабирующей функции опишите две процедуры нахождения материнского вейвлета $\psi(x)$ по заданному набору коэффициентов $\{g_n = (-1)^n h_{M-1-n}\}_{n=0}^{M-1}$ масштабирующих уравнений (7.37) и (7.38).

Отметим, что для большинства практических приложений ЦОС построение конкретных масштабирующих функций и вейвлетов реально не требуется, достаточно лишь задать импульсные характеристики КЗФ.

7.7. Вейвлеты Добеши

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 $\delta_{m,n} = \begin{cases} 1, & m=n \\ 0, & m \neq n \end{cases}$, получим $\forall m \in \mathbb{Z}$:

$$\delta_{0,m} = \langle \phi(x), \phi(x-m) \rangle = 2 \int_{-\infty}^{\infty} \sum_{j \in \mathbb{Z}} h_j \phi(2x-j) \sum_{k \in \mathbb{Z}} h_k \phi(2(x-m)-k) dx =$$

$$\begin{aligned}
 &= \sum_{j \in \mathbb{Z}} h_j \sum_{k \in \mathbb{Z}} h_k \int_{-\infty}^{\infty} \phi(2x-j)\phi(2x-(2m+k))d(2x) = \\
 &= \sum_{j \in \mathbb{Z}} h_j \sum_{k \in \mathbb{Z}} h_k \langle \phi(x-j), \phi(x-(2m+k)) \rangle = \sum_{j \in \mathbb{Z}} h_j \sum_{k \in \mathbb{Z}} h_k \delta_{j,2m+k} = \sum_{j \in \mathbb{Z}} h_j h_{j-2m}.
 \end{aligned}$$

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 алгебраического полинома (по возможности более высокого порядка N) при разложении по базису в пространстве V_0 . Чем выше степень N произвольного 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 аппроксимации 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 использования 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 эквивалентное 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) и (7.38).

Лемма 7.4. Для того чтобы выполнялось условие (7.45), необходимо и достаточно выполнения требования¹

$$\sum_{n=0}^{M-1} n^m g_n = 0, \quad m = 0, \dots, N. \quad (7.46)$$

¹ Считаем, что при $n = m = 0$: $n^m = 1$.

◀ Воспользовавшись соотношением (7.38), получаем для интеграла

$$I_m = \int_{-\infty}^{+\infty} t^m \psi(t) dt = \sqrt{2} \int_{-\infty}^{+\infty} t^m \sum_{n=0}^{M-1} g_n \phi(2t-n) dt = 2^{-m-\frac{1}{2}} \sum_{n=0}^{M-1} g_n \int_{-\infty}^{+\infty} (x+n)^m \phi(x) dx,$$

где последний переход осуществлен в результате замены переменной $x = 2t - n$. Учитывая, что $\phi(x) = 0$ при $x \notin [0; M-1]$, получаем далее¹:

$$I_m = 2^{-m-\frac{1}{2}} \sum_{n=0}^{M-1} g_n \int_0^{M-1} (x+n)^m \phi(x) dx = 2^{-m-\frac{1}{2}} \sum_{n=0}^{M-1} g_n \int_0^{M-1} \left(\sum_{k=0}^m C_m^k x^{m-k} n^k \right) \phi(x) dx.$$

Окончательно,

$$I_m = \int_{-\infty}^{+\infty} t^m \psi(t) dt = 2^{-m-\frac{1}{2}} \sum_{k=0}^m \left(C_m^k \int_0^{M-1} x^{m-k} \phi(x) dx \sum_{n=0}^{M-1} n^k g_n \right).$$

Из полученного соотношения видно, что выполнение требований (7.46) влечет выполнение соотношений (7.45), т. е. условия (7.46) являются достаточными для (7.45).

Докажем необходимость (7.46) для выполнения (7.45). Пусть условия (7.45) выполнены, т. е. $I_m = 0$, $m = 0, \dots, N$. Тогда

$$I_0 = \frac{1}{\sqrt{2}} \left(\sum_{n=0}^{M-1} g_n \right) \int_0^{M-1} \phi(x) dx = 0,$$

откуда с учетом (7.4): $\sum_{n=0}^{M-1} g_n = 0$, т. е. условие (7.46) для $m = 0$ выполнено. Далее рассмотрим последовательно равенства $I_m = 0$ для $m = 1, \dots, N$:

$$I_m = 2^{-m-\frac{1}{2}} \sum_{k=0}^m C_m^k \left(\int_0^{M-1} x^{m-k} \phi(x) dx \underbrace{\sum_{n=0}^{M-1} n^k g_n}_{=0 \text{ при } k=0, \dots, m-1} \right) = 2^{-m-\frac{1}{2}} \underbrace{\int_0^{M-1} \phi(x) dx}_{=0} \sum_{n=0}^{M-1} n^m g_n = 0,$$

что означает выполнение условия (7.46) и для $m = 1, \dots, N$. Таким образом, выполнение требований (7.45) влечет выполнение соотношений (7.46). ▶

Лемма 7.5. Условие (7.46) эквивалентно требованию

$$\sum_{n=0}^{M-1} n^m (-1)^n h_n = 0, \quad m = 0, \dots, N. \quad (7.47)$$

¹ Напомним формулу бинома Ньютона: $(x+y)^N = \sum_{k=0}^N C_N^k x^{N-k} y^k$, где $C_N^k = \frac{N!}{k!(N-k)!}$ — число сочетаний из N по k .

◀ 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:

$$\begin{aligned} S_m &= \sum_{n=0}^{M-1} n^m (-1)^n h_n = \sum_{n=0}^{M-1} n^m (-1)^n g_{M-1-n} (-1)^{1-n} = - \sum_{n=0}^{M-1} n^m g_{M-1-n} = \\ &= - \sum_{n=0}^{M-1} (M-1-n)^m g_n = (-1)^{m+1} \sum_{n=0}^{M-1} (n+1-M)^m g_n = \\ &= (-1)^{m+1} \sum_{n=0}^{M-1} g_n \sum_{j=0}^m C_m^j n^j (1-M)^{m-j} = (-1)^{m+1} \sum_{j=0}^m \left(C_m^j (1-M)^{m-j} \sum_{n=0}^{M-1} g_n n^j \right), \end{aligned}$$

или, окончательно,

$$S_m = \sum_{n=0}^{M-1} n^m (-1)^n h_n = (-1)^{m+1} \sum_{j=0}^m C_m^j (1-M)^{m-j} \sum_{n=0}^{M-1} g_n n^j,$$

откуда следует достаточность условий (7.46) для выполнения (7.47).

Докажем необходимость (7.46) для выполнения (7.47). Пусть условия (7.47) выполнены, т.е. $S_m = 0$, $m = 0, \dots, N$. Тогда $S_0 = \sum_{n=0}^{M-1} (-1)^n h_n = - \sum_{n=0}^{M-1} g_n = 0$, т.е. условие (7.46) для $m = 0$ выполнено. Далее рассмотрим последовательно равенства $S_m = 0$ для $m = 1, \dots, N$:

$$S_m = (-1)^{m+1} \sum_{j=0}^m C_m^j (1-M)^{m-j} \underbrace{\sum_{n=0}^{M-1} g_n n^j}_{0, j=0, \dots, m-1} = (-1)^{m+1} \sum_{n=0}^{M-1} g_n n^m = 0,$$

что означает выполнение условия (7.46) и для $m = 1, \dots, N$. Таким образом, выполнение требований (7.47) влечет выполнение соотношений (7.46). ▶

Из лемм 7.4 и 7.5 непосредственно вытекает следующая теорема.

Теорема 7.1. Для выполнения условий (7.45) необходимо и достаточно выполнение соотношений (7.47).

Таким образом, мы определили необходимые (обязательные) требования, которым должны отвечать коэффициенты масштабирования уравнения — равенства (7.28), (7.29), (7.43), и дополнительные — условия (7.47). Заметим также, что условие (7.29) эквивалентно (7.47) для $m = 0$. Поэтому будем искать коэффициенты масштабирования уравнений (7.37) и (7.38), исходя из требований (7.28), (7.29) и (7.43), выбирая необходимое количество недостающих условий из (7.47) при $m = 1, 2, \dots$

Пример 7.6. Вейвлет Хаара — «Добеши-1»¹. В примере 7.1 возможное решение уравнения (7.1) было «угадано». Найти данное решение можно по рассмотренной выше методике.

◀ Если положить $M = 2$ (в масштабирующем уравнении — два ненулевых коэффициента: h_0, h_1), то условия (7.28), (7.29) примут вид системы уравнений

$$\begin{cases} h_0 + h_1 = \sqrt{2} \\ h_0 - h_1 = 0 \end{cases},$$

которая имеет единственное решение: $h_0 = h_1 = 1/\sqrt{2}$. Оно удовлетворяет и необходимому условию (7.43). ▶

Пример 7.7. Вейвлет «Добеши-2».

◀ Положим $M = 4$. Необходимые условия (7.43), (7.28), (7.29) дают четыре уравнения:

$$\begin{cases} h_0^2 + h_1^2 + h_2^2 + h_3^2 = 1 \\ h_0 h_2 + h_1 h_3 = 0 \\ h_0 + h_1 + h_2 + h_3 = \sqrt{2} \\ h_0 - h_1 + h_2 - h_3 = 0 \end{cases}.$$

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

¹ Ингрид Добеши — математик, которой принадлежат многие основополагающие результаты теории вейвлет-преобразований и КМА.

$$\begin{cases} -h_1 + 2h_2 - 3h_3 = 0 \\ h_0 h_2 + h_1 h_3 = 0 \\ h_0 + h_2 = 1/\sqrt{2} \\ h_1 + h_3 = 1/\sqrt{2} \end{cases}$$

имеет два «зеркальных» решения (убедитесь!):

$$\left(h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}} \right),$$

$$\left(h_0 = \frac{1-\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3-\sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1+\sqrt{3}}{4\sqrt{2}} \right).$$

Обычно вейвлетом «Добеши-2» называют функцию, порождаемую первым из двух приведенных наборов решений. Для построения графиков масштабирующей функции и вейвлета (рис. 7.13) можно воспользоваться одним из двух методов, описанных в разделе 7.6. ►

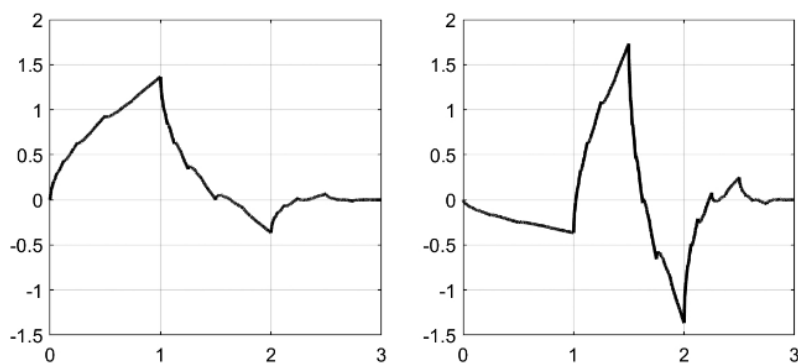


Рис. 7.13. Графики масштабирующей функции (слева) и материнского вейвлета (справа), соответствующие КЗФ «Добеши-2»

Увеличивая число $M = 2L$, можно построить другие вейвлеты и масштабирующие функции семейства «Добеши- L ». На рис. 7.14 приведен пример масштабирующей функции и вейвлета, полученных для $L = 10$.

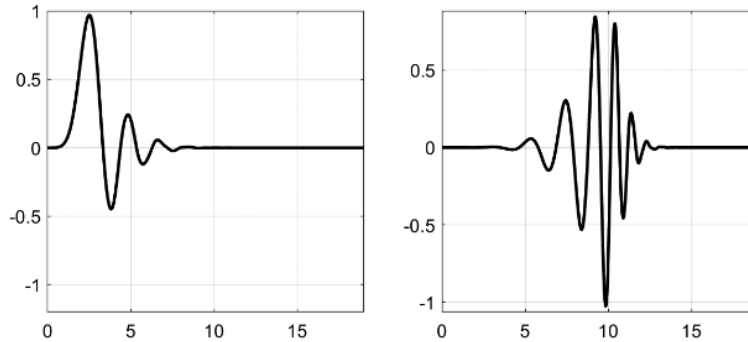


Рис. 7.14. Графики масштабирующей функции (слева) и материнского вейвлета (справа), соответствующие КЗФ «Добеши-10»

Упражнение. Запишите систему уравнений, определяющую значения коэффициентов $\{h_n\}$ для КМА «Добеши-3». Обратите внимание: при $m=0$ уравнение (7.43) вновь является избыточным и его нужно исключить.

Конечно, рассмотренный подход к построению вейвлетов, основанный на решении системы уравнений (7.28), (7.29), (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

Ортогональные базисы вейвлетов с ограниченным носителем не всегда оказываются удобными для представления сигналов. Более широкое распространение на практике получили так называемые *биортогональные* вейвлет-базисы с «ослабленными» требованиями к ортогональности. В этом случае по-прежнему $V_{m+1} = V_m \oplus W_m$, но функциональные базисы $\{\phi_{m,n}\}_{n \in \mathbb{Z}} \subset V_m$ и $\{\psi_{m,n}\}_{n \in \mathbb{Z}} \subset W_m$ уже не являются ортогональными в этих подпространствах. Однако при этом в некоторых дуальных подпространствах \tilde{V}_m и \tilde{W}_m должна

существовать пара базисов (также не обязательно ортогональных) $\{\check{\phi}_{m,n}\}_{n \in \mathbb{Z}} \subset \check{V}_m$ и $\{\check{\psi}_{m,n}\}_{n \in \mathbb{Z}} \subset \check{W}_m$, таких что $\langle \check{\phi}_{m,n}, \psi_{m,k} \rangle = \langle \check{\psi}_{m,n}, \phi_{m,k} \rangle = 0$ и $\langle \check{\phi}_{m,n}, \phi_{m,k} \rangle = \delta_{n,k}$, $\langle \check{\psi}_{m,n}, \psi_{m,k} \rangle = \delta_{n,k}$, 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 базиса. Так, если $f(x) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} \check{\psi}_{m,n}(x)$, то $c_{m,n} = \langle f, \psi_{m,n} \rangle$ (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 и определяется соотношением, аналогичным (7.16):

$$\sum_{n \in \mathbb{Z}} a_{m,n} \check{\phi}_{m,n}(x) = \underbrace{\sum_{n \in \mathbb{Z}} a_{m-1,n} \check{\phi}_{m-1,n}(x)}_{\check{J}_{m-1} \in \check{V}_{m-1}} + \underbrace{\sum_{n \in \mathbb{Z}} c_{m-1,n} \check{\psi}_{m-1,n}(x)}_{\check{Y}_{m-1} \in \check{W}_{m-1}}. \quad (7.48)$$

Соответствующие вычислительные процедуры вновь можно трактовать как обработку сигнала по схеме рис. 7.10.

Упражнение. Покажите, что формулы (7.19) и (7.20) получаются в результате скалярного умножения уравнения (7.48) на функцию $\phi_{m-1,k}(x) \in V_{m-1}$ или $\psi_{m-1,k}(x) \in W_{m-1}$, с последующей подстановкой соотношений (7.17) или (7.18).

Упражнение. Убедитесь, что обращающая (7.19) и (7.20) формула восстановления (7.22) принимает вид:

$$a_{m,k} = \sum_{n \in \mathbb{Z}} a_{m-1,n} \check{h}_{k-2n} + \sum_{n \in \mathbb{Z}} c_{m-1,n} \check{g}_{k-2n}.$$

Указание: умножив скалярно равенство (7.48) на функцию $\phi_{m,k}(x) \in V_m$, в полученное выражение подставьте затем представления (7.17) и (7.18), записанные для функций $\check{\phi}_{m-1,k}(x) \in \check{V}_{m-1}$ и $\check{\psi}_{m-1,k}(x) \in \check{W}_{m-1}$.

В общем случае для биортогональных преобразований коэффициенты в формулах анализа и синтеза $\check{h}_n \neq h_n$, $\check{g}_n \neq g_n$. Более того, число ненулевых коэффициентов в ИХ $\{\check{h}(n) = h_{-n}\}$, $\{\check{g}(n) = g_{-n}\}$, $\{h(n) = \check{h}_n\}$, $\{g(n) = \check{g}_n\}$ может быть различным. Для того чтобы определить биортогональное вейвлет-преобразование, необходимо задать уже два набора коэффициентов: $\{h_n\}$ и $\{\check{h}_n\}$ (или $\{g_n\}$ и $\{\check{g}_n\}$).

Этим наборам (в нашем рассмотрении – конечным!) на основании (7.1) и (7.3) соответствует пара функций $\phi(x) = \sqrt{2} \sum_n h_n \phi(2x - n)$ и $\psi(x) = \sqrt{2} \sum_n g_n \phi(2x - n)$, используемых для анализа (разложения), и две функции $\check{\phi}(x) = \sqrt{2} \sum_n \check{h}_n \check{\phi}(2x - n)$, $\check{\psi}(x) = \sqrt{2} \sum_n \check{g}_n \check{\phi}(2x - n)$, которые используются для синтеза сигнала. Подробнее вопросы связи характеристик биортогональных фильтров будут рассмотрены в разделе 7.14; сейчас отметим, что точное восстановление сигнала при обработке по схеме рис. 7.10 обеспечивается выбором коэффициентов $g_n = \alpha(-1)^n \check{h}_{1-n-2k}$ и $\check{g}_n = \alpha(-1)^n h_{1-n-2k}$, где числа $\alpha \in \{-1; 1\}$ и $K \in \mathbb{Z}$ принимаются одинаковыми для всех значений индекса $n \in \mathbb{Z}$.

Поскольку ортогональность целочисленных сдвигов масштабирующих функций в подпространствах V_0 и \check{V}_0 уже не является обязательной, то требование (7.43) к коэффициентам уравнения (7.1) снимается, и вместо него возникает условие $\sum_n h_n \check{h}_{n-2k} = \delta_{k,0}$ (обоснуйте его по аналогии с доказательством соотношения (7.43) для ортогонального случая).

Таким образом, требования к фильтрам декомпозиции (рис. 7.6) и реконструкции (рис. 7.8) ослабевают, и возможности выбора преобразований для разложения сигналов существенно расширяются. Недостатком биортогональных базисов является потеря возможности анализа квадратичной ошибки представления сигнала в области преобразований, т. к. равенство Парсевалю уже не выполняется, и поэтому для (7.48) в общем случае $\sum_{n \in \mathbb{Z}} a_{m,n}^2 \neq \sum_{n \in \mathbb{Z}} a_{m-1,n}^2 + \sum_{n \in \mathbb{Z}} c_{m-1,n}^2$. Однако чаще стремятся использовать такие биортогональные базисы, для которых в этом неравенстве с приемлемой для практики точностью все же можно записать приближенное равенство:

$$\sum_{n \in \mathbb{Z}} a_{m,n}^2 \approx \sum_{n \in \mathbb{Z}} a_{m-1,n}^2 + \sum_{n \in \mathbb{Z}} c_{m-1,n}^2.$$

На сегодняшний день предложено большое количество биортогональных базисов и соответствующих им фильтров, которые реализованы, в том числе, в различных программных средах. В качестве примера приведем набор коэффициентов биортогональных фильтров, который получил широкое распространение в приложениях цифровой обработки сигналов.

Пример 7.8. Биортогональные сплайн-вейвлеты 9/7 [31].

◀ Импульсная характеристика $\{\check{h}(n) = h_{-n}\}$ 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 $\{h(n) = \tilde{h}_n\}$ НЧ-фильтра синтеза также симметрична и имеет семь ненулевых коэффициентов:

$$\tilde{h}_0 = 0,788486, \quad \tilde{h}_{\pm 1} = 0,418092, \quad \tilde{h}_{\pm 2} = -0,040689, \quad \tilde{h}_{\pm 3} = -0,064539.$$

Коэффициенты ИХ ВЧ-фильтров анализа (декомпозиции) и синтеза (реконструкции) находятся по формулам $g_n = (-1)^n \tilde{h}_{-n}$ и $\tilde{g}_n = (-1)^n h_{-n}$.

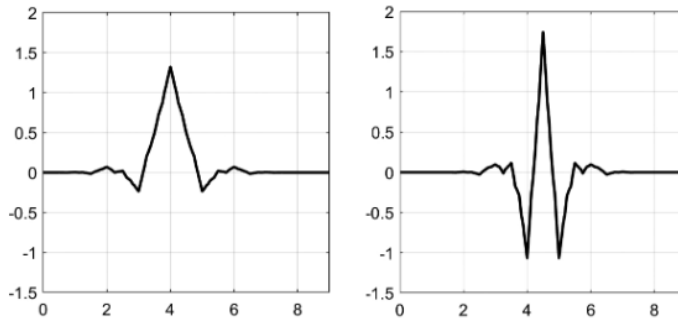


Рис. 7.15. Графики масштабирующей функции $\phi(x)$ (слева) и материнского вейвлета $\psi(x)$ (справа), соответствующие фильтрам анализа (декомпозиции) биортогонального преобразования 9/7

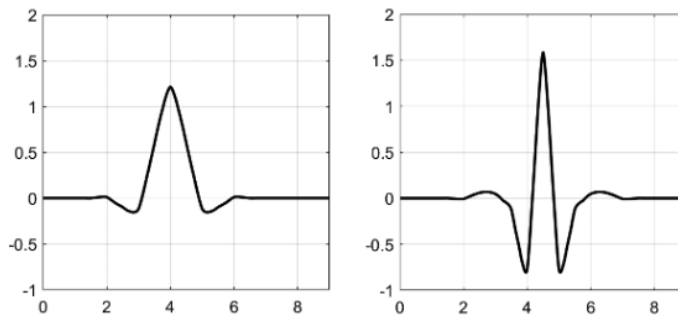


Рис. 7.16. Графики масштабирующей функции $\tilde{\phi}(x)$ (слева) и материнского вейвлета $\tilde{\psi}(x)$ (справа), соответствующие фильтрам синтеза (композиции) биортогонального преобразования 9/7

Графики функций анализа и синтеза, которые соответствуют приведенным значениям коэффициентов фильтров, изображены на рис. 7.15 и 7.16. ►

Рассмотренные в примере 7.8 фильтры применяются, в частности, в схеме компрессии данных, которую определяет стандарт сжатия фотографических изображений JPEG 2000.

7.9. Применение дискретных вейвлет-преобразований для сжатия сигналов

Сжатие (компрессия) сигналов с потерей информации является наиболее важной областью применения вейвлет-преобразований. Основная идея многих алгоритмов вейвлет-компрессии состоит в следующем.

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. 7.3, 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 некоррелированные 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 $\{c_{m,n}\}$ разложения обрабатываемого сигнала удобно упорядочить в виде древовидной структуры (рис. 7.17).

При таком упорядочивании каждый узел $\psi_{j,m}$ дерева (за исключением листьев) имеет по два непосредственных (ближайших) потомка $\{\psi_{j+1,2m}, \psi_{j+1,2m+1}\}$, причем носители

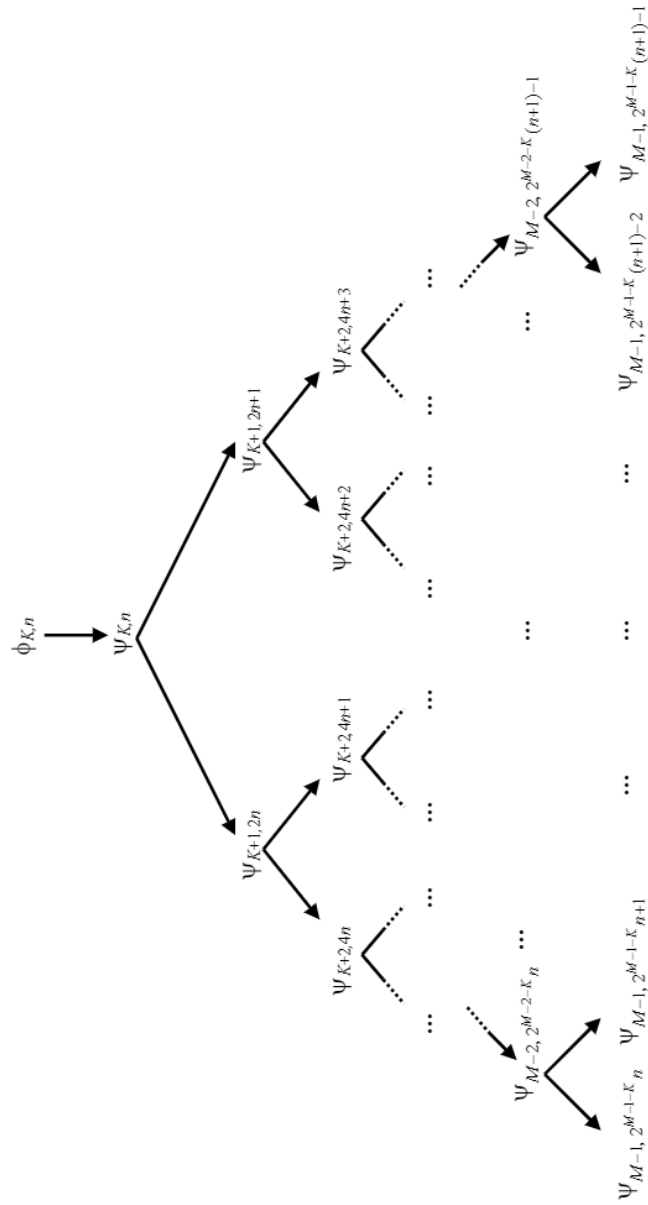


Рис. 7.17. Схема упорядочивания базисных функций в виде древовидной структуры

$\Delta_{j+1,2m} = \text{supp}\psi_{j+1,2m}(t)$ 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 $\Delta_{0,0} = \text{supp}\psi(t) = [0; N]$. (Заметим, что $N \geq 1$, так как в противном случае целочисленные сдвиги $\{\psi(t-n)\}_{n \in \mathbb{Z}}$ не покроют полностью числовую ось $t \in \mathbb{R}$, причем число N — целое, см. раздел 7.6.) Тогда

$$\Delta_{j,m} = \left[\frac{m}{2^j}; \frac{m}{2^j} + \frac{N}{2^j} \right],$$

$$\Delta_{j+1,2m} = \left[\frac{2m}{2^{j+1}}; \frac{2m+N}{2^{j+1}} \right] = \left[\frac{m}{2^j}; \frac{m}{2^j} + \frac{N}{2^{j+1}} \right] \subset \Delta_{j,m},$$

$$\Delta_{j+1,2m+1} = \left[\frac{2m+1}{2^{j+1}}; \frac{2m+1+N}{2^{j+1}} \right] = \left[\frac{m}{2^j} + \frac{1}{2^{j+1}}; \frac{m}{2^j} + \frac{N+1}{2^{j+1}} \right] \subset \Delta_{j,m},$$

так как для $N \geq 1$ справедливо: $\frac{N+1}{2^{j+1}} \leq \frac{N}{2^j}$.

Схематически локализация базисных функций на плоскости время×частота отражена на рис. 7.18. Вейвлет-базис обладает переменным время-частотным разрешением базисных функций (см. также раздел 6.9 и рис. 6.15).

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, 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:

$$\|f_M - \tilde{f}_M\|^2 = \sum_n |a_{M,n} - \tilde{a}_{M,n}|^2 = \sum_{c_{k,l} \in T_{j,m}} c_{k,l}^2,$$

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. При этом во многих случаях используются идеи RD-оптимизации параметров алгоритмов сжатия с потерями, описанные в разделе 6.7.

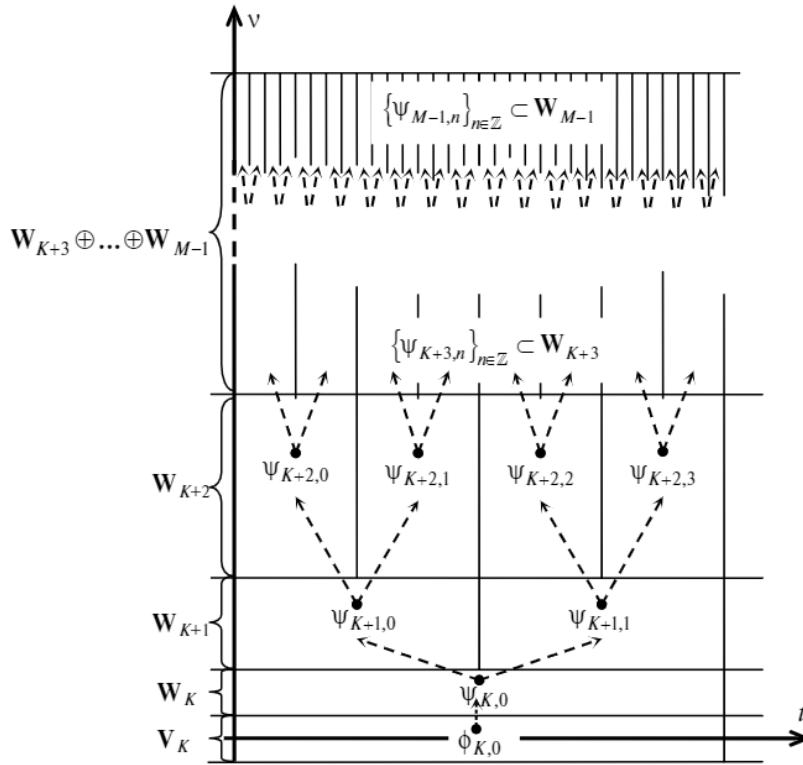


Рис. 7.18. Время-частотная локализация базисных функций, используемых для представления сигнала в виде (7.21). (Сравните с рис. 6.15.) Пунктиром обозначены логические связи базисных функций «родитель-потомки», определяемые графом рис. 7.17

После квантования (скалярного или векторного) оставшихся вейвлет-коэффициентов на этапе статистического кодирования «неподрезанных» ветвей необходимо учитывать имеющиеся между коэффициентами статистические зависимости, оставшиеся после дискретного вейвлет-преобразования. Так, помимо отмеченных статистических связей вида «родитель-потомки», имеется и зависимость между соседними вейвлет-коэффициентами одного уровня j , $c_{j,m}$ и $c_{j,m+1}$. Данная зависимость обусловлена тем, что соответствующие базисные функции $\psi_{j,m}(t)$ и $\psi_{j,m+1}(t)$ имеют близкую пространственную локализацию в области аргумента t : носители

$$\Delta_{j,m} = \text{supp}\psi_{j,m}(t) = \left[\frac{m}{2^j}; \frac{m}{2^j} + \frac{N}{2^j} \right],$$

$$\Delta_{j,m+1} = \text{supp}\psi_{j,m+1}(t) = \left[\frac{m+1}{2^j}; \frac{m+1}{2^j} + \frac{N}{2^j} \right]$$

в общем случае частично перекрываются (при $N=1$, т.е. в случае вейвлетов Хаара — примыкают друг к другу). Если на области аргумента $t \in (\Delta_{j,m} \cup \Delta_{j,m+1})$ сигнал $f(t)$ можно считать близким к стационарному, то для коэффициентов $c_{j,m} = \langle f(t), y_{j,m}(t) \rangle$ и $c_{j,m+1} = \langle f(t), y_{j,m+1}(t) \rangle$ статистики оказываются близкими, что также используется для их моделирования, наряду с учетом имеющихся зависимостей типа «родитель-потомки».

7.10. Подавление шумов фильтрацией в базисе дискретных вейвлет-преобразований

Рассмотрим случайный сигнал $X(n) = f(n) + X_{\text{ш}}(n)$, где $f(n)$ — полезный (неслучайный) информационный сигнал, а $X_{\text{ш}}(n)$ — случайная шумовая помеха. Чаще всего предполагается, что значения $X_{\text{ш}}(n)$ являются отсчетами белого шума (некоррелированы). Задача очистки сигнала от шума состоит в подавлении (т.е. уменьшении дисперсии) аддитивной компоненты $X_{\text{ш}}(n)$, при этом возможные искажения, вносимые в сигнал $f(n)$, должны быть минимальны (незначительны).



Для фильтрации шума в базисе дискретного вейвлет-преобразования сначала выполняется декомпозиция сигнала $X(n)$, т. е. выполняется разложение сигнала на несколько уровней по формулам (7.19) и (7.20), в результате чего получаем представление сигнала в форме (7.21). Как было отмечено в разделе 6.8, для широкого класса сигналов (в частности, мультимедийных) энергия полезного сигнала $f(n)$ распределяется неравномерно и в большей степени локализована в области нижних частот. Для белого шума, как мы знаем (см. раздел 1.8), распределение энергии по частотам является равномерным. Поэтому доля шумовой составляющей сигнала в коэффициентах вейвлет-разложения больше для высокочастотных компонент разложения (7.21). Если энергия шума $X_{\text{ш}}(n)$ примерно равна энергии полезного сигнала $f(n)$, то можно говорить о том, что вейвлет-коэффициенты, соответствующие высокочастотным компонентам разложения (7.21), в большей степени характеризуют шумовую составляющую сигнала, а низкочастотные коэффициенты в основном несут информацию о полезном сигнале.

Простейший способ удаления шума состоит в том, чтобы обнулить те значения коэффициентов $\{c_{m,n}\}$ разложения (7.21), которые меньше по модулю некоторого порогового значения T . Такая процедура называется *пороговой фильтрацией*. В силу сделанных выше замечаний очевидным является то, что обнуляемые вейвлет-коэффициенты в основном соответствуют высокочастотным компонентам разложения (7.21). Таким образом, процедура фильтрации шума в вейвлет-базисе сводится к следующей последовательности шагов.

1. Выполнение дискретного вейвлет-преобразования сигнала (декомпозиция с использованием формул (7.19), (7.20)).
2. Пороговая фильтрация вейвлет-коэффициентов.
3. Восстановление по отфильтрованному вейвлет-спектру дискретного сигнала в результате обратного дискретного вейвлет-преобразования (композиция на основе формулы (7.22)).

На практике используются два способа пороговой фильтрации вейвлет-коэффициентов разложения (7.21): *жесткая* и *мягкая* фильтрация. Жесткая пороговая фильтрация определяется функцией $y = y(x)$:

$$y = \begin{cases} x, & \text{если } |x| > T, \\ 0, & \text{если } |x| \leq T. \end{cases}$$

При мягкой фильтрации, наряду с обнулением малых коэффициентов, происходит «подтягивание» к нулю тех коэффициентов, которые не были обнулены. Соответствующая мягкой пороговой фильтрации функция обработки $y = y(x)$ имеет вид:

$$y = \begin{cases} x - T, & \text{если } x > T, \\ x + T, & \text{если } x < -T, \\ 0, & \text{если } |x| \leq T. \end{cases}$$

Реализовать жесткую пороговую фильтрацию проще, однако мягкая фильтрация позволяет во многих случаях устранить появление в восстановленном сигнале артефактов, ухудшающих субъективное восприятие аудио и визуальных сигналов.

Как жесткая, так и мягкая пороговая фильтрация вейвлет-коэффициентов может производиться в различных вариантах.

- Универсальная (общая) фильтрация, когда порог T выбирается общим для всех уровней m вейвлет-коэффициентов $\{c_{m,n}\}$ разложения (7.21).
- Многоуровневая фильтрация, когда пороги $T = T_m$ выбираются различными для разных уровней разложения m . Например, высокочастотные компоненты фильтруются более грубо, т. е. с бóльшим пороговым уровнем.
- Локально-адаптивная фильтрация, когда пороги фильтрации коэффициентов $\{c_{m,n}\}$ могут быть различными не только для различных уровней m , но и для различных номеров n , причем соответствующие пороги $T = T_{m,n}$ могут меняться по определенному правилу в процессе обработки в зависимости от оценок статистики уже обработанных коэффициентов.

Отметим, что описанные подходы к удалению шумов часто используются как составляющие алгоритмов компрессии сигналов на основе вейвлет-преобразований.

Описанные способы и варианты пороговой фильтрации, так же как и методы сжатия сигналов, могут проводиться в различных вейвлет-базисах. Выбирая конкретный вейвлет-базис для обработки сигнала, необходимо учитывать следующее: более гладкие

вейвлеты позволяют получить и более гладкую аппроксимацию полезного сигнала, тогда как менее гладкие короткие вейвлеты (имеющие малое количество ненулевых коэффициентов h_n в масштабирующем уравнении (7.1)) лучше выделяют локальные особенности (пики) сигнала.

7.11. Двумерные дискретные вейвлет-преобразования

В двумерном случае вместо подпространства функций одной переменной $V_M \in L_2(\mathbb{R})$ рассмотрим множество функций двух переменных $V_M^2 = V_M \otimes V_M$, которое представляет собой прямое (тензорное) произведение соответствующих подпространств функций одной переменной. То есть поставим во взаимно однозначное соответствие всем парам элементов $f_{1,M}(x) \in V_M \subset L_2(\mathbb{R})$, $f_{2,M}(x) \in V_M \subset L_2(\mathbb{R})$ элементы $f_M(x_1, x_2) \in V_M^2 \subset L_2(\mathbb{R}^2)$ по следующему правилу:

$$(f_{1,M}(x) \in V_M, f_{2,M}(x) \in V_M) \leftrightarrow f_M(x_1, x_2) = f_{1,M}(x_1)f_{2,M}(x_2) \in V_M^2.$$

Полученные подпространства функций двух переменных, очевидно, вложены: $V_M^2 \subset V_{M+1}^2 \subset L_2(\mathbb{R}^2)$, $V_M^2 \xrightarrow{M \rightarrow +\infty} L_2(\mathbb{R}^2)$.

Базисом введенного подпространства V_M^2 являются функции

$$\left\{ \phi_{M,k,l}(x_1, x_2) = \phi_{M,k}(x_1)\phi_{M,l}(x_2) = 2^M \phi(2^M x_1 - k)\phi(2^M x_2 - l) \right\}_{k,l \in \mathbb{Z}} \subset V_M^2.$$

А именно, $\forall f_M(x_1, x_2) \in V_M^2$ существует представление:

$$\begin{aligned} f_M(x_1, x_2) &= f_{1,M}(x_1)f_{2,M}(x_2) = \sum_{k \in \mathbb{Z}} a_{M,k}(f_{1,M})\phi_{M,k}(x_1) \sum_{l \in \mathbb{Z}} a_{M,l}(f_{2,M})\phi_{M,l}(x_2) = \\ &= \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} a_{M,k,l}\phi_{M,k,l}(x_1, x_2), \end{aligned} \quad (7.49)$$

где

$$\begin{aligned} a_{M,k,l} &= a_{M,k}(f_{1,M})a_{M,l}(f_{2,M}) = \int_{-\infty}^{+\infty} f_{1,M}(x_1)\phi_{M,k}(x_1)dx_1 \int_{-\infty}^{+\infty} f_{2,M}(x_2)\phi_{M,l}(x_2)dx_2 = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_M(x_1, x_2)\phi_{M,k,l}(x_1, x_2)dx_1 dx_2. \end{aligned}$$

Функцию $\phi(x_1, x_2) = \phi(x_1)\phi(x_2)$ назовем двумерной масштабирующей функцией — она порождает базис пространства V_M^2 :

$$\left\{ \phi_{M,k,l}(x_1, x_2) = 2^M \phi(2^M x_1 - k, 2^M x_2 - l) \right\}_{k,l \in \mathbb{Z}}.$$

Так как любой элемент $f_{M+1}(x) \in V_{M+1}$ единственным образом представим в виде $f_{M+1}(x) = f_M(x) + y_M(x)$, где $f_M(x) \in V_M$ и $y_M(x) \in W_M$, то для произвольного элемента $f_{M+1}(x_1, x_2) \in V_{M+1}^2$ получаем следующее (единственное) представление:

$$\begin{aligned} f_{M+1}(x_1, x_2) &= \underbrace{(f_{1,M}(x_1) + y_{1,M}(x_1))}_{f_{1,M+1}(x_1)} \underbrace{(f_{2,M}(x_2) + y_{2,M}(x_2))}_{f_{2,M+1}(x_2)} = \\ &= \underbrace{f_{1,M}(x_1)f_{2,M}(x_2)}_{f_M(x_1, x_2)} + \underbrace{y_{1,M}(x_1)f_{2,M}(x_2)}_{\text{обозначим } y_M^H(x_1, x_2)} + \underbrace{f_{1,M}(x_1)y_{2,M}(x_2)}_{\text{обозначим } y_M^V(x_1, x_2)} + \underbrace{y_{1,M}(x_1)y_{2,M}(x_2)}_{\text{обозначим } y_M^D(x_1, x_2)} = \\ &= f_M(x_1, x_2) + y_M^V(x_1, x_2) + y_M^H(x_1, x_2) + y_M^D(x_1, x_2). \end{aligned} \quad (7.50)$$

Упражнение. Покажите, что слагаемые $f_M(x_1, x_2)$, $y_M^V(x_1, x_2)$, $y_M^H(x_1, x_2)$, $y_M^D(x_1, x_2)$ в представлении (7.50) являются функциями, ортогональными в пространстве $L_2(\mathbb{R}^2)$, где скалярное произведение вещественных элементов $u(x_1, x_2)$ и $v(x_1, x_2)$ определяется как $\langle u, v \rangle = \iint_{\mathbb{R}^2} u(x_1, x_2)v(x_1, x_2)dx_1dx_2$.

Таким образом, в случае функций двух переменных подпространство $V_{M+1}^2 = V_{M+1} \otimes V_{M+1}$ разлагается на ортогональную сумму четырех подпространств:

$$\begin{aligned} V_{M+1}^2 &= (V_M \oplus W_M) \otimes (V_M \oplus W_M) = \\ &= \underbrace{V_M \otimes V_M}_{V_M^2} \oplus \underbrace{W_M \otimes V_M}_{\text{Обозн. } W_M^H} \oplus \underbrace{V_M \otimes W_M}_{\text{Обозн. } W_M^V} \oplus \underbrace{W_M \otimes W_M}_{\text{Обозн. } W_M^D} = V_M^2 \oplus W_M^V \oplus W_M^H \oplus W_M^D. \end{aligned}$$

Три подпространства $W_M^H = W_M \otimes V_M$, $W_M^V = V_M \otimes W_M$, $W_M^D = W_M \otimes W_M$ назовем *подпространствами двумерных вейвлетов*. Базисами в этих подпространствах являются системы функций:

$$\begin{aligned} \{\psi_{M,k,l}^H(x_1, x_2) = \psi_{M,k}(x_1)\phi_{M,l}(x_2) = 2^M \psi(2^M x_1 - k)\phi(2^M x_2 - l)\}_{k,l \in \mathbb{Z}} &\subset W_M^H, \\ \{\psi_{M,k,l}^V(x_1, x_2) = \phi_{M,k}(x_1)\psi_{M,l}(x_2) = 2^M \phi(2^M x_1 - k)\psi(2^M x_2 - l)\}_{k,l \in \mathbb{Z}} &\subset W_M^V, \\ \{\psi_{M,k,l}^D(x_1, x_2) = \psi_{M,k}(x_1)\psi_{M,l}(x_2) = 2^M \psi(2^M x_1 - k)\psi(2^M x_2 - l)\}_{k,l \in \mathbb{Z}} &\subset W_M^D. \end{aligned}$$

Данные базисы и соответствующие подпространства порождаются тремя типами двумерных материнских вейвлетов:

$$\psi^H(x_1, x_2) = \psi(x_1)\phi(x_2), \quad \psi^V(x_1, x_2) = \phi(x_1)\psi(x_2), \quad \psi^D(x_1, x_2) = \psi(x_1)\psi(x_2).$$



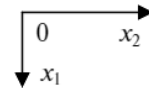
Ранее для «запуска» процедуры дискретного вейвлет-преобразования одномерный сигнал $f(x) \in L_2(\mathbb{R})$ мы изначально описывали набором коэффициентов $a_{M,n} = \langle f, \phi_{M,n} \rangle$ его проекции $f_M(x) = A_M(f(x)) = \sum_{n \in \mathbb{Z}} a_{M,n} \phi_{M,n}(x)$. Причем, как мы отмечали в разделе 7.1, точность описания сигнала $f(x) \in L_2(\mathbb{R})$ его проекцией $f_M(x) \in V_M$ тем выше, чем больше значение индекса M соответствующего подпространства V_M , а сами коэффициенты $\{a_{M,n}\}_{n \in \mathbb{Z}}$ можно считать масштабированными отсчетами сигнала $f(x)$ (см. (7.12)).

По аналогии с одномерным случаем, в качестве исходного объекта вейвлет-анализа возьмем сигнал $f_M(x_1, x_2) \in V_M^2$, представленный в форме (7.49), т. е. заданный двумерным массивом (матрицей) коэффициентов $\{a_{M,k,l}\}_{k,l}$. Отметим, что при достаточно большом значении индекса M вновь можно считать эти коэффициенты определенным образом пронормированными отсчетами двумерного сигнала. Выполнение одного уровня двумерного дискретного вейвлет-преобразования означает, что от формы представления сигнала (7.49) необходимо перейти к форме (см. (7.50)):

$$\begin{aligned}
 f_M(x_1, x_2) = & \underbrace{\sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} a_{M-1,k,l} \phi_{M-1,k,l}(x_1, x_2)}_{f_{M-1}(x_1, x_2) \in V_{M-1}^2} + \\
 & + \underbrace{\sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} c_{M-1,k,l}^H \Psi_{M-1,k,l}^H(x_1, x_2)}_{y_{M-1}^H(x_1, x_2) \in W_{M-1}^H} + \underbrace{\sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} c_{M-1,k,l}^V \Psi_{M-1,k,l}^V(x_1, x_2)}_{y_{M-1}^V(x_1, x_2) \in W_{M-1}^V} + \\
 & + \underbrace{\sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} c_{M-1,k,l}^D \Psi_{M-1,k,l}^D(x_1, x_2)}_{y_{M-1}^D(x_1, x_2) \in W_{M-1}^D}. \tag{7.51}
 \end{aligned}$$

Фактически, из матрицы $\{a_{M,k,l}\}_{k,l}$ необходимо получить четыре матрицы коэффициентов: $\{a_{M-1,k,l}\}_{k,l}$, $\{c_{M-1,k,l}^H\}_{k,l}$, $\{c_{M-1,k,l}^V\}_{k,l}$, $\{c_{M-1,k,l}^D\}_{k,l}$. Отметим, что для сигналов, наблюдаемых на конечных интервалах, и базисных функций с ограниченными носителями границы индексов суммирования в соотношении (7.51) и количество коэффициентов в указанных матрицах будут конечными.

В рассматриваемых матрицах $\{a_{M,k,l}\}_{k,l}$, $\{a_{M-1,k,l}\}_{k,l}$, $\{c_{M-1,k,l}^H\}_{k,l}$, $\{c_{M-1,k,l}^V\}_{k,l}$, $\{c_{M-1,k,l}^D\}_{k,l}$ номером строки считаем индекс k , который соответствует переменной x_1 . Индекс l , соответствующий переменной x_2 , для указанных матриц будем считать номером столбца. Соответствующая ориентация осей координат изображена на рисунке.



Выполнить разложение (7.51) сигнала, представленного в форме (7.49), можно следующим образом. Запишем сигнал (7.49) в виде:

$$f_M(x_1, x_2) = \sum_{k \in \mathbb{Z}} \phi_{M,k}(x_1) \underbrace{\sum_{l \in \mathbb{Z}} a_{M,k,l} \phi_{M,l}(x_2)}_{\text{Обозн. } F_{M,k}(x_2) \in V_M}. \quad (7.52)$$

Для всех необходимых индексов k выполним одномерное вейвлет-разложение сигналов

$$F_{M,k}(x_2) = \sum_{l \in \mathbb{Z}} a_{M,k,l} \phi_{M,l}(x_2) = \sum_{l \in \mathbb{Z}} A_{M-1,k,l} \phi_{M-1,l}(x_2) + \sum_{l \in \mathbb{Z}} C_{M-1,k,l} \psi_{M-1,l}(x_2),$$

для чего воспользуемся соотношениями (7.19) и (7.20), которые в текущих обозначениях принимают, соответственно, вид:

$$A_{M-1,k,l} = \sum_{n \in \mathbb{Z}} a_{M,k,n} h_{n-2l}, \quad C_{M-1,k,l} = \sum_{n \in \mathbb{Z}} a_{M,k,n} g_{n-2l}.$$

Подставим полученное разложение $F_{M,k}(x_2)$ в соотношение (7.52):

$$\begin{aligned} f_M(x_1, x_2) &= \sum_{k \in \mathbb{Z}} \phi_{M,k}(x_1) \left(\sum_{l \in \mathbb{Z}} A_{M-1,k,l} \phi_{M-1,l}(x_2) + \sum_{l \in \mathbb{Z}} C_{M-1,k,l} \psi_{M-1,l}(x_2) \right) = \\ &= \sum_{l \in \mathbb{Z}} \phi_{M-1,l}(x_2) \underbrace{\sum_{k \in \mathbb{Z}} A_{M-1,k,l} \phi_{M,k}(x_1)}_{\text{Обозн. } U_{M,l}(x_1) \in V_M} + \sum_{l \in \mathbb{Z}} \psi_{M-1,l}(x_2) \underbrace{\sum_{k \in \mathbb{Z}} C_{M-1,k,l} \phi_{M,k}(x_1)}_{\text{Обозн. } Y_{M,l}(x_1) \in V_M}. \end{aligned}$$

Далее выполним вейвлет-декомпозицию функций

$$\begin{aligned} U_{M,l}(x_1) &= \sum_{k \in \mathbb{Z}} A_{M-1,k,l} \phi_{M,k}(x_1) = \sum_{k \in \mathbb{Z}} a_{M-1,k,l} \phi_{M-1,k}(x_1) + \sum_{k \in \mathbb{Z}} c_{M-1,k,l}^H \psi_{M-1,k}(x_1), \\ Y_{M,l}(x_1) &= \sum_{k \in \mathbb{Z}} C_{M-1,k,l} \phi_{M,k}(x_1) = \sum_{k \in \mathbb{Z}} c_{M-1,k,l}^V \phi_{M-1,k}(x_1) + \sum_{k \in \mathbb{Z}} c_{M-1,k,l}^D \psi_{M-1,k}(x_1), \end{aligned}$$

где на основании (7.19) и (7.20):

$$\begin{aligned} a_{M-1,k,l} &= \sum_{n \in \mathbb{Z}} A_{M-1,n,l} h_{n-2k}, & c_{M-1,k,l}^H &= \sum_{n \in \mathbb{Z}} A_{M-1,n,l} g_{n-2k}, \\ c_{M-1,k,l}^V &= \sum_{n \in \mathbb{Z}} C_{M-1,n,l} h_{n-2k}, & c_{M-1,k,l}^D &= \sum_{n \in \mathbb{Z}} C_{M-1,n,l} g_{n-2k}. \end{aligned}$$

Подставив полученные разложения $U_{M,l}(x_1)$ и $Y_{M,l}(x_1)$ в последнее выражение для $f_M(x_1, x_2)$, окончательно получим представление (7.51).

Описанную выше схему вычислений можно интерпретировать следующим образом: сначала матрица $\{a_{M,k,l}\}_{k,l}$ подвергается построчной обработке (формулы (7.19), (7.20)) НЧ и ВЧ фильтрами декомпозиции с последующим прореживанием по схеме рис. 7.6. В результате такой обработки получают две матрицы $\{A_{M-1,k,l}\}_{k,l}$, $\{C_{M-1,k,l}\}_{k,l}$, каждая из которых содержит по сравнению с исходной матрицей $\{a_{M,k,l}\}_{k,l}$ примерно вдвое меньшее количество столбцов за счет выполненного после фильтрации прореживания. Затем две

полученные матрицы $\{A_{M-1,k,l}\}_{k,l}$ и $\{C_{M-1,k,l}\}_{k,l}$ обрабатываются по аналогичной схеме рис. 7.6 вдоль столбцов и каждая матрица вновь разделяется на две. Сказанное иллюстрирует рис. 7.19. Тот же результат — разложение (7.51) — был бы получен, если бы сначала была выполнена фильтрация вдоль столбцов, а затем вдоль строк матрицы $\{a_{M,k,l}\}_{k,l}$.

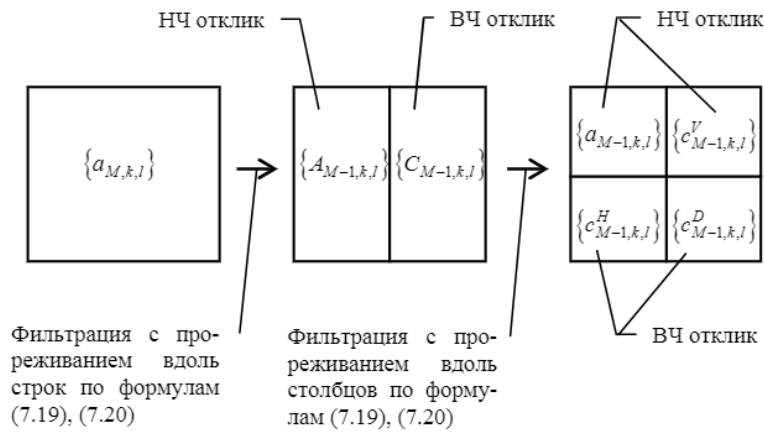


Рис. 7.19. Схема реализации одного уровня двумерного дискретного вейвлет-преобразования при помощи одномерной фильтрации. Используется следующая индексация элементов матриц: k — номер строки, l — номер столбца

Верхние индексы V, H, D в обозначениях подпространств вейвлетов и базисных функций из этих подпространств являются сокращениями от английских слов vertical, horizontal, diagonal. Поясним содержательную сущность такой индексации. Так, по определению, коэффициенты

$$c_{M,k,l}^V = \langle f(x_1, x_2), \psi_{M,k,l}^V(x_1, x_2) \rangle = \iint_{\mathbb{R}^2} f(x_1, x_2) \phi_{M,k}(x_1) \psi_{M,l}(x_2) dx_1 dx_2$$

характеризуют резкие изменения (ВЧ-составляющую) сигнала по переменной x_2 (горизонтальное направление координат) и сглаженный сигнал (НЧ-составляющую) по переменной x_1 (вертикальное направление координат). Выделение резких изменений сигнала по направлению горизонтальной оси и сглаживание

(подавление резких изменений) вдоль вертикальной оси соответствует выявлению вертикально ориентированных «фронтов» сигнала $f(x_1, x_2)$. Например, если функция $f(x_1, x_2)$ описывает значения яркости в точках изображения (x_1, x_2) , то коэффициенты $\{c_{M,k,l}^V\}_{k,l}$ будут принимать большие значения для базисных функций $\Psi_{M,k,l}^V(x_1, x_2) = \phi_{M,k}(x_1)\psi_{M,l}(x_2)$, локализация которых в пространственной области соответствует вертикально ориентированным контурам изображения. Аналогично, коэффициенты

$$c_{M,k,l}^H = \langle f(x_1, x_2), \Psi_{M,k,l}^H(x_1, x_2) \rangle = \iint_{\mathbb{R}^2} f(x_1, x_2) \psi_{M,k}(x_1) \phi_{M,l}(x_2) dx_1 dx_2$$

характеризуют горизонтально ориентированные, а

$$c_{M,k,l}^D = \langle f(x_1, x_2), \Psi_{M,k,l}^D(x_1, x_2) \rangle = \iint_{\mathbb{R}^2} f(x_1, x_2) \psi_{M,k}(x_1) \psi_{M,l}(x_2) dx_1 dx_2$$

— диагонально ориентированные перепады фронтов сигнала.

Упражнение. Опишите процедуру обращения одноуровневого двумерного дискретного вейвлет-преобразования, позволяющую получить коэффициенты $\{a_{M,k,l}\}_{k,l}$ представления (7.49) по заданным коэффициентам $\{a_{M-1,k,l}\}_{k,l}$, $\{c_{M-1,k,l}^H\}_{k,l}$, $\{c_{M-1,k,l}^V\}_{k,l}$, $\{c_{M-1,k,l}^D\}_{k,l}$ представления (7.51).

Введенные нами двумерные вейвлет-преобразования, вычисление которых сводится к поочередной одномерной обработке вдоль координат x_1 , x_2 , называются *разделимыми*. Такая схема вычислений оказывается возможной благодаря специфической структуре введенных подпространств КМА, которые определены через прямые произведения подпространств КМА функций одной переменной: $V_M^2 = V_M \otimes V_M$. В более общем случае возможно построение КМА для функций двух (и более) переменных, когда подпространства КМА не могут быть представлены через прямые произведения подпространств КМА функций одной переменной. Однако изучение неразделимых многомерных вейвлет-преобразований выходит за рамки нашего рассмотрения.

Матрицы $\{a_{M-1,k,l}\}_{k,l}$, $\{c_{M-1,k,l}^H\}_{k,l}$, $\{c_{M-1,k,l}^V\}_{k,l}$, $\{c_{M-1,k,l}^D\}_{k,l}$ называют *суббэндами* (от англ. *subband* — частотный поддиапазон). Низкочастотный суббэнд $\{a_{M-1,k,l}\}_{k,l}$ характеризует сглаженную (т.е. менее точную, более грубую) проекцию $f_{M-1}(x_1, x_2) \in V_{M-1}^2$ двумерного сигнала, остальные суббэнды отвечают за высокочастотные

компоненты разложения сигнала. Помимо приведенной на рис. 7.19, часто используется и другая схема взаимного расположения суббэндов, при которой суббэнды $\{c_{M-1,k,l}^H\}_{k,l}$ и $\{c_{M-1,k,l}^V\}_{k,l}$ меняются местами — такое упорядочивание используется, например, в программной среде MATLAB.

Описанную процедуру одноуровневого двумерного дискретного преобразования можно повторно применить к НЧ-суббэндам $\{a_{M-1,k,l}\}_{k,l}$, $\{a_{M-2,k,l}\}_{k,l}$ и т.д. При N -кратном ее повторении получим N -уровневое преобразование. На рис. 7.20 приведены этапы двухуровневого дискретного преобразования, выполненного при помощи программного пакета MATLAB над одним из стандартных тестовых изображений. Обратите внимание на взаимное



Рис. 7.20. Двумерное дискретное вейвлет-преобразование: а) исходное изображение, б) результат выполнения одноуровневого преобразования, в) результат выполнения двухуровневого преобразования

расположение вертикальных и горизонтальных саббэндов, которые на рис. 7.20б помечены типами соответствующих базисных функций.

Аналогично одномерному случаю (см. раздел 7.9), базисные функции двумерного вейвлет-спектра также обладают важным свойством, широко используемым в 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 вейвлета-родителя имеется уже по четыре потомка. Так, если $\psi(x, y)$ — родительский вейвлет (типа H, V, D), то им порождаются следующие четыре функции-потомка:

$$\psi(x, y) \rightarrow \left\{ \begin{array}{l} 2\psi(2x, 2y), \quad 2\psi(2x-1, 2y), \\ 2\psi(2x, 2y-1), \quad 2\psi(2x-1, 2y-1) \end{array} \right\}. \quad (7.53)$$

Схема связей «родитель-потомки» отражена на рис. 7.21. Видим, что коэффициенты спектра могут быть упорядочены в древовидную структуру, как это показано на следующем рис. 7.22. Родителем трех первых вейвлетов по определению считается единственная масштабирующая функция (начало дерева), дальнейшие связи устанавливаются в соответствии с (7.53).

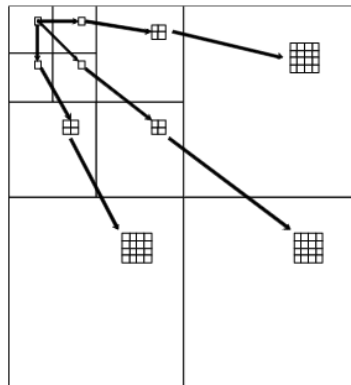


Рис. 7.21. Структура связей «родитель-потомки» между базисными функциями двумерного вейвлет-спектра на примере трехуровневого преобразования

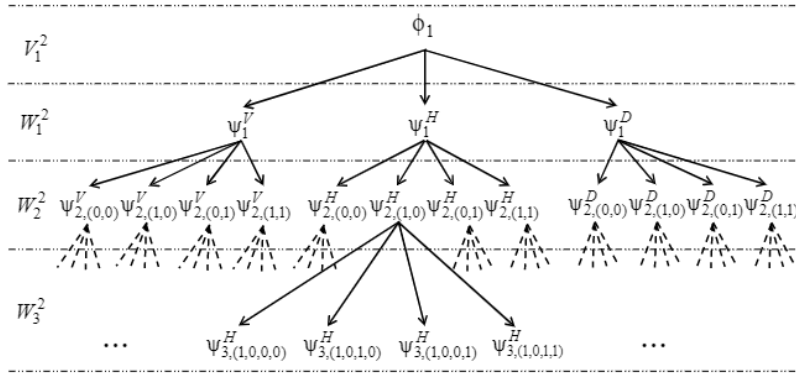


Рис. 7.22. Способ упорядочивания двумерных базисных функций в древовидную структуру на примере трехуровневого дискретного вейвлет-преобразования

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. Точно так же применимы для обработки двумерных сигналов и описанные в разделе 7.10 методы пороговой фильтрации в области вейвлет-преобразований.

На двумерном дискретном вейвлет-преобразовании с биортонормальными фильтрами из примера 7.8 (см. раздел 7.8) основан стандарт сжатия цифровых изображений JPEG 2000.

7.12. Метод сжатия цифровых изображений JPEG 2000

Идея метода JPEG 2000 так же, как и метода JPEG, укладывается в общую схему обработки, представленную ранее на рис. 6.1. Рассмотрим ключевые моменты реализации JPEG 2000 на примере обработки полутонового (монокроматического) изображения.

Для обработки изображений в рассматриваемом методе применяется дискретное вейвлет-преобразование. Хотя ортогональность

является весьма удобным и полезным свойством (в частности, евклидову норму ошибки, вносимой в исходный дискретный сигнал, можно контролировать в области преобразования), на практике приходится использовать вейвлет-преобразования с «ослабленной» ортогональностью, для которых равенство Парсевала выполняется приближенно — биортогональные вейвлеты (см. раздел 7.8). В JPEG 2000 возможно применение двух типов биортогональных фильтров: 5/3 и 9/7 (цифры означают количество ненулевых отсчетов в конечных импульсных характеристиках фильтров прямых/обратных преобразований). Первый набор фильтров 5/3 допускает точную целочисленную реализацию процедур анализа (декомпозиции) и синтеза (композиции) сигнала, второй набор 9/7 имеет рациональные коэффициенты, но дает вейвлет-спектр, дальнейшее квантование и кодирование которого показывает более высокие характеристики обработки изображения в терминах «ошибка-сжатие». Значения коэффициентов биортогональных фильтров 9/7 приведены в примере 7.8. Применяется двумерное вейвлет-преобразование, схема вычислений которого рассмотрена в предыдущем разделе 7.11.

После вычисления двумерного ДВП изображения полученные коэффициенты разложения скалярно квантуются. В методе JPEG 2000 используется *равномерное квантование с мертвой зоной*, при котором каждый коэффициент Y вейвлет-спектра преобразуется в целое число \tilde{Y} (*индекс квантования*) по следующему правилу:

$$\tilde{Y} = \text{sign}(Y) \lceil |Y|/q \rceil, \quad (7.54)$$

где q — дискрет квантования, а обозначение $\lceil x \rceil$ используется для целой части числа x . Смысл термина «равномерное квантование с мертвой зоной» становится понятен, если изобразить график зависимости индекса квантования \tilde{Y} от величины Y .

Упражнение. Изобразите графически зависимость $\tilde{Y}(Y)$ (7.54).

Если обозначить Δ_Y — интервал, при попадании в который величина Y преобразуется в целое число \tilde{Y} , то длина интервала, содержащего 0: $|\Delta_0| = 2q$, а остальные интервалы имеют длину q . Расширенная зона Δ_0 вокруг нуля и называется «мертвой» — попавшие в нее коэффициенты вейвлет-преобразования обнуляются. Отметим, что дискрет квантования q в стандарте JPEG 2000

не обязательно является целым числом и может выбираться различным для разных саббэндов изображения.

Деквантование, т. е. приближенное восстановление коэффициента $\hat{Y} \approx Y$ по индексу квантования \tilde{Y} , при декодировании изображения, проводится следующим образом:

$$\hat{Y} = \begin{cases} (\tilde{Y} + \gamma)q & \text{при } \tilde{Y} > 0 \\ (\tilde{Y} - \gamma)q & \text{при } \tilde{Y} < 0, \\ 0 & \text{при } \tilde{Y} = 0 \end{cases} \quad (7.55)$$

где параметр γ может быть задан для декодера любым из диапазона $0 \leq \gamma < 1$. Минимизация ошибки $M(|Y - \tilde{Y}|^2)$ для равномерно распределенной величины Y даст значение $\gamma = 0,5$. Если же, как это имеет место на практике, распределение вейвлет-коэффициентов имеет моду при $Y = 0$ и функция плотности убывает при удалении от 0, то параметр γ следует выбрать меньшим. Так, часто выбирается значение $\gamma = 0,375$.

Важнейшим этапом схемы компрессии JPEG 2000 является статистическое (энтропийное) кодирование проквантованных коэффициентов вейвлет-спектра изображения. Каждый саббэнд кодируется независимо от других, т. е. статистические зависимости родитель-потомок между коэффициентами (см. разделы 7.9 и 7.11) не используются. Это позволяет уйти от необходимости обязательного декодирования коэффициентов-родителей перед декодированием коэффициентов-потомков. В JPEG 2000 используется идеология «вложенного» кодирования (embedded encoding), когда в начале файла сжатого изображения (т. е. в начале декодируемого потока данных) содержится информация, позволяющая грубо восстановить изображение, а последующие данные позволяют его уточнить, последовательно уменьшая ошибку. При этом чем больше данных передано декодеру, тем точнее будет восстановлено изображение, а саму передачу данных можно прервать в любой момент. Поэтому сначала могут передаваться грубо проквантованные коэффициенты всего спектра, а затем все коэффициенты уточняются. Последовательность передачи данных, при которой сначала передаются коэффициенты-родители, а затем — потомки, уже не является обязательной.

При обработке каждого саббэнда последний разбивается на неперекрывающиеся прямоугольные *кодовые блоки* индексов квантования (фактически, матрицы из целых чисел $\{I(n) = \tilde{Y}(n)\}$), которые кодируются независимо друг от друга, полностью автономно. Линейные размеры (число строк, столбцов) кодового блока должны быть степенью числа 2, но по высоте не меньше 4, а общее количество коэффициентов разложения в каждом блоке не должно превышать 4096. Рассмотрим процедуру кодирования вейвлет-спектра на примере кодирования одного кодового блока \mathbf{V} .

Пусть N — наименьшее число битов, которого достаточно для представления каждого из целых чисел $I(n)$ кодового блока $\mathbf{V} = \{I(n)\}$, полученных из оригинальных вейвлет-коэффициентов в результате применения к ним процедуры (7.54). Индексы квантования представим в прямом коде: $I = s b_{N-2} b_{N-3} \dots b_0$, где s — бит знака, остальные биты — значащие, т. е.

$$I = (1 - 2s) \sum_{j=0}^{N-2} b_j 2^j, \quad 0 \leq |I| = \sum_{j=0}^{N-2} b_j 2^j \leq 2^{N-1} - 1.$$

Будем называть k -й *битовой плоскостью* для кодового блока $\mathbf{V} = \{I(n)\}$ множество битов $P_k = \{b_k(n)\}$. Кодер обрабатывает кодовый блок послойно, начиная с самой значимой битовой плоскости P_{N-2} , см. рис. 7.23.

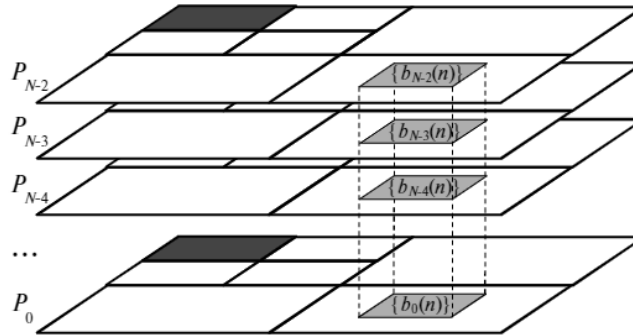


Рис. 7.23. Послойное представление целых чисел блока $\mathbf{V} = \{I(n)\}$ битовыми плоскостями

Первый единичный бит $b_k(n) = 1$, появившийся в представлении некоторого числа $I(n)$ при сканировании k -й битовой плоскости,

называется *битом значимости*, после появления такого единичного бита соответствующий коэффициент $I(n)$ считается *значимым*.

Кодер формирует *битовый* поток данных для статистического кодирования, сканируя коэффициенты в битовой плоскости в порядке, определяемом рис. 7.24, из которого становится понятным, почему по высоте кодовый блок не должен иметь меньше четырех строк. Статистическое кодирование реализуется побитно многомоделным адаптивным арифметическим кодером, входной алфавит которого содержит только два символа. Выбор конкретной модели распределения вероятностей для кодирования очередного бита производится арифметическим кодером *контекстно*, т. е. в зависимости от определенной предыстории кодирования.

Сканирование каждого слоя и кодирование битов в общем случае проводится в три прохода (этапа), которые называются *распространение значимости* (significance propagation), *уточнение значения* (magnitude refinement) и *зачистка* (clean-up).

Рассмотрим процедуру арифметического кодирования k -го слоя ($0 \leq k < N-2$).

На первом проходе (распространение значимости) кодируются биты $b_k(n)$ только тех пока еще не значимых индексов $I(n)$, среди непосредственных соседей которых в кодовом блоке (вертикальных — v , горизонтальных — h , или диагональных — d , см. рис. 7.25) есть хотя бы один уже ставший ранее значимым индекс квантования. Соседство со значимыми индексами повышает вероятность появления нового значимого элемента, и эту априорную информацию необходимо учитывать, чтобы в первую очередь кодировать бóльшие по абсолютной величине индексы квантования. Для кодирования бита значимости используется девять статистических моделей распределения вероятностей. Номер модели определяется количеством и расположением значимых соседей (рис. 7.25).

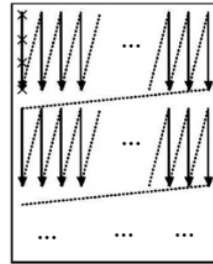


Рис. 7.24. Порядок сканирования битовой плоскости кодового блока В

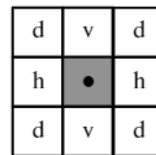


Рис. 7.25

Как только очередной индекс квантования $I(n)$ стал значимым, после кодирования признака (бита) значимости $b_k(n) = 1$ сразу же кодируется знаковый бит $s(n)$ этого индекса. Кодирование знакового бита производится с помощью одной из пяти статистических моделей, из которых определенным образом выбирается одна, по значениям знаков вертикальных и горизонтальных соседей.

На втором проходе (уточнение значения) арифметически кодируются очередные биты индексов кодирования, ставших значимыми ранее (на предыдущих битовых слоях). При этом используются три модели распределения вероятностей. Если коэффициент $I(n)$ стал значимым «только что» (на предшествующем слое P_{k+1}), то наблюдаются некоторые статистические зависимости между текущим кодируемым битом $b_k(n)$ и битами значимости соседей рис. 7.25, поэтому для кодирования $b_k(n)$ выбирается одна из двух статистических моделей, выбираемых в зависимости от количества значимых соседей. Если же коэффициент $I(n)$ стал значимым ранее, чем на предыдущем слое P_{k-1} , то появление 0 или 1 для бита $b_k(n)$ практически равновероятно и для кодирования $b_k(n)$ используется третья модель.

На третьем проходе (зачистка) кодируются очередные биты $b_k(n)$ тех индексов $I(n)$, которые еще не кодировались на двух предыдущих проходах плоскости P_k . С большой вероятностью эти биты окажутся нулевыми, т. е. сканируемые на третьем проходе коэффициенты не окажутся значимыми. Поэтому для кодирования на этапе зачистки предусмотрен специальный режим *прогона* (run mode), который ориентирован на объединение серий идущих друг за другом незначимых коэффициентов. Режим прогона включается, если все четыре коэффициента в очередном столбце скана (см. рис. 7.24) не имеют значимых соседей. Тогда значимость всей этой четверки коэффициентов кодируется *одним* бинарным символом-признаком S в специальной модели арифметического кодера. Если все коэффициенты в текущей четверке незначимы, то кодируется значение $S = 0$ и производится переход на следующий четырехэлементный столбец. Если же в текущей четверке хотя бы один коэффициент оказался значимым, то кодируется признак $S = 1$, затем в отдельной модели арифметического кодера (соответствующей равновероятному появлению входных бинарных символов) двумя битами кодируется позиция первого значимого коэффициента в столбце.



Как только на этапе зачистки появился значимый коэффициент и его позиция в скане закодирована, дальнейшая обработка этого коэффициента и последующих элементов производится по алгоритму этапа распространения значимости (т. е. как на первом проходе), с использованием тех же моделей контекстного кодирования значимости и знака. Кодирование по алгоритму этапа распространения значимости продолжается до тех пор, пока не окажется, что очередной столбец из четырех элементов скана не имеет значимых соседей — тогда произойдет переход в режим прогона.

При сканировании первого слоя P_{N-2} , с которого начинается кодирование кодового блока, проводится только этап зачистки, так как перед началом обработки блока *все* индексы кодирования считаются незначимыми.

Всего многомодельный бинарный арифметический кодер, используемый для обработки кодовых блоков, имеет 18 адаптивных и одну неадаптивную модель (с равной вероятностью появления 0 и 1). Девять моделей с номерами 0...8 используются для контекстного кодирования значимости коэффициентов, модель 9 используется в режиме прогонки, пять моделей 10–14 предназначены для контекстного кодирования знака, три модели 15–17 используются на втором проходе сканирования (этапе уточнения значения). Отметим также, что используемый в JPEG 2000 статистический кодер представляет собой модификацию «стандартного» арифметического кодера. Данная модификация — это так называемый *MQ-кодер*, который ориентирован на обработку двоичного входного алфавита и позволяет избежать при кодировании символов и адаптации статистических моделей операций умножения-деления, что, естественно, существенно повышает производительность обработки (ценой некоторого снижения эффективности сжатия данных).

Используемое в JPEG 2000 побитовое кодирование индексов квантования $I(n)$ в порядке от наиболее значимого (старшего) бита к младшему позволяет проводить деквантование вейвлет-коэффициентов и при неполном декодировании значения индекса (например, когда передача данных была прервана из-за сбоя аппаратуры). Конечно, восстановление коэффициента по неполным данным влечет большую ошибку квантования.

Пример 7.9. Из пяти значащих битов индекса квантования I восстановлено два старших: 10; восстановлен бит знака 0. Сравнить оценки дисперсии ошибки квантования для данного случая и для случая полного восстановления индекса квантования, если использовался шаг квантования q .

◀ Из условий примера следует, что индекс квантования I : $10000_{(2)} = 16_{(10)} \leq I \leq 23_{(10)} = 10111_{(2)}$. Квантуемый вейвлет-коэффициент Y лежал достаточно далеко от нулевой зоны (и заметно отличался от нулевого значения), ошибку квантования с достаточной точностью можно считать равномерно распределенной случайной величиной. Если известен точный индекс квантования I , дисперсия, как известно: $D = q^2/12$.

Из (7.54) следует, что для $16 \leq I \leq 23$: $Y \in [16q; 24q)$. Восстановленное значение вейвлет-коэффициента следует положить равным $20q$. При этом дисперсия ошибки в предположении ее равномерного распределения $D = (24q - 16q)^2/12 = 16q^2/3$. ▶

7.13. Вейвлет-пакеты

Рассмотрим возможность расширения класса функциональных базисов подпространств $V_j \subset L_2(\mathbb{R})$ за счет применения процедур рекуррентной декомпозиции не только для подпространств КМА $V_{j-1} \subset V_j$, но и для подпространств вейвлетов $W_{j-1} \subset V_j$, где $W_{j-1} \oplus V_{j-1} = V_j$. Эти процедуры основаны на следующей лемме.

Лемма 7.6. Пусть функция $\theta_0(t)$ такова, что система $\{\theta_0(t-n)\}_{n \in \mathbb{Z}}$ образует ортонормированный базис в некотором подпространстве U . Определим функции

$$\theta_{-1}^0(t) = \sum_{j \in \mathbb{Z}} h_j \theta_0(t-j), \quad \theta_{-1}^1(t) = \sum_{j \in \mathbb{Z}} g_j \theta_0(t-j), \quad (7.56)$$

где коэффициенты $\{h_n\}$, $\{g_n\}$ заданы уравнениями (7.1) и (7.3) соответственно. Тогда система функций

$$\{\theta_{-1,n}^0(t) = \theta_{-1}^0(t-2n), \quad \theta_{-1,n}^1(t) = \theta_{-1}^1(t-2n)\}_{n \in \mathbb{Z}}$$

образует ортонормированный базис подпространства U .

◀ В силу ортогональности сдвигов материнского вейвлета и масштабирующей функции, используя уравнения (7.1) и (7.3), имеем $\forall m, n$:

$$\begin{aligned}
 \langle \psi(t-m), \phi(t-n) \rangle &= \int_{-\infty}^{+\infty} \sqrt{2} \sum_k g_k \phi(2(t-m)-k) \sqrt{2} \sum_l h_l \phi(2(t-n)-l) dt = \\
 &= \sum_k g_k \sum_l h_l \int_{-\infty}^{+\infty} \phi(2t-2m-k) \phi(2t-2n-l) d(2t) = \\
 &= \sum_k g_k \sum_l h_l \underbrace{\langle \phi(x-2m-k), \phi(x-2n-l) \rangle}_{1 \text{ при } 2n+l=2m+k, \text{ иначе } 0} = \sum_k g_k h_{2(m-n)+k} = 0,
 \end{aligned}$$

ПОЭТОМУ

$$\begin{aligned}
 \langle \theta_{-1,n}^0(t), \theta_{-1,m}^1(t) \rangle &= \langle \sum_l h_l \theta_0(t-2n-l), \sum_k g_k \theta_0(t-2m-k) \rangle = \\
 &= \sum_k g_k \sum_l h_l \underbrace{\langle \theta_0(x-2m-k), \theta_0(x-2n-l) \rangle}_{1 \text{ при } 2n+l=2m+k, \text{ иначе } 0} = \sum_k g_k h_{2(m-n)+k} = 0.
 \end{aligned}$$

Из ортонормированности сдвигов вейвлета

$$\begin{aligned}
 \langle \psi(t-m), \psi(t-n) \rangle &= \int_{-\infty}^{+\infty} \sqrt{2} \sum_k g_k \phi(2(t-m)-k) \sqrt{2} \sum_l g_l \phi(2(t-n)-l) dt = \\
 &= \sum_k g_k \sum_l g_l \int_{-\infty}^{+\infty} \phi(2t-2m-k) \phi(2t-2n-l) d(2t) = \\
 &= \sum_k g_k \sum_l g_l \langle \phi(x-2m-k), \phi(x-2n-l) \rangle = \sum_k g_k g_{2(m-n)+k} = \begin{cases} 1 & \text{при } m=n \\ 0 & \text{при } m \neq n \end{cases}
 \end{aligned}$$

следует, что

$$\begin{aligned}
 \langle \theta_{-1,n}^1(t), \theta_{-1,m}^0(t) \rangle &= \langle \sum_l g_l \theta_0(t-2n-l), \sum_k h_k \theta_0(t-2m-k) \rangle = \\
 &= \sum_k h_k \sum_l g_l \langle \theta_0(x-2m-k), \theta_0(x-2n-l) \rangle = \sum_k h_k g_{2(m-n)+k} = \begin{cases} 1 & \text{при } m=n \\ 0 & \text{при } m \neq n \end{cases}.
 \end{aligned}$$

Аналогично, на основании (7.43) $\forall m, n$:

$$\begin{aligned}
 \langle \theta_{-1,n}^0(t), \theta_{-1,m}^0(t) \rangle &= \langle \sum_l h_l \theta_0(t-2n-l), \sum_k h_k \theta_0(t-2m-k) \rangle = \\
 &= \sum_k h_k \sum_l h_l \langle \theta_0(x-2m-k), \theta_0(x-2n-l) \rangle = \sum_k h_k h_{2(m-n)+k} = \begin{cases} 1 & \text{при } m=n \\ 0 & \text{при } m \neq n \end{cases}.
 \end{aligned}$$

Рассмотрев все возможные скалярные произведения элементов, мы установили, что система $\{\theta_{-1,n}^0(t), \theta_{-1,n}^1(t)\}_{n \in \mathbb{Z}} \subset U$ является ортонормированной. Докажем, что она образует базис в U . Для этого убедимся, что любой элемент заданного функционального базиса $\{\theta_0(t-n)\}_{n \in \mathbb{Z}} \subset U$ можно выразить как линейную комбинацию элементов ортонормированной системы $\{\theta_{-1,n}^0(t), \theta_{-1,n}^1(t)\}_{n \in \mathbb{Z}}$, которая поэтому также является базисом в U . А именно, покажем справедливость представления

$$\theta_0(t-n) = \sum_k (h_{n-2k} \theta_{-1,k}^0(t) + g_{n-2k} \theta_{-1,k}^1(t)).$$

Для доказательства последнего равенства убедимся, что

$$\|a-b\|^2 = \langle a-b, a-b \rangle = \|a\|^2 - 2\langle a, b \rangle + \|b\|^2 = 0,$$

где обозначено $a = \theta_0(t-n)$, $b = \sum_k (h_{n-2k} \theta_{-1,k}^0(t) + g_{n-2k} \theta_{-1,k}^1(t))$. В силу ортонормированности систем $\{\theta_0(t-n)\}_{n \in \mathbb{Z}}$ и $\{\theta_{-1,n}^0(t), \theta_{-1,n}^1(t)\}_{n \in \mathbb{Z}}$ имеем $\|a\|^2 = \|\theta_0(t-n)\|^2 = 1$, а на основании равенства Парсеваля, (7.34) и (7.43):

$$\|b\|^2 = \sum_k h_{n-2k}^2 + \sum_k g_{n-2k}^2 = \underbrace{\sum_k h_{n-2k}^2}_{\text{все четные или нечетные индексы}} + \underbrace{\sum_k h_{1-n+2k}^2}_{\text{все нечетные или четные индексы}} = \sum_k h_k^2 = 1.$$

Осталось показать, что $\langle a, b \rangle = 1$. Имеем:

$$\begin{aligned} \langle a, b \rangle &= \left\langle \theta_0(t-n), \sum_k (h_{n-2k} \sum_j h_j \theta_0(t-2k-j) + g_{n-2k} \sum_j g_j \theta_0(t-2k-j)) \right\rangle = \\ &= \sum_k \left(h_{n-2k} \sum_j h_j \langle \theta_0(t-n), \theta_0(t-2k-j) \rangle + g_{n-2k} \sum_j g_j \langle \theta_0(t-n), \theta_0(t-2k-j) \rangle \right) = \\ &= \sum_k (h_{n-2k}^2 + g_{n-2k}^2) = \sum_k h_{n-2k}^2 + \sum_k h_{1-n+2k}^2 = \sum_k h_k^2 = 1. \quad \blacktriangleright \end{aligned}$$

Изданной леммы следует, что вместо базиса $\{\psi_{0,n}(t) = \psi_{0,0}(t-n)\}_{n \in \mathbb{Z}}$ в подпространстве вейвлетов W_0 можно использовать ортогональный базис

$$\left\{ \psi_{-1,n}^0(t) = \sum_{j \in \mathbb{Z}} h_j \psi(t-2n-j), \quad \psi_{-1,n}^1(t) = \sum_{j \in \mathbb{Z}} g_j \psi(t-2n-j) \right\}_{n \in \mathbb{Z}}.$$

Фактически это означает возможность разложения подпространства W_0 в ортогональную сумму двух подпространств: $W_0 = W_{-1}^0 \oplus W_{-1}^1$, базисами в которых являются системы функций

$$\left\{ \psi_{-1,n}^0(t) = \psi_{-1}^0(t-2n) \right\}_{n \in \mathbb{Z}} \subset W_{-1}^0, \quad \left\{ \psi_{-1,n}^1(t) = \psi_{-1}^1(t-2n) \right\}_{n \in \mathbb{Z}} \subset W_{-1}^1,$$

определяемые соотношениями (7.56).

Аналогично, подпространство W_M можно разложить на сумму ортогональных подпространств $W_M = W_{M-1}^0 \oplus W_{M-1}^1$, имеющих, соответственно, базисы

$$\begin{aligned} \left\{ \psi_{M-1,n}^0(t) = \psi_{M-1}^0(t-2n) = \sum_{j \in \mathbb{Z}} h_j \psi_{M,0}(t-2n-j) \right\}_{n \in \mathbb{Z}} &\subset W_{M-1}^0, \\ \left\{ \psi_{M-1,n}^1(t) = \psi_{M-1}^1(t-2n) = \sum_{j \in \mathbb{Z}} g_j \psi_{M,0}(t-2n-j) \right\}_{n \in \mathbb{Z}} &\subset W_{M-1}^1. \end{aligned}$$

Пример 7.10. Для вейвлета Хаара (см. пример 7.1)

$$\psi(t) = \begin{cases} 1, & t \in [0; 1/2) \\ -1, & t \in [1/2; 1) \\ 0, & t \notin [0; 1) \end{cases}$$

построить функции $\psi_{-1,0}^0(t) = \psi_{-1}^0(t)$ и $\psi_{-1,0}^1(t) = \psi_{-1}^1(t)$, сдвиги которых образуют ортонормированный базис в подпространствах W_{-1}^0, W_{-1}^1 соответственно.

◀ Воспользовавшись результатами решения примера 7.1, на основании соотношений (7.56) получаем:

$$\psi_{-1,0}^0(t) = \psi_{-1}^0(t) = \frac{1}{\sqrt{2}}(\psi(t) + \psi(t-1)) = \begin{cases} 1/\sqrt{2}, & t \in [0; 1/2) \cup [1; 3/2) \\ -1/\sqrt{2}, & t \in [1/2; 1) \cup [3/2; 2) \\ 0, & t \notin [0; 2) \end{cases}$$

$$\psi_{-1,0}^1(t) = \psi_{-1}^1(t) = \frac{1}{\sqrt{2}}(\psi(t) - \psi(t-1)) = \begin{cases} 1/\sqrt{2}, & t \in [0; 1/2) \cup [3/2; 2) \\ -1/\sqrt{2}, & t \in [1/2; 3/2) \\ 0, & t \notin [0; 2) \end{cases}.$$

Постройте графики полученных функций. ▶

Определим правило, которое позволяет по заданному представлению функции $y_M(x) = \sum_{n \in \mathbb{Z}} c_{M,n} \psi_{M,n}(x)$, $y_M(x) \in W_M$, получить разложение

$$y_M(x) = \underbrace{\sum_{n \in \mathbb{Z}} c_{M-1,n}^0 \psi_{M-1,n}^0(x)}_{y_{M-1}^0(x) \in W_{M-1}^0} + \underbrace{\sum_{n \in \mathbb{Z}} c_{M-1,n}^1 \psi_{M-1,n}^1(x)}_{y_{M-1}^1(x) \in W_{M-1}^1}.$$

Умножим скалярно равенство

$$\sum_{n \in \mathbb{Z}} c_{M-1,n}^0 \psi_{M-1,n}^0(x) + \sum_{n \in \mathbb{Z}} c_{M-1,n}^1 \psi_{M-1,n}^1(x) = \sum_{n \in \mathbb{Z}} c_{M,n} \psi_{M,n}(x)$$

на функцию $\psi_{M-1,k}^0(x)$. Имеем $\forall k \in \mathbb{Z}$:

$$\begin{aligned} c_{M-1,k}^0 &= \langle y_M(x), \psi_{M-1,k}^0 \rangle = \sum_{n \in \mathbb{Z}} c_{M,n} \langle \psi_{M,n}(x), \psi_{M-1,k}^0 \rangle = \\ &= \sum_{n \in \mathbb{Z}} c_{M,n} \left\langle \psi_{M,n}(x), \sum_{j \in \mathbb{Z}} h_j \psi_{M,0}(x - 2k - j) \right\rangle = \\ &= \sum_{n \in \mathbb{Z}} c_{M,n} \sum_{j \in \mathbb{Z}} h_j \langle \psi_{M,n}(x), \psi_{M,2k+j}(x) \rangle = \sum_{n \in \mathbb{Z}} c_{M,n} h_{n-2k}. \end{aligned}$$

Упражнение. Покажите, что $\forall k \in \mathbb{Z}: c_{M-1,k}^1 = \sum_{n \in \mathbb{Z}} c_{M,n} g_{n-2k}$.

Как видим, для разложения подпространства $W_M = W_{M-1}^0 \oplus W_{M-1}^1$ используются уже знакомые нам формулы (7.19) и (7.20), которые можно интерпретировать как фильтрацию при помощи НЧ и ВЧ фильтров по схеме рис. 7.6. Обратная к обработке по формулам (7.19) и (7.20) процедура, как было установлено ранее, реализуется посредством формулы (7.22), которая принимает в используемых в данном разделе обозначениях вид: $c_{M,k} = \sum_{n \in \mathbb{Z}} c_{M-1,n}^0 h_{k-2n} + \sum_{n \in \mathbb{Z}} c_{M-1,n}^1 g_{k-2n}$. Интерпретировать последнюю формулу удобно как обработку при помощи фильтров по схеме рис. 7.8.

Блоки фильтров рис. 7.6 можно далее применять (или не применять) к откликам ВЧ-фильтров повторно. Допустив возможность дальнейшей обработки по схеме рис. 7.6 не только для прореженного отклика НЧ-фильтра $\tilde{H}(z)$, но и для прореженного отклика ВЧ-фильтра $\tilde{G}(z)$ на каждом уровне разложения дискретного сигнала, мы получаем возможность более гибкого выбора базиса представления сигнала, исходно заданного некоторой проекцией $f_M(x) \in V_M \subset L_2(\mathbb{R})$. Полученный в результате базис называется *базисом вейвлет-пакетов*. Для того чтобы описать конкретный используемый базис вейвлет-пакетов, необходимо указать, к откликам каких ВЧ-фильтров применялась (если применялась) обработка фильтрами по схеме рис. 7.6. Такое описание структуры базиса вейвлет-пакетов удобно отображать в виде двоичного графа-дерева.

В качестве примера приведем «каркас» дерева, описывающего возможные базисы трехуровневого пакетного преобразования (рис. 7.26), где сплошными линиями обозначена схема вычислений классического вейвлет-преобразования (определяющего минимальное дерево и минимальный объем вычислений), а пунктиром — возможные пакетные модификации базисов. Любой узел такого дерева, описывающего базис пакетов, может либо порождать два узла, либо не порождать ни одного. Данный граф определяет порядок применения формул (7.19) и (7.20) к дискретной последовательности коэффициентов $\{a_{M,n}\}_{n \in \mathbb{Z}}$: переход по левому ребру, выходящему из какого-либо узла, означает применение формулы (7.19), переход по правому — применение формулы (7.20). На рис. 7.27

приведен пример графа, описывающего один из других возможных пакетных вейвлет-базисов для трехуровневого разложения.

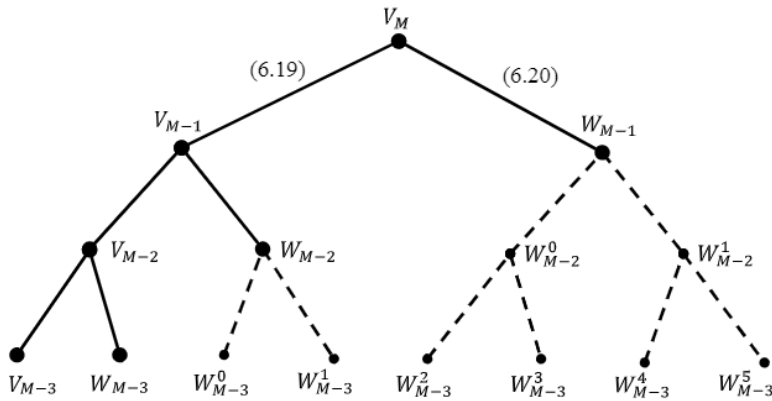


Рис. 7.26. Варианты построения базисов вейвлет-пакетов трехуровневого преобразования. Пунктирные линии определяют возможные (необязательные) дополнительные разложения подпространств вейвлетов

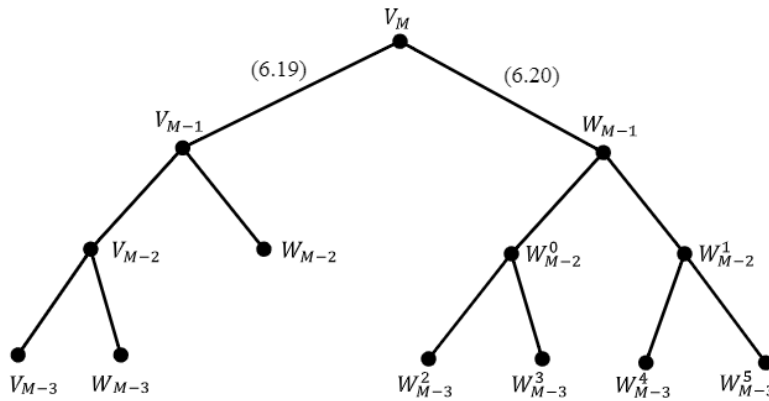


Рис. 7.27. Пример возможного базиса вейвлет-пакетов для трехуровневого преобразования

Пример 7.11. Пусть вектор $X=(x_0, \dots, x_7)$ представляет собой конечный набор коэффициентов $\{x_k = a_{M,k}\}_{k=0}^7$ разложения (7.5) сигнала $f_M(t) \in V_M$. Найти выражение для коэффициентов

полного, т.е. с использованием всех возможных декомпозиций вейвлет-пространств, трехуровневого разложения (рис. 7.26) сигнала $f_M(t)$ в базисе вейвлет-пакетов, порожденных вейвлетом Хаара (см. пример 7.1).

◀ Воспользовавшись результатами решения примера 7.4, получаем схему вычислений и выражения для коэффициентов полного разложения сигнала $f_M(t)$ по базису вейвлет-пакетов (вектора \mathbf{Y}), которые приведены на рис. 7.28.

Соответствующее преобразование можно записать в матричном виде:

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix}.$$

Несложно убедиться (проделайте это!), что данное дискретное преобразование представляет собой 8-точечное дискретное преобразование Уолша (см. раздел 2.9). ▶

Упражнение. Убедитесь, что минимальное трехуровневое разложение сигнала из примера 7.11 по базису пакетов на основе вейвлета Хаара (см. рис. 7.29) определяет 8-точечное дискретное преобразование Хаара, введенное в разделе 2.10.

Распространяя результаты решения примера 7.11 и последнего упражнения на случай обработки векторов большей размерности, можно сказать, что ДПУ и ДПХ являются вариантами дискретной реализации пакетного вейвлет-преобразования на основе фильтров из примера 7.4.

Упражнение. По аналогии с рис. 7.28 и 7.29 изобразите схему вычисления трехуровневого пакетного вейвлет-преобразования 8-компонентного вектора с фильтрами из примера 7.4, которую определяет структура пакетного базиса рис. 7.27.

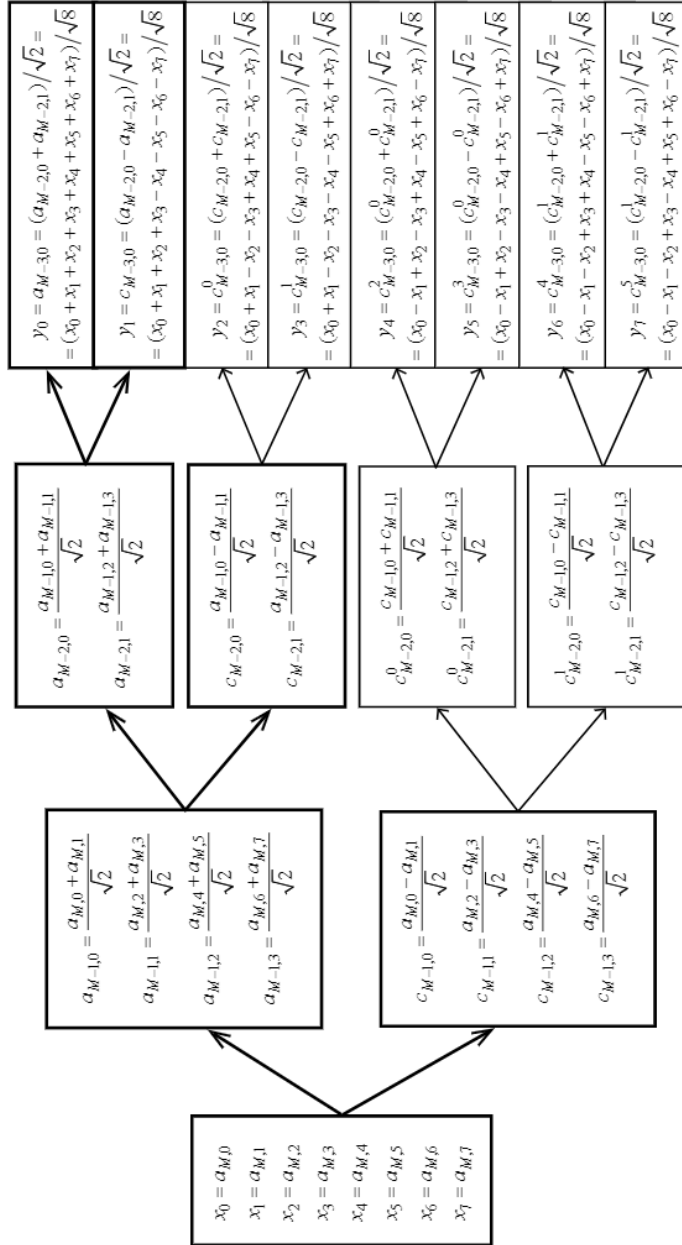


Рис. 7.28. Схема полных вычислений трехуровневого пакетного преобразования на основе фильтров из примера 7.4 для 8-точечного входного вектора. Утолщенными линиями обозначены минимально необходимые вычисления для классического вейвлет-преобразования, не использующего разложения подпространств вейвлетов (см. также рис. 7.26)

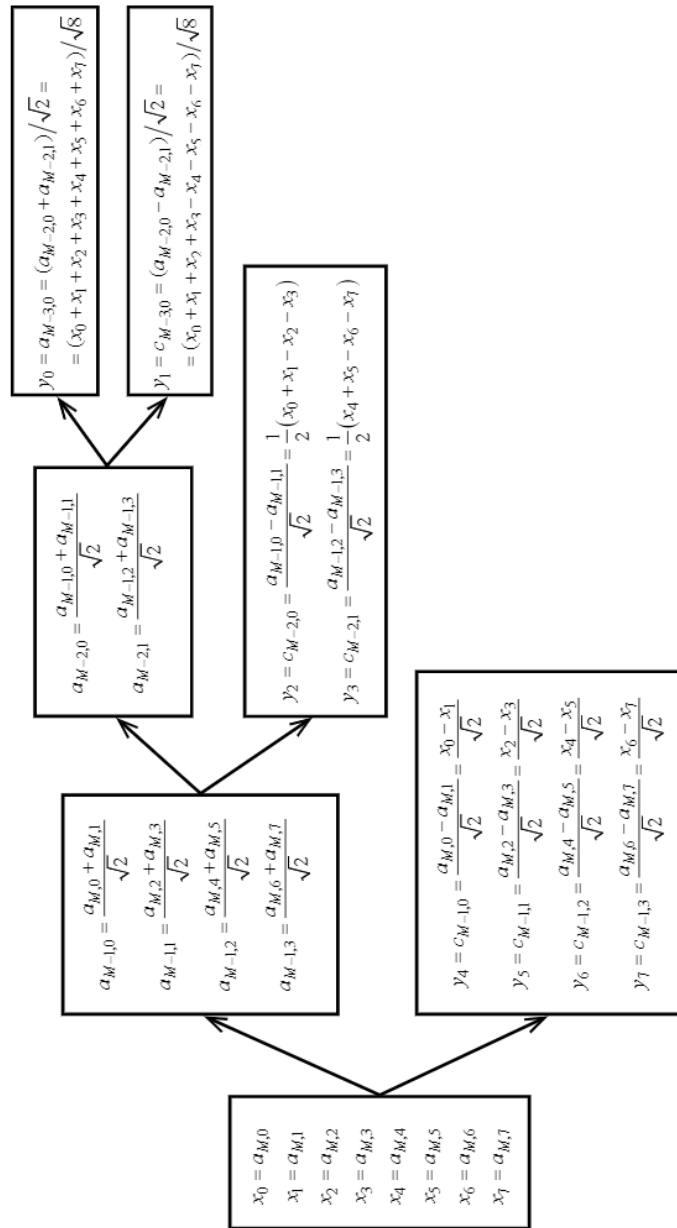


Рис. 7.29. Схема трехуровневого пакетного преобразования на основе фильтров из примера 7.4 для 8-точечного входного вектора с минимальными вычислениями (сравните с рис. 7.28)

Укажем на общие для различных вейвлетов свойства временной локализации базисов пакетов. Пусть масштабирующая функция $\phi(t)$ и вейвлет $\psi(t)$ имеют ограниченные носители: $\text{supp}\phi(t) \subset [0; M-1]$, $\text{supp}\psi(t) \subset [0; M-1]$, где M — целое число. Это означает, что наборы ненулевых коэффициентов в масштабирующих уравнениях (7.1) и (7.3) конечны: $\{h_n\}_{n=0}^{M-1}$, $\{g_n\}_{n=0}^{M-1}$ (см. раздел 7.6 и формулы (7.37), (7.38)). Поэтому полученные из материнского вейвлета базисные функции вейвлет-пакетов

$$\left\{ \psi_{-1,n}^0(t) = \sum_{j=0}^{M-1} h_j \psi(t-2n-j), \quad \psi_{-1,n}^1(t) = \sum_{j=0}^{M-1} g_j \psi(t-2n-j) \right\}_{n \in \mathbb{Z}},$$

где $\psi(t-2n-j) = 0$ при $t \notin [2n+j; 2n+j+M-1]$, обладают носителями $\Omega_{-1,n} = [2n; 2n+2(M-1)]$ вдвое большей длины по сравнению с носителями материнского вейвлета $\Omega_{0,0} = [0; M-1]$. В частном случае, для вейвлета Хаара, мы наблюдали данное явление в примере 7.10.

Таким образом, разложение вейвлет-пространства $W_0 = W_{-1}^0 \oplus W_{-1}^1$ приводит к базисным системам

$$\left\{ \psi_{-1,n}^0(t) = \psi_{-1}^0(t-2n) \right\}_{n \in \mathbb{Z}} \subset W_{-1}^0, \quad \left\{ \psi_{-1,n}^1(t) = \psi_{-1}^1(t-2n) \right\}_{n \in \mathbb{Z}} \subset W_{-1}^1,$$

функции которых имеют худшую временную локализацию. Однако в спектральной области частотная локализация соответствующих систем базисных функций повышается.

Пример 7.12. Оценить, как локализованы в частотной области спектры базисных функций подпространств W_{-1}^0 , W_{-1}^1 и $W_0 = W_{-1}^0 \oplus W_{-1}^1$.

◀ Применяя преобразование Фурье к уравнениям (7.1) и (7.3), в циклических частотах можно представить спектры масштабирующей функции и материнского вейвлета (см. также (7.25)) соответственно:

$$S_\phi(\omega) = \frac{1}{\sqrt{2}} K_H(\omega/2) S_\phi(\omega/2), \quad S_\psi(\omega) = \frac{1}{\sqrt{2}} K_G(\omega/2) S_\phi(\omega/2),$$

где $K_H(\omega) = \sum_{k=0}^{M-1} h_k e^{-i\omega k}$, $K_G(\omega) = \sum_{k=0}^{M-1} g_k e^{-i\omega k}$. Для $\omega \geq 0$ областью частотной локализации основной энергии спектра $S_\phi(\omega) = \mathcal{F}\{\phi(t)\}$ можно считать частотную полосу $0 \leq \omega \leq \pi$, а для спектра $S_\psi(\omega) = \mathcal{F}\{\psi(t)\}$ — полосу $\pi \leq \omega \leq 2\pi$. Действительно (вспомним изученные в разделе 7.5 свойства ЧХ КЗФ), тогда из спектра

$S_\psi(\omega/2) = \mathcal{F}\{2\phi(2t)\}$ с частотной полосой $0 \leq \omega \leq 2\pi$ множитель $K_H(\omega/2)$ выделяет нижние $0 \leq \omega \leq \pi$, а $K_G(\omega/2)$ — верхние $\pi \leq \omega \leq 2\pi$ частоты (см. также пример 7.5). Отметим, что такое расщепление спектра на две половины соответствует частотным свойствам базисных функций подпространств КМА и подпространств вейвлетов, рассмотренным в разделе 7.1, хотя и является весьма условным: реально частотные полосы масштабирующей функции $\phi(t)$ и вейвлета $\psi(t)$ заметно шире и частично перекрываются. (Для простейшего КМА из примера 7.1 это перекрытие весьма существенно, см. рис. 7.2, где линейная частота $\nu = 1$ соответствует на рис. 7.30 циклической частоте $\omega = 2\pi$.)

Найдем спектры базисных функций подпространств W_{-1}^0 и W_{-1}^1 :

$$S_n^0(\omega) = \mathcal{F}\left\{\psi_{-1,n}^0(t) = \sum_{j=0}^{M-1} h_j \psi(t - 2n - j)\right\} = e^{-2i\omega n} K_H(\omega) S_\psi(\omega),$$

$$S_n^1(\omega) = \mathcal{F}\left\{\psi_{-1,n}^1(t) = \sum_{j=0}^{M-1} g_j \psi(t - 2n - j)\right\} = e^{-2i\omega n} K_G(\omega) S_\psi(\omega)$$

(убедитесь!). Учитывая свойства ЧХ $K_H(\omega)$ и $K_G(\omega)$ (см. раздел 7.5), видим, что каждая из частотных полос спектров $S_n^0(\omega) = \mathcal{F}\{\psi_{-1,n}^0(t)\}$ и $S_n^1(\omega) = \mathcal{F}\{\psi_{-1,n}^1(t)\}$ покрывает около половины полосы $\pi \leq \omega \leq 2\pi$ спектра $S_\psi(\omega)$, что схематично отражено на рис. 7.30. Таким образом, по сравнению с подпространством W_0 частотная локализация базисных функций подпространств W_{-1}^0 и W_{-1}^1 повышается вдвое. ►

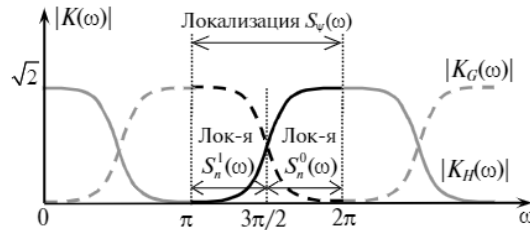


Рис. 7.30. Расщепление спектров базисных функций при разложении подпространства вейвлетов $W_0 = W_{-1}^0 \oplus W_{-1}^1$

Дополнительное разложение вейвлет-пространства $W_0 = W_{-1}^0 \oplus W_{-1}^1$, которое снижает временную, но повышает частотную локализацию базисных функций, может дать более эффективное представление сигнала. Лучшая аппроксимация возможна в том случае, если на интервале, определяемом носителем базисных

функций во временной области, анализируемый сигнал можно считать стационарным (см. также обсуждение данного вопроса в разделе 6.8).

После расширения, для заданного набора коэффициентов $\{h_n\}$ и $\{g_n\}$ масштабирующих уравнений (7.1) и (7.3), возможностей построения базисов разложения сигналов за счет использования вейвлет-пакетов необходимо найти ответ на вопрос о том, по какому критерию следует принимать решение о целесообразности применения (или неприменения) формул (7.19), (7.20) к откликам ВЧ-фильтров. То есть необходимо определить правило выбора базиса вейвлет-пакетов из их многообразия. Для этого предложены различные критерии, основанные на минимизации *энтропии сигнала относительно базиса*.

Общая идея используемых критериев состоит в том, что они характеризуют качество аппроксимации по концентрации распределения энергии исходного дискретного сигнала в наборе коэффициентов-трансформант. При этом предполагается, что сравниваемые базисы ортогональны, т.е. евклидова норма преобразованного вектора сохраняется (см. лемму 2.1 и пример 2.16).

Пусть дискретный сигнал представлен вещественным вектором $\mathbf{X} = (x_0, \dots, x_{N-1})$, а вектор $\mathbf{Y} = (y_0, \dots, y_{N-1})$ (также вещественный) получен в результате дискретного ортогонального преобразования (например, пакетного вейвлет-преобразования) вектора \mathbf{X} . В хорошо подобранном базисе энергия сигнала $E = \sum_{k=0}^{N-1} x_k^2 = \sum_{k=0}^{N-1} y_k^2$ должна иметь выраженную локализацию, т.е. как можно более неравномерное распределение по компонентам: малое количество коэффициентов-трансформант вектора \mathbf{Y} должно нести основную часть энергии дискретного сигнала \mathbf{X} .

Характеризовать неравномерность распределения энергии в векторе \mathbf{Y} можно, например, по следующему параметру [31]:

$$H = - \sum_{k=0}^{N-1} y_k^2 \log y_k^2, \quad (7.57)$$

где общее для всех слагаемых основание логарифма может быть выбрано произвольно. При фиксированном значении $E = \sum_{k=0}^{N-1} y_k^2$ максимум величины (7.57) достигается при равномерном распределении энергии по трансформантам, т.е. когда $y_0^2 = y_1^2 = \dots = y_{N-1}^2$ (это доказывается аналогично доказательству свойства 3° энтропии

(5.2), см. раздел 5.1). Чем сильнее выражена неравномерность распределения энергии в векторе Y , тем меньше значение параметра (7.57), который носит название *энтропии Шеннона*.

Таким образом, величину (7.57) можно использовать в качестве критерия выбора ортогональных базисов дискретных преобразований: для заданного сигнала X лучшим из некоторого набора исследуемых базисов является тот, для которого значение параметра (7.57) будет наименьшим. Критерий (7.57) получил наибольшее распространение на практике для выбора базисов вейвлет-пакетов.

Другая популярная форма для энтропии сигнала относительно базиса — *логарифмическая*:

$$H = \sum_{k=0}^{N-1} \log y_k^2 = \log \prod_{k=0}^{N-1} y_k^2,$$

которая, как несложно показать, также принимает максимальное значение при равномерном распределении энергии сигнала по трансформантам. Похожая форма для критерия эффективности декоррелирующих преобразований уже встречалась нам ранее, см. (6.8), (6.9).

Как фильтрация шумов, так и сжатие сигналов в базисах вейвлет-пакетов обычно дают лучшие результаты по сравнению с «классическим» вейвлет-базисом. Однако нужно отметить, что вычислительная сложность обработки сигналов в базисе вейвлет-пакетов значительно возрастает, так как процесс подбора базиса требует выполнения большего количества процедур фильтрации (7.19) и (7.20), а также вычисления оценок (7.57).

7.14. Вычисление ДВП по схеме лифтинга

Рассмотрим выполнение прямого и обратного одномерного ДВП на один уровень по схеме вычислений рис. 7.10. В отличие от раздела 7.5, все дальнейшие рассуждения будем проводить для более общего случая биортогональных ДВП, когда ПФ фильтров анализа $\tilde{H}(z) = \sum_n \tilde{h}(n)z^{-n}$ и $\tilde{G}(z) = \sum_n \tilde{g}(n)z^{-n}$ определяются конечными ИХ $\{\tilde{h}(n) = h_{-n}\}$ и $\{\tilde{g}(n) = g_{-n}\}$, а ПФ фильтров синтеза $H(z) = \sum_n h(n)z^{-n}$ и $G(z) = \sum_n g(n)z^{-n}$ — конечными ИХ $\{h(n) = \tilde{h}_n\}$ и $\{g(n) = \tilde{g}_n\}$, причем, вообще говоря, $h_n \neq \tilde{h}_n$, $g_n \neq \tilde{g}_n$ (см. раздел 7.8).

Следовательно, $\tilde{H}(z) \neq H(z^{-1})$, $\tilde{G}(z) \neq G(z^{-1})$ и условия точного восстановления сигнала при обработке по схеме рис. 7.10 для биортогональных фильтров примут более общий, по сравнению с (7.30), вид (убедитесь!):

$$\begin{cases} \tilde{H}(z)H(z) + \tilde{G}(z)G(z) = 2 \\ \tilde{H}(-z)H(z) + \tilde{G}(-z)G(z) = 0 \end{cases} \quad (7.58)$$

Все передаточные функции в (7.58) имеют вид *полиномов Лорана*, т.е. представляют собой многочлены вида $L(z) = \sum_{k=M}^N a_k z^{-k}$, где целые числа $N \geq M$. Отметим, что любой полином Лорана можно также выразить в виде $L(z) = z^{-J} P_K(z)$, где $J \in \mathbb{Z}$, а $P_K(z) = \sum_{n=0}^K b_n z^n$ представляет собой «традиционный» алгебраический многочлен порядка $K \geq 0$. Очевидно, что результат сложения, перемножения полиномов Лорана также является полиномом Лорана.

Введем *матрицы модуляции* для фильтров рис. 7.10 следующим образом:

$$\mathbf{M}(z) = \begin{bmatrix} H(z) & H(-z) \\ G(z) & G(-z) \end{bmatrix}, \quad \tilde{\mathbf{M}}(z) = \begin{bmatrix} \tilde{H}(z) & \tilde{H}(-z) \\ \tilde{G}(z) & \tilde{G}(-z) \end{bmatrix}.$$

Тогда (7.58) можно переписать в виде

$$\tilde{\mathbf{M}}(z)^T \mathbf{M}(z) = 2\mathbf{E}, \quad (7.59)$$

где \mathbf{E} — единичная матрица размерности 2×2 .

Обозначив $H_0(z) = \sum_k h(2k)z^{-k}$, $H_1(z) = \sum_k h(2k+1)z^{-k}$, $\tilde{H}_0(z) = \sum_k \tilde{h}(2k)z^{-k}$, $\tilde{H}_1(z) = \sum_k \tilde{h}(2k-1)z^{-k}$ (все суммы — конечные), представим ПФ

$$H(z) = H_0(z^2) + z^{-1}H_1(z^2), \quad \tilde{H}(z) = \tilde{H}_0(z^2) + z\tilde{H}_1(z^2).$$

Аналогично: $G(z) = G_0(z^2) + z^{-1}G_1(z^2)$ и $\tilde{G}(z) = \tilde{G}_0(z^2) + z\tilde{G}_1(z^2)$. Назовем *полифазными матрицами*

$$\mathbf{P}(z) = \begin{bmatrix} H_0(z) & G_0(z) \\ H_1(z) & G_1(z) \end{bmatrix}, \quad \tilde{\mathbf{P}}(z) = \begin{bmatrix} \tilde{H}_0(z) & \tilde{G}_0(z) \\ \tilde{H}_1(z) & \tilde{G}_1(z) \end{bmatrix}.$$

Тогда (убедитесь!)

$$\mathbf{M}(z) = \mathbf{P}(z^2)^T \begin{bmatrix} 1 & 1 \\ z^{-1} & -z^{-1} \end{bmatrix}, \quad \tilde{\mathbf{M}}(z) = \tilde{\mathbf{P}}(z^2)^T \begin{bmatrix} 1 & 1 \\ z & -z \end{bmatrix}$$

и подстановка полученных представлений в (7.59) дает уравнение $\tilde{\mathbf{P}}(z^2)\mathbf{P}(z^2)^T = \mathbf{E}$. Так как величина z^2 может принимать любое комплексное значение, полученное соотношение запишем в виде

$$\tilde{\mathbf{P}}(z)\mathbf{P}(z)^T = \mathbf{E}. \quad (7.60)$$

Равенство (7.60) определяет, что $\det \mathbf{P}(z) = 1/\det \tilde{\mathbf{P}}(z)$. Это возможно только в случае, когда определители, которые являются полиномами Лорана, являются одночленами, т.е. $\det \mathbf{P}(z) = Cz^{-J}$ и $\det \tilde{\mathbf{P}}(z) = z^J/C$, где $J \in \mathbb{Z}$, а $C \neq 0$. Ограничение

$$\det \mathbf{P}(z) = H_0(z)G_1(z) - H_1(z)G_0(z) = Cz^{-J}$$

означает, что, умножая на некоторое число только ИХ $g(n)$ и соответствующий полином $G(z) = \sum_n g(n)z^{-n}$, всегда можно добиться выполнения условия $C = 1$. Выполняя при необходимости сдвиг ИХ на N отсчетов: $g(n) \rightarrow g(n - N)$, который соответствует умножению полинома $G(z)$ на величину z^{-N} , можно также получить значение $J = 0$. Таким образом, если задана пара КИХ-фильтров $H(z)$ и $G(z)$, путем умножения на число и сдвига только ИХ $g(n)$ всегда можно добиться для полифазной матрицы $\mathbf{P}(z)$ выполнения условия $\det \mathbf{P}(z) = 1$. Тогда также $\det \tilde{\mathbf{P}}(z) = 1$. Поэтому везде далее мы будем считать, что $\det \mathbf{P}(z) = \det \tilde{\mathbf{P}}(z) = 1$.

Упражнение. Как нужно модифицировать ИХ $g(n)$ и $\tilde{g}(n) = g(-n)$ ВЧ-фильтров из примера 7.4, чтобы получить $\det \mathbf{P}(z) = \det \tilde{\mathbf{P}}(z) = 1$?

Если задана одна из полифазных матриц $\mathbf{P}(z)$ или $\tilde{\mathbf{P}}(z)$, то вторая сразу же находится из решения уравнения (7.60). Например, задав $\mathbf{P}(z)$, имеем:

$$\tilde{\mathbf{P}}(z) = \begin{bmatrix} \tilde{H}_0(z) & \tilde{G}_0(z) \\ \tilde{H}_1(z) & \tilde{G}_1(z) \end{bmatrix} = (\mathbf{P}(z)^T)^{-1} = \begin{bmatrix} H_0(z) & H_1(z) \\ G_0(z) & G_1(z) \end{bmatrix}^{-1} = \begin{bmatrix} G_1(z) & -H_1(z) \\ -G_0(z) & H_0(z) \end{bmatrix}.$$

Отсюда получаем соотношения, связывающие ПФ квадратурных фильтров:

$$\begin{aligned} \tilde{H}(z) &= \sum_n \tilde{h}(n)z^{-n} = \tilde{H}_0(z^2) + z\tilde{H}_1(z^2) = G_1(z^2) - zG_0(z^2) = \\ &= \sum_k g(2k+1)z^{-2k} - z\sum_k g(2k)z^{-2k} = z\left(\sum_k g(2k+1)z^{-(2k+1)} - \sum_k g(2k)z^{-2k}\right) = \\ &= -z\sum_k g(k)(-z)^{-k} = -zG(-z). \end{aligned}$$

Аналогично, $\tilde{G}(z) = \tilde{G}_0(z^2) + z\tilde{G}_1(z^2) = -H_1(z^2) + zH_0(z^2) = \dots = zH(-z)$.

Упражнение. Покажите, что условия $\tilde{H}(z) = -zG(-z)$ и $\tilde{G}(z) = zH(-z)$ эквивалентны требованиям $\tilde{h}(n) = (-1)^n g(n+1)$ и $\tilde{g}(n) = (-1)^{n+1} h(n+1)$.

Чтобы определить КИХ-фильтры, соответствующие некоторому биортогональному ДВП, фактически необходимо лишь задать

полифазную матрицу $P(z)$ (или $\tilde{P}(z)$), элементами которой являются многочлены Лорана, а определитель равен единице.

Используя полифазное представление ИХ (см. раздел 3.10) для квадратурных фильтров, схему обработки сигнала, соответствующую рис. 7.10, можно изобразить как на рис. 7.31. Количество арифметических операций, необходимых для вычисления сверток в фильтрах, снижается при этом примерно в два раза.

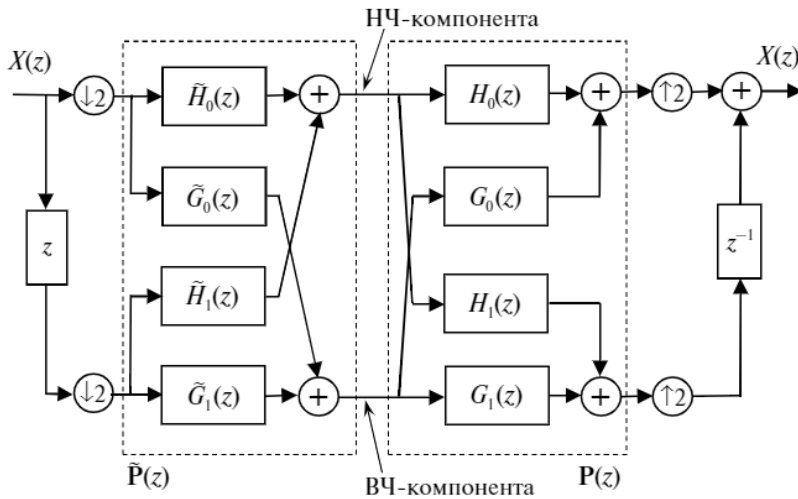


Рис. 7.31. Полифазная схема обработки сигнала, эквивалентная рис. 7.10

Упражнение. Простейшие матрицы модуляции — единичные, когда $P(z) = \tilde{P}(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Разложение сигнала на НЧ- и ВЧ-компоненты в данном случае реально не выполняется, так как полифазные компоненты фактически не обрабатываются. Такое преобразование называют «ленивым»¹. Изобразите, как упростится и будет выглядеть в этом случае схема обработки рис. 7.31.

Определение. Фильтры с ПФ $H(z)$ и $G(z)$ будем называть *комплементарной парой* (или *комплементарными фильтрами*), если

¹ В оригинале *lazy transform* [60, 61].

соответствующая полифазная матрица $\mathbf{P}(z)$ имеет определитель, равный единице: $\det \mathbf{P}(z) = 1$.

Как следует из соотношения (7.60), комплементарная пара фильтров синтеза $(H(z), G(z))$ порождает комплементарную пару фильтров анализа $(\tilde{H}(z), \tilde{G}(z))$, и наоборот. Установим необходимые и достаточные условия, выполнение которых позволяет преобразовывать фильтры, сохраняя их комплементарность.

Теорема 7.2. Для произвольной комплементарной пары $(H(z), G(z))$ новый фильтр с ПФ $G^{new}(z)$ образует комплементарную пару $(H(z), G^{new}(z))$ тогда и только тогда, когда ПФ имеет вид

$$G^{new}(z) = G(z) + S(z^2)H(z),$$

где $S(z)$ — произвольный полином Лорана.

◀ Представим новую ПФ в виде $G^{new}(z) = G(z) + D(z)$. Из условия

$$\det \underbrace{\begin{bmatrix} H_0(z) & G_0^{new}(z) \\ H_1(z) & G_1^{new}(z) \end{bmatrix}}_{\mathbf{P}^{new}(z)} = \det \underbrace{\begin{bmatrix} H_0(z) & G_0(z) \\ H_1(z) & G_1(z) \end{bmatrix}}_{\mathbf{P}(z)} + \det \begin{bmatrix} H_0(z) & D_0(z) \\ H_1(z) & D_1(z) \end{bmatrix} = 1$$

следует, что $H_0(z)D_1(z) = H_1(z)D_0(z)$. Обозначим $S(z) = D_1(z)/H_1(z) = D_0(z)/H_0(z)$. Для того чтобы при произвольных многочленах $H_0(z)$ и $H_1(z)$ функции $D_0(z) = H_0(z)S(z)$ и $D_1(z) = H_1(z)S(z)$ являлись полиномами Лорана, необходимо и достаточно в качестве $S(z)$ (в общем случае, дробно-рациональной функции) взять какой-либо (любой!) многочлен Лорана. Тогда

$$D(z) = D_0(z^2) + z^{-1}D_1(z^2) = S(z^2)(H_0(z^2) + z^{-1}H_1(z^2)) = S(z^2)H(z)$$

и теорема доказана. ▶

Преобразование комплементарной пары $(H(z), G(z))$ в $(H(z), G^{new}(z))$, которое определяет теорема 7.2, в терминах полифазных матриц имеет вид:

$$\mathbf{P}^{new}(z) = \begin{bmatrix} H_0(z) & G_0^{new}(z) \\ H_1(z) & G_1^{new}(z) \end{bmatrix} = \mathbf{P}(z) \begin{bmatrix} 1 & S(z) \\ 0 & 1 \end{bmatrix}. \quad (7.61)$$

Учитывая условие (7.60), несложно показать (проделайте это самостоятельно), что изменение полифазной матрицы фильтров синтеза, которое определяет соотношение (7.61), влечет следующую модификацию полифазной матрицы фильтров анализа:

$$\tilde{\mathbf{P}}^{new}(z) = \tilde{\mathbf{P}}(z) \begin{bmatrix} 1 & 0 \\ -S(z) & 1 \end{bmatrix}. \quad (7.62)$$

Применение (7.62) к комплементарным фильтрам анализа $(\tilde{H}(z), \tilde{G}(z))$ порождает новую комплементарную пару фильтров $(\tilde{H}^{new}(z), \tilde{G}(z))$, где

$$\tilde{H}^{new}(z) = \tilde{H}(z) - S(z^2)\tilde{G}(z).$$

Процедуры преобразования комплементарных фильтров анализа и синтеза с каким-либо полиномом Лорана $S(z)$ в соответствии с (7.61) и (7.62) называют *лифтингом*.

Упражнение. Покажите, что для произвольной комплементарной пары $(H(z), G(z))$ любая новая пара $(H^{new}(z), G(z))$ является комплементарной тогда и только тогда, когда новый фильтр пары имеет ПФ вида

$$H^{new}(z) = H(z) + T(z^2)G(z),$$

где $T(z)$ — какой-либо многочлен Лорана. Убедитесь, что соответствующее преобразование полифазной матрицы примет вид:

$$\mathbf{P}^{new}(z) = \mathbf{P}(z) \begin{bmatrix} 1 & 0 \\ T(z) & 1 \end{bmatrix}. \quad (7.63)$$

Для фильтров анализа преобразование (7.63) влечет следующее изменение полифазной матрицы, вытекающее из условия (7.60):

$$\tilde{\mathbf{P}}^{new}(z) = \tilde{\mathbf{P}}(z) \begin{bmatrix} 1 & -T(z) \\ 0 & 1 \end{bmatrix}, \quad (7.64)$$

которое порождает новый фильтр

$$\tilde{G}^{new}(z) = \tilde{G}(z) - T(z^2)\tilde{H}(z).$$

Пара соотношений (7.63) и (7.64) определяет процедуру *дуального лифтинга*.

Применение лифтинга и дуального лифтинга можно представить как модификацию схемы обработки рис. 7.10, приведенную на рис. 7.32.

Процедуры лифтинга и дуального лифтинга позволяют разложить на множители полифазные матрицы $\mathbf{P}(z)$, $\tilde{\mathbf{P}}(z)$ и построить их факторизованное представление, которое влечет сокращение числа вычислительных операций, необходимых для реализации прямого или обратного ДВП. Для пояснения методики такой факторизации

сначала необходимо рассмотреть специфику операции деления многочленов Лорана.

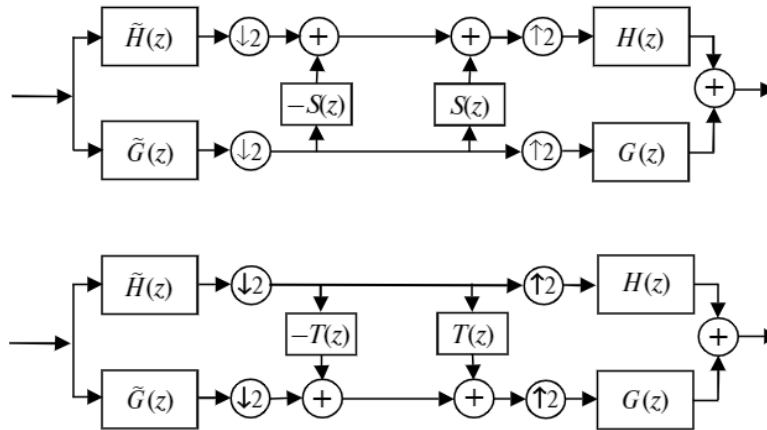


Рис. 7.32. Интерпретация лифтинга (вверху) и дуального лифтинга (внизу) как изменение схемы обработки квадратурными фильтрами

Определение. Порядком полинома Лорана $L(z) = \sum_{k=M}^N a_k z^{-k}$, где $a_M \neq 0$ и $a_N \neq 0$, назовем величину $|L(z)| = N - M$. Для нулевого полинома $L(z) = 0$ считаем порядок $|0| = -\infty$.

Замечание. Любой одночлен вида $L(z) = c z^{-k}$ имеет порядок $|L(z)| = 0$.

Для двух произвольных полиномов Лорана $A(z)$ и $B(z) \neq 0$, $|A(z)| \geq |B(z)|$, всегда существуют такие полиномы $Q(z)$ (частное, $|Q(z)| = |A(z)| - |B(z)|$) и $R(z)$ (остаток от деления, $|R(z)| < |B(z)|$), что

$$A(z) = B(z)Q(z) + R(z).$$

Для операций нахождения частного и остатка от деления полиномов будем использовать обозначения $Q(z) = A(z) \setminus B(z)$, $R(z) = A(z) \% B(z)$ соответственно. В общем случае деление полиномов Лорана без остатка (точное деление, когда $R(z) = 0$) невозможно, а результат деления

$$\frac{A(z)}{B(z)} = Q(z) + \frac{R(z)}{B(z)},$$

где $|R(z)| < |B(z)|$, не является единственным.

Пример 7.13. Найти результат деления полинома $A(z) = z^{-1} + 6 + z$ на полином $B(z) = 1 + z$.

◀ Так как $|A(z)|=2$ и $|B(z)|=1$, то порядки полиномов для частного $|Q(z)|=|A(z)|-|B(z)|=1$, а для остатка от деления $|R(z)|=0 < |B(z)|=1$. То есть остаток от деления есть одночлен $R(z) = A(z) - B(z)Q(z) = cz^j$. Это означает, что полином-произведение $B(z)Q(z)$ будет совпадать с полиномом $A(z) = z^{-1} + 6 + z$ для двух его слагаемых, а для одного слагаемого будет несовпадение, которое и даст остаток от деления в виде одночлена $R(z) = cz^j$. Выбирая соответствующее значение $j \in \{-1, 0, 1\}$ по своему усмотрению, мы получим тогда три разных результата (проверьте!) операции деления $A(z)/B(z) = Q(z) + (R(z)/B(z))$:

$$\frac{z^{-1} + 6 + z}{1 + z} = (5z^{-1} + 1) + \frac{-4z^{-1}}{1 + z} = (z^{-1} + 1) + \frac{4}{1 + z} = (z^{-1} + 5) + \frac{-4z}{1 + z}. \quad \blacktriangleright$$

Обобщая результаты примера 7.13, отметим, что если $Q(z) = A(z) \setminus B(z)$, то полином $B(z)Q(z)$ должен совпадать с $A(z)$ по крайней мере для $|A(z)| - |B(z)| + 1$ слагаемых. Несовпадающие не более чем $|B(z)|$ слагаемых формируют остаток деления $R(z) = A(z) \% B(z)$ и могут быть произвольно выбраны в исходном представлении $A(z) = \sum_{j=M}^N a_j z^{-j}$ для $|B(z)|$ последовательно идущих значений индекса j .

Рассмотрим важную для построения метода факторизации полифазных матриц задачу поиска наибольшего общего делителя многочленов Лорана.

Определение. Пусть $A(z)$ и $B(z)$ — полиномы Лорана. *Наибольшим общим делителем* (НОД) назовем для них такой полином Лорана $D(z)$ максимально возможного порядка $|D(z)|$, для которого $A(z) \% D(z) = 0$ и $B(z) \% D(z) = 0$. Если $|D(z)| = 0$, то многочлены $A(z)$ и $B(z)$ называют *взаимно простыми*.

Замечание. НОД $D(z)$ полиномов Лорана определяется с точностью до множителя cz^k . Действительно, если $A(z) = D(z)Q_A(z)$ и $B(z) = D(z)Q_B(z)$, где частные $Q_A(z) = A(z) \setminus D(z) = A(z) / D(z)$ и $Q_B(z) = B(z) \setminus D(z) = B(z) / D(z)$, то изменение делителя $D^{new}(z) = cz^k D(z)$ не изменит его порядка $|D(z)| = |D^{new}(z)|$ и свойств делимости без остатка полиномов $A(z)$ и $B(z)$, а лишь модифицирует частные от деления: $Q_A^{new}(z) = Q_A(z)z^k / c$, $Q_B^{new}(z) = Q_B(z)z^k / c$.

Для поиска НОД многочленов, как и для поиска НОД целых чисел, используется *алгоритм Евклида*. Отличие для случая полиномов Лорана заключается в том, что решение задачи поиска НОД может быть неединственным в силу приведенного выше замечания.

Теорема 7.3. (Алгоритм Евклида для полиномов Лорана.) Пусть заданы два полинома Лорана $A(z)$ и $B(z) \neq 0$, причем $|A(z)| \geq |B(z)|$. Положив $A_0(z) = A(z)$ и $B_0(z) = B(z)$, будем выполнять для $k = 1, 2, \dots, n$ итерационную процедуру:

$$\begin{aligned} A_k(z) &= B_{k-1}(z), \\ B_k(z) &= A_{k-1}(z) \% B_{k-1}(z). \end{aligned} \quad (7.65)$$

Процедура прекращается, как только для некоторого значения $k = n$ получаем $B_n(z) = 0$. Тогда полином Лорана $A_n(z) = B_{n-1}(z)$ будет НОД многочленов $A(z)$ и $B(z)$.

Опустим несложное доказательство данной теоремы¹ и рассмотрим аспекты ее практического применения. Вводя обозначение $Q_k(z) = A_{k-1}(z) \setminus B_{k-1}(z)$, одну итерацию процедуры (7.65) можно представить в виде

$$\begin{bmatrix} A_k(z) \\ B_k(z) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -Q_k(z) \end{bmatrix} \begin{bmatrix} A_{k-1}(z) \\ B_{k-1}(z) \end{bmatrix},$$

а последовательность шагов алгоритма Евклида до достижения условия $B_n(z) = 0$ будет тогда определяться следующим соотношением:

$$\begin{bmatrix} A_n(z) \\ 0 \end{bmatrix} = \prod_{k=n}^1 \begin{bmatrix} 0 & 1 \\ 1 & -Q_k(z) \end{bmatrix} \begin{bmatrix} A(z) \\ B(z) \end{bmatrix}.$$

Отсюда, учитывая, что

$$\begin{bmatrix} 0 & 1 \\ 1 & -Q_k(z) \end{bmatrix}^{-1} = \begin{bmatrix} Q_k(z) & 1 \\ 1 & 0 \end{bmatrix},$$

получаем представление

$$\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = \prod_{k=1}^n \begin{bmatrix} Q_k(z) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} A_n(z) \\ 0 \end{bmatrix}, \quad (7.66)$$

где $A_n(z)$ — НОД полиномов Лорана $A(z)$ и $B(z)$. Если $A_n(z)$ — одночлен, т. е. $|A_n(z)| = 0$, то $A(z)$ и $B(z)$ являются взаимно простыми.

¹ Для алгебраических полиномов с неотрицательными степенями доказательство см., например, [22]. В точности то же доказательство можно привести и в терминах полиномов Лорана.

Пример 7.14. Найти представление (7.66) для полиномов $A(z) = z^{-1} + 6 + z$ и $B(z) = 1 + z$.

◀ Выполняя процедуру (7.65) при $k = 1$, для первого шага алгоритма Евклида имеем, как один из трех возможных вариантов (см. решение примера 7.13):

$$A_1(z) = 1 + z, \quad B_1(z) = 4, \quad Q_1(z) = z^{-1} + 1.$$

Следующий шаг для $k = 2$ дает

$$A_2(z) = 4, \quad B_2(z) = 0, \quad Q_2(z) = \frac{1}{4} + \frac{1}{4}z$$

и оказывается последним. Представление (7.66) получаем тогда в виде:

$$\begin{bmatrix} z^{-1} + 6 + z \\ 1 + z \end{bmatrix} = \begin{bmatrix} z^{-1} + 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} + \frac{1}{4}z & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \end{bmatrix}.$$

Так как найденный НОД $A_2(z) = 4$ является одночленом, $|A_2(z)| = 0$, то полиномы Лорана $A(z) = z^{-1} + 6 + z$ и $B(z) = 1 + z$ являются взаимно простыми. ▶

Упражнение. Убедитесь, что в условиях примера 7.14 два других возможных результата выполнения первого шага алгоритма Евклида (см. решение примера 7.13) также приводят к НОД в виде одночлена, т.е. в любом случае $|A_2(z)| = 0$ в силу того, что $A(z) = z^{-1} + 6 + z$ и $B(z) = 1 + z$ являются взаимно простыми полиномами.

Применим алгоритм Евклида для полифазных компонент $H_0(z) = \sum_k h(2k)z^{-k}$ и $H_1(z) = \sum_k h(2k+1)z^{-k}$ представления $H(z) = H_0(z^2) + z^{-1}H_1(z^2)$. Если для $H_0(z)$ и $H_1(z)$ существует многочлен $D(z)$ (общий делитель) такой, что $H_0(z) = D(z)A(z)$, $H_1(z) = D(z)B(z)$, где $A(z)$ и $B(z)$ также являются полиномами Лорана, то условие $\det \mathbf{P}(z) = 1$ означает, что

$$\det \mathbf{P}(z) = H_0(z)G_1(z) - H_1(z)G_0(z) = D(z) \underbrace{(A(z)G_1(z) - B(z)G_0(z))}_{\text{полином Лорана}} = 1.$$

Следовательно, $D(z)$ является одночленом, т.е. $|D(z)| = 1$, и полиномы $H_0(z)$ и $H_1(z)$ являются взаимно простыми. Для взаимно простых многочленов можно считать НОД равным некоторой константе K (см. замечание к определению НОД полиномов Лорана, данное выше). Тогда с помощью алгоритма Евклида по аналогии с (7.66) мы сможем получить представление

$$\begin{bmatrix} H_0(z) \\ H_1(z) \end{bmatrix} = \prod_{k=1}^n \begin{bmatrix} Q_k(z) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} K \\ 0 \end{bmatrix}, \quad (7.67)$$

однако для этого (см. теорему 7.3) необходимо выполнение условия $|A(z)| = |H_0(z)| \geq |B(z)| = |H_1(z)|$. Возможное нарушение данного условия (если исходно $|H_0(z)| < |H_1(z)|$) можно устранить двумя способами.

Способ 1. Получив разложение вида (7.66) для $A(z) = H_1(z)$, $B(z) = H_0(z)$, где $|A(z)| \geq |B(z)|$, затем переставим компоненты,

$$\begin{bmatrix} H_0(z) \\ H_1(z) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} A(z) \\ B(z) \end{bmatrix},$$

что означает добавление в (7.67) одной матрицы-сомножителя с $Q_1(z) = 0$.

Способ 2. Модифицируем ИХ $h(n)$, выполнив ее сдвиг на один отсчет: $h^{new}(n) = h(n+1)$. Тогда ПФ $H^{new}(z) = zH(z)$, а в новой ИХ четные и нечетные компоненты поменяются местами: $h^{new}(2n) = h(2n+1)$, $h^{new}(2n+1) = h(2n+2)$. Получим $|A(z)| = |H_0^{new}(z)| \geq |B(z)| = |H_1^{new}(z)|$ (убедитесь). Отметим, что для сохранения свойства комплементарности пары фильтров $(H^{new}(z), G^{new}(z))$ модификация ПФ $H^{new}(z) = zH(z)$ потребует также изменения ПФ $G^{new}(z) = -z^{-1}G(z)$.

Упражнение. Покажите, что изменение ИХ $h^{new}(n) = h(n+1)$ влечет следующее изменение ИХ соответствующего комплементарного фильтра: $g^{new}(n) = -g(n-1)$. *Указание:* выразите элементы новой полифазной матрицы $P^{new}(z)$ через компоненты исходной полифазной матрицы $P(z)$ и убедитесь, что если $\det P(z) = 1$, то и $\det P^{new}(z) = 1$.

Применение способов «переворачивания» компонент $H_0(z)$ и $H_1(z)$ к (7.67) либо добавляет (способ 1), либо не добавляет (способ 2) матрицу-сомножитель в факторизованное представление (7.67). Поэтому будем считать далее, что число матриц-сомножителей n в представлении (7.67) всегда является четным: если перестановка компонент $H_0(z)$ и $H_1(z)$ не требуется, а число n — нечетное, то способы 1 и 2 применяются совместно. Отметим, что выполняемые в способе 2 сдвиги импульсных характеристик комплементарных фильтров не изменяют их АЧХ.

Получив для некоторого заданного фильтра с ПФ $H(z) = H_0(z^2) + z^{-1}H_1(z^2)$ представление (7.67), всегда можно сразу же сформировать и ПФ $G^0(z) = G_0^0(z^2) + z^{-1}G_1^0(z^2)$ одного из возможных комплементарных ему фильтров, если построить соответствующую полифазную матрицу следующим образом:

$$\mathbf{P}^0(z) = \begin{bmatrix} H_0(z) & G_0^0(z) \\ H_1(z) & G_1^0(z) \end{bmatrix} = \prod_{k=1}^n \begin{bmatrix} Q_k(z) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix}. \quad (7.68)$$

Действительно, поскольку число n — четное и определитель произведения матриц равен произведению определителей матриц-сомножителей, то всегда получим в (7.68) $\det \mathbf{P}^0(z) = 1$. Все прочие комплементарные фильтры $G(z)$ могут быть получены в результате лифтинга (7.61):

$$\mathbf{P}(z) = \mathbf{P}^0(z) \begin{bmatrix} 1 & S(z) \\ 0 & 1 \end{bmatrix}, \quad (7.69)$$

где в качестве $S(z)$ может быть выбран любой полином Лорана, см. теорему 7.2.

Подводя итоги рассмотрения вопросов факторизации полифазных матриц, сформулируем их в виде следующей теоремы.

Теорема 7.4. Пусть задана какая-то комплементарная пара фильтров $(H(z), G(z))$, определяемая полифазной матрицей $\mathbf{P}(z)$. Тогда обязательно найдутся такие полиномы Лорана $S_k(z)$, $T_k(z)$, $k \in \{1, 2, \dots, m\}$, и константа $K \neq 0$, что

$$\mathbf{P}(z) = \prod_{k=1}^m \begin{bmatrix} 1 & S_k(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ T_k(z) & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix}. \quad (7.70)$$

◀ Используем для нечетных и четных значений индекса k в разложении (7.68) представления

$$\begin{bmatrix} Q_k(z) & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & Q_k(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{и} \quad \begin{bmatrix} Q_k(z) & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ Q_k(z) & 1 \end{bmatrix}$$

соответственно. Тогда учитывая, что число n является четным, можем переписать (7.68) в следующем виде:

$$\mathbf{P}^0(z) = \prod_{k=1}^{n/2} \begin{bmatrix} 1 & Q_{2k-1}(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ Q_{2k}(z) & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix} = \prod_{k=1}^{m-1} \begin{bmatrix} 1 & S_k(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ T_k(z) & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix},$$

где $S_k(z) = Q_{2k-1}(z)$, $T_k(z) = Q_{2k}(z)$, $m-1 = n/2$. Чтобы получить требуемое представление (7.70), полученное разложение для $\mathbf{P}^0(z)$

необходимо дополнить парой матриц-сомножителей, положив $S_m(z) = K^2 S(z)$ и $T_m(z) = 0$, так как (см. (7.68), (7.69))

$$\begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix} \begin{bmatrix} 1 & S(z) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & K^2 S(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix}. \quad \blacktriangleright$$

Имея факторизацию (7.70) полифазной матрицы фильтров синтеза, на основании (7.60) можно сразу же записать разложение полифазной матрицы для фильтров анализа:

$$\tilde{\mathbf{P}}(z) = (\mathbf{P}(z)^T)^{-1} = \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix} \prod_{k=m}^1 \begin{bmatrix} 1 & 0 \\ -T_k(z) & 1 \end{bmatrix} \begin{bmatrix} 1 & -S_k(z) \\ 0 & 1 \end{bmatrix}. \quad (7.71)$$

Каждая из матриц факторизованных представлений в (7.70) и (7.71) имеет единичный детерминант и определяет некоторую модификацию полифазной матрицы, представляющей собой произведение предыдущих матриц-сомножителей. Начальной можно считать матрицу «ленивого» преобразования $\mathbf{P}_1(z) = \tilde{\mathbf{P}}_1(z) = \mathbf{E}$ (единичная матрица размерности 2×2). Тогда дискретная декомпозиция сигнала, определяемая парой комплементарных фильтров $(\tilde{H}(z), \tilde{G}(z))$ и соответствующей им факторизацией (7.71), может быть изображена в виде схемы обработки, приведенной на рис. 7.33а. На рис. 7.33б представлена схема восстановления дискретного сигнала, которая определяется факторизацией полифазной матрицы (7.70) фильтров синтеза $(H(z), G(z))$.

Для биортогональных ДВП имеется бóльшая, по сравнению с ортогональными ДВП, вариативность применения схемы лифтинга для факторизации полифазных матриц и построения вычислительных схем рис. 7.33. Так, сначала можно получить разложение (7.67) для фильтров синтеза, применяя алгоритм Евклида к паре $A(z) = H_0(z)$ и $B(z) = H_1(z)$, а затем по найденным полиномам $S_k(z) = Q_{2k-1}(z)$, $T_k(z) = Q_{2k}(z)$ и константе K записать представления (7.70) и (7.71). Другой вариант факторизации полифазных матриц $\mathbf{P}(z)$ и $\tilde{\mathbf{P}}(z)$ можно получить, если алгоритм Евклида применить к фильтрам анализа $\tilde{H}_0(z)$ и $\tilde{H}_1(z)$. Рассмотрим это на примере.

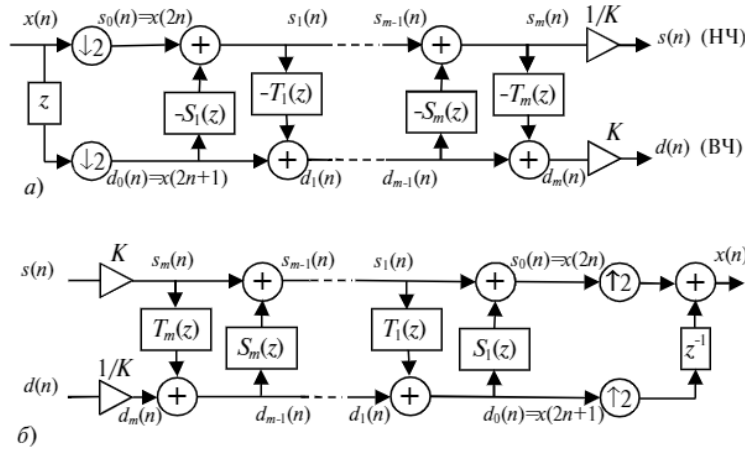


Рис. 7.33. Реализация прямого (а) и обратного (б) биортогонального ДВП, соответствующая факторизованному представлению полифазных матриц (7.71) и (7.70) соответственно

Пример 7.15. Найти разложение полифазных матриц и привести соответствующую рис. 7.33 вычислительную схему для биортогональных фильтров из примера 7.8.

◀ Найдем сначала разложение

$$\begin{bmatrix} \tilde{H}_0(z) \\ \tilde{H}_1(z) \end{bmatrix} = \prod_{k=1}^n \begin{bmatrix} Q_k(z) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} K \\ 0 \end{bmatrix}$$

для полиномов $\tilde{H}_0(z) = h_4 z^{-2} + h_2 z^{-1} + h_0 + h_2 z + h_4 z^2$ и $\tilde{H}_1(z) = h_3 z^{-2} + h_1 z^{-1} + h_1 + h_3 z$, коэффициенты которых заданы в примере 7.8 и определяют НЧ-фильтр анализа с ПФ $\tilde{H}(z) = \tilde{H}_0(z^2) + z\tilde{H}_1(z^2)$.

Положив $A_0(z) = \tilde{H}_0(z)$ и $B_0(z) = \tilde{H}_1(z)$, применим алгоритм Евклида (см. теорему 7.3). Как возможный вариант (с учетом неоднозначности операции деления полиномов) получаем последовательно:

$$\begin{aligned} Q_1(z) &= A_0(z) \setminus B_0(z) = \alpha(1+z), & \text{где } \alpha &= h_4/h_3, \\ B_1(z) &= A_0(z) - Q_1(z)B_0(z) = r_1 z^{-1} + r_0 + r_1 z, \\ & \text{где } r_1 &= h_2 - h_4 - (h_1 h_4/h_3), \quad r_0 = h_0 - 2(h_1 h_4/h_3), \\ A_1(z) &= h_3 z^{-2} + h_1 z^{-1} + h_1 + h_3 z; \end{aligned}$$

$$\begin{aligned} Q_2(z) &= A_1(z) \setminus B_1(z) = \beta(z^{-1} + 1), & \text{где } \beta &= h_3/r_1, \\ B_2(z) &= A_1(z) - Q_2(z)B_1(z) = s_0z^{-1} + s_0, & \text{где } s_0 &= h_1 - h_3 - h_3(r_0/r_1), \\ A_2(z) &= r_1z^{-1} + r_0 + r_1z; \end{aligned}$$

$$\begin{aligned} Q_3(z) &= A_2(z) \setminus B_2(z) = \gamma(1 + z), & \text{где } \gamma &= r_1/s_0, \\ B_3(z) &= A_2(z) - Q_3(z)B_2(z) = K, & \text{где } K &= r_0 - 2r_1 \\ A_3(z) &= s_0z^{-1} + s_0; \end{aligned}$$

$$\begin{aligned} Q_4(z) &= A_3(z) \setminus B_3(z) = \delta(z^{-1} + 1), & \text{где } \delta &= s_0/K, \\ B_4(z) &= A_3(z) - Q_4(z)B_3(z) = 0, \\ A_4(z) &= K. \end{aligned}$$

Дополним фильтр $\tilde{H}(z)$ комплементарным фильтром $\tilde{G}(z)$, построив полифазную матрицу в соответствии с (7.68):

$$\tilde{\mathbf{P}}(z) = \prod_{k=1}^4 \begin{bmatrix} Q_k(z) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix}.$$

Тогда обозначая $S_k(z) = Q_{2k-1}(z)$, $T_k(z) = Q_{2k}(z)$, $k \in \{1, 2\}$, на основании теоремы 7.4 получим факторизацию

$$\begin{aligned} \tilde{\mathbf{P}}(z) &= \begin{bmatrix} 1 & S_1(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ T_1(z) & 1 \end{bmatrix} \begin{bmatrix} 1 & S_2(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ T_2(z) & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix} = \\ &= \begin{bmatrix} 1 & \alpha(1+z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta(z^{-1}+1) & 1 \end{bmatrix} \begin{bmatrix} 1 & \gamma(1+z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \delta(z^{-1}+1) & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix}, \end{aligned}$$

которая определяет вычисление как прямого, так и обратного ДВП по схеме лифтинга — см. рис. 7.33, где нужно положить $m = 2$. Введенные выше параметры имеют здесь следующие числовые значения, найденные по заданному набору $\{h_n\}$:

$$\alpha \approx -1,58613; \beta \approx -0,0529801; \gamma \approx 0,882911; \delta \approx 0,443507; K \approx 1,1496.$$

Каждый из четырех фильтров с ПФ $S_1(z) = \alpha(1+z)$, $T_1(z) = \beta(z^{-1}+1)$, $S_2(z) = \gamma(1+z)$, $T_2(z) = \delta(z^{-1}+1)$ для вычисления очередного отсчета сигнала выполняет одно сложение и одно умножение. Кроме того, для каждого значения индекса n выполняется 4 сложения в сумматорах и 2 умножения на нормировочные константы K и $1/K$, см. рис. 7.33. Итого, вычисление одной пары значений $(d(n), s(n))$ при прямом ДВП требует 8 сложений и 6 умножений, или 14 арифметических операций. Такое же количество вычисли-



тельных операций необходимо для обработки каждой пары чисел $(d(n), s(n))$ при обратном ДВП. ►

Упражнение. Положим, что для ИХ фильтра анализа $\tilde{g}(n) = (-1)^{n+1} h(n+1)$. Покажите, что для вычисления каждой пары НЧ- и ВЧ-компонент $(d(n), s(n))$ при непосредственной реализации прямого ДВП по схеме рис. 7.31 с фильтрами из примера 7.8 необходимо выполнить 9 умножений и 14 сложений, т. е. 23 арифметические операции. *Указание:* учтите, что ИХ фильтры из примера 7.8 являются симметричными.

Приведенное упражнение и пример 7.15 показывают, что использование схемы лифтинга для вычисления биортогонального вейвлет-преобразования, используемого в стандарте JPEG 2000, позволяет сократить число арифметических операций примерно в $23/14 \approx 1,64$ раза. В общем случае с увеличением длины ИХ фильтров ДВП (с ростом числа ненулевых компонент в КИХ) выигрыш в сокращении количества арифметических операций от использования схемы лифтинга приближается к двукратному [60].

Итак, получаемая в результате лифтинга вычислительная схема реализации биортогональных ДВП полностью определяется разложением (7.70) полифазной матрицы фильтров синтеза, или связанной с ней факторизацией (7.71) полифазной матрицы фильтров анализа. Ортогональные ДВП являются частным случаем биортогональных преобразований с более жесткими требованиями к фильтрам.

Пример 7.16. Найти разложение полифазных матриц (7.70) и (7.71) для ортогональных фильтров Добеши-2 (см. пример 7.7).

◀ Найденные в примере 7.7 коэффициенты фильтра анализа

$$h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, \quad h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$$

определяют ПФ фильтра синтеза $H(z) = \sum_{k=0}^3 h_k z^{-k}$ и ПФ фильтра анализа $\tilde{H}(z) = \sum_{k=0}^3 h_k z^k = H(z^{-1})$. Находя разложение (7.67) по алгоритму Евклида, получим следующее факторизованное представление вида (7.70) для полифазной матрицы фильтров синтеза:

$$\mathbf{P}(z) = \begin{bmatrix} 1 & -\sqrt{3} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{\sqrt{3}}{4} + \frac{\sqrt{3}-2}{4}z^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}+1}{\sqrt{2}} & 0 \\ 0 & \frac{\sqrt{3}-1}{\sqrt{2}} \end{bmatrix}. \quad (7.72)$$

Этому представлению соответствует ВЧ фильтр синтеза с ПФ (убедитесь!) $G(z) = -h_3z^2 + h_2z - h_1 + h_0z^{-1}$, т.е. для записи $G(z) = \sum_k g_k z^{-k}$ имеем $g_k = (-1)^{1-k} h_{1-k}$.

Взяв за основу факторизацию (7.72), для построения вычислений по схеме лифтинга (рис. 7.33), получим в обозначениях (7.70) $m = 2$ и $S_1(z) = -\sqrt{3}$, $T_1(z) = \frac{\sqrt{3}}{4} + \frac{\sqrt{3}-2}{4}z^{-1}$, $S_2(z) = z$, $T_2(z) = 0$, $K = \frac{\sqrt{3}+1}{\sqrt{2}}$.

Так как вейвлеты Добеши являются ортогональными и соответствующая четверка фильтров анализа и синтеза образует КЗФ (это означает, что $\tilde{H}(z) = H(z^{-1})$ и $\tilde{G}(z) = G(z^{-1})$), то $\tilde{\mathbf{P}}(z) = \mathbf{P}(z^{-1})$, и на основании (7.72) можем записать:

$$\tilde{\mathbf{P}}(z) = \begin{bmatrix} 1 & -\sqrt{3} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{\sqrt{3}}{4} + \frac{\sqrt{3}-2}{4}z & 1 \end{bmatrix} \begin{bmatrix} 1 & z^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}+1}{\sqrt{2}} & 0 \\ 0 & \frac{\sqrt{3}-1}{\sqrt{2}} \end{bmatrix}. \quad (7.73)$$

Поскольку $\mathbf{P}(z) = (\tilde{\mathbf{P}}^T(z))^{-1}$ (см. (7.60)), то из (7.73) получаем также следующую факторизацию (убедитесь!):

$$\mathbf{P}(z) = \begin{bmatrix} 1 & 0 \\ \sqrt{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{\sqrt{3}}{4} - \frac{\sqrt{3}-2}{4}z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -z^{-1} & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}-1}{\sqrt{2}} & 0 \\ 0 & \frac{\sqrt{3}+1}{\sqrt{2}} \end{bmatrix},$$

которая в обозначениях (7.70) дает следующие параметры лифтинга для $m = 2$:

$$S_1(z) = 0, \quad T_1(z) = \sqrt{3}, \quad S_2(z) = -\frac{\sqrt{3}}{4} - \frac{\sqrt{3}-2}{4}z^{-1}, \quad T_2(z) = -z^{-1}, \quad K = \frac{\sqrt{3}-1}{\sqrt{2}}. \quad \blacktriangleright$$

Упражнение. Покажите, что обе полученные в примере 7.16 схемы вычисления ортогонального ДВП Добеши-2 имеют одну и ту же сложность в 9 арифметических операций на вычисление одной пары НЧ- и ВЧ-компонент $(d(n), s(n))$. Непосредственная



реализация данного ДВП по схеме рис. 7.31 сводится к 14 арифметическим операциям на пару $(d(n), s(n))$ (убедитесь!), т. е. сокращение объема вычислений от использования лифтинга в данном случае $14/9 \approx 1,56$ раза.

Упражнение. Рассмотрим H -фильтр ДВП Добеши-2, отличающийся от рассмотренного в примере 7.16 сдвигом ИХ на один отсчет и имеющий ПФ $H(z) = h_0 z + h_1 + h_2 z^{-1} + h_3 z^{-2} = z \sum_{k=0}^3 h_k z^{-k}$, коэффициенты которой (h_0, h_1, h_2, h_3) найдены в примере 7.7. Получите для $H(z)$ разложение (7.67) и соответствующее факторизованное представление (7.70). Сравните количество арифметических операций, необходимое для реализации ДВП по полученной лифтинг-схеме, с результатом из примера 7.16.

Помимо сокращения объема вычислений, которое достигается при использовании рассмотренного способа факторизации полифазных матриц, схема лифтинга может использоваться также и для построения новых ДВП. Взяв в качестве начальной полифазной матрицы единичную матрицу $P(z) = E$, далее можно добавлять новые сомножители в (7.70), выбирая по какому-либо принципу новые полиномы $S_k(z)$ и $T_k(z)$. Таким образом, лифтинг может применяться не только для уменьшения объема вычислений при реализации ДВП, но и для синтеза новых ДВП. Построенные таким образом биортогональные вейвлеты были названы *вейвлетами второго поколения* [61].

Заключение

Закономерно, что цифровая форма представления сигналов повсеместно заменяет аналоговое представление в самых различных системах обработки, передачи и хранения сигналов. Надежность и качество — вот основные преимущества цифровой формы представления. Современные технические возможности цифровой аппаратуры позволяют использовать эти преимущества все более полно. С основными подходами, применяемыми для эффективного представления цифровых сигналов, мы постарались познакомить читателя наряду с изложением общих теоретических основ ЦОС. Другой современный аспект ЦОС заключается в использовании вейвлет-преобразований, теории и практике применения которых полностью посвящена заключительная глава.

Не вызывает сомнения тот факт, что прогресс в области систем обработки сигналов во многом связан с дальнейшим внедрением алгоритмов и методов ЦОС. Теория ЦОС продолжает развиваться и усложняться, и, конечно, данное учебное пособие нужно рассматривать только как первый шаг к знакомству с ней. Для дальнейшего изучения теории ЦОС и ее приложений может помочь приводимый далее список литературы, которая была использована при написании данного пособия.

Литература

1. *Айфичер Э., Джервис Б.* Цифровая обработка сигналов: практический подход.— 2-е изд. — М.: Вильямс, 2008. — 992 с.
2. *Блейхут Р.* Быстрые алгоритмы цифровой обработки сигналов. — М.: Мир, 1989. — 448 с.
3. *Блаттер К.* Вейвлет-анализ. Основы теории — М.: Техносфера, 2004. — 280 с.
4. *Вариченко Л. В., Лабунец В. Г., Раков М. А.* Абстрактные алгебраические системы и цифровая обработка сигналов. — Киев: Наукова думка, 1986. — 248 с.
5. Методы сжатия данных / Ватолин Д., Ратушняк А., Смирнов М., Юкин В. — Устройство архиваторов, сжатие изображений и видео. — М.: ДИАЛОГ-МИФИ, 2002. — 384 с.
6. *Вентцель Е. С.* Теория вероятностей: Учебник для студ. вузов. — М.: КноРус, 2010. — 664 с.
7. *Волков Е. А.* Численные методы: Уч. пособие для вузов. — СПб.: Лань, 2008. — 256 с.
8. *Гадзиковский В. И.* Цифровая обработка сигналов. — М.: СОЛОН-ПРЕСС, 2013. — 766 с.
9. *Гольденберг Л. М., Матюшкин Б. Д., Поляк М. Н.* Цифровая обработка сигналов. — М.: Радио и связь, 1990. — 256 с.
10. *Гоноровский И. С.* Радиотехнические цепи и сигналы: Учебник для вузов. — М.: Дрофа, 2006. — 720 с.
11. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. — М.: Техносфера, 2012. — 1104 с.
12. *Дворкович В. П., Дворкович А. В.* Цифровые видеоинформационные системы (теория и практика). — М.: Техносфера, 2012. — 1008 с.
13. *Дворкович В. П., Дворкович А. В.* Оконные функции для гармонического анализа сигналов. — М.: Техносфера, 2016. — 208 с.
14. *Джиган В. И.* Адаптивная фильтрация сигналов: теория и алгоритмы. — М.: Техносфера, 2013. — 528 с.
15. *Дмитриев В. И.* Прикладная теория информации: Учебник для вузов. — М.: Высшая школа, 1989. — 320 с.
16. *Добеши И.* Десять лекций по вейвлетам. — Ижевск: Регулярная и хаотическая динамика, 2004. — 464 с.
17. *Ефимов А. В., Золотарев Ю. Г., Терпигорева В. М.* Математический анализ (специальные разделы). — М.: Высшая школа, 1980. Ч. 1 и 2. — 279 и 295 с.

18. *Иванов М. Т. и др.* Теоретические основы радиотехники: Учеб. пособие / М. Т. Иванов, А. Б. Сергиенко, В. Н. Ушаков; под ред. В. Н. Ушакова. — М.: Высш. шк., 2002. — 306 с.
19. *Каппелини В., Константинодис А., Эмилиани П.* Цифровые фильтры и их применение. — М.: Энергоатомиздат, 1983. — 360 с.
20. *Адаптивная фильтрация* / Под ред. К. Ф. Н. Коуэна и П. М. Гранта. — М.: Мир, 1988. — 392 с.
21. *Купер Дж., Макгиллем К.* Вероятностные методы анализа сигналов и систем. — М.: Мир, 1989. — 376 с.
22. *Курош А. Г.* Курс высшей алгебры. — СПб.: Лань, 2013. — 432 с.
23. *Лайонс Р.* Цифровая обработка сигналов. — М.: Бином-Пресс, 2011. — 656 с.
24. *Лесин В. В., Лисовец Ю. П.* Основы методов оптимизации: Учебное пособие. — 3-е изд., испр. — СПб.: Лань, 2011. — 352 с.
25. *Лэй Э.* Цифровая обработка сигналов для инженеров и технических специалистов: практическое руководство. — М.: ООО «Группа ИДТ», 2007. — 336 с.
26. *Малла С.* Вейвлеты в обработке сигналов. — М.: Мир, 2005. — 671 с.
27. *Марпл-мл. С.Л.* Цифровой спектральный анализ и его приложения. — М.: Мир, 1990. — 584 с.
28. *Нуссбаумер Г.* Быстрое преобразование Фурье и алгоритмы вычисления сверток. — М.: Радио и связь, 1985. — 248 с.
29. *Оппенгейм А., Шафер М.* Цифровая обработка сигналов. — М.: Техносфера, 2012. — 1048 с.
30. *Панин В. В.* Основы теории информации. — М.: БИНОМ. Лаборатория знаний, 2009. — 438 с.
31. *Петухов А. П.* Введение в теорию базисов всплесков. — СПб.: Изд-во СПбГТУ, 1999. — 132 с.
32. *Привалов И. И.* Введение в теорию функций комплексного переменного. — М.: Высшая школа, 1999. — 432 с.
33. *Птачек М.* Цифровое телевидение. Теория и техника. — М.: Радио и связь, 1990. — 528 с.
34. *Пугачев В. С.* Теория вероятностей и математическая статистика. — М.: ФИЗМАТЛИТ, 2002. — 496 с.
35. *Рабинер П., Гоулд Б.* Теория и применения цифровой обработки сигналов. — М.: Мир, 1978. — 848 с.
36. *Розанов Ю. А.* Случайные процессы. — М.: Наука, 1979. — 184 с.
37. *Ричардсон Я.* Видеокодирование. H.264 и MPEG-4 — стандарты нового поколения. — М.: Техносфера, 2005. — 368 с.

38. *Свешников А. Г., Тихонов А. Н.* Теория функций комплексной переменной. — М.: ФИЗМАТЛИТ, 2005. — 336 с.
39. *Сиберт У. М.* Цепи, сигналы, системы. — М.: Мир, 1988. Ч. 1 и 2. — 336 и 360 с.
40. *Смит С.* Цифровая обработка сигналов. Практическое руководство для инженеров и научных работников. — М.: Додэка-XXI, 2012. — 720 с.
41. *Смоленцев Н. К.* Основы теории вейвлетов. Вейвлеты в MATLAB. — М.: ДМК Пресс, 2014. — 628 с.
42. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера. — 2-е изд., испр. — М.: ФИЗМАТЛИТ, 2003. — 784 с.
43. Основы цифровой обработки сигналов: Курс лекций / Солонина А. И., Улахович Д. А., Арбузов С. М., Соловьева Е. Б. — СПб.: БХВ-Петербург, 2005. — 768 с.
44. *Сэлмон Д.* Сжатие данных, изображений и звука. — М.: Техносфера, 2004. — 432 с.
45. *Треногин В. А.* Функциональный анализ. — М.: ФИЗМАТЛИТ, 2007. — 488 с.
46. *Уидроу Б., Стирнз С.* Адаптивная обработка сигналов. — М.: Радио и связь, 1989. — 440 с.
47. *Улахович Д. А.* Основы теории линейных электрических цепей. — СПб.: БХВ-Петербург, 2012. — 816 с.
48. *Умняшкин С. В.* О модификации дискретного косинусного преобразования // Изв. Тульского гос. ун-та. Сер. Математика. Механика. Информатика. — Тула, 1998. — Т. 4. — Вып. 1. — С. 143–147.
49. *Умняшкин С. В.* Анализ эффективности применения ортогональных преобразований для кодирования дискретных сигналов с коррелированными отсчетами // Цифровая обработка сигналов. — № 4. — 2008. — С. 15–18.
50. *Умняшкин С. В.* Теоретические основы цифровой обработки и представления сигналов. — М.: Техносфера, 2012. — 368 с.
51. *Уэлстид С.* Фракталы и вейвлеты для сжатия изображений в действии. — М.: Триумф, 2003. — 320 с.
52. *Фихтенгольц Г. В.* Курс дифференциального и интегрального исчисления. — СПб.: Лань, 2009. — Т.3. — 656 с.
53. *Фрейзер М.* Введение в вейвлеты в свете линейной алгебры. — М.: БИНОМ. Лаборатория знаний, 2007. — 487 с.
54. *Харкевич А. А.* Спектры и анализ. — М.: Либроком, 2009. — 240 с.
55. *Чуи Ч.* Введение в вейвлеты. — М.: Мир, 2001. — 412 с.
56. *Шапиро Л., Стокман Дж.* Компьютерное зрение. — М.: БИНОМ, 2006. — 752 с.

57. Шовенгердт *P. А.* Дистанционное зондирование. Модели и методы обработки изображений. — М.: Техносфера, 2013. — 560 с.
58. Штарк *Г.-Г.* Применение вейвлетов для ЦОС. — М.: Техносфера, 2007. — 192 с.
59. *Elliott D. F., Rao K. R.* Fast transforms: algorithms, analyses, applications. — London: Academic Press inc., 1982. — 488 p.
60. *Daubechies I., Sweldens W.* Factoring wavelet transforms into lifting steps // J. Fourier Anal. Appl., 1998. — Vol. 4, No. 3. — P. 247–269.
61. *Sweldens W.* The Lifting Scheme: A Construction of Second Generation Wavelets // SIAM J. Math. Analysis, 1998. — Vol. 29, No. 2. — P. 511–546.
62. *Taubman D., Marcellin M.* JPEG 2000: Image Compression Fundamentals, Standards and Practice. — Springer, 2002. — 793 p.
63. The JPEG 2000 suite / P. Schelkens, A. Skodras, N. Ebrahimi (Eds). — John Wiley & Sons Ltd., 2009. — 544 p.



**КНИГИ ИЗДАТЕЛЬСТВА
«ТЕХНОСФЕРА» МОЖНО ПРИОБРЕСТИ:**

В магазинах:

г. Москва

Сеть магазинов «Новый книжный»
www.nk1.ru

Торговый дом «Библио-Глобус»,
ст. м. Лубянка, ул. Мясницкая, 6
тел. 8(495) 781-19-00, 624-46-80

«Московский дом книги»,
ст. м. Арбатская, ул. Новый Арбат, 8
тел. 8(495) 789-35-91

«Дом книги на Спартаковской»,
ст. м. Бауманская, ул. Спартаковская, 23
тел. 8(499) 400-41-06

«Молодая гвардия», ст.м. Полянка,
ул. Б. Полянка, 28
тел. 8(499) 238-50-01

«Дом технической книги»,
ст. м. Ленинский проспект,
Ленинский проспект, 40
тел. 8(499) 137-60-38

«Дом медицинской книги»,
Комсомольский проспект, 25
тел. 8(495) 789-35-91,
(495)789-31-14

МГУ, химический факультет (корп. 3)
киоск ПБОЮЛ Макарова О.В.
ГСП-1, Ленинские горы

Ближнее зарубежье:

г. Минск

ИП Юзвук Наталья Николаевна
тел. 375-17-294-54-65

г. Харьков

Гуманитарный центр
«Литера Нова»
тел. 057-731-41-69

В городах России:

Сеть магазинов «Новый книжный»
www.nk1.ru

г. Санкт-Петербург

«Санкт-Петербургский дом книги»,
(Дом Зингера) Невский пр.,28

Книготорговая сеть
«Академическая литература»
тел. (812) 329-10-29

г. Новосибирск

Книжный магазин «Консул»
ул. Разъездная, 16
тел. (383)217-45-40

г. Красноярск

Магазин АКАДЕМКНИГА
просп. Мира, 66
academbook@ksc.krasn.ru

Информация о новинках:
www.technosfera.ru

— наложенным платежом
(заказы принимаются
по e-mail, по почте)

— интернет-эквайринг

— по безналичному расчету
(заказы принимаются по e-mail,
по факсу с указанием полных
реквизитов юридического лица)

Как заказать наши книги?

По почте: 125319, г. Москва, а/я 91
По факсу: +7(495) 956-33-46
E-mail: knigi@technosfera.ru
sales@technosfera.ru