

Министерство образования и науки Российской Федерации
Ярославский государственный университет им. П. Г. Демидова

**Ю. А. БРЮХАНОВ, А. Л. ПРИОРОВ,
В. И. ДЖИГАН, В. В. ХРЯЩЕВ**

ОСНОВЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ

Учебное пособие

Рекомендовано

*Научно-методическим советом университета для студентов,
обучающихся по направлениям «Радиотехника»
и «Инфокоммуникационные технологии и системы связи»*

Ярославль
ЯрГУ
2013

УДК 621.37(075.8)
ББК 3811.3я73
О 75

*Рекомендовано
Редакционно-издательским советом университета
в качестве учебного издания. План 2013 года.*

Рецензенты:
доктор технических наук В. В. Витязев;
кафедра основ радиотехники
Московского энергетического института
(технического университета)

О 75 Основы цифровой обработки сигналов: учебное
пособие / Ю. А. Брюханов, А. Л. Приоров, В. И. Джиган,
В. В. Хрящев; Яросл. гос. ун-т им. П. Г. Демидова.
— Ярославль : ЯрГУ, 2013. — 344 с.
ISBN 978-5-8397-0933-1

В учебном пособии описаны современные методы цифровой обработки сигналов.

Предназначено для студентов, обучающихся по направлениям 210400.62 Радиотехника и 210700.62 Инфокоммуникационные технологии и системы связи (дисциплина «Основы цифровой обработки сигналов», цикл Б3), очной формы обучения.

Материал может быть использован также при подготовке студентами курсовых и выпускных квалификационных работ.

Ил. 134. Табл. 4. Библиогр.: 39 назв.

УДК 621.37(075.8)
ББК 3811.3я73

ISBN 978-5-8397-0933-1

© ЯрГУ, 2013

ОГЛАВЛЕНИЕ

ПРИНЯТЫЕ СОКРАЩЕНИЯ.....	9
ВВЕДЕНИЕ	11
1. МАТЕМАТИЧЕСКИЙ АППАРАТ	14
1.1. Последовательности.....	14
1.2. Дискретный ряд Фурье	16
1.3. Z-преобразование	16
1.3.1. Последовательности конечной длины	17
1.3.2. Физически реализуемые последовательности	17
1.3.3. Нереализуемые последовательности	17
1.3.4. Z-преобразование некоторых последовательностей.....	18
1.4. Соотношение между z-преобразованием и Фурье- преобразованием последовательности.....	19
1.5. Обратное z-преобразование	21
1.6. Свойства z-преобразования	23
1.6.1. Линейность.....	23
1.6.2. Задержка	23
1.6.3. Свертка последовательностей.....	23
1.6.4. Перемножение последовательностей.....	24
1.6.5. Одностороннее z-преобразование. Задержка физически реализуемых последовательностей	25
1.7. Дискретное преобразование Фурье.....	26
1.7.1. Связь между дискретным преобразованием Фурье и z-преобразованием.....	28
1.7.2. Дискретное преобразование Фурье последовательностей конечной длины	32
1.7.3. Эффекты наложения	33
1.8. Свойства дискретного преобразования Фурье	35
1.8.1. Линейность.....	36
1.8.2. Сдвиг.....	36
1.8.3. Свойства симметрии	37
1.9. Свертка периодических последовательностей	39
1.10. Линейная свертка конечных последовательностей.....	41
1.11. Разностные уравнения	45
1.12. Решение разностных уравнений с применением одностороннего z-преобразования.....	47

1.13. Контрольные вопросы	49
1.14. Литература	50
2. ЦИФРОВЫЕ СИГНАЛЫ И ИХ СПЕКТРЫ.....	51
2.1. Взаимная связь непрерывного сигнала и полученного из него дискретизированного сигнала	51
2.2. Спектр дискретизированного сигнала	54
2.3. Соотношения между спектрами непрерывных и дискретных сигналов	57
2.4. Восстановление непрерывного сигнала из дискретного	60
2.5. Типовые периодические сигналы.....	63
2.5.1. Последовательность показательных импульсов.....	64
2.5.2. Косинусоидальное колебание	64
2.5.3. Последовательность прямоугольных импульсов	66
2.5.4. Последовательность прямоугольных радиоимпульсов.....	67
2.6. Типовые непериодические сигналы.....	69
2.6.1. Показательный импульс	69
2.6.2. Прямоугольный импульс.....	70
2.6.3. Прямоугольный радиоимпульс.....	71
2.6.4. Цифровой единичный импульс	73
2.6.5. Единичный скачок.....	74
2.7. Контрольные вопросы	75
2.8. Литература	77
3. ЦИФРОВЫЕ ЦЕПИ	78
3.1. Цифровые линейные цепи с постоянными параметрами	78
3.2. Физическая реализуемость. Устойчивость.....	82
3.3. Частотная характеристика. Геометрическая интерпретация	84
3.4. Способы построения цифровых цепей	87
3.5. Структурные схемы цепей без полюсов	96
3.6. Нерекурсивная цепь второго порядка.....	98
3.6.1. Частотные свойства.....	98
3.6.2. Временные характеристики	109
3.7. Рекурсивная цепь второго порядка	113
3.7.1. Частотные свойства.....	114
3.7.2. Временные характеристики	123
3.8. Контрольные вопросы	130

3.9. Литература	131
4. ЭФФЕКТЫ КВАНТОВАНИЯ В ЦИФРОВЫХ СИГНАЛАХ.....	133
4.1. Представление чисел	133
4.2. Усечение и округление	138
4.3. Квантование сигнала.....	141
4.4. Аналогово-цифровое преобразование	145
4.5. Цифроаналоговое преобразование сигналов	146
4.6. Контрольные вопросы	150
4.7. Литература	150
5. ЭФФЕКТЫ КВАНТОВАНИЯ В ЦИФРОВЫХ ЦЕПЯХ	152
5.1. Реакция идеальной цифровой цепи на шум квантования.....	152
5.2. Квантование арифметических операций в БИХ-цепях.....	155
5.3. Ограничения, вызванные переполнением, и отношение сигнал/шум	159
5.4. Квантование арифметических операций в КИХ-цепях	161
5.5. Предельные циклы и пульсации, вызванные переполнением.....	162
5.6. Квантование коэффициентов.....	165
5.7. Техническая реализация цифровых цепей	167
5.8. Контрольные вопросы	171
5.9. Литература	173
6. БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ.....	174
6.1. Алгоритмы БПФ по основанию 2.....	174
6.2. БПФ с прореживанием по времени	175
6.3. БПФ с прореживанием по частоте	181
6.4. Контрольные вопросы	184
6.5. Литература	185
7. СИНТЕЗ ЦИФРОВЫХ КИХ-ФИЛЬТРОВ МЕТОДОМ ОКОН.....	186
7.1. Основные определения	186
7.2. Синтез цифровых фильтров	187
7.2.1. Задание требований к цифровым фильтрам.....	188
7.2.2. Типы частотно-избирательных фильтров и задание требований к ним	189

7.3. Синтез КИХ-фильтров методом окон	192
7.3.1. Общая характеристика задачи синтеза КИХ-фильтров методом окон	192
7.3.2. Явление Гиббса	195
7.3.3. Окна и их основные параметры	198
7.3.4. Методика синтеза КИХ-фильтров на основе оконных функций	203
7.3.5. Особенности использования метода окон	204
7.4. Контрольные вопросы	205
7.5. Литература	206
8. СИНТЕЗ ЦИФРОВЫХ БИХ-ФИЛЬТРОВ МЕТОДОМ БИЛИНЕЙНОГО Z-ПРЕОБРАЗОВАНИЯ	207
8.1. Основные определения	207
8.2. Аналоговые фильтры-прототипы нижних частот	208
8.2.1. Фильтры Баттерворта	210
8.2.2. Фильтры Чебышёва	214
8.2.3. Эллиптические фильтры (фильтры Золотарева-Кауэра)	219
8.3. Билинейное z -преобразование	221
8.3.1. Определение билинейного z -преобразования	221
8.3.2. Свойства билинейного z -преобразования	222
8.3.3. Частотные преобразования	228
8.3.4. Методика синтеза цифровых фильтров на основе билинейного z -преобразования	231
8.3.5. Особенности использования метода билинейного z -преобразования	232
8.4. Контрольные вопросы	233
8.5. Литература	234
9. СИНТЕЗ ОПТИМАЛЬНЫХ ПО ЧЕБЫШЁВУ ЦИФРОВЫХ КИХ-ФИЛЬТРОВ	235
9.1. Основные определения	235
9.2. КИХ-фильтры с линейной ФЧХ	236
9.2.1. Фильтры типа I – симметричная импульсная характеристика, N – четное	236
9.2.2. Фильтры типа II – симметричная импульсная характеристика, N – нечетное	237

9.2.3. Фильтры типа III – антисимметричная импульсная характеристика, N – четное.....	238
9.2.4. Фильтры типа IV – антисимметричная импульсная характеристика, N – нечетное.....	239
9.3. Теорема чередования	240
9.4. Алгоритм замены Ремеза.....	245
9.5. Контрольные вопросы	248
9.6. Литература	250
10. МНОГОСКОРОСТНАЯ ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ	251
10.1. Основные операторы изменения частоты дискретизации.....	251
10.1.1. Временные соотношения.....	251
10.1.2. Соотношения в частотной области	254
10.1.3. Многоскоростные тождества	259
10.1.5. Основные структуры.....	261
10.1.6. Связь входного и выходного сигналов	263
10.1.7. Требования к цифровым фильтрам	264
10.1.8. Банки цифровых фильтров.....	266
10.2. Контрольные вопросы	268
10.3. Литература	269
11. ОБРАБОТКА СИГНАЛОВ НА ОСНОВЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ	270
11.1. Основные определения	270
11.2. Введение в теорию вейвлет-преобразования	271
11.2.1. Непрерывное вейвлет-преобразование (НВП).....	272
11.2.2. Признаки вейвлета	274
11.2.3. Примеры вейвлетообразующих функций (материнских вейвлетов).....	275
11.2.4. Свойства вейвлет-анализа	278
11.2.5. Дискретное вейвлет-преобразование	279
11.2.6. Быстрое вейвлет-преобразование.....	280
11.2.7. Частотный подход к вейвлет-преобразованию.....	283
11.3. Удаление шумов из сигнала с использованием вейвлет-преобразования	285
11.4. Контрольные вопросы	288
11.5. Литература	289

12. ОСНОВЫ АДАПТИВНОЙ ОБРАБОТКИ СИГНАЛОВ	291
12.1. Вводные замечания	291
12.2. Требования к адаптивным фильтрам	292
12.3. Критерии функционирования адаптивных фильтров	295
12.4. Приложения адаптивных фильтров	296
12.5. Корреляционная матрица	301
12.6. Собственные числа и собственные векторы корреляционной матрицы.....	303
12.7. Винеровская фильтрация.....	307
12.8. Поверхность среднеквадратической ошибки.....	310
12.9. Алгоритм Ньютона.....	313
12.10. Постоянные времени алгоритма Ньютона	316
12.11. Алгоритм наискорейшего спуска	317
12.12. Постоянные времени алгоритма наискорейшего спуска.....	319
12.13. LMS-алгоритм.....	322
12.14. Переходные процессы в LMS-алгоритме	324
12.15. Качество адаптивной фильтрации с помощью LMS-алгоритма.....	326
12.16. NLMS-алгоритм.....	327
12.17. Задача наименьших квадратов и ее решение	329
12.18. Рекурсивная задача наименьших квадратов	331
12.19. Решение рекурсивной задачи наименьших квадратов	332
12.20. Качество адаптивной фильтрации с помощью RLS-алгоритма.....	334
12.21. Заключительные замечания по адаптивной обработке сигналов	335
12.22. Контрольные вопросы	337
12.23. Литература	341
Заключение.....	343

ПРИНЯТЫЕ СОКРАЩЕНИЯ

- АА – адаптивный алгоритм
ААР – адаптивная антенная решетка
АКФ – автокорреляционная функция
АФ – адаптивный фильтр
АЧХ – амплитудно-частотная характеристика
АЦП – аналогово-цифровой преобразователь
БВП – быстрое вейвлет-преобразование
БГШ – белый гауссовский шум
БИХ – бесконечная импульсная характеристика
БПФ – быстрое преобразование Фурье
ВК – весовые коэффициенты
ДВП – дискретное вейвлет-преобразование
ДН – диаграмма направленности
ДПФ – дискретное преобразование Фурье
ИХ – импульсная характеристика
КИХ – конечная импульсная характеристика
КМА – кратномасштабный анализ
ЛПП – линейная (цепь) с постоянными параметрами
ЛИС – линейная инвариантная к сдвигу (цепь)
ЛС – линейная стационарная (цепь)
МСИ – межсимвольная интерференция
НВП – непрерывное вейвлет-преобразование
НПФ – непрерывное преобразование Фурье
ОДПФ – обратное дискретное преобразование Фурье
ОПФ – оконное преобразование Фурье
ПЗ – полоса задерживания фильтра
ПЗУ – постоянное запоминающее устройство
ПП – полоса пропускания фильтра

ПФ – полосовой фильтр
ПХ – переходная характеристика
РФ – режекторный фильтр
СКО – среднеквадратическое отклонение
ФВЧ – фильтр верхних частот
ФНЧ – фильтр нижних частот
ФЧХ – фазочастотная характеристика
ЧХ – частотная характеристика
ЦАП – цифроаналоговый преобразователь
ЦОС – цифровая обработка сигналов
ЦФ – цифровой фильтр

LS (Least Squares) – наименьшие квадраты
LMS (Least Mean Square) – алгоритм по критерию наименьшего среднего квадрата
MSE-функция (Means Square Error) – функция среднеквадратической ошибки
NLMS (Normalized LMS) – нормализованный LMS-алгоритм
QR RLS-алгоритм – RLS-алгоритм на основе QR-разложения матрицы
RG – регистр
RLS (Recursive Least Squares) – рекурсивный алгоритм по критерию наименьших квадратов
WLS (Weighted Least Squares) – взвешенные, наименьшие квадраты

ВВЕДЕНИЕ

*Дай, господи, желания и воли,
чтобы науку трудную постичь*

Состояние и перспективы развития информационных технологий в начале XXI века характеризуются широким практическим использованием техники цифровой обработки сигналов (ЦОС) – одной из самых динамичных и быстро развивающихся технологий в мире инфокоммуникаций и радиотехники. Методы и техника ЦОС вызывают повышенный интерес ученых и специалистов, работающих в различных областях, таких как связь и системы управления, радиотехника и электроника, акустика и сейсмология, радиовещание и телевидение, измерительная техника и приборостроение. Ежегодно в мире проводятся десятки международных научно-технических конференций и семинаров, посвященных решению актуальных проблем ЦОС. Издаются многочисленные книги и журналы, освещающие последние достижения в области теории и практики обработки цифровых сигналов. Постоянно растет число предприятий, организаций и научных центров, использующих в своих разработках методы и технологии ЦОС.

Вся история развития теории и техники ЦОС напрямую связана с достижениями в области дискретной схмотехники и компьютерных технологий. Дискретизация непрерывной информации во времени и квантование по уровню являются основой ее эффективного кодирования, преобразования, передачи и архивации. С появлением в 1940-х годах первых ЭВМ стало реальным возникновение нового фундаментального научного направления вычислительной математики, одним из разделов которого можно считать машинные алгоритмы цифровой обработки данных. Однако ограниченные вычислительные ресурсы используемых в те годы ЭВМ не позволяли проводить обработку данных в реальном времени.

Положение стало радикально меняться с появлением в 1960-х годах класса малых ЭВМ, ориентированных преимущественно на решение задач управления и обработки

данных в реальном времени. С этого момента формируется круг проблем и задач теории ЦОС как самостоятельного научного направления, которому в дальнейшем предстоит свой многоэтапный путь становления и развития.

Первый крупный вклад в теорию ЦОС внесли американские ученые Б. Голд и Ч. Рэйдер. Их монография стала первой настольной книгой специалистов по ЦОС. Фундаментальной работой, подводящей итоги начального становления теории ЦОС как нового научного направления, является книга Л. Рабинера, Б. Гоулда. Практически одновременно появляется одно из первых учебных пособий по ЦОС авторов А. Оппенгейма и Р. Шафера. Заметный вклад в развитие теории ЦОС внесли также и отечественные ученые.

Дисциплина «Основы цифровой обработки сигналов» является базовой для изучения связистами и радиотехниками. Здесь рассматриваются спектральная теория периодических и непериодических сигналов, линейная теория цепей дискретного времени, способы построения цифровых цепей, частотные и временные характеристики базовых звеньев первого и второго порядков, изучаются эффекты квантования и переполнения в цифровых сигналах и цепях. Большую часть этого материала можно найти в известных книгах Э. Айфичера, Б. Джервиса, С. Митры и М. Хайеса.

Пособие состоит из двенадцати разделов. Оно обобщает многолетний опыт авторов и их коллег по преподаванию данного материала студентам радиинженерных специальностей и направлений физического факультета Ярославского государственного университета им. П. Г. Демидова. Материал может быть использован при решении задач курсового и дипломного проектирования, а также для выполнения студенческих научных работ.

Содержание учебного пособия распределено между соавторами следующим образом: введение написано А. Л. Приоровым, разделы 1–5 – Ю. А. Брюхановым, разделы 6–11 – А. Л. Приоровым и В. В. Хрящевым при активном участии В. В. Волохова, раздел 12 – В. И. Джиганом.

Выражаем глубокую благодарность члену-корреспонденту РАН Ю. Б. Зубареву, профессорам В. П. Дворковичу, А. В. Дворковичу, В. В. Витязеву, А. А. Петровскому, Б. А. Алпатотову, Ю. С. Бехтину, В. Г. Карташеву, В. Г. Евстигнееву, М. А. Щербакову, А. К. Бернюкову, кандидату технических наук А. М. Лаврентьеву, которые оказали значительное влияние на формирование взглядов авторов в данном научном направлении. Эти взгляды формировались также в совместной работе с нашими коллегами с кафедры динамики электронных систем ЯргУ, среди которых особенно хочется отметить А. Н. Тараканова, И. В. Апалькова, И. С. Мочалова, М. А. Дубова, А. И. Топникова, С. А. Новоселова, Е. В. Сергеева. Всем им авторы очень признательны и за совместную исследовательскую работу, и за помощь в обеспечении учебного процесса.

Отдельные слова благодарности профессору М. К. Чобану и старшему преподавателю, кандидату технических наук В. А. Волохову за предоставленные материалы, а также ведущему электронику Ю. А. Лукашевичу и студентам кафедры ДЭС А. Носкову, В. Антипову и К. Туманову за активное участие в создании оригинал-макета пособия.

1. МАТЕМАТИЧЕСКИЙ АППАРАТ

Теория цифровых сигналов и линейных систем связана с описанием и обработкой временных последовательностей. В настоящем разделе и большей части книги будем считать, что квантование элементов последовательности по уровню отсутствует. Это предположение о бесконечно малом шаге квантования, относящееся как к отсчетам сигнала, так и к коэффициентам линейных цепей (систем), будет использовано при изучении общей теории дискретных (по времени, но не по уровню) сигналов и цепей. После этого будут рассмотрены различные эффекты, возникающие в цифровых сигналах и цепях с определенной точностью квантования по уровню из-за конечной длины машинного слова.

1.1. Последовательности

Цифровые последовательности определяются лишь для дискретных значений независимой переменной (времени). Обычно время квантуется равномерно, т. е. $t = kT$, где T – интервал между отсчетами. Математически цифровые сигналы представляются в виде непрерывной последовательности чисел. Для описания последовательностей может быть использовано одно из следующих обозначений:

$$\{x(n)\}, N_1 \leq n \leq N_2, \quad (1.1a)$$

$$\{x(nT)\}, N_1 \leq n \leq N_2, \quad (1.1б)$$

$$x(n), N_1 \leq n \leq N_2, \quad (1.1в)$$

$$x(nT), N_1 \leq n \leq N_2. \quad (1.1г)$$

Обозначения (1.1a) и (1.1в) могут применяться и при неравномерном расположении отсчетов, тогда как (1.1б) и (1.1г) явно предполагают их равномерное размещение. Если отсчеты

расположены равномерно, то такая последовательность называется решетчатой функцией.

Последовательность может быть получена несколькими способами. Проще всего взять набор чисел и расположить их в виде последовательности. Например, числа $0, 1, 2, \dots, (N-1)$ образуют «пилообразную» последовательность $x(n)=n, 0 \leq n \leq N-1$. Другой способ состоит в использовании некоторого рекуррентного соотношения. Например, $x(n)=x(n-1)/2$ с начальным условием $x(0)=1$ дает последовательность

$$x(n)=(1/2)^n, 0 \leq n < \infty.$$

Еще один способ: взять равноотстоящие отсчеты непрерывного колебания и из их величин образовать последовательность, т. е. положить $x(nT)=x(t)/_{t=nT}, -\infty \leq n \leq \infty$, где T – интервал дискретизации. Физически для получения $x(n)$ этим способом используются дискретизаторы, а для получения цифровых последовательностей – аналогово-цифровые преобразователи (АЦП).

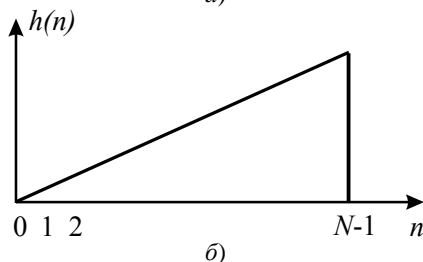
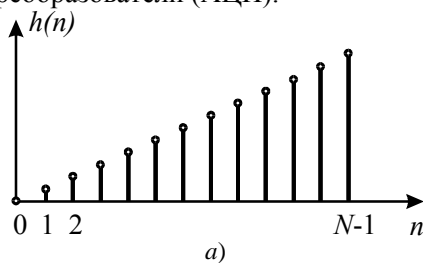


Рис. 1.1. Примеры изображения последовательностей:

а) в виде отрезков соответствующей длины; б) – в виде огибающей

Первые два способа получения последовательностей не связаны со временем, тогда как третий существенно от него зависит. Отсюда видно, что для описания последовательностей пригодны в том или ином смысле все обозначения (1.1).

Часто полезным и информативным является графическое изображение последовательностей. Для получения графического изображения (рис.1.1) будем использовать два способа.

Покажем это на примере изображения последовательности $x(n)=n$, $0 \leq n \leq N-1$. При использовании первого способа n_0 -й элемент последовательности изображается отрезком соответствующей длины, проведенным от оси абсцисс из точки $n=n_0$ (рис.1.1а). В некоторых случаях нет смысла изображать каждую выборку, достаточно провести только огибающую последовательности (рис.1.1б).

1.2. Дискретный ряд Фурье

Если комплексный спектр произвольной последовательности $x(n)$ может быть выражен функцией

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}, \quad (1.2)$$

то существует и обратное преобразование, позволяющее выразить $x(n)$ через $X(e^{j\omega})$,

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega. \quad (1.3)$$

Здесь и ниже, за исключением особо оговоренных случаев, $\omega = \omega T$.

1.3. Z-преобразование

Одним из наиболее полезных методов представления последовательностей и работы с ними является z -преобразование. Для последовательности $x(n)$, заданной при всех n , оно определяется следующим образом:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n}, \quad (1.4)$$

где z – комплексная переменная. Ясно, что комплексная функция (1.4) определена лишь для тех значений z , при которых степенной ряд сходится.

1.3.1. Последовательности конечной длины

Если $x(n)$ отлична от нуля только в интервале $N_1 \leq n \leq N_2$ ($N_1 < N_2$), где N_1 и N_2 конечны, то $X(z)$ сходится в z -плоскости везде, за исключением, может быть, точек $z=0$ или $z=\infty$.

На последовательностях конечной длины основан важный класс цифровых цепей. Типичная последовательность $\{x(n)\}$ конечной длины имеет следующий вид (рис. 1.2):

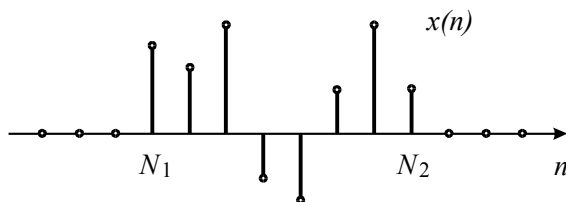


Рис. 1.2. Типичная последовательность конечной длины

Последовательности бесконечной длины составляют основу другого большого класса цифровых цепей.

1.3.2. Физически реализуемые последовательности

Физически реализуемые последовательности – если $x(n)$ отличается от нуля только при $0 \leq n < \infty$. При этом ряд $X(z)$ сходится везде вне круга радиуса R_1 . Величина R_1 зависит от положения полюсов функции $X(z)$. Ниже будет показано, что при $R_1 < 1$ соответствующая система является устойчивой. Физически реализуемые последовательности весьма важны, так как на их основе строится большинство реальных систем.

1.3.3. Нереализуемые последовательности

Физически нереализуемые последовательности – если $x(n)$ имеет ненулевые значения в области $-\infty < n < N_1 \leq 0$. При этом ряд $X(z)$ сходится во всех точках, лежащих в круге радиуса R_1 , причем R_1 определяется положением полюсов $X(z)$. В практических задачах нереализуемые последовательности обычно

не встречаются, но при рассмотрении некоторых теоретических вопросов они могут представлять интерес.

1.3.4. Z-преобразование некоторых последовательностей

а. Цифровой единичный импульс

$$x(n) = \delta(n) = 1, n=0; x(n) = 0, n \neq 0.$$

На основании (1.4) имеем $X(z) = 1$. Эта функция сходится на всей z -плоскости, так как $x(n)$ является последовательностью конечной длины.

б. Единичный скачок

$$x(n) = 1(n) = 1, n \geq 0; x(n) = 0, n < 0.$$

Поскольку $x(n) = 0$ везде, кроме $n \geq 0$, где $x(n) = 1$, то

$$X(z) = \sum_{n=-\infty}^{\infty} z^{-n} = \frac{1}{1 - z^{-1}},$$

причем $X(z)$ сходится при $|z| > 1$, так как имеет единственный полюс $z = 1$.

в. Комплексная экспонента

$$x(n) = 0, n < 0; x(n) = e^{j\omega n}, n \geq 0.$$

Вычисляя z -преобразование, получим

$$X(z) = \sum_{n=-\infty}^{\infty} e^{j\omega n} z^{-n} = \sum_{n=-\infty}^{\infty} (e^{j\omega} z^{-1})^n = \frac{1}{1 - z^{-1} e^{j\omega}},$$

причем $X(z)$ сходится при $|z| > 1$, так как единственным полюсом $X(z)$ является $z = e^{j\omega}$.

г. Простая показательная последовательность

$$x(n) = 0, n < 0; x(n) = a^n, n \geq 0.$$

Подставив $x(n)$ в (1.4), получим

$$X(z) = \sum_{n=-\infty}^{\infty} a^n z^{-n} = \sum_{n=-\infty}^{\infty} (az^{-1})^n = \frac{1}{1 - az^{-1}}.$$

Здесь $X(z)$ сходится при $|z| > a$, так как имеет только один полюс $z = a$.

1.4. Соотношение между z -преобразованием и Фурье-преобразованием последовательности

Z -преобразование последовательности можно рассматривать как способ ее однозначного представления в комплексной z -плоскости. Из определения (1.4) видно, что z -преобразование, вычисленное на единичной окружности, т. е. при $z=e^{j\omega}$, дает

$$X(z)|_{z=e^{j\omega}} = X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n},$$

что совпадает с преобразованием Фурье исходной последовательности (см. формулу (1.2)). Поэтому единичная окружность в z -плоскости играет существенную роль.

Обычным способом графического изображения информации, содержащейся в z -преобразовании, является задание особых точек – полюсов и нулей функции $X(z)$. Так, например, z -преобразование простой показательной последовательности может быть представлено следующим образом (рис. 1.3а).

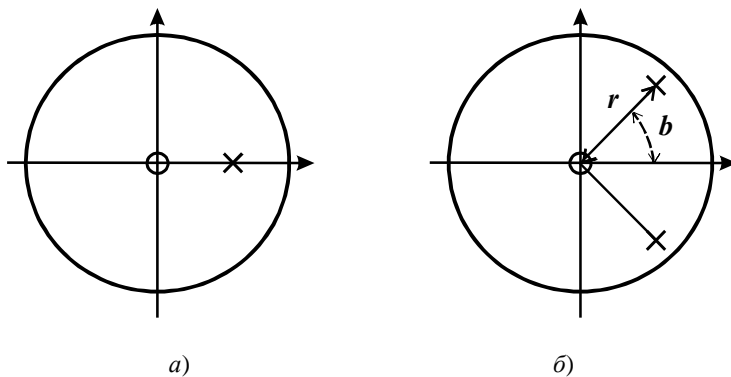


Рис. 1.3. Изображения z -преобразования:
а) простой показательной последовательности;
б) последовательности с парой комплексно сопряженных полюсов

Здесь и ниже крестиком изображен полюс, а кружком - нуль функции $X(z)$. С помощью такого изображения расположения нулей и полюсов, а также при использовании дополнительного предположения о физической реализуемости системы можно однозначно (с точностью до постоянного множителя) восстановить z -преобразование.

Пример. Найдём z -преобразование последовательности

$$x(n) = \begin{cases} \frac{r^n \sin[(n+1)b]}{\sin b}, & n \geq 0, \\ 0 & , n < 0. \end{cases}$$

Решение. Используя определение z -преобразования, получим

$$\begin{aligned} X(z) &= \sum_{n=0}^{\infty} r^n \left\{ \frac{\sin[(n+1)b]}{\sin b} \right\} z^{-n} = \sum_{n=0}^{\infty} \frac{r^n z^{-n}}{\sin b} \left\{ \frac{e^{j(n+1)b} - e^{-j(n+1)b}}{2j} \right\} = \\ &= \sum_{n=0}^{\infty} (rz^{-1}e^{jb})^n \frac{e^{jb}}{2j \sin b} - \sum_{n=0}^{\infty} (rz^{-1}e^{-jb})^n \frac{e^{-jb}}{2j \sin b} = \\ &= \frac{1}{2j \sin b} \left(\frac{e^{jb}}{1 - rz^{-1}e^{jb}} - \frac{e^{-jb}}{1 - rz^{-1}e^{-jb}} \right), \\ X(z) &= \frac{1}{1 - 2r(\cos b)z^{-1} + r^2z^{-2}}. \end{aligned}$$

Здесь $X(z)$ сходится при $|z| > r$. Расположение нулей и полюсов такой функции в z -плоскости показано на рис. 1.3б. Она имеет пару комплексно сопряженных полюсов в точках $z = re^{\pm jb}$ и двойной нуль при $z=0$.

Как уже упоминалось, зная расположение нулей и полюсов функции $X(z)$, ее можно восстановить. Так, если известно, что функция $X(z)$ имеет N полюсов в точках $z=p_1, p_2, \dots, p_N$

и M нулей в точках $z=z_1, z_2, \dots, z_M$, то она может быть выписана в виде отношения произведений

$$X(z) = A \frac{\prod_{i=1}^M (1 - z_i z^{-1})}{\prod_{i=1}^N (1 - p_i z^{-1})}, \quad (1.5)$$

где A – произвольная постоянная. Перемножив сомножители, получим, что наиболее общей формой $X(z)$ является дробно-рациональная функция от z^{-1} , т. е.

$$X(z) = \frac{\sum_{i=0}^M a_i z^{-i}}{1 + \sum_{i=1}^N b_i z^{-i}}. \quad (1.6)$$

Полученное выражение часто используется при синтезе цепей.

1.5. Обратное z -преобразование

Важно уметь перейти не только от последовательности к ее z -преобразованию, но и обратно от z -преобразования к последовательности. Способ обратного перехода называется обратным z -преобразованием и формально определяется соотношением

$$x(n) = \frac{1}{2\pi j} \oint_{C_1} X(z) z^{n-1} dz. \quad (1.7)$$

В правой части этого равенства стоит контурный интеграл в z -плоскости по любому замкнутому контуру в области сходимости, охватывающему начало координат. Например, контуром интегрирования может быть окружность радиуса $C_1 > R_1$, где R_1 – радиус сходимости z -преобразования (т. е. предполагается, что последовательность $x(n)$ физически реализуема).

Обратное z -преобразование можно найти несколькими способами:

- 1) прямым вычислением интеграла (1.7) с использованием теоремы о вычетах;
- 2) разложением $X(z)$ на простые дроби;
- 3) обычным делением числителя $X(z)$ на его знаменатель;
- 4) разложением в степенной ряд.

Первый способ основан на известной теореме из теории функций комплексного переменного, утверждающей, что контурный интеграл (1.7) может быть вычислен непосредственно через вычеты

$$x(n) = \sum_{p_i} \operatorname{res} X(z) z^{n-1} \Big|_{p_i \text{ внутри } C_1} \quad (1.8)$$

Рассмотрим пример $X(z) = 1/(1-az^{-1})$. Из равенства (1.8) при $n \geq 0$ получим

$$x(n) = \operatorname{res} \frac{z^{n-1}}{1-az^{-1}} \Big|_{p=a}$$

т. е. $x(n) = a^n$, $n \geq 0$. При $n < 0$ кратный полюс z -преобразования находится в точке $z=0$. Прямое вычисление вычета в полюсе $z=0$ дает $x(n) = 0$ при $n < 0$.

При использовании второго способа z -преобразование записывают в виде дроби (1.5) и представляют суммой

$$X(z) = \sum_{i=1}^N \frac{\alpha_i}{1-p_i z^{-1}}.$$

С учетом того, что каждое слагаемое $\alpha_i/(1-p_i z^{-1})$ имеет обратное z -преобразование вида $\alpha_i p_i^n$, получим

$$x(n) = \begin{cases} \sum_{i=1}^N \alpha_i p_i^n, & n \geq 0, \\ 0, & n < 0. \end{cases}$$

Способы 3 и 4 рассматривать не будем. С ними можно познакомиться в пособиях по z -преобразованию.

1.6. Свойства z -преобразования

Z -преобразование весьма полезно при исследовании цифровых сигналов и линейных цепей. Чтобы полностью использовать его возможности, необходимо знать основные свойства z -преобразования, связанные с линейностью, задержкой последовательностей, сверткой, перемножением последовательностей, задержкой физически реализуемых последовательностей.

1.6.1. Линейность

Z -преобразование линейно. Это означает, что если $X_1(z)$ и $X_2(z)$ являются z -преобразованиями последовательностей $x_1(n)$ и $x_2(n)$, то при любых действительных a и b z -преобразование последовательности $ax_1(n) + bx_2(n)$ равно $aX_1(z) + bX_2(z)$.

1.6.2. Задержка

Если последовательность $x(n)$ имеет z -преобразование $X(z)$, то z -преобразование последовательности $x(n-n_0)$ при любых n_0 равно $z^{-n_0} X(z)$. Это свойство z -преобразования особенно полезно при переходе от представления линейной системы разностным уравнением к представлению ее z -преобразованием и наоборот.

1.6.3. Свертка последовательностей

Если $x(n)$ и $y(n)$ связаны зависимостью

$$y(n) = \sum_{m=-\infty}^{\infty} x(m)h(n-m) = \sum_{m=-\infty}^{\infty} h(m)x(n-m),$$

то

$$Y(z) = X(z)H(z), \quad (1.9)$$

где $X(z)$, $H(z)$ и $Y(z)$ являются соответственно z -преобразованиями последовательностей $x(n)$, $h(n)$ и $y(n)$. Таким образом, операция свертки последовательностей приводит к перемножению их z -преобразований.

Нетрудно заметить, что $H(z)$ можно выразить из соотношения (1.9) в виде

$$H(z) = Y(z)/X(z). \quad (1.10)$$

Равенство (1.9) важно как практическое средство, позволяющее без вычисления свертки найти выходную последовательность системы по ее импульсной характеристике (ИХ) и входной последовательности. Рассчитывая отклик $y(n)$ путем перемножения двух z -преобразований и вычисления обратного z -преобразования, часто удается свести сложную задачу к более простой.

1.6.4. Перемножение последовательностей

Если последовательности $x_1(n)$ и $x_2(n)$ имеют z -преобразования $X_1(z)$ и $X_2(z)$, то последовательность $x_3(n) = x_1(n)x_2(n)$ имеет z -преобразование:

$$X_3(z) = \frac{1}{2\pi j} \oint_C X_1(v)X_2(z/v)v^{-1} dv. \quad (1.11)$$

В область сходимости $X_3(z)$ входят все z , для которых справедливо следующее условие: если некоторая точка v принадлежит области сходимости $X_1(z)$, то z/v принадлежит области сходимости $X_2(z)$. В формуле (1.11) контур интегрирования является замкнутой кривой, лежащей внутри пересечения областей сходимости функций $X_1(v)$ и $X_2(z/v)$.

Соотношение (1.11) называют теоремой о комплексной свертке, так как оно представляет z -преобразование произведения $[x_1(n) \times x_2(n)]$ в виде комплексной свертки z -преобразований соответствующих последовательностей. Воспользовавшись подстановками $z = e^{j\omega}$ и $v = e^{j\theta}$, выразим преобразование Фурье от произведения последовательностей через преобразование Фурье от каждой из них. Оно имеет вид

$$X_3(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_1(e^{j\theta}) X_2(e^{j(\omega-\theta)}) d\theta$$

и является широко известной сверткой двух преобразований Фурье. Это соотношение требуется при проектировании цифровых фильтров методом весовых функций.

1.6.5. Одностороннее z -преобразование. Задержка физически реализуемых последовательностей

При решении большинства практических задач обычно имеют дело с физически реализуемыми последовательностями, поэтому полезно ввести «одностороннее» z -преобразование, определяемое как

$$X(z) = \sum_{n=0}^{\infty} x(n) z^{-n}. \quad (1.12)$$

При этом предполагается, что поведение последовательности $x(n)$ до точки $n=0$ неизвестно и его можно не учитывать. Для многих последовательностей свойства одностороннего z -преобразования аналогичны свойствам обычного z -преобразования. Основным исключением является свойство, связанное со сдвигом (задержкой) последовательностей.

Рассмотрим, например, последовательность $x_1(n)$ с односторонним z -преобразованием $X_1(z)$ и задержанную последовательность $x_2(n) = x_1(n-1)$. Одностороннее z -преобразование от $x_2(n)$ равно

$$X_2(z) = \sum_{n=0}^{\infty} x_2(n) z^{-n} = \sum_{n=0}^{\infty} x_1(n-1) z^{-n}.$$

Положив $m = n-1$, получим равенство

$$X_2(z) = \sum_{m=-1}^{\infty} x_1(m) z^{-m} z^{-1},$$

которое может быть переписано следующим образом:

$$X_2(z) = z^{-1} \left[x_1(-1)z + \sum_{m=0}^{\infty} x_1(m)z^{-m} \right] = z^{-1} [X_1(z)] + x_1(-1). \quad (1.13)$$

Задержка на один отсчет по-прежнему приводит к умножению одностороннего z -преобразования на z^{-1} , но при этом необходимо учесть значения последовательности $x_1(n)$ при $n < 0$, т. е. важную роль начинают играть начальные условия.

В качестве другого примера рассмотрим z -преобразование последовательности $x_2(n) = x_1(n-2)$, равное

$$X_2(z) = z^{-2} [X_1(z)] + x_1(-2) + x_1(-1)z^{-1}. \quad (1.14)$$

Из выражений (1.13) и (1.14) можно получить формулу для случая задержки последовательности на произвольное число n_0 отсчетов ($n_0 > 0$). Она имеет вид

$$X_2(z) = z^{-n_0} [X_1(z)] + x_1(-n_0) + x_1(-n_0+1)z^{-1} + \dots + x_1(-1)z^{-n_0+1},$$

причем $x_2(n) = x_1(n-n_0)$.

1.7. Дискретное преобразование Фурье

Рассмотрим периодическую последовательность $x_p(n)$ с периодом N , т. е. $x_p(n) = x_p(n+kN)$ для любого целого значения k . Такие последовательности не могут быть представлены z -преобразованием, так как не существует ни одного значения z , для которого бы сходилось z -преобразование такой последовательности. Однако их можно представить рядом Фурье, т. е. суммой синусоидальных и косинусоидальных последовательностей с частотами, кратными основной частоте $2\pi/N$ периодической последовательности.

В противоположность рядам Фурье непрерывных периодических функций здесь достаточно только N различных комплексных экспонент с периодом, равным целой части основного периода N . Это является следствием того, что комплексная экспонента

$$e_k(n) = e^{j\frac{2\pi}{N}kn} \quad (1.15)$$

периодична по k с периодом N . Так, $e_0(n) = e_N(n)$, $e_1(n) = e_{N+1}(n)$ и так далее.

Следовательно, множество N комплексных экспонент (1.15) с $k=0, 1, 2, \dots, N-1$ определяет все различные комплексные экспоненты с частотами, кратными $\frac{2\pi}{N}$. Поэтому представление периодической последовательности $x_p(n)$ рядом Фурье содержит только N этих комплексных экспонент и, следовательно, имеет вид

$$x_p(n) = \frac{1}{N} \sum_{k=0}^{N-1} X_p(k) e^{j\frac{2\pi}{N}nk}. \quad (1.16)$$

Множитель $\frac{1}{N}$ введён для удобства и не влияет на характер представления.

Чтобы выразить коэффициенты $X_p(k)$ через $x_p(n)$, умножим обе части равенства (1.16) на $e^{-j(2\pi/N)mn}$ и просуммируем результаты по n :

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

Меняя в правой части этого выражения порядок суммирования и используя формулу

$$\sum_{n=0}^{N-1} e^{j\frac{2\pi}{N}n(k-m)} = \begin{cases} N, & \text{при } k = m, \\ 0, & \text{при } k \neq m, \end{cases} \quad (1.17)$$

получим

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

или (после перестановки левой и правой частей этого равенства и замены индекса m на k)

$$X_p(k) = \sum_{n=0}^{N-1} x_p(n) e^{-j\frac{2\pi}{N}nk}. \quad (1.18)$$

Соотношение (1.18) носит название дискретного преобразования Фурье (ДПФ), а (1.16) – обратного ДПФ (ОДПФ).

1.7.1. Связь между дискретным преобразованием Фурье и z -преобразованием

Из определений (1.16) и (1.18) видно, что обе последовательности $x_p(n)$ и $X_p(k)$ периодичны с периодом в N отсчетов. Ясно также (см. (1.18)), что $X_p(k)$ полностью определяется одним периодом $x_p(n)$. Отсюда возникает интересный вопрос: как связаны z -преобразование конечной последовательности, образованной из одного периода периодической последовательности, и ДПФ всей периодической последовательности? Иначе говоря, рассмотрим последовательность конечной длины

$$x(n) = \begin{cases} x_p(n), & \text{при } 0 \leq n \leq N-1, \\ 0, & \text{при других } n, \end{cases} \quad (1.19)$$

причем последовательность $x_p(n)$ имеет период в N отсчетов, т. е. $x(n)$ представляет собой один период периодической последовательности $x_p(n)$. Z -преобразование $x(n)$ равно

$$X(z) = \sum_{n=0}^{N-1} x(n) z^{-n}.$$

Вычисляя эту сумму при $z = e^{j(2\pi/N)k}$, т. е. в точке на единичной окружности с полярным углом $2\pi k/N$, находим

$$X(z) \Big|_{z = e^{j\frac{2\pi}{N}k}} = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}nk}.$$

Сравнивая последнюю сумму с (1.18) и учитывая, что $x_p(n)=x(n)$ на интервале $0 \leq n \leq N-1$, получаем

$$X_p(k) = X(e^{j\frac{2\pi}{N}k}).$$

Итак, коэффициенты ДПФ всей периодической последовательности равны значениям z -преобразования последовательности конечной длины в N точках, равномерно распределенных по единичной окружности.

Для иллюстрации приведенных положений рассмотрим периодическую последовательность с периодом N (рис. 1.4а), определяемую как

$$x_p(n) = a^n, \quad 0 \leq n \leq N-1,$$

$$x_p(n+mN) = x_p(n), \quad m = \pm 1, \pm 2, \dots$$

Согласно определению (1.18), ее ДПФ равно

$$X_p(k) = \sum_{n=0}^{N-1} a^n e^{-j\frac{2\pi}{N}nk} = \sum_{n=0}^{N-1} \left(a e^{-j\frac{2\pi}{N}n} \right)^n = \frac{1 - a^N}{1 - a e^{-j\frac{2\pi}{N}k}},$$

$$0 \leq k \leq N-1.$$

Модули и фазы элементов последовательности $X_p(k)$ для случая $a=0.9$, $N=16$ изображены на рис. 1.4б.

Последовательность конечной длины $x(n)$ (рис. 1.4в) определяется как

$$x(n) = \begin{cases} a^n & \text{при } 0 \leq n \leq N-1, \\ 0 & \text{при других } n \end{cases},$$

т. е. она состоит из одного периода последовательности $x_p(n)$. Ее z -преобразование равно

$$X(z) = \sum_{n=0}^{N-1} a^n z^{-n} = \frac{1 - a^N z^{-N}}{1 - a z^{-1}}.$$

Вычисляя $X(z)$ на единичной окружности, получим

$$X(e^{j\omega}) = \frac{1 - a^N e^{-j\omega N}}{1 - a e^{-j\omega}}.$$

Модуль и фаза этой функции для $0 \leq \omega \leq 2\pi$ изображены на рис.1.4з. Значения $X_p(k)$ и $X(e^{j\omega})$ в точках $\omega_k = 2\pi k/N$ совпадают.

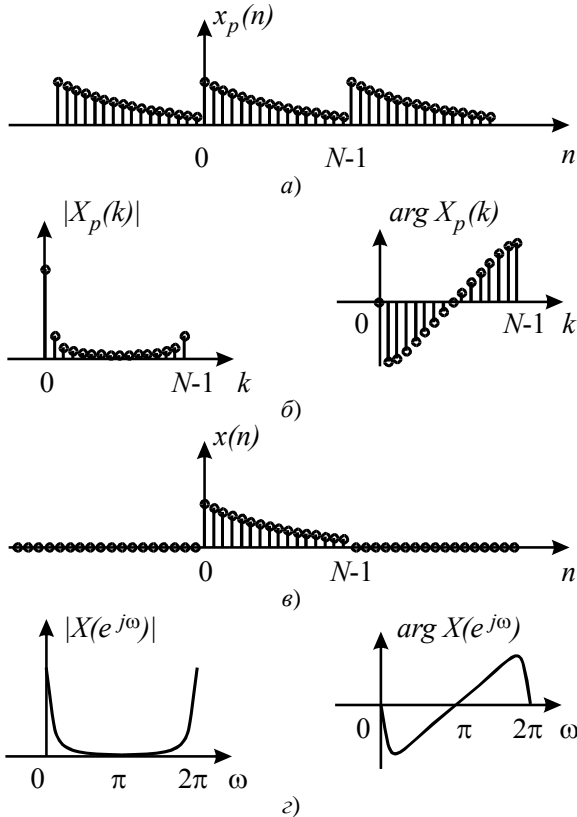


Рис. 1.4. Связь между ДПФ и z -преобразованием:
 а) последовательность с периодом N ; б) модули и фазы элементов последовательности $X_p(k)$; в) последовательность конечной длины;
 г) модуль и фаза функции $X(z)$ на единичной окружности

Найдем z -преобразование последовательности конечной длины через коэффициенты ДПФ периодической последовательности. Из соотношений (1.16), (1.19) и определения z -преобразования получаем

$$\begin{aligned} X(z) &= \sum_{n=0}^{N-1} x(n)z^{-n} = \sum_{n=0}^{N-1} \frac{1}{N} \sum_{k=0}^{N-1} X_p(k) e^{j\frac{2\pi}{N}kn} z^{-n} = \\ &= \sum_{k=0}^{N-1} X_p(k) \frac{1}{N} \sum_{n=0}^{N-1} \left(e^{j\frac{2\pi}{N}k} z^{-1} \right)^n = \sum_{k=0}^{N-1} \frac{X_p(k)}{N} \frac{1 - z^{-N}}{1 - z^{-1} e^{j\frac{2\pi}{N}k}}. \end{aligned}$$

Вычисляя $X(z)$ на единичной окружности, получим

$$X(e^{j\omega}) = \frac{1 - a^N e^{-j\omega N}}{1 - a e^{-j\omega}}.$$

Это равенство показывает, что z -преобразование последовательности непосредственно связано с коэффициентами её ДПФ. Для точек на единичной окружности оно принимает вид

$$X(e^{j\omega}) = \sum_{k=0}^{N-1} \frac{X_p(k)}{N} \frac{e^{-j\omega\frac{N-1}{2}} \sin\left(\frac{\omega N}{2}\right)}{e^{-j\omega\frac{\pi k}{N}} \sin\left(\frac{\omega}{2} - \frac{\pi k}{N}\right)},$$

где функции вида $\sin\left(\frac{\omega N}{2}\right) / \sin\left(\frac{\omega}{2} - \frac{\pi k}{N}\right)$ интерполируют значения коэффициентов ДПФ $X_p(k)$ на всю ось частот. Следовательно, с помощью последней формулы по коэффициентам ДПФ последовательности конечной длины можно найти ее непрерывный частотный спектр.

1.7.2. Дискретное преобразование Фурье последовательностей конечной длины

Из выражений (1.18) и (1.19) следует, что коэффициенты ДПФ последовательности конечной длины однозначно представляют саму последовательность, так как по ним можно точно восстановить исходную последовательность, используя ОДПФ. Итак, хотя ДПФ и ОДПФ вводятся для периодических последовательностей, важно, что через них можно представлять последовательности конечной длины.

Представление конечных последовательностей с помощью ДПФ удобно также для получения значений преобразования Фурье в L точках, равномерно распределенных на единичной окружности. Для получения требуемого частотного разрешения L может быть выбрано значительно большим, чем N .

Рассмотрим конечную последовательность $\{x(n)\}$, $0 \leq n \leq N-1$ с z -преобразованием на единичной окружности

$$X(e^{j\omega}) = \sum_{n=0}^{N-1} x(n)e^{-j\omega n}.$$

Вычисляя $X(e^{j\omega})$ на частотах $\omega_l = (2\pi/L)l$, $l=0, 1, \dots, L-1$, получим

$$X(e^{j\frac{2\pi}{L}l}) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{L}nl}. \quad (1.20)$$

Введем новую последовательность $\hat{x}(n)$ длины L точек ($L > N$):

$$x(n) = \begin{cases} x(n), & 0 \leq n \leq N-1, \\ 0, & N \leq n \leq L-1 \end{cases}$$

и найдем ее L -точечное ДПФ:

$$\hat{X}(z) = \sum_{n=0}^{L-1} \hat{x}(n)e^{-j\frac{2\pi}{L}nk}.$$

Поскольку $\hat{x}(n) = 0$ при $n \geq N$, это равенство можно записать в виде

$$\hat{X}(k) = \sum_{n=0}^{N-1} x(n) e^{-j \frac{2\pi}{L} nk}. \quad (1.21)$$

Сравнивая (1.21) и (1.20), получим

$$\hat{X}(k) = X(e^{j \frac{2\pi}{L} k}).$$

Таким образом, простое дополнение последовательности конечной длины нулевыми отсчетами позволяет достичь любого разрешения при расчете преобразования Фурье этой последовательности для совокупности точек, равномерно распределенных по единичной окружности. (Вместе с тем следует иметь в виду, что частотное разложение зависит только от длительности сигнала N . Выбор $L > N$ лишь улучшает условия различения синусоидальных компонент). При спектральном анализе конечных последовательностей эта несложная операция является одной из наиболее полезных.

Итак, было показано, что ДПФ однозначно представляет последовательность конечной длины, содержащую N элементов, причем коэффициенты ДПФ равны значениям z -преобразования последовательности в N точках, равномерно распределенных по единичной окружности. Аналогично z -преобразование любой (в том числе и бесконечной) последовательности однозначно представляет эту последовательность.

1.7.3. Эффекты наложения

Рассмотрим, какая получится последовательность, если в качестве коэффициентов ДПФ взять значения z -преобразования произвольной последовательности, вычисленного в N точках, равномерно распределенных по единичной окружности.

Для большей ясности предположим, что последовательность $x(n)$ (не обязательно конечная) имеет z -преобразование

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}.$$

Определим набор коэффициентов ДПФ как

$$X_p(k) = X(z) \Big|_{z = e^{j\frac{2\pi}{N}k}} = \sum_{n=-\infty}^{\infty} x(n)e^{-j\frac{2\pi}{N}nk}. \quad (1.22)$$

По этим коэффициентам можно найти периодическую последовательность $x_p(n)$, равную

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

Подставляя значения коэффициентов (1.22) в последнюю формулу и заменяя индекс суммирования на m , получим

$$\begin{aligned} x_p(n) &= \frac{1}{N} \sum_{k=0}^{N-1} \left[\sum_{m=-\infty}^{\infty} x(m)e^{-j\frac{2\pi}{N}mk} \right] e^{j\frac{2\pi}{N}nk} = \\ &= \sum_{m=-\infty}^{\infty} x(m) \left[\frac{1}{N} \sum_{k=0}^{N-1} e^{-j\frac{2\pi}{N}k(m-n)} \right] = \sum_{m=-\infty}^{\infty} x(m) \sum_{r=-\infty}^{\infty} \delta(m-n+rN), \\ x_p(n) &= \sum_{r=-\infty}^{\infty} x(n-rN). \end{aligned}$$

Это соотношение является весьма важным. Оно показывает, что периодическая последовательность, получаемая из ОДПФ набора значений z -преобразования непериодической последовательности, вычисленных в N точках, равномерно распределенных по единичной окружности, состоит из сдвинутых и наложенных копий исходной непериодической последовательности. Если длина последовательности $x(n)$ не

превышает N отсчетов, то наложение в $x_p(n)$ фактически отсутствует. Это равенство также показывает, что искажения, связанные с наложением, которые возникают при описании бесконечной последовательности конечным числом N коэффициентов ДПФ, уменьшаются при увеличении N .

На рис. 1.5 изображены последовательности $x_1(n)$ и $x_2(n)$ и соответствующие им N -точечные периодические эквиваленты. В первом примере длина последовательности $x_1(n)$ близка к N , поэтому $x_p(n)$ повторяет ее почти без искажений. Во втором примере длина $x(n)$ значительно больше N , поэтому периодическая последовательность заметно отличается от исходной.

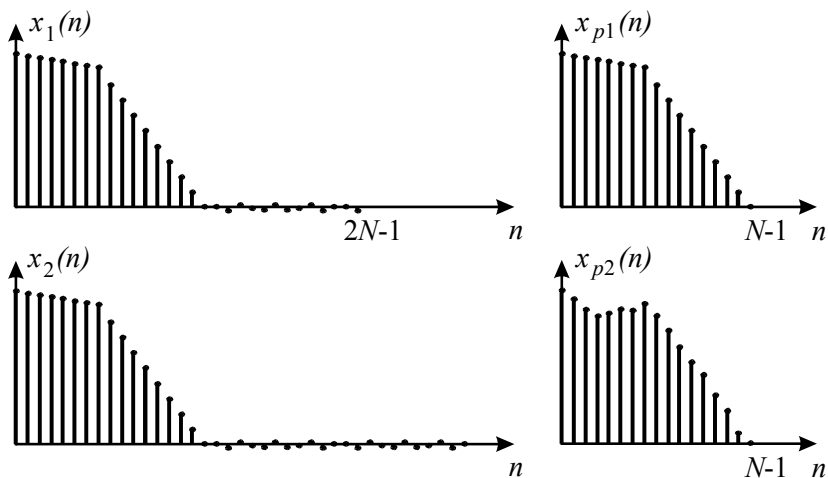


Рис. 1.5. Две последовательности разной длины и соответствующие им N -точечные периодические эквиваленты

1.8. Свойства дискретного преобразования Фурье

Некоторые свойства ДПФ играют в практических вопросах обработки сигналов важную роль. В основном они будут перечислены, а детали будут рассматриваться только в случае необходимости.

1.8.1. Линейность

Если $x_p(n)$ и $y_p(n)$ – периодические последовательности (с периодом в N отсчетов каждая), а $X_p(k)$ и $Y_p(k)$ – их ДПФ, то при любых действительных a и b ДПФ последовательности $ax_p(n)+by_p(n)$ равно $aX_p(k)+bY_p(k)$. Это положение справедливо и для последовательности конечной длины.

1.8.2. Сдвиг

Если последовательность $x_p(n)$ периодическая с периодом в N отсчетов, а ее ДПФ равно $X_p(k)$, то ДПФ периодической последовательности вида $x_p(n-n_0)$ будет равно $X_p(k)e^{-j(2\pi/N)n_0k}$. При анализе последовательностей конечной длины необходимо учитывать специфический характер временного сдвига последовательности (рис. 1.6).

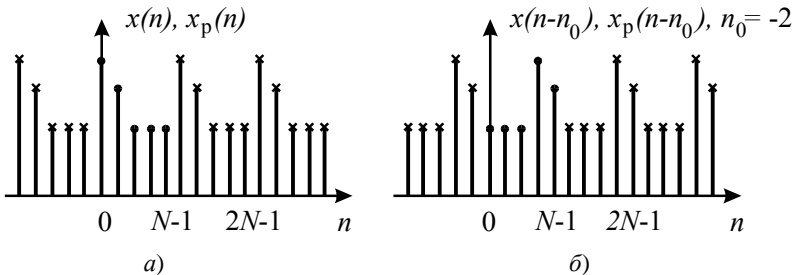


Рис. 1.6. Специфический характер временного сдвига конечной последовательности $x(n)$ и эквивалентной периодической последовательности $x_p(n)$:

а) исходные последовательности; б) сдвинутые последовательности

Так, на рис. 1.6а кружочками изображена конечная последовательность $x(n)$ длиной в 5 отсчетов. Здесь же крестиками отмечены отсчеты эквивалентной периодической последовательности $x_p(n)$, имеющей то же ДПФ, что и $x(n)$. Чтобы найти ДПФ сдвинутой последовательности $x(n-n_0)$, причем $n_0 < N$ (рис. 1.6б), следует рассмотреть сдвинутую периодическую последовательность $x_p(n-n_0)$ и в качестве эквивалентной сдвинутой конечной последовательности

(имеющей ДПФ $X(k)e^{-j(2\pi/N)n_0k}$) принять отрезок последовательности $x_p(n-n_0)$ в интервале $0 \leq n \leq N-1$. Таким образом, с точки зрения ДПФ последовательность $x(n-n_0)$ получается путем кругового сдвига элементов последовательности $x_p(n)$ на n_0 отсчетов.

1.8.3. Свойства симметрии

Если периодическая последовательность $x_p(n)$ с периодом в N отсчетов является действительной, то ее ДПФ $X_p(k)$ удовлетворяет следующим условиям симметрии:

$$\operatorname{Re}[X_p(k)] = \operatorname{Re}[X_p(N-k)], \quad \operatorname{Im}[X_p(k)] = -\operatorname{Im}[X_p(N-k)], \quad (1.23)$$

$$|X_p(k)| = |X_p(N-k)|, \quad \arg X_p(k) = -\arg X_p(N-k).$$

Следовательно, имеет место симметрия спектра относительно $\frac{N}{2}$. Используя условие симметрии, последовательность $x_p(n)$ можно представить в виде конечной суммы:

$$\begin{aligned} x_p(n) = & \frac{1}{N} X_p(0) + \frac{2}{N} |X_p(1)| \cos \left[\frac{2\pi}{N} n + \arg X_p(1) \right] + \\ & + \frac{2}{N} |X_p(2)| \cos \left[\frac{4\pi}{N} n + \arg X_p(2) \right] + \dots \\ & \dots + \frac{1}{N} |X_p(N/2)| \cos \left[\pi n + \arg X_p(N/2) \right]. \end{aligned}$$

Аналогичные равенства справедливы и для конечной действительной последовательности $x(n)$, имеющей N -точечное ДПФ $X(k)$. Если ввести дополнительное условие симметрии последовательности $x_p(n)$, т. е. считать, что

$$x_p(n) = x_p(N-n),$$

то окажется, что $X_p(k)$ может быть только действительной.

Поскольку чаще всего приходится иметь дело с действительными последовательностями, то, вычислив одно ДПФ, можно получить ДПФ двух последовательностей, используя свойства симметрии (1.23). Рассмотрим действительные периодические последовательности $x_p(n)$ и $y_p(n)$ с периодами в N отсчетов и N -точечными ДПФ $X_p(k)$ и $Y_p(k)$ соответственно. Введем комплексную последовательность $z_p(n)$ вида

$$z_p(n) = x_p(n) + jy_p(n).$$

Ее ДПФ равно

$$Z_p(k) = \sum_{n=0}^{N-1} [x_p(n) + jy_p(n)] e^{-j\frac{2\pi}{N}nk},$$

$$Z_p(k) = X_p(k) + jY_p(k).$$

Выделяя действительную и мнимую части последнего равенства, получим

$$\operatorname{Re}[Z_p(k)] = \operatorname{Re}[X_p(k)] - \operatorname{Im}[Y_p(k)],$$

$$\operatorname{Im}[Z_p(k)] = \operatorname{Im}[X_p(k)] + \operatorname{Re}[Y_p(k)].$$

Действительные части $X_p(k)$ и $Y_p(k)$ симметричны, а мнимые асимметричны, поэтому их легко разделить, используя операции сложения и вычитания

$$\operatorname{Re}[X_p(k)] = \frac{\operatorname{Re}[Z_p(k)] + \operatorname{Re}[Z_p(N-k)]}{2},$$

$$\operatorname{Im}[X_p(k)] = \frac{\operatorname{Im}[Z_p(k)] - \operatorname{Im}[Z_p(N-k)]}{2},$$

$$\operatorname{Re}\left[Y_p(k)\right] = \frac{\operatorname{Im}\left[Z_p(k)\right] + \operatorname{Im}\left[Z_p(N-k)\right]}{2},$$

$$\operatorname{Im}\left[Y_p(k)\right] = \frac{-\operatorname{Re}\left[Z_p(k)\right] + \operatorname{Re}\left[Z_p(N-k)\right]}{2}.$$

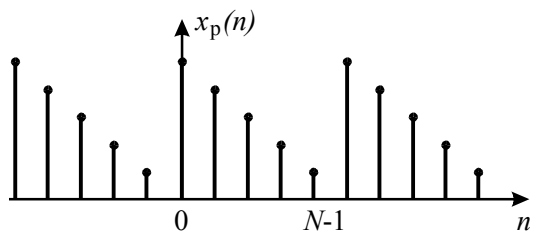
Итак, вычисляя одно N -точечное ДПФ $Z_p(k)$, удается преобразовать сразу две действительные последовательности $x_p(n)$ и $y_p(n)$ длиной по N отсчетов. Если эти последовательности являются еще и симметричными, то число операций, необходимых для получения их ДПФ, можно сократить еще больше.

1.9. Свертка периодических последовательностей

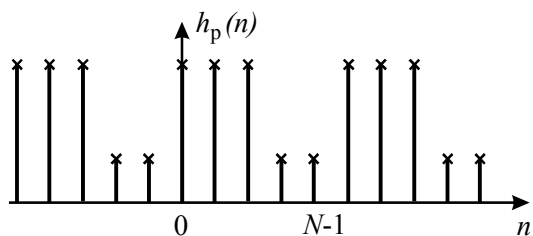
Рассмотрим две периодические последовательности $x_p(n)$ и $h_p(n)$ с периодами по N отсчетов. Сверткой таких последовательностей (круговой, циклической, периодической сверткой) называют последовательность $y_p(n)$, определяемую соотношением

$$y_p(n) = \sum_{l=0}^{N-1} x_p(l)h_p(n-l). \quad (1.24)$$

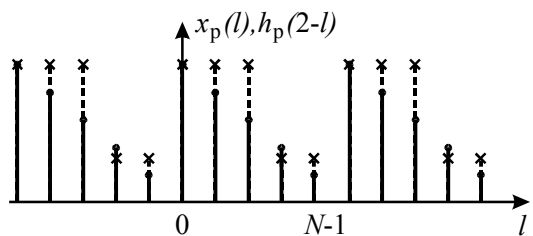
Поясним понятие круговой свертки. На рис. 1.7а и 1.7б изображены периодические последовательности $x_p(n)$ и $h_p(n)$, а на рис. 1.7в показано, как вычисляется значение круговой свертки (1.24) при $n=2$. В силу периодичности последовательностей $x_p(l)$ и $h_p(n-l)$ достаточно рассматривать их на интервале $0 \leq l \leq N-1$. С изменением n последовательность $h_p(n-l)$ смещается относительно $x_p(l)$. Когда отсчет $h_p(n-l)$ выходит за точку $l=N-1$, точно такой же отсчет появляется в точке $l=0$. Поэтому круговая свертка определяет свертку двух последовательностей, заданных по окружности.



a)



б)



в)

Рис. 1.7. Круговая свертка: а) последовательность $x_p(n)$; б) последовательность $h_p(n)$; в) вычисление значения круговой свертки при $n=2$

Согласно выражению (1.18) последовательности $x_p(n)$ и $h_p(n)$ имеют соответствующие ДПФ:

$$X_p(k) = \sum_{n=0}^{N-1} x_p(n) e^{-j\frac{2\pi}{N}nk},$$

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

Определим ДПФ круговой свертки:

$$\begin{aligned} Y_p(k) &= \sum_{n=0}^{N-1} y_p(n) e^{-j\frac{2\pi}{N}kn} = \sum_{n=0}^{N-1} \left[\sum_{l=0}^{N-1} x_p(l) h_p(n-l) \right] e^{-j\frac{2\pi}{N}kn} = \\ &= \sum_{l=0}^{N-1} x_p(l) \underbrace{\left[\sum_{n=0}^{N-1} h_p(n-l) e^{-j\frac{2\pi}{N}k(n-l)} \right]}_{H_p(k)} e^{-j\frac{2\pi}{N}kl} = \\ &= H_p(k) \underbrace{\sum_{l=0}^{N-1} x_p(l) e^{-j\frac{2\pi}{N}kl}}_{X_p(k)} = X_p(k) H_p(k). \end{aligned} \quad (1.25)$$

Таким образом, перемножая ДПФ двух периодических последовательностей и находя ОДПФ произведения, получаем такой же результат, как при круговой свертке таких последовательностей.

Формула (1.25) справедлива и для конечных последовательностей, если рассматривать $x_p(n)$ и $h_p(n)$ как эквивалентные им периодические последовательности с теми же ДПФ. Однако для конечных последовательностей обычно нужна линейная (ее называют аperiodической), а не круговая свертка, поэтому в приведенные формулы следует внести уточнения.

1.10. Линейная свертка конечных последовательностей

Рассмотрим две конечные последовательности $x(n)$ и $h(n)$ длины по N_1 и N_2 отсчетов, т. е. $x(n)$ отлична от нуля при $0 \leq n \leq N_1 - 1$, а $h(n)$ — при $0 \leq n \leq N_2 - 1$. Линейной, или аperiodической, сверткой этих последовательностей называют последовательность $y(n)$, определяемую соотношением

$$y(n) = \sum_{m=0}^n x(m)h(n-m), \quad (1.26)$$

где $x(m)$ и $h(n-m)$ равны нулю вне соответствующих интервалов. На рис. 1.8 приведены примеры последовательностей $x(n)$, $h(n)$ и $y(n)$. Последовательность $y(n)$ является конечной и имеет длину (N_1+N_2-1) отсчетов. Выше было доказано, что, перемножая ДПФ двух периодических последовательностей и находя ОДПФ произведения, получим такой же результат, как при круговой свертке таких последовательностей.

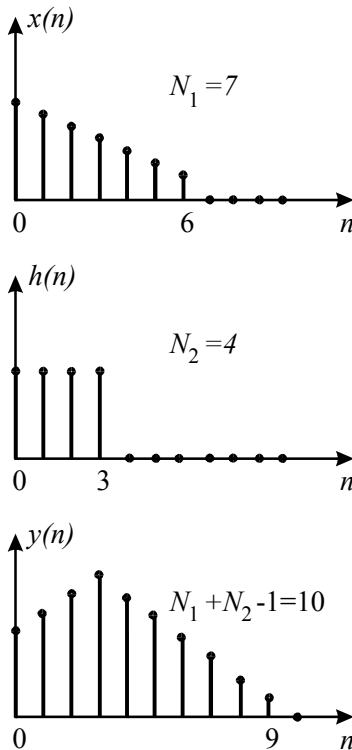


Рис. 1.8. Линейная свертка последовательностей $x(n)$ и $h(n)$

Исходя из этого (см. также пример на рис. 1.7) можно довольно просто получить линейную свертку двух конечных последовательностей. Свертка периодических последовательностей периодична и имеет тот же период, что и сами последовательности. Поскольку длительность свертки $y(n)$ (рис. 1.8) равна (N_1+N_2-1) отсчетам, то для получения такого периода при круговой свертке необходимо, чтобы $x(n)$ и $h(n)$ содержали по (N_1+N_2-1) отсчетов, что достигается дополнением каждой из двух последовательностей соответствующим числом нулевых отсчетов. После этого можно найти (N_1+N_2-1) -точечные ДПФ дополненных последовательностей, перемножить их и выполнить ОДПФ произведения. В результате получается искомая свертка $y(n)$.

На рис. 1.9, иллюстрирующем эти операции, изображены эквивалентные периодические последовательности, используемые при вычислении круговой свертки. Дополнение исходных последовательностей конечной длины $x(n)$ и $h(n)$ нулевыми отсчетами доводит период до нужной величины и позволяет устранить круговые наложения, характерные для круговой свертки. В результате каждый период последовательности $y_p(n)$ (рис. 1.9) совпадает с $y(n)$ (рис. 1.8).

Рассмотренный метод вычисления свертки двух конечных последовательностей с применением алгоритма ДПФ называется быстрой сверткой в противоположность методу прямого вычисления суммы (1.26), называемому прямой или медленной сверткой. Термин «быстрая» употребляется потому, что ДПФ можно вычислить быстро и эффективно, используя один из алгоритмов быстрого преобразования Фурье (БПФ). Можно показать, что даже при умеренных величинах (N_1+N_2-1) (например, около 30) быстрая свертка оказывается эффективнее прямой. Поэтому рассмотренная методика является важным вычислительным средством при обработке сигналов.

Для практических приложений важно отметить, что в приведенном выше примере размер ДПФ не обязательно ограничивать величиной (N_1+N_2-1) . Выполнять ДПФ можно по любому числу отсчетов L , удовлетворяющему условию $L \geq N_1+N_2-1$. Если это условие удовлетворяется, то, в отличие от

описанной методики, последовательности $x(n)$ и $h(n)$ дополняются другим числом нулевых отсчетов. В результате эквивалентная периодическая последовательность $y_p(n)$ будет иметь в конце периодов $(L-N_1-N_2+1)$ нулей. Ясно, что эти отличия никак не искажают желаемого результата.

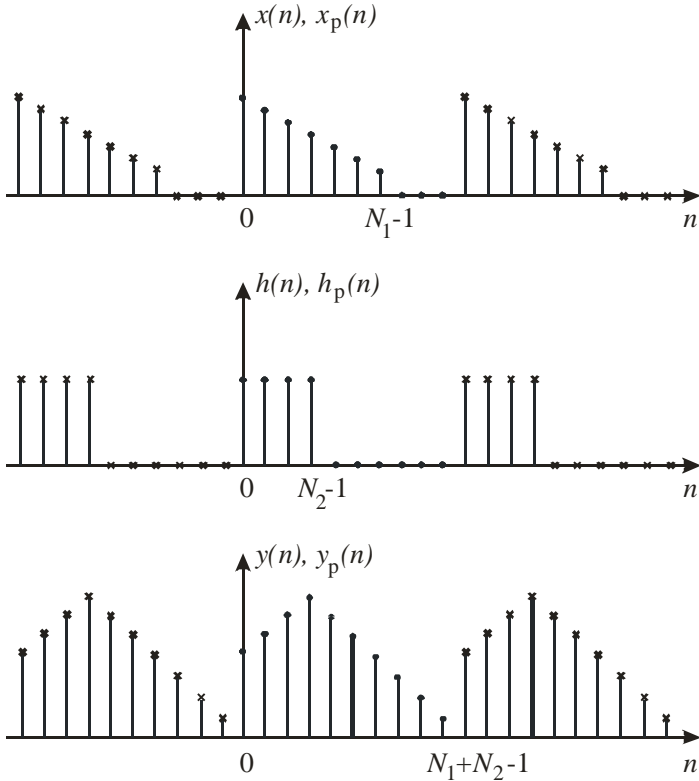


Рис. 1.9. Вычисление линейной свертки последовательностей с помощью круговой свертки

Возможность произвольного выбора L существенна, поскольку практические алгоритмы вычисления ДПФ при разных L имеют неодинаковую эффективность. Так, например, для некоторых алгоритмов необходимо, чтобы L равнялось степени 2.

В этом случае в качестве L приходится выбирать число, равное степени 2 и не меньшее, чем $N_1 + N_2 - 1$.

1.11. Разностные уравнения

Разностные уравнения широко используются для исследования процессов в цифровых системах. Описание линейных систем разностными уравнениями очень важно, так как оно часто позволяет найти эффективные способы построения таких систем. Более того, по разностному уравнению можно определить многие характеристики рассматриваемой системы, включая собственные частоты, порядок системы, частоты, соответствующие нулевому коэффициенту передачи и т. д.

В самом общем случае линейное неоднородное разностное уравнение M -го порядка с постоянными коэффициентами, относящееся к физически реализуемой системе, имеет вид

$$y(n) = \sum_{i=0}^M a_i x(n-i) - \sum_{i=1}^M b_i y(n-i), \quad n \geq 0, \quad (1.27)$$

где коэффициенты a_i и b_i описывают конкретную систему, причем $b_M \neq 0$.

Уравнение (1.27) записано в виде, удобном для решения методом прямой подстановки. Имея набор начальных условий (например, $x(i)$, $y(i)$ для $i = -1, -2, \dots, -M$) и последовательность $x(n)$, по формуле (1.27) можно непосредственно вычислить последовательность $y(n)$ для $n \geq 0$. Например, разностное уравнение

$$y(n) = x(n) - 3y(n-1) \quad (1.28)$$

с начальным условием $y(-1) = 0$ и $x(n) = n^2 + n$ можно решить подстановкой, что дает

$$y(0) = x(0) - 3y(-1) = 0,$$

$$y(1) = x(1) - 3y(0) = 2,$$

$$y(2) = x(2) - 3y(1) = 0,$$

$$y(3)=x(3)-3y(2)=12,$$

$$y(4)=x(4)-3y(3)=-16,$$

$$y(5)=x(5)-3y(4)=78,$$

$$y(6)=x(6)-3y(5)=-192,$$

.....

Хотя решение разностных уравнений прямой подстановкой и целесообразно в некоторых случаях, значительно полезнее получить решение уравнения в явном (аналитическом) виде. При этом основная идея сводится к получению двух решений разностного уравнения: однородного и частного.

Однородное решение $y_0(n)$ получается путем подстановки нулей вместо всех членов, содержащих элементы последовательности $x(n)$, и нахождения решения $y_0(n)$ при нулевой последовательности $x(n)$. Именно этот класс решений описывает основные свойства заданной системы. Частное решение получают, подбирая вид последовательности $y(n)$ при последовательности $x(n)$. Для определения произвольных постоянных однородного решения используются начальные условия.

В качестве примера решим этим методом уравнение (1.28). Однородное уравнение имеет вид

$$y(n)+3y(n-1)=0.$$

Известно, что характеристическими решениями однородных уравнений, соответствующих линейным разностным уравнениям первого порядка с постоянными коэффициентами, являются решения вида Aq^n . Поэтому, подставив в последнее уравнение Aq^n вместо $y(n)$, получим

$$Aq^n+3Aq^{n-1}=0, Aq^{n-1}(q+3)=0,$$

$$q=-3, y_0(n)=A(-3)^n.$$

Частное решение, соответствующее последовательности $x(n)=n^2+n$, попробуем найти в виде

$$y_c(n) = Bn^2 + Cn + D.$$

Из уравнения (1.28) получаем

$$Bn^2 + Cn + D + 3B(n-1)^2 + 3C(n-1) + 3D = n^2 + n.$$

Поскольку коэффициенты при равных степенях n должны совпадать, B , C и D должны быть равны:

$$B = 1/4, \quad C = 5/8, \quad D = 9/32.$$

Таким образом, общее решение имеет вид

$$y(n) = n^2/4 + 5n/8 + 9/32 + A(-3)^n.$$

Коэффициент A определяется из начального условия $y(-1) = 0$, откуда $A = -9/32$ и

$$y(n) = n^2/4 + 5n/8 + 9/32 [1 - (-3)^n].$$

Выборочная проверка этого решения при $n \geq 0$ показывает полное его совпадение с приведенным выше прямым решением. Очевидное преимущество аналитического решения состоит в том, что оно позволяет весьма просто определить $y(n)$ для любого конкретного $n = n_0$.

Методы нахождения решений подобных уравнений подробно освещены в литературе по разностным уравнениям. Большое значение разностных уравнений состоит в том, что они непосредственно определяют способ построения цифровой системы.

1.12. Решение разностных уравнений с применением одностороннего z -преобразования

Разностные уравнения обычно определены при $n \geq 0$ и имеют набор начальных условий. Поэтому нетрудно понять, каким образом можно использовать одностороннее z -преобразование для решения таких уравнений.

В качестве примера рассмотрим разностное уравнение первого порядка

$$y(n) = x(n) - by(n-1)$$

с начальным условием $y(-1)=K$. Пусть $x(n)=e^{j\omega n}$. Чтобы найти одностороннее z -преобразование $y(n)$, умножим обе части уравнения на z^{-n} и просуммируем от 0 до ∞ :

$$\sum_{n=0}^{\infty} y(n)z^{-n} = \sum_{n=0}^{\infty} x(n)z^{-n} - b \sum_{n=0}^{\infty} y(n-1)z^{-n}.$$

Воспользовавшись свойством, связанным с задержкой последовательности, имеем

$$Y(z) = X(z) - bz^{-1}Y(z) - by(-1),$$

откуда

$$Y(z) = \frac{X(z) - by(-1)}{1 + bz^{-1}}.$$

Поскольку

$$X(z) = \frac{1}{1 - e^{j\omega} z^{-1}},$$

то

$$Y(z) = \frac{-bK}{1 + bz^{-1}} + \frac{1}{(1 + bz^{-1})(1 - e^{j\omega} z^{-1})}.$$

Вычислим обратное z -преобразование:

$$y(n) = \left(K - \frac{1}{b + e^{j\omega}}\right)(-b)^{n+1} + \frac{e^{j\omega(n+1)}}{b + e^{j\omega}}.$$

Первое слагаемое представляет собой свободную составляющую отклика, определяемую начальными условиями, а второе – вынужденные колебания системы. При $|b| < 1$ свободные колебания экспоненциально убывают.

Изложенное нетрудно обобщить на системы более высокого порядка. Например, разностное уравнение L -го порядка может иметь вид

$$y(n) = \sum_{i=0}^L a_i x(n-i) - \sum_{i=1}^L b_i y(n-i)$$

с начальными условиями $\{y(-1), y(-2), \dots, y(-L)\}$.

Замечание. Здесь предполагается, что последовательность $x(n)=0$ при $n<0$.

Вычисляя односторонние z -преобразования от обеих частей последнего уравнения, получим

$$Y(z) = \sum_{i=0}^L a_i z^{-i} X(z) - \sum_{i=1}^L b_i [z^{-i} Y(z) + y(-i) + y(-i+1)z^{-1} + \dots + y(-1)z^{-(i-1)}].$$

Теперь можно получить выражение для $Y(z)$ через $X(z)$ и начальные условия и, взяв обратное z -преобразование, найти $y(n)$.

1.13. Контрольные вопросы

1. Для каких значений независимой переменной (времени) определяются цифровые последовательности?
2. Где в z -плоскости сходится $X(z)$, если $x(n)$ отлична от нуля только на интервале $N_1 \leq n \leq N_2$, причем N_1 и N_2 конечны?
3. Когда последовательность $x(n)$ будет физически реализуемой?
4. Z -преобразование, вычисленное на единичной окружности...
 - а) совпадает с преобразованием Фурье;
 - б) переводит последовательность во временную область;
 - в) совпадает с обратным преобразованием Фурье;
 - г) инвертирует знак у значений отсчетов исходной последовательности.
5. Какое z -преобразование имеет последовательность $x(n-n_0)$, если z -преобразование последовательности $x(n)$ равно $X(z)$?
6. ДПФ и ОДПФ вводятся для периодических последовательностей...
 - а) поэтому через них нельзя представить последовательности конечной длины;
 - б) но только для последовательностей $x(n)>0$, при $n>0$;
 - в) но через них можно представить последовательности конечной длины;
 - г) поэтому для работы с последовательностями конечной длины необходимо использовать z -преобразование.

7. Исходные конечные последовательности имеют длины N_1 и N_2 . Чему равна длина u последовательности, полученной в результате свертки исходных последовательностей?

1.14. Литература

1. Брюханов, Ю. А. Цифровые цепи и сигналы: учеб. пособие / Ю. А. Брюханов. – Ярославль: ЯрГУ, 2005.
2. Солонина, А. И. Основы цифровой обработки сигналов: курс лекций / А. И. Солонина и др. – СПб.: БХВ-Петербург, 2005.
3. Рабинер, Л. Теория и применение цифровой обработки сигналов: пер. с англ. / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978.
4. Оппенгейм, А. В. Цифровая обработка сигналов: пер. с англ. / А. В. Оппенгейм, Р. В. Шафер. – М.: Связь, 1979.
5. Каптелини, В. Цифровые фильтры и их применение: пер. с англ. / В. Каптелини, А. Дж. Константиноидис, П. Эмилиани. – М.: Энергоатомиздат, 1983.
6. Баскаков, С. И. Радиотехнические цепи и сигналы / С. И. Баскаков. – М.: Высшая школа, 2005.
7. Гоноровский, И. С. Радиотехнические цепи и сигналы / И. С. Гоноровский, М. П. Демин. – М.: Радио и связь, 1994.
8. Карташев, В. Г. Основы теории дискретных сигналов и цифровых фильтров / В. Г. Карташев. – М.: Высшая школа, 1982.
9. Гольденберг, Л. М. Цифровая обработка сигналов: справочник / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1985.
10. Гольденберг, Л. М. Цифровая обработка сигналов: учеб. пособие / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1990.
11. Биркган С. Е. Разностные уравнения: учеб. пособие / С. Е. Биркган, Ю. А. Брюханов; Яросл. гос. ун-т им. П. Г. Демидова. – Ярославль: ЯрГУ, 1994.

2. ЦИФРОВЫЕ СИГНАЛЫ И ИХ СПЕКТРЫ

Известно, что все сигналы, определённые на конечном или бесконечном интервалах, можно классифицировать как непрерывные и дискретные.

Непрерывные сигналы $x(t)$ характеризуются непрерывным изменением аргумента t и поэтому определены на бесконечном количестве временных точек любого, даже конечного, интервала $[0, T_c]$. Такие сигналы также называются континуальными или аналоговыми. Непрерывные сигналы, полученные на выходе реального устройства – дискретизатора, будем называть дискретизированными. Дискретные сигналы, в отличие от непрерывных, определены в фиксированные моменты времени nT и, следовательно, на конечном интервале имеют конечное число точек определения, а на бесконечном – бесконечное. Цифровой сигнал от дискретного отличается тем, что его значения аппроксимируются конечным числом двоичных разрядов.

2.1. Взаимная связь непрерывного сигнала и полученного из него дискретизированного сигнала

Рассмотрим процесс преобразования непрерывного сигнала $x_{вх}(t)$ в дискретизированный $x_{вых}(t)$, который физически реализуется при помощи устройства, называемого ключом (рис. 2.1).

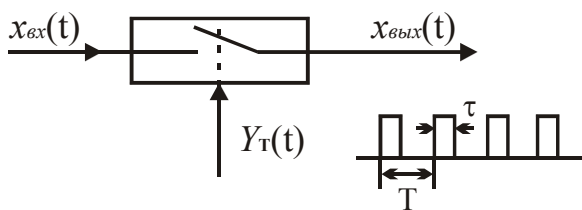


Рис. 2.1. Дискретизация непрерывного сигнала

Сигнал, подвергаемый преобразованию, подаётся на вход ключа. На управляющий вход ключа поступает последовательность дискретизирующих импульсов $Y_T(t)$,

обеспечивающих замыкание ключа на время действия импульса. Если значение $Y_T(t)$ в паузе равно нулю, а амплитуда импульса – единице, то сигнал на выходе ключа можно представить в виде

$$x_{\text{вых}}(t) = x_{\text{вх}}(t)Y_T(t). \quad (2.1)$$

Рассмотрим временные диаграммы сигналов. На рис. 2а изображен непрерывный сигнал $x_{\text{вх}}(t)$, определённый на интервале $[0, T_C]$ и поступающий на вход ключа. Выходной сигнал зависит от длительности импульса дискретизирующей последовательности τ (рис. 2.2б), в течение которого ключ замкнут, и отношения T_C/T . Остановимся на частном случае, когда $T_C/T = N$, где N – некоторое целое число. Сигнал на выходе ключа, соответствующий этому случаю, изображен на рис. 2.2в. Выходной сигнал на интервале $[0, T]$ можно записать через *rect*-функцию – математическую модель дискретизирующего импульса:

$$x_{\text{вых1}}(t) = x_{\text{вх}}(t) \sum_{n=0}^{N-1} \text{rect}\left(\frac{t - \tau/2 - nT}{\tau}\right). \quad (2.2)$$

На интервалах времени $[nT, nT + \tau]$, т. е. когда ключ замкнут, мгновенное значение выходного сигнала совпадает с соответствующими значениями входного.

Если $0 < \tau < T$, то ключ, а следовательно, и выходной сигнал (2.2), физически реализуем. При $\tau \rightarrow 0$ имеем $\text{rect}(\tau/t) \rightarrow u_0(t)$. Здесь $u_0(t)$ – единичный импульс, определяемый выражением

$$u_0(\tau) = \begin{cases} 1 & \text{при } \tau = 0 \\ 0 & \text{при } \tau \neq 0 \end{cases}.$$

Тогда (2.2) примет вид

$$x_{\text{вых2}}(t) = x_{\text{вх}}(t) \sum_{n=0}^{N-1} u_0(t - nT). \quad (2.3)$$

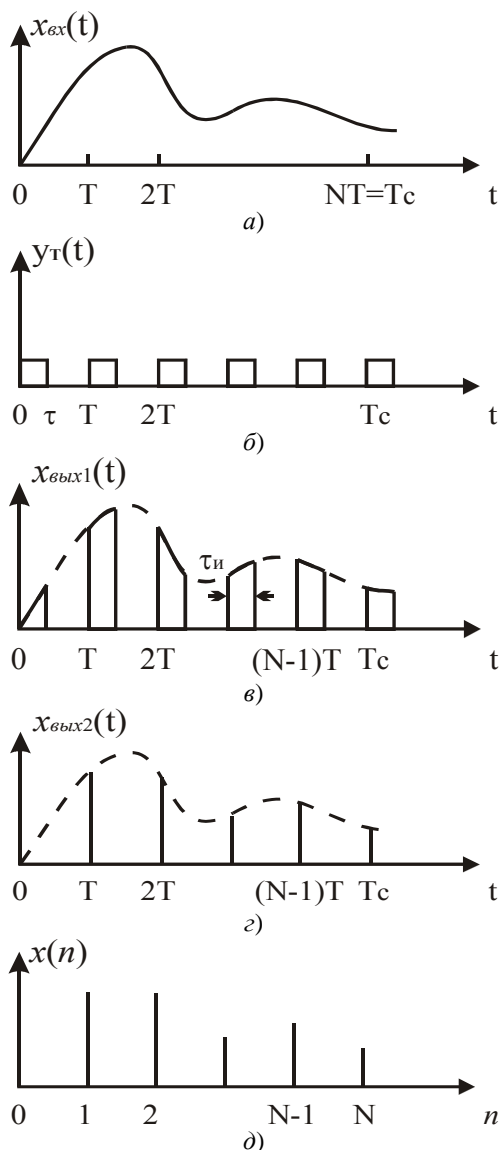


Рис. 2.2. Преобразование непрерывного сигнала в дискретный:
 а) непрерывный сигнал; б) дискретизирующая последовательность;
 в) сигнал на выходе ключа; г) абстрактный сигнал; д) выходная последовательность

В (2.2) и (2.3) $x_{\text{вых1}}(t)$, $x_{\text{вых2}}(t)$ – дискретизированные сигналы, определенные в любой точке интервала $[0, T_C]$. Однако, в отличие от (2.2), выражение (2.3) является математической моделью физически нереализуемого сигнала. Все известные физические процессы имеют протяженность во времени, поэтому импульс с нулевой длительностью является математической абстракцией. Абстрактен и сам сигнал $x_{\text{вых2}}(t)$, определяемый выражением (2.3) и условно изображенный на рис. 2.2г.

Замена непрерывного сигнала конечной числовой последовательностью получила распространение в теории и технике дискретных и цифровых цепей. При обработке числовых последовательностей интегрирование заменяется суммированием, а дифференцирование – вычислением конечных разностей. Математическая модель дискретного сигнала получается из выражения (2.3):

$$x(nT) = x_{\text{ex}}(t)u_0(t - nT). \quad (2.4)$$

В выражении (2.4) формально можно перейти к безразмерному времени n , имеющему смысл номера выборки или порядкового номера числа в последовательности:

$$x(n) = x_{\text{ex}}(t)u_0(t - nT). \quad (2.5)$$

Последовательность $x(n)$ изображена на рис. 2.2д.

Процесс преобразования непрерывного сигнала в дискретный называется дискретизацией, а устройство, его реализующее – дискретизатором. Понятно, что дискретизатор, выполняющий преобразования (2.4), (2.5), физически не реализуем, поскольку на его выходе может присутствовать только дискретизированный сигнал с конечной длительностью импульсов. Однако цифровой сигнал, получаемый на выходе АЦП, реально существует.

2.2. Спектр дискретизированного сигнала

Запишем периодическую последовательность дискретизирующих импульсов $Y_T(t)$ в виде ряда Фурье:

$$Y_T(t) = \frac{\tau}{T} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\omega_1\tau}{2}\right) \cos(n\omega_1 t + \phi_n),$$

где $\omega_1=2\pi/T$, ϕ_n характеризуют задержку последовательности импульсов $Y_T(t)$ на $0,5\tau$.

Учитывая, что $n\omega_1\tau/2=n\pi\tau/T$, а также имея в виду равенство $(1/n\pi)\sin(n\pi\tau/T)=(\tau/T)\text{sinc}(n\pi\tau/T)$, получим

$$Y_T(t) = \frac{\tau}{T} \left[1 + 2 \sum_{n=1}^{\infty} \text{sinc}\left(\frac{n\pi\tau}{T}\right) \cos(n\omega_1 t + \phi_n) \right].$$

Тогда выражение (2.1) примет вид

$$x_{\text{былх}}(t) = \frac{\tau}{T} \left[x_{\text{гх}}(t) + 2x_{\text{гх}}(t) \sum_{n=1}^{\infty} \text{sinc}\left(\frac{n\pi\tau}{T}\right) \cos(n\omega_1 t + \phi_n) \right].$$

Первому слагаемому в правой части соответствует спектр $X_{\text{БХ}}(\omega)$ исходного непрерывного сигнала, а каждому из произведений $x_{\text{гх}}(t)\cos(n\omega_1 t + \phi_n)$ – спектр $\frac{1}{2}[X_{\text{БХ}}(\omega - n\omega_1) + X_{\text{БХ}}(\omega + n\omega_1)]$. Следовательно, искомый спектр определяется выражением

$$X_{\text{былх}}(\omega) = \frac{\tau}{T} \left[X_{\text{гх}}(\omega) + \sum_{n=1}^{\infty} \text{sinc}\left(\frac{n\pi\tau}{T}\right) X_{\text{гх}}(\omega - n\omega_1) - \sum_{n=1}^{\infty} \text{sinc}\left(\frac{n\pi\tau}{T}\right) X_{\text{гх}}(\omega + n\omega_1) \right].$$

Поскольку $\text{sinc}(0)=1$, то окончательно получим

$$X_{\text{былх}}(\omega) = \frac{\tau}{T} \sum_{n=-\infty}^{\infty} \text{sinc}\left(\frac{n\pi\tau}{T}\right) X_{\text{гх}}(\omega - n\omega_1). \quad (2.6)$$

Графики модулей функций $X_{\text{гх}}(\omega)$ и $X_{\text{былх}}(\omega)$ приведены на рис. 2.3а и б, в соответственно.

Итак, спектр $X_{\text{былх}}(\omega)$ дискретизированного сигнала, полученного из непрерывного, представляет собой последовательность спектров $X_{\text{гх}}(\omega)$ исходного сигнала $x_{\text{гх}}(t)$,

сдвинутых один относительно другого на $\omega_1=2\pi/T$ и убывающих по закону $\text{sinc}(n\pi\tau/T)$.

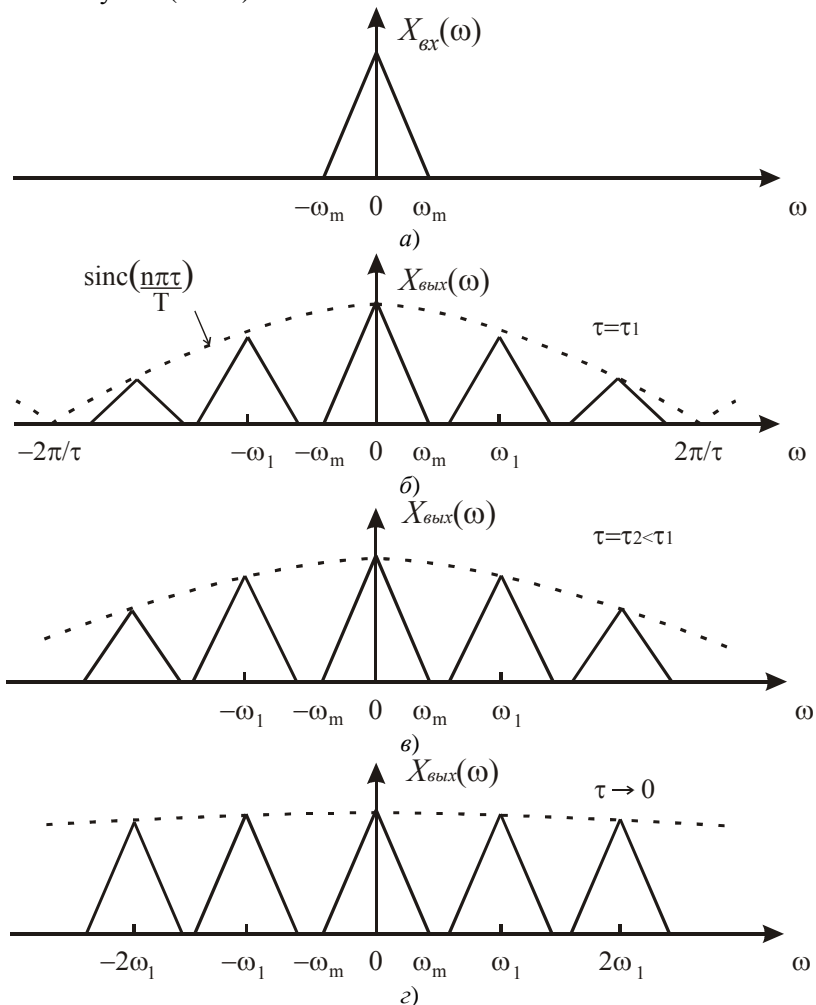


Рис. 2.3. Связь между спектрами непрерывного и дискретизированных сигналов:

- а) модуль спектра непрерывного сигнала;
- б) модуль спектра дискретизированного сигнала при $\tau = \tau_1$;
- в) модуль спектра дискретизированного сигнала при $\tau = \tau_2$;
- з) модуль спектра дискретизированного сигнала при $\tau \rightarrow 0$

Множитель τ/T в (2.6) отражает уменьшение модуля спектра $X_{\text{выл}}(\omega)$ сигнала $x_{\text{выл}}(t)$ по сравнению с модулем спектра $X_{\text{вх}}(\omega)$ сигнала $x_{\text{вх}}(t)$, поскольку энергия сигнала $x_{\text{выл}}(t)$ меньше энергии сигнала $x_{\text{вх}}(t)$ в τ/T раз. Поэтому изображенные на рис. 2.3 модули спектров можно рассматривать как нормированные.

С уменьшением отношения τ/T лепестки спектра убывают медленнее, и в пределе при $\tau \rightarrow 0$ спектр приобретет строго периодическую структуру с нулевыми уровнями лепестков, что соответствует спектру дискретного сигнала.

На рис. 2.3б-г приведены спектры сигналов для трех последовательно уменьшающихся значений τ .

2.3. Соотношения между спектрами непрерывных и дискретных сигналов

Рассмотрим математическую модель дискретного сигнала $x(nT)$. Установим связь между спектром $X(e^{j\omega T})$ последовательности $x(nT)$ и преобразованием Фурье $X_n(j\Omega)$ непрерывного сигнала $x(t)$ и обсудим следствия, вытекающие из нее.

Пара преобразований Фурье для непрерывного сигнала $x(t)$ имеет вид

$$X_n(j\Omega) = \int_{-\infty}^{\infty} x(t) e^{-j\Omega t} dt,$$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_n(j\Omega) e^{j\Omega t} d\Omega. \quad (2.7)$$

Аналогичные соотношения для дискретного сигнала имеют вид

$$X(e^{j\omega T}) = \sum_{n=-\infty}^{\infty} x(nT) e^{-j\omega n T},$$

$$x(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X(e^{j\omega T}) e^{j\omega nT} d\omega. \quad (2.8)$$

Поскольку $x(nT) = x(t)|_{t=nT}$, то можно связать $X_n(j\Omega)$ и $X(e^{j\omega T})$, вычислив интеграл (2.7) для $t=nT$, причем интеграл с бесконечными пределами следует заменить бесконечной суммой интегралов на интервалах длиной $2\pi/T$.

Таким образом

$$x(nT) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{(2m-1)\pi/T}^{(2m+1)\pi/T} X_n(j\Omega) e^{j\Omega nT} d\Omega.$$

Заменяв Ω на $\omega + \frac{2\pi}{T}m$ и изменив порядок действий, получим

$$x(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \left[\frac{1}{T} \sum_{m=-\infty}^{\infty} X_n(j\omega + j\frac{2\pi}{T}m) \right] e^{j\omega nT} d\omega.$$

Сравнив это выражение с (2.8), получим искомое соотношение:

$$X(e^{j\omega T}) = \frac{1}{T} \sum_{m=-\infty}^{\infty} X_n(j\omega + j\frac{2\pi}{T}m). \quad (2.9)$$

Из этой формулы видно, что периодическая спектральная функция последовательности состоит из суммы бесконечного числа спектральных компонентов непрерывного сигнала.

Если спектр непрерывного сигнала ограничен по полосе диапазоном частот $|\Omega| \leq \pi/T$, т. е. $X_n(j\Omega) = 0$ при $|\Omega| > \pi/T$, из соотношения (2.9) следует, что в диапазоне частот $|\omega| \leq \pi/T$ имеем

$$X(e^{j\omega T}) = \frac{1}{T} X_n(j\Omega). \quad (2.10)$$

В этом случае спектр последовательности связан со спектром непрерывного сигнала следующим образом (рис. 2.4).

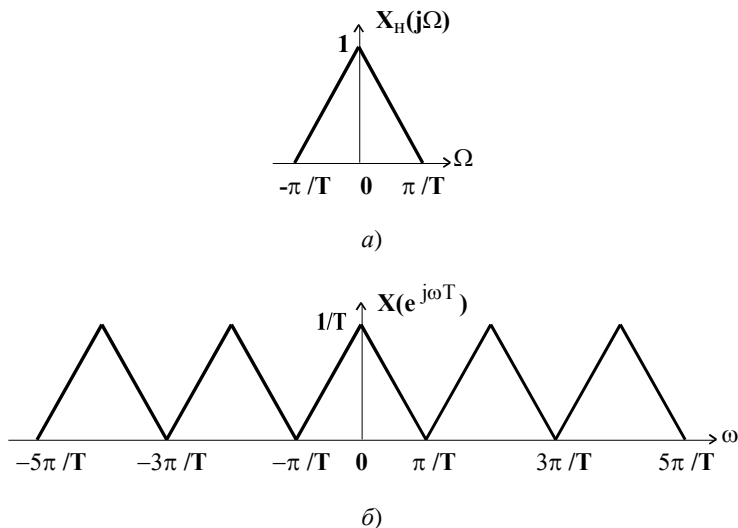


Рис. 2.4. Связь между спектрами непрерывного и дискретного сигналов: а) модуль спектра непрерывного сигнала; б) дискретного сигнала

Если же $X_n(j\Omega)$ не ограничен диапазоном $|\Omega| \leq \pi/T$, то соотношение между спектрами дискретного и непрерывного сигналов оказывается более сложным. Покажем типичный пример (рис. 2.5а–в). Спектр непрерывного сигнала (рис. 2.5а) ограничен полосой $|\Omega| \leq 3\pi/2T$.

Из формулы (2.9) следует, что члены с $m=0, \pm 1$ дают вклад в $X(e^{j\omega T})$ в диапазоне частот $|\omega| \leq \pi/T$ (рис. 2.5б). Поэтому в отличие от предыдущего примера спектр последовательности (рис. 2.5в) связан со спектром исходного сигнала значительно более сложным образом. Причина заключается в том, что частота дискретизации $1/T$ была недостаточно большой и высокочастотные составляющие спектра $X_n(j\Omega)$ попали в область более низких частот в спектре $X(e^{j\omega T})$. Такое смещение спектральных составляющих из одного диапазона в другой

называют наложением спектров, а последовательность, соответствующую спектру рис. 2.5в, – представлением сигнала $x(t)$ с наложением. Наложения можно избежать, дискретизируя непрерывные сигналы с достаточно высокой частотой.

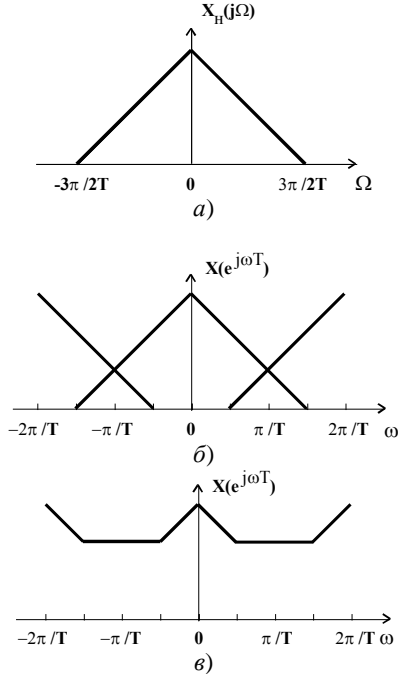


Рис. 2.5. Связь между спектрами непрерывного и дискретного сигналов при наложении спектров. Модули спектров: а) непрерывного сигнала; б) членов разложения с $m=0, \pm 1$; в) дискретного сигнала

2.4. Восстановление непрерывного сигнала из дискретного

Из рис. 2.4б видно, что если $\Omega_s \leq \pi/T$, т. е. если частота дискретизации, по крайней мере, вдвое больше наивысшей частоты спектра $X_n(j\Omega)$, то $X(e^{j\omega T})$ совпадает с $X_n(j\Omega)$ на интервале $-\pi/T \leq \omega \leq \pi/T$. Эта частота дискретизации иногда называется частотой Найквиста (Котельникова). В этом случае можно ожидать, что функция $x(t)$ может быть восстановлена по

выборкам $x(nT)$ при помощи подходящей интерполяционной формулы.

Чтобы вывести интерполяционную формулу, предположим, что $\Omega_c \leq \pi T$ (рис. 2.4б), тогда согласно (2.10) имеем

$$X(e^{j\Omega T}) = \frac{1}{T} X_H(j\Omega), \quad -\pi/T \leq \Omega \leq \pi/T.$$

В соответствии с преобразованием Фурье в непрерывном времени

$$x(t) = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} X_H(j\Omega) e^{j\Omega t} d\Omega.$$

Объединяя два последних выражения, можно записать:

$$x(t) = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} TX(e^{j\Omega T}) e^{j\Omega t} d\Omega.$$

Используя формулу дискретного ряда Фурье

$$X(e^{j\Omega T}) = \sum_{n=-\infty}^{\infty} x(nT) e^{-j\Omega nT},$$

запишем:

$$x(t) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \left(\sum_{n=-\infty}^{\infty} x(nT) e^{-j\Omega nT} \right) e^{j\Omega t} d\Omega$$

или, изменяя порядок суммирования и интегрирования,

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT) \left[\frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} e^{j\Omega(t-nT)} d\Omega \right].$$

Выполнив интегрирование, получим

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT) \left\{ \frac{\sin\left(\frac{\pi}{T}(t-nT)\right)}{\frac{\pi}{T}(t-nT)} \right\}. \quad (2.11)$$

Выражение (2.11) дает интерполяционную формулу для восстановления аналогового сигнала $x(t)$ по его выборкам. Заметим, что это выражение описывает свертку входного сигнала и ИХ фильтра нижних частот (ФНЧ) с прямоугольной амплитудно-частотной характеристикой (АЧХ). Представление аналогового сигнала в виде (2.11) справедливо только для функций с ограниченным спектром при достаточно малом T , т. е. при отсутствии эффекта наложения.

Выражение (2.11) можно понимать как разложение сигнала с непрерывным временем в ряд вида

$$x(t) = \sum_{n=-\infty}^{\infty} C_n \Phi_n(t), \quad (2.12)$$

где коэффициенты C_n и функции $\Phi_n(t)$ определяются выражениями

$$C_n = x(nT) \quad (2.13a)$$

и

$$\Phi_n(t) = \frac{\sin\left(\frac{\pi}{T}(t - nT)\right)}{\frac{\pi}{T}(t - nT)}. \quad (2.13b)$$

Существует много классов функций $\Phi_n(t)$, которые можно использовать для представления функции с непрерывным временем в виде (2.12), включая синусоидальные функции, функции Лагерра и полиномы Лежандра. При любом таком представлении последовательность коэффициентов C_n можно рассматривать как дискретный сигнал, представляющий аналоговый сигнал $x(t)$. Однако не все такие представления в равной степени полезны. Большим преимуществом выбора функции $\Phi_n(t)$ вида (2.13b) является то, что коэффициенты легко получаются путем дискретизации сигнала с непрерывным временем. Примеры дискретного и восстановленного сигналов приведены на рис. 2.6 а и б соответственно.

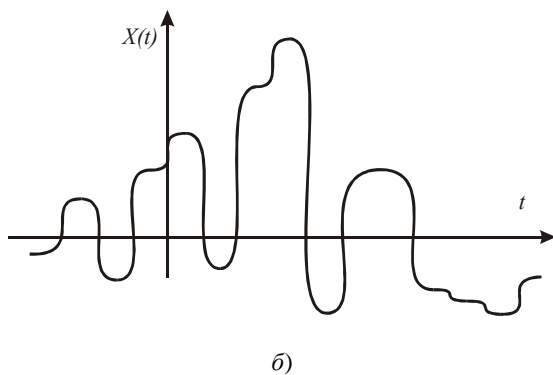
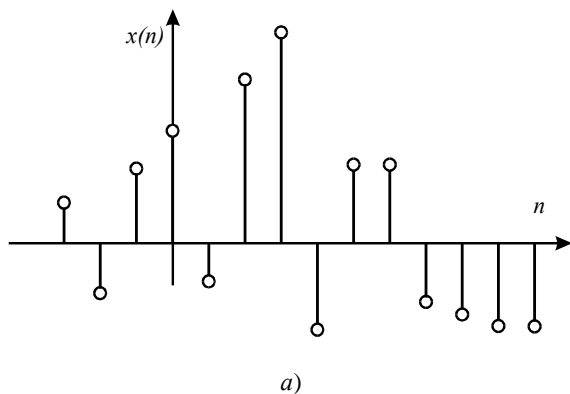
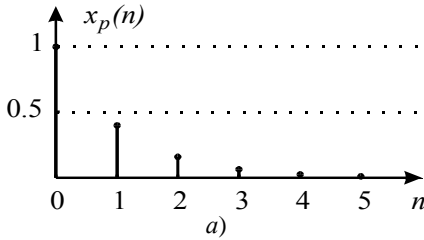


Рис. 2.6. Восстановление непрерывного сигнала:
a) исходный дискретный сигнал;
б) восстановленный непрерывный сигнал

2.5. Типовые периодические сигналы

Рассмотрим ряд часто встречающихся цифровых периодических сигналов вместе с их спектрами, получаемыми с помощью ДПФ.

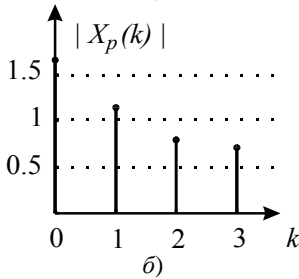
2.5.1. Последовательность показательных импульсов



Цифровой сигнал на одном периоде T выражается зависимостью

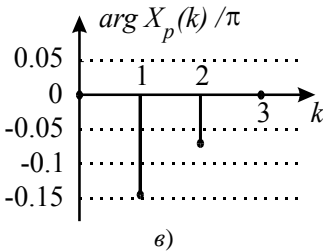
$$x_p(n) = a^n \text{ при } n \in [0; T-1].$$

График сигнала на этом интервале при $a=0.4$, $T=6$ показан на рис. 2.7а. Согласно формуле (1.18) спектральная функция равна



$$X_p(k) = \sum_{n=0}^{T-1} a^n e^{-j\frac{2\pi}{T}kn} =$$

$$= \sum_{n=0}^{T-1} \left[a e^{-j\frac{2\pi}{T}k} \right]^n = \frac{1 - a^T}{1 - a e^{-j\frac{2\pi}{T}k}}$$



$$, \quad k \in [0; T-1].$$

Графики модуля и аргумента этой функции для сигнала с указанными параметрами на интервале $k \in [0; T/2]$ показаны на рис. 2.7б и 2.7в соответственно.

Рис. 2.7. Последовательность показательных импульсов: а) график сигнала на одном периоде; б) амплитудный спектр; в) фазовый спектр

2.5.2. Косинусоидальное колебание

Пусть $x_p(n) = \cos \frac{2\pi}{T} n$. График этой функции на одном периоде при $T=6$ показан на рис. 2.8а. Для определения спектра представим

$$\cos \frac{2\pi}{T} n = \frac{1}{2} \left(e^{j\frac{2\pi}{T} n} + e^{-j\frac{2\pi}{T} n} \right).$$

Найдем спектральную функцию:

$$\begin{aligned} X_p(k) &= \frac{1}{2} \sum_{n=0}^{T-1} \left(e^{j\frac{2\pi}{T} n} + e^{-j\frac{2\pi}{T} n} \right) e^{-j\frac{2\pi}{T} kn} = \\ &= \frac{1}{2} \sum_{n=0}^{T-1} e^{-j\frac{2\pi}{T} n(k-1)} + \frac{1}{2} \sum_{n=0}^{T-1} e^{-j\frac{2\pi}{T} n(k+1)}. \end{aligned} \quad (2.14)$$

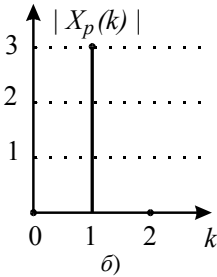
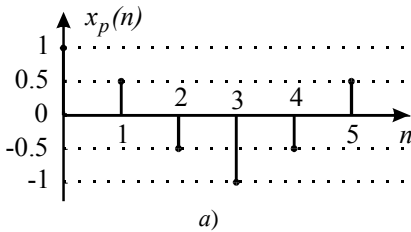


Рис. 2.8. Косинусоидальное колебание:
а) график сигнала на одном периоде;
б) амплитудный спектр

На основании формулы (1.17) имеем

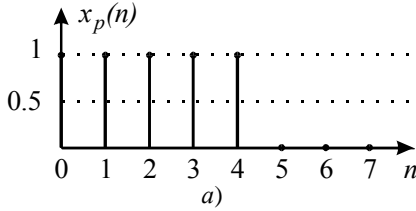
$$\sum_{n=0}^{T-1} e^{j\frac{2\pi}{T} n(k-m)} = \begin{cases} T & \text{при } k = m \\ 0 & \text{при } k \neq m. \end{cases}$$

Первая сумма в (2.14) равна T при $k=1$, а вторая сумма равна T при $k=T-1$. Следовательно, спектральная функция вещественна и равна:

$$X_p(k) = \frac{T}{2} [\delta(k-1) + \delta(k-T+1)]$$

График этой функции на интервале $k \in [0; T/2]$ при $T=6$ показан на рис. 2.8б.

2.5.3. Последовательность прямоугольных импульсов



Цифровой сигнал на одном периоде T запишем в виде функции

$$x_p(n) = \begin{cases} 1 & \text{при } n \in [0; n_u - 1], \\ 0 & \text{при } n \in [n_u; T - 1]. \end{cases}$$

График этой функции на одном периоде при $n_u=5$, $T=8$ показан на рис. 2.9а.

Найдем спектральную функцию

$$X_p(k) = \sum_{n=0}^{n_u-1} e^{-j\frac{2\pi}{T}kn}.$$

При $k=0$ имеем

$$X_p(0) = \sum_{n=0}^{n_u-1} x_p(n) = n_u.$$

При $k \neq 0$ получим

$$\begin{aligned} X_p(k) &= \frac{1 - e^{-j\frac{2\pi}{T}kn_u}}{1 - e^{-j\frac{2\pi}{T}k}} = \\ &= \frac{e^{-j\frac{2\pi}{T}k\frac{n_u}{2}}}{e^{-j\frac{2\pi}{T}k\frac{n_u}{2}}} \frac{e^{j\frac{2\pi}{T}k\frac{n_u}{2}} - e^{-j\frac{2\pi}{T}k\frac{n_u}{2}}}{e^{j\frac{2\pi}{T}k\frac{n_u}{2}} - e^{-j\frac{2\pi}{T}k\frac{n_u}{2}}} = \\ &= e^{-j\pi k\frac{n_u-1}{T}} \frac{\sin \frac{\pi kn_u}{T}}{\sin \frac{\pi k}{T}}. \end{aligned} \quad (2.15)$$

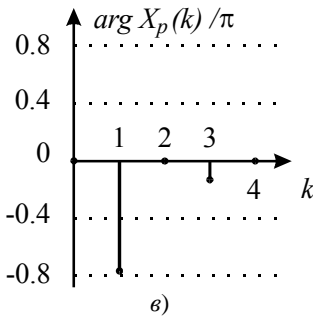
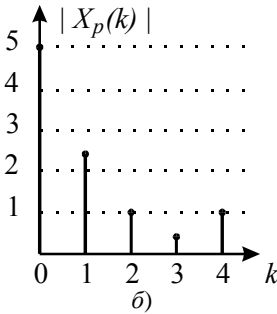


Рис. 2.9. Последовательность прямоугольных импульсов:
а) график сигнала на одном периоде;
б) амплитудный спектр;
в) фазовый спектр

Первый переход через нуль огибающей спектра происходит при $k=T/n_u$. Графики модуля и аргумента спектральной функции

на интервале $k \in [0; T/2]$ сигнала с указанными параметрами показаны на рис. 2.9б и 2.9в соответственно.

2.5.4. Последовательность прямоугольных радиоимпульсов

Цифровой сигнал на одном периоде T запишем в виде функции

$$x_p(n) = \begin{cases} \cos\left(\frac{2\pi}{n_s}n + \phi_0\right) & \text{при } n \in [0; n_u - 1], \\ 0 & \text{при } n \in [n_u; T - 1]. \end{cases}$$

Здесь n_s – период заполнения, n_u – длительность импульса. Осциллограмма сигнала на одном периоде при $n_s=6$, $n_u=13$, $T=24$, $\phi_0=0$ показана на рис. 2.10а.

Найдем спектральную функцию:

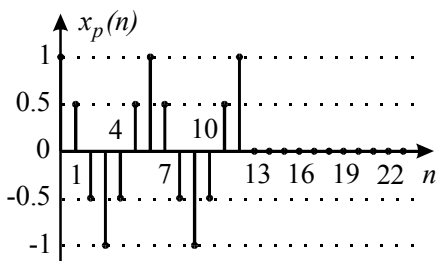
$$\begin{aligned} X_p(k) &= \sum_{n=0}^{n_u-1} \cos\left(\frac{2\pi}{n_s}n + \phi_0\right) e^{-j\frac{2\pi}{T}kn} = \\ &= \frac{1}{2} \sum_{n=0}^{n_u-1} \left[e^{j\left[\frac{2\pi}{n_s}n + \phi_0\right]} + e^{-j\left[\frac{2\pi}{n_s}n + \phi_0\right]} \right] e^{-j\frac{2\pi}{T}kn} = \\ &= X_{p1}(k) + X_{p2}(k). \end{aligned}$$

Получим выражение для первой из этих функций:

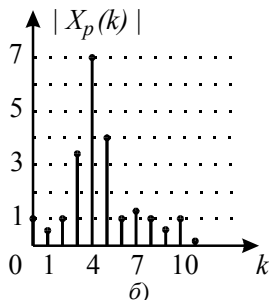
$$X_{p1}(k) = \frac{1}{2} \sum_{n=0}^{n_u-1} e^{j\left[\frac{2\pi}{n_s}n + \phi_0\right]} e^{-j\frac{2\pi}{T}kn} = \frac{1}{2} e^{j\phi_0} \frac{1 - e^{j\left(\frac{2\pi}{n_s} - \frac{2\pi}{T}k\right)n_u}}{1 - e^{j\left(\frac{2\pi}{n_s} - \frac{2\pi}{T}k\right)}}.$$

Введем $\alpha_k = \pi\left(\frac{1}{n_s} - \frac{k}{T}\right)$, тогда

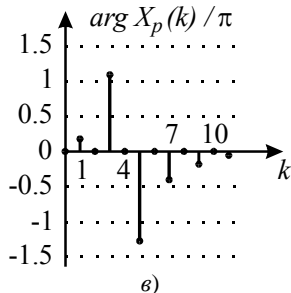
$$X_{p1}(k) = \frac{1}{2} e^{j\phi_0} \frac{e^{j\alpha_k n_u}}{e^{j\alpha_k}} \frac{\sin(\alpha_k n_u)}{\sin \alpha_k} = \frac{1}{2} e^{j\phi_0} e^{j\alpha_k (n_u-1)} \frac{\sin(\alpha_k n_u)}{\sin \alpha_k}.$$



а)



б)



в)

Рис. 2.10. Последовательность прямоугольных радиоимпульсов:
а) график сигнала на одном периоде;
б) амплитудный спектр;
в) фазовый спектр

Аналогичные преобразования проделаем и для функции $X_{p2}(k)$:

$$\begin{aligned}
 X_{p2}(k) &= \\
 &= \frac{1}{2} \sum_{n=0}^{n_u-1} e^{-j\left[\frac{2\pi}{n_s}n+\phi_0\right]} e^{-j\frac{2\pi}{T}kn} = \\
 &= \frac{1}{2} e^{-j\phi_0} \sum_{n=0}^{n_u-1} e^{-j\left(\frac{2\pi}{n_s} + \frac{2\pi}{T}k\right)n} = \\
 &= \frac{1}{2} e^{-j\phi_0} \frac{1 - e^{-j\left(\frac{2\pi}{n_s} + \frac{2\pi}{T}k\right)n_u}}{1 - e^{-j\left(\frac{2\pi}{n_s} + \frac{2\pi}{T}k\right)}}.
 \end{aligned}$$

Введем $\beta_k = \pi \left(\frac{1}{n_s} + \frac{k}{T} \right)$,

тогда

$$\begin{aligned}
 X_{p2}(k) &= \\
 &= \frac{1}{2} e^{-j\phi_0} \frac{e^{-j\beta_k n_u}}{e^{-j\beta_k}} \frac{\sin(\beta_k n_u)}{\sin \beta_k} = \\
 &= \frac{1}{2} e^{-j\phi_0} e^{-j\beta_k (n_u-1)} \frac{\sin(\beta_k n_u)}{\sin \beta_k}.
 \end{aligned}$$

Результирующий спектр равен

$$X_p(k) = \frac{1}{2} \left[e^{j(\alpha_k (n_u-1) + \phi_0)} \frac{\sin(\alpha_k n_u)}{\sin \alpha_k} + e^{-j(\beta_k (n_u-1) + \phi_0)} \frac{\sin(\beta_k n_u)}{\sin \beta_k} \right]. \quad (2.16)$$

Графики амплитудного и фазового спектров на интервале $k \in [0; T/2]$ сигнала с указанными параметрами показаны на рис. 2.10б и 2.10в соответственно. Обратите внимание на отсутствие симметрии спектра относительно несущей частоты (в расчетном случае это $k=T/n_s=4$). Симметрия относительно $k=T/2$ сохраняется, так как сигнал вещественный.

2.6. Типовые непериодические сигналы

Рассмотрим основные цифровые сигналы конечной длительности, а также единичный скачок вместе с их спектрами.

2.6.1. Показательный импульс

Задана последовательность конечной длины

$$x(n) = \begin{cases} a^n & \text{при } n \in [0; N-1], \\ 0 & \text{при других } n, \end{cases}$$

т. е. она состоит из одного периода соответствующей последовательности $x_p(n)$. График показательного импульса для $a=0.4$, $N=6$ показан на рис. 2.11а.

Значение преобразования Фурье в L точках ($L > N$), равномерно распределенных по единичной окружности на z -плоскости, находится по формуле (1.21), т. е.

$$\hat{X}(k) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{L}kn} = \sum_{n=0}^{N-1} \left(ae^{-j\frac{2\pi}{L}k} \right)^n = \frac{1 - \left(ae^{-j\frac{2\pi}{L}k} \right)^N}{1 - ae^{-j\frac{2\pi}{L}k}},$$

$$k \in [0; L-1].$$

Амплитудный и фазовый спектры импульса с указанными параметрами на интервале $k \in [0; L/2]$ для $L=12$ приведены на рис. 2.11б и 2.11в соответственно.

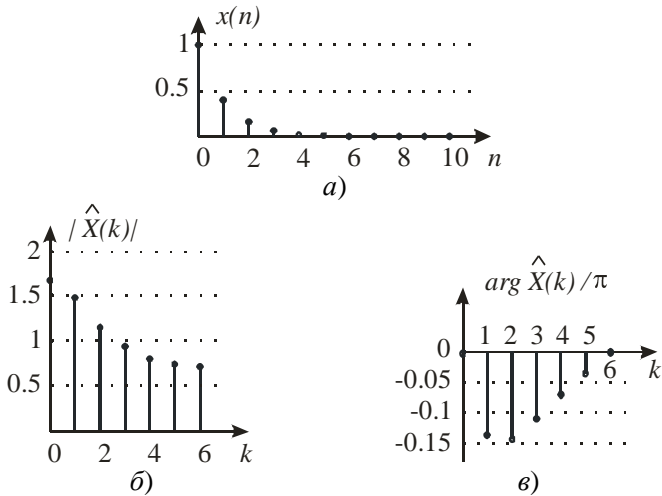


Рис. 2.11. Показательный импульс:
 а) график сигнала; б) амплитудный спектр; в) фазовый спектр

2.6.2. Прямоугольный импульс

Одиночный прямоугольный импульс длительностью n_u выражается зависимостью

$$x(n) = \begin{cases} 1 & \text{при } n \in [0; n_u - 1], \\ 0 & \text{при других } n. \end{cases}$$

Осциллограмма импульса при $n_u = 5$ показана на рис. 2.12а.

Спектральная функция в L точках единичной окружности на z -плоскости находится по формуле (1.21), т. е.

$$\hat{X}(k) = \sum_{n=0}^{n_u-1} e^{-j\frac{2\pi}{L}kn}.$$

По аналогии с выражением (2.15) при $k \neq 0$ запишем

$$\hat{X}(k) = e^{-j\pi k \frac{n_u-1}{L}} \frac{\sin[\pi k n_u / L]}{\sin(\pi k / L)}, \quad k \in [1; L-1].$$

При $k=0$ имеем $\hat{X}(k) = n_u$.

Первый переход через нуль огибающей спектра происходит при $k=L/n_u$. Изменяя величину L , при расчете спектра можно достичь любого разрешения. Графики модуля и аргумента спектральной функции импульсадлительностью $n_u=5$ на интервале $k \in [0; L/2]$ при $L=12$ представлены на рис. 2.12б и 2.12в соответственно.

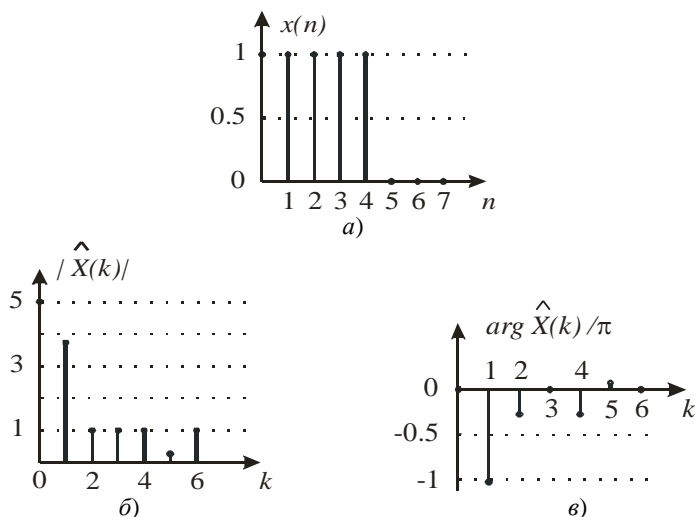


Рис. 2.12. Прямоугольный импульс:
 а) график сигнала; б) амплитудный спектр; в) фазовый спектр

2.6.3. Прямоугольный радиоимпульс

Аналитическое выражение этого цифрового сигнала длительностью n_u имеет вид

$$x(n) = \begin{cases} \cos\left(\frac{2\pi}{n_s}n + \phi_0\right) & \text{при } n \in [0; n_u - 1], \\ 0 & \text{при других } n. \end{cases}$$

Здесь n_s – период заполнения. Осциллограмма сигнала при $n_s=6$, $n_u=13$ изображена на рис. 2.13а.

С помощью формулы (1.21) найдем значения преобразования Фурье в L точках единичной окружности:

$$\hat{X}(k) = \sum_{n=0}^{n_u-1} \cos\left(\frac{2\pi}{n_s}n + \phi_0\right) e^{-j\frac{2\pi}{L}kn}.$$

Выполнив преобразования, как и в п. 2.5.4, и обозначив

$$\hat{\alpha}_k = \pi\left(\frac{1}{n_s} - \frac{k}{L}\right),$$

$$\hat{\beta}_k = \pi\left(\frac{1}{n_s} + \frac{k}{L}\right),$$

получим по аналогии с (2.3)

$$\hat{X}(k) = \frac{1}{2} \left[e^{j(\hat{\alpha}_k(n_u-1)+\phi_0)} \frac{\sin(\hat{\alpha}_k n_u)}{\sin \hat{\alpha}_k} + e^{-j(\hat{\beta}_k(n_u-1)+\phi_0)} \frac{\sin(\hat{\beta}_k n_u)}{\sin \hat{\beta}_k} \right].$$

Несущей частоте соответствуют $k \in \{L/n_s; L-L/n_s\}$. На рис. 2.13б и 2.13в показаны соответственно графики амплитудного и фазового спектров сигнала с указанными параметрами на интервале $k \in [0; L/2]$ при $L=36$.

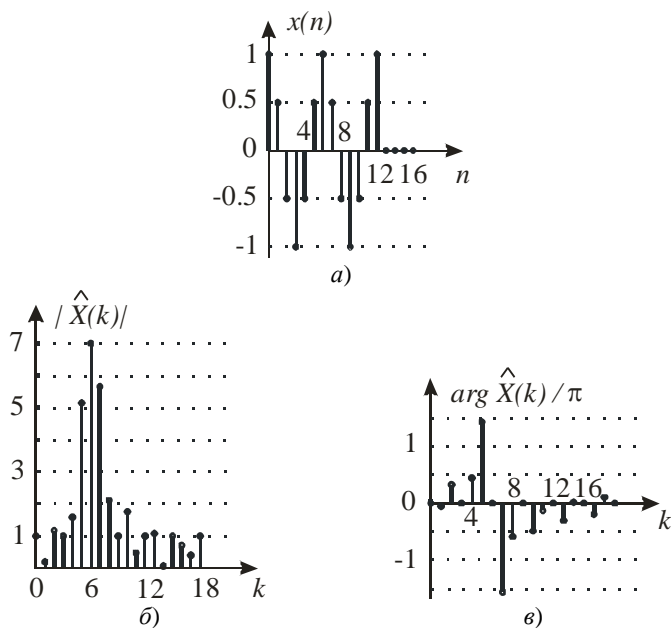


Рис. 2.13. Прямоугольный радиоимпульс:
 а) график сигнала; б) амплитудный спектр; в) фазовый спектр

2.6.4. Цифровой единичный импульс

Цифровой единичный импульс (или единичный отсчет) задается следующим образом:

$$\delta(n) = \begin{cases} 1 & \text{при } n = 0, \\ 0 & \text{при } n \neq 0. \end{cases}$$

В цифровых системах этот импульс играет такую же роль, как аналоговый единичный импульс (или дельта-функция Дирака) $\delta(t)$ в аналоговых системах. Важное различие между ними состоит в том, что первый реализуется в виде конечного числа, тогда как второй рассматривается только как обобщенная функция.

Осциллограмма цифрового единичного импульса представлена на рис. 2.14а.

Спектральная функция этого сигнала находится по формуле

$$\hat{X}(k) = \sum_{n=0}^{L-1} x(n)e^{-j\frac{2\pi}{L}kn} = \sum_{n=0}^{L-1} \delta(n)e^{-j\frac{2\pi}{L}kn} = 1(k), \quad k \in [0; L-1].$$

График этой функции на интервале $k \in [0; L/2]$ при $L=5$ изображен на рис. 2.14б.

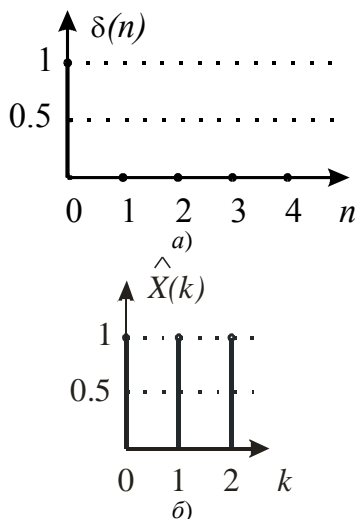


Рис. 2.14. Цифровой единичный импульс: а) график сигнала; б) амплитудный спектр

2.6.5. Единичный скачок

Это непериодический сигнал, выражающийся зависимостью

$$x(n) = 1(n) = \begin{cases} 1 & \text{при } n \geq 0, \\ 0 & \text{при } n < 0. \end{cases}$$

Осциллограмма единичного скачка представлена на рис. 2.15а.

Спектральную функцию этого сигнала будем находить по формуле конечной последовательности:

$$\hat{X}(k) = \sum_{n=0}^{L-1} e^{-j\frac{2\pi}{L}kn}$$

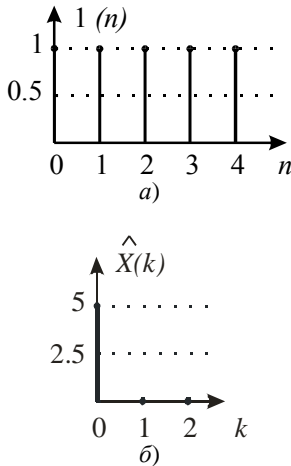


Рис. 2.15. Цифровой единичный импульс: а) осциллограмма; б) амплитудный спектр

На основании соотношений (1.17) получим

$$X(k) = L\delta(k).$$

График этой функции на интервале $k \in [0; L/2]$ при $L=5$ показан на рис. 2.15б. Следовательно, в отличие от спектра непрерывного единичного скачка ДПФ цифрового единичного скачка содержит только нулевую (постоянную) составляющую, т. е. этот сигнал аппаратом ДПФ «воспринимается» как постоянный ток. Последнее обстоятельство обусловлено тем, что ДПФ изначально предназначено для анализа

спектров периодических последовательностей (сигналов), т. е. полагается, что исследуемый сигнал имеет период N (или L).

2.7. Контрольные вопросы

1. Дайте определение непрерывного, дискретизированного, дискретного и цифрового сигналов. Приведите примеры.
2. В чем заключается процедура дискретизации аналоговых сигналов во временной и частотной областях?
3. Что означает выражение «спектр сигнала локализован в некоторой области»? Чем определяется верхняя частота спектра? Что такое ширина спектра? Приведите примеры.
4. Какие непрерывные сигналы имеют сплошной, а какие – дискретный спектры? Объясните почему.

5. Как связан спектр дискретизированного сигнала со спектрами исходного непрерывного сигнала и дискретизирующей последовательности? Приведите примеры.
6. Чем определяется закон огибающей спектра дискретизированного сигнала?
7. Как спектр дискретизированного сигнала зависит от частоты дискретизации и длительности дискретизирующих импульсов? Приведите примеры.
8. Как спектр дискретизированного сигнала зависит от параметров исходного континуального сигнала (периодический, непериодический)? Приведите примеры.
9. Какие ограничения налагают на дискретизируемый сигнал?
10. Как изменится спектр дискретизированного сигнала, если в качестве последовательности дискретизирующих импульсов использовать импульсы формы, отличной от прямоугольной? Имеет ли это смысл?
11. Сформулируйте теорему Котельникова (Найквиста) в частотной и временной областях.
12. Чему равно минимально возможное число дискретизирующих импульсов на периоде исходного континуального сигнала, чтобы можно было его восстановить?
13. Получите выражение для спектра дискретизированного сигнала. Приведите примеры.
14. Чем определяется точность восстановления непрерывного сигнала?
15. Запишите формулу ОДПФ, выведите формулу прямого ДПФ.
16. Покажите связь между ДПФ и z-преобразованием.
17. Как найти ДПФ последовательности конечной длины?
18. Поясните основные свойства ДПФ.
19. Выведите формулы спектров периодических дискретных сигналов:
 - последовательность экспоненциальных импульсов;
 - косинусоидальное колебание;
 - последовательность прямоугольных импульсов;
 - меандр;
 - сигнал пилообразной формы;
 - сигнал треугольной формы;

- последовательность прямоугольных радиоимпульсов.
20. Выведите формулы спектров непериодических дискретных сигналов:
- цифровой единичный импульс;
 - экспоненциальный импульс;
 - прямоугольный импульс;
 - прямоугольный радиоимпульс;
 - единичный скачок.

2.8. Литература

1. Брюханов, Ю. А. Цифровые цепи и сигналы: учеб. пособие / Ю. А. Брюханов. – Ярославль: ЯрГУ, 2005.
2. Солонина, А. И. Основы цифровой обработки сигналов: курс лекций / А. И. Солонина и др. – СПб.: БХВ-Петербург, 2005.
3. Ланнэ, А. А. Основы цифровой обработки сигналов. Ч. 1: учеб. пособие / А. А. Ланнэ, Б. Д. Матюшкин, Д. А. Улахович. – СПб.: Военная академия связи, 1995.
4. Рабинер, Л. Теория и применение цифровой обработки сигналов: пер. с англ. / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978.
5. Оппенгейм, А. В. Цифровая обработка сигналов: пер. с англ. / А. В. Оппенгейм, Р. В. Шафер. – М.: Связь, 1979.
6. Каппелини, В. Цифровые фильтры и их применение: пер. с англ. / В. Каппелини, А. Дж. Константинович, П. Эмилиани. – М.: Энергоатомиздат, 1983.
7. Баскаков, С. И. Радиотехнические цепи и сигналы / С. И. Баскаков. – М.: Высшая школа, 2005.
8. Гоноровский, И. С. Радиотехнические цепи и сигналы / И. С. Гоноровский, М. П. Демин. – М.: Радио и связь, 1994.
9. Карташев, В. Г. Основы теории дискретных сигналов и цифровых фильтров / В. Г. Карташев. – М.: Высшая школа, 1982.
10. Гольденберг, Л. М. Цифровая обработка сигналов: учеб. пособие / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1990.

3. ЦИФРОВЫЕ ЦЕПИ

Цифровыми называются цепи дискретного времени, предназначенные для обработки цифровых сигналов. Процессы в таких цепях описываются разностными уравнениями.

3.1. Цифровые линейные цепи с постоянными параметрами

Цифровая цепь по существу осуществляет преобразование одной последовательности (называемой входной) в другую (называемую выходной).

Простое представление цифровой цепи выглядит следующим образом (рис. 3.1)



Рис. 3.1. Общее представление цифровой цепи

Входная последовательность обозначена через $x(n)$, а выходная – через $y(n)$. Функционально они связаны соотношением

$$y(n) = F[x(n)],$$

где вид оператора $F(\cdot)$ зависит от свойств конкретной цепи.

Линейная цепь определяется следующим образом. Если $x_1(n)$ и $x_2(n)$ – некоторые входные сигналы (последовательности, воздействия), а $y_1(n)$ и $y_2(n)$ – соответствующие им выходные сигналы (последовательности, отклики, реакции) линейной цепи, то при подаче на вход сигнала $ax_1(n) + bx_2(n)$, на выходе образуется сигнал $ay_1(n) + by_2(n)$ (a и b – произвольные постоянные).

Цепь с постоянными параметрами (инвариантная к сдвигу) характеризуется тем, что если входной последовательности $x(n)$ соответствует выходная последовательность $y(n)$, то входной последовательности $x(n-n_0)$ при любом n_0 соответствует на выходе последовательность $y(n-n_0)$.

Покажем теперь, что в линейной цепи с постоянными параметрами входная и выходная последовательности связаны соотношением типа свертки. Допустим, что $x(n)$ – входная, а $y(n)$ – выходная последовательности и пусть $h(n)$ – отклик цепи на цифровой единичный импульс.

Последовательность $h(n)$ называют также импульсной характеристикой цепи. Линейную цепь, импульсная характеристика которой является последовательностью конечной длины, называют цепью с конечной импульсной характеристикой (КИХ), или, что то же самое, КИХ-цепью. Последовательности бесконечной длины соответствует цепь с бесконечной импульсной характеристикой (БИХ), или БИХ-цепь.

Функцию $x(n)$ можно записать в виде

$$x(n) = \sum_{m=-\infty}^{\infty} x(m)\delta(n-m).$$

Поскольку $h(n)$ является откликом цепи на последовательность $\delta(n)$, а параметры цепи постоянны, то $h(n-m)$ будет откликом на последовательность $\delta(n-m)$.

Из свойств линейности следует, что откликом на последовательность $x(m)\delta(n-m)$ должна быть последовательность $x(m)h(n-m)$. Поэтому отклик на $x(n)$ будет равен

$$y(n) = \sum_{m=-\infty}^{\infty} x(m)h(n-m). \quad (3.1a)$$

Он имеет вид свертки, что и требовалось доказать.

Простой заменой переменных равенство (3.1a) может быть преобразовано к виду

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m). \quad (3.1б)$$

Таким образом, ИХ $h(n)$ полностью описывает линейную цепь с постоянными параметрами, что отражается на рис. 3.2.

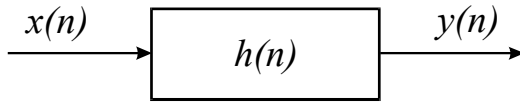


Рис. 3.2. Представление цифровой цепи с помощью импульсной характеристики

На рис. 3.3 показано, как процесс вычисления линейной свертки осуществляется на практике:

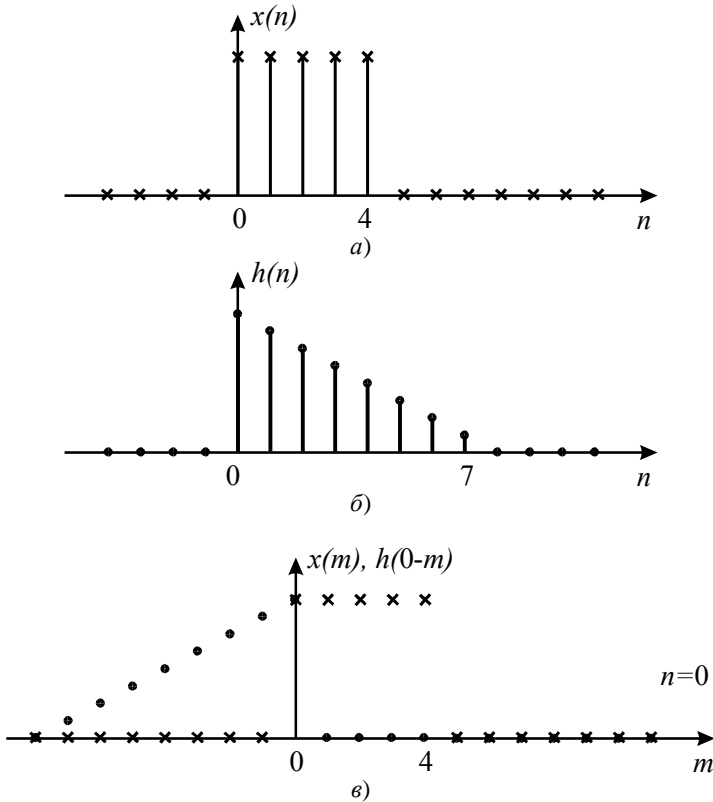


Рис. 3.3. Процесс вычисления линейной свертки:
 а) входная последовательность; б) ИХ цепи;
 в) вычисление свертки при $n=0$

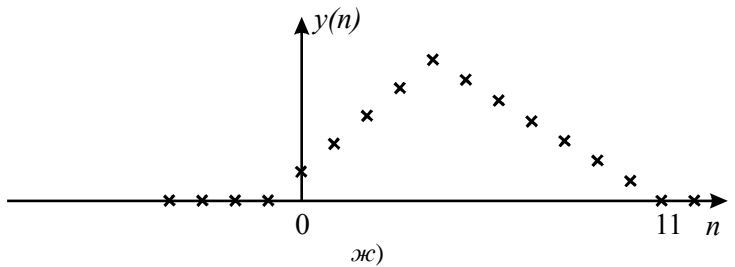
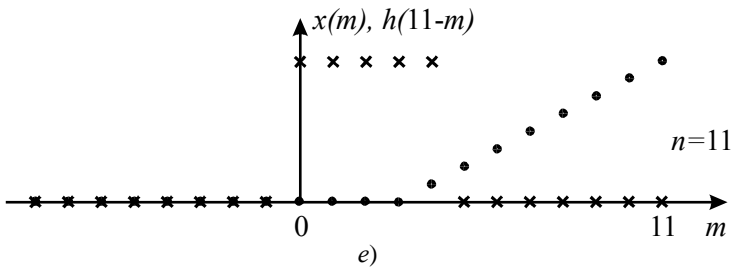
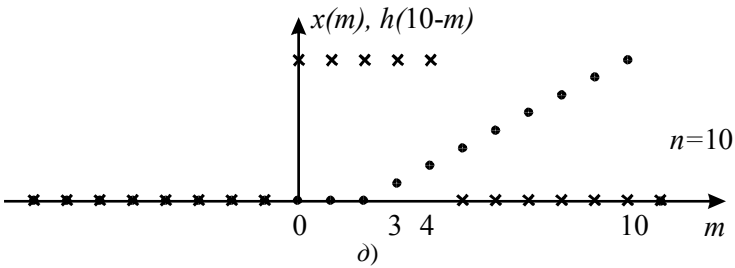
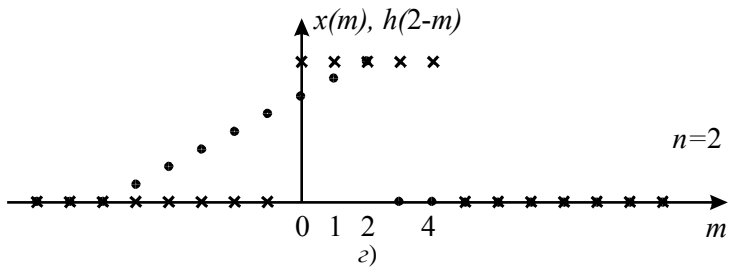


Рис. 3.3. Процесс вычисления линейной свертки (продолжение):
 з) вычисление свертки при $n=2$; д) вычисление свертки при $n=10$;
 е) вычисление свертки при $n=11$; ж) выходная последовательность

На рис. 3.3а изображена входная последовательность $x(n)$, отличная от нуля при $0 \leq n \leq 4$. На рис. 3.3б приведен пример ИХ $h(n)$, отличной от нуля при $0 \leq n \leq 7$. На рис. 3.3в – 3.3е представлены $x(m)$ и $h(n-m)$ для $n=0, 2, 10$ и 11 . Очевидно, что при $n < 0$ и $n > 11$ последовательности $x(m)$ и $h(n-m)$ не перекрываются и $y(n)$ равно нулю. На рис. 3.3ж приведена последовательность $y(n)$, являющаяся искомой сверткой.

3.2. Физическая реализуемость. Устойчивость

Линейную цепь с постоянными параметрами (ЛПП, линейная инвариантная к сдвигу (ЛИС), линейная стационарная (ЛС)) называют физически реализуемой, если величина отклика при $n=n_0$ зависит только от отсчетов входной последовательности с номерами $n \leq n_0$. Для линейной цепи это означает, что ИХ $h(n)$ равна нулю при $n < 0$.

Линейная цепь называется устойчивой, если при любой ограниченной входной последовательности выходная последовательность также ограничена. Необходимым и достаточным условием устойчивости цепи является следующее требование к ее ИХ:

$$\sum_{n=-\infty}^{\infty} |h(n)| < \infty. \quad (3.2)$$

Необходимость и достаточность условия (3.2) нетрудно доказать.

Предположим сначала, что условие (3.2) не удовлетворяется, т. е.

$$\sum_{n=-\infty}^{\infty} |h(n)| = \infty.$$

Рассмотрим ограниченную последовательность

$$x(n) = \begin{cases} +1 & \text{при } h(-n) \geq 0, \\ -1 & \text{при } h(-n) < 0. \end{cases}$$

Согласно формуле (3.1a) при $n=0$ отклик равен

$$y(0) = \sum_{m=-\infty}^{\infty} x(m)h(-m) = \sum_{m=-\infty}^{\infty} |h(-m)| = \sum_{m=-\infty}^{\infty} |h(m)| = \infty.$$

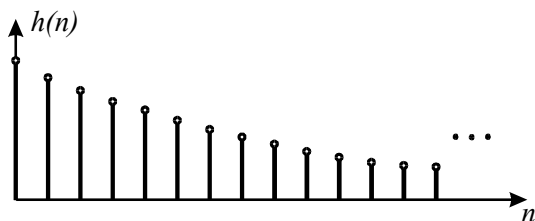
Таким образом, последовательность $y(0)$ не ограничена, так что неравенство (3.2) является необходимым условием устойчивости цепи.

Для доказательства достаточности предположим, что условие (3.2) выполняется, а на вход поступает ограниченная последовательность $x(n)$, т. е.

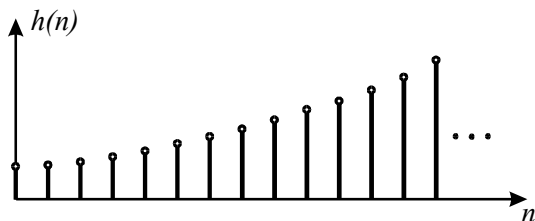
$$|x(n)| \leq M.$$

Из формулы (3.1a) получаем

$$\begin{aligned} y(n) &= \left| \sum_{m=-\infty}^{\infty} x(m)h(n-m) \right| \leq \sum_{m=-\infty}^{\infty} |x(m)||h(n-m)| \leq \\ &\leq M \sum_{m=-\infty}^{\infty} |h(n-m)| < \infty. \end{aligned}$$



a)



б)

Рис. 3.4. Импульсная характеристика:
a) устойчивой цепи; б) неустойчивой цепи

Таким образом, последовательность $y(n)$ ограничена, поэтому цепь устойчива, что и требовалось доказать.

Приведем примеры ИХ устойчивой и неустойчивой цепей. Импульсная характеристика, приведенная на рис. 3.4а, имеет вид $h(n)=a^n 1(n)$, причем $0 < a < 1$, поэтому условие (3.2) удовлетворяется и цепь устойчива. Выражение для ИХ на рис. 3.4б имеет тот же вид, но $a > 1$, поэтому условие (3.2) не выполняется и цепь неустойчива.

3.3. Частотная характеристика. Геометрическая интерпретация

Рассмотрим класс входных последовательностей

$$x(n) = e^{j\omega n}, \quad -\infty < n < \infty.$$

Если такая последовательность поступает на вход линейной цепи с ИХ $h(n)$, то на выходе (см. (3.1б) и (1.2)) появится последовательность

$$y(n) = \sum_{m=-\infty}^{\infty} h(m) e^{j\omega(n-m)} = e^{j\omega n} \sum_{m=-\infty}^{\infty} h(m) e^{-j\omega m} = x(n) H(e^{j\omega}).$$

Таким образом, для выбранного класса входных последовательностей отклик совпадает с входной последовательностью с точностью до комплексного множителя $H(e^{j\omega})$, который выражается через ИХ следующим образом:

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n) e^{-j\omega n}. \quad (3.3)$$

Поскольку последовательность вида $e^{j\omega n}$ функционально эквивалентна дискретизированной синусоиде с частотой ω , то множитель $H(e^{j\omega})$ называют частотной характеристикой (ЧХ) цепи, т. к. он представляет коэффициент передачи цепи для каждого значения круговой частоты ω . Модуль $H(e^{j\omega})$ называется амплитудно-частотной характеристикой, а аргумент – фазочастотной характеристикой цепи.

Вычислим в качестве примера ЧХ линейной цепи с ИХ $h(n) = a^n 1(n)$. Она имеет вид

$$H(e^{j\omega}) = \sum_{n=-0}^{\infty} a^n e^{-j\omega n} = \sum_{n=-0}^{\infty} (ae^{-j\omega})^n.$$

Так как $|a| < 1$, то сумма геометрической прогрессии будет равна

$$H(e^{j\omega}) = \frac{1}{1 - ae^{-j\omega}}.$$

Покажем графики ИХ $h(n)$, а также модуля и аргумента (фазы) ЧХ $H(e^{j\omega})$ как функции частоты ω в диапазоне $0 \leq \omega \leq 2\pi$ (рис. 3.5).

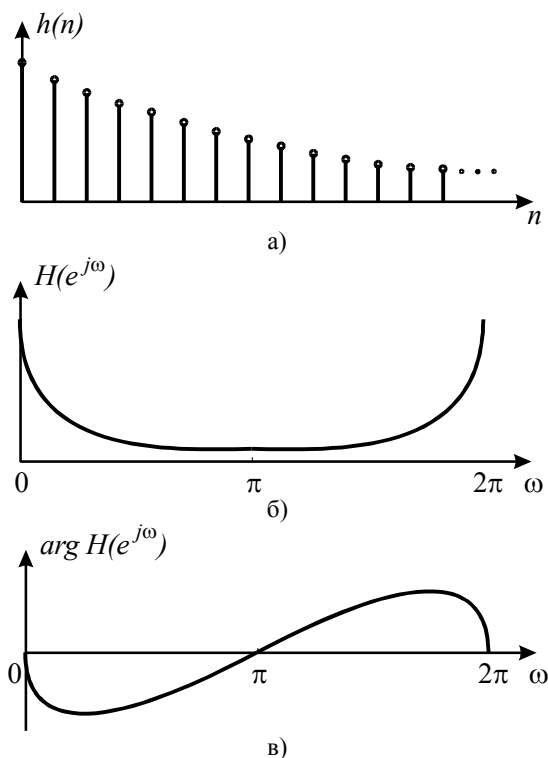


Рис. 3.5. Характеристики цепи:
а) импульсная; б) амплитудно-частотная; в) фазочастотная

Отметим некоторые свойства ЧХ. Она является периодической функцией ω , причем ее период равен 2π . Эта периодичность связана со спецификой дискретизированного колебания: входная последовательность с частотой $(\omega + 2m\pi)$ ($m = \pm 1, \pm 2, \dots$) не отличается от входной последовательности с частотой ω , т. е.

$$\hat{x}(n) = e^{j(\omega + 2m\pi)n} = e^{j\omega n} = x(n).$$

Поскольку $H(e^{j\omega})$ – периодическая функция, то для полного описания достаточно задать ее на любом интервале длиной 2π . Обычно для этой цели используют интервал $0 \leq \omega \leq 2\pi$.

Другим важным свойством ЧХ является то, что для действительных ИХ $h(n)$ (как обычно и бывает на практике) модуль $H(e^{j\omega})$ симметричен, а аргумент асимметричен относительно $\omega = \pi$ на интервале $0 \leq \omega \leq 2\pi$. Аналогично действительная часть $H(e^{j\omega})$ симметрична, а мнимая – асимметрична на том же интервале. Поэтому при действительных ИХ интервал частот, на котором задают ЧХ, обычно сокращают до $0 \leq \omega \leq \pi$.

Функции (3.3) в z -плоскости соответствует $H(z)$, называемая передаточной функцией (системной функцией) цепи. Выражение для нее в общем случае имеет вид правой части (1.5), (1.6). В п. 1.4 отмечалось, что для графического изображения информации, содержащейся в z -преобразовании, используется изображение на z -плоскости нулей и полюсов, называемое нуль-полюсной диаграммой.

С помощью замены $X = H, z = e^{j\omega}$ функция (1.5) приводится к виду

$$H(e^{j\omega}) = A \frac{\prod_{i=1}^M (e^{j\omega} - z_i)}{\prod_{i=1}^N (e^{j\omega} - p_i)}.$$

Последнее выражение используется для геометрической интерпретации частотных свойств цифровой цепи. При этом на z -плоскости изображаются единичная окружность, нули и полюсы. Для оценки свойств цепи на частоте ω на плоскость наносится вектор $e^{j\omega}$; из точки $z=e^{j\omega}$, находящейся на единичной окружности, во все нули и полюсы проводятся векторы. По их величине определяется модуль передаточной функции на заданной частоте ω , а по углам – аргумент (фаза).

Для определения характеристики на всех частотах $0 \leq \omega \leq \pi$ необходимо поворачивать вектор $e^{j\omega}$ против часовой стрелки от точки $z=+1$ до точки $z=-1$. При этом следует иметь в виду, что наименьшие значения $|e^{j\omega} - z_i|$ или $|e^{j\omega} - p_i|$ соответствуют частотам ω , когда вектор $e^{j\omega}$ пересекает нуль или полюс соответственно.

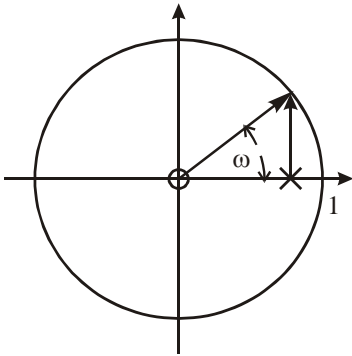


Рис. 3.6. Нуль-полюсная диаграмма цепи

В продолжение рассматриваемого примера запишем выражение

$$H(z) = \frac{z}{z - a}.$$

Эта функция имеет нуль в начале координат и вещественный полюс $p_1 = a$. Геометрическая интерпретация свойств этой цепи на частоте ω выполнена на рис. 3.6.

3.4. Способы построения цифровых цепей

При рассмотрении разностных уравнений в п. 1.11 отмечалось, что они непосредственно определяют способ построения цифровой цепи. Покажем это на следующих примерах. Разностное уравнение первого порядка самого общего вида

$$y(n) = -b_1 y(n-1) + a_0 x(n) + a_1 x(n-1)$$

можно реализовать с помощью схемы, изображенной на рис. 3.7. Блок «задержка» осуществляет задержку цифрового сигнала на один отсчет.

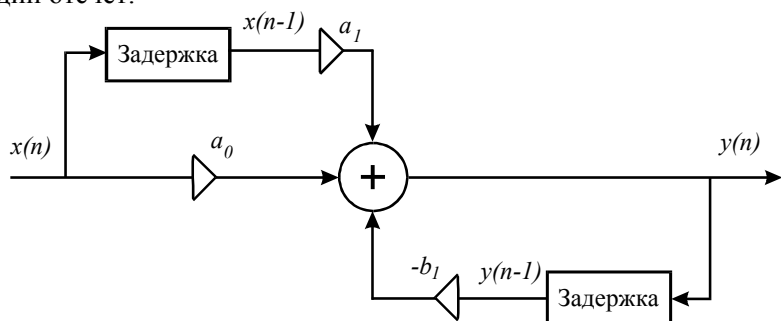


Рис. 3.7. Структурная схема цифровой цепи первого порядка

Рассмотренная форма построения цепи, в которой для входной и выходной последовательностей используются отдельные элементы задержки, называется прямой формой I. Ниже будут обсуждаться различные методы построения этой и других цифровых цепей.

Разностное уравнение второго порядка самого общего вида

$$y(n) = -b_1 y(n-1) - b_2 y(n-2) + a_0 x(n) + a_1 x(n-1) + a_2 x(n-2)$$

можно реализовать с помощью следующей схемы (рис. 3.8).

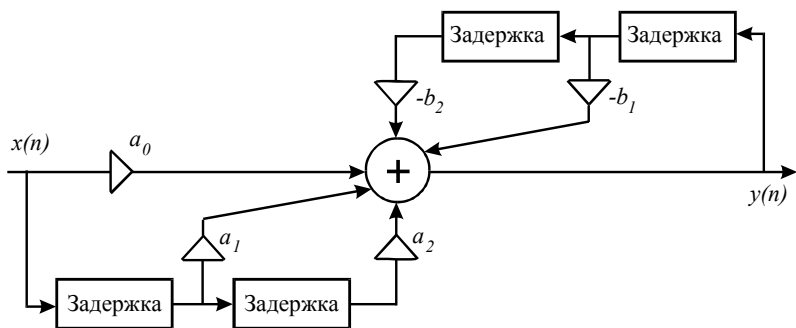


Рис. 3.8. Структурная схема цифровой цепи второго порядка

В этой схеме для входной и выходной последовательностей также используются отдельные элементы задержки.

Цифровые цепи первого и второго порядков могут быть использованы при реализации цепей более высокого порядка, так как последние могут быть представлены в виде последовательно или параллельно соединенных цепей первого и второго порядка.

Цифровые цепи с заданной передаточной функцией можно построить различными способами. Прежде всего, все цифровые цепи можно разделить на два больших класса: рекурсивные и нерекурсивные. Для рекурсивных цепей соотношение между входной последовательностью $\{x(n)\}$ и откликом цепи $\{y(n)\}$ может быть записано следующим образом: $y(n)=F[y(n-1), y(n-2), \dots, x(n), x(n-1), \dots]$, т. е. текущий отсчет отклика $y(n)$ определяется не только текущим и предшествующими значениями входной последовательности, но и предшествующими значениями отклика. В нерекурсивных цепях связь между входной последовательностью и откликом имеет вид $y(n)=F[x(n), x(n-1), \dots]$, т. е. текущий отсчет отклика зависит от текущего и предшествующего значений входной последовательности.

Приведем несколько возможных способов построения цифровых цепей. Как уже отмечалось, z -преобразование (передаточную функцию), соответствующее цифровой цепи, можно выразить в виде дробно-рационального полинома от переменной z^{-1} (смотри формулу (1.6)), т. е.

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{i=0}^N a_i z^{-i}}{\sum_{i=0}^N b_i z^{-i}}, \quad (3.4)$$

где a_i, b_i – коэффициенты (параметры цепи), причем $b_0=1$. (Предполагается, что степени числителя и знаменателя одинаковы, хотя в общем случае они могут быть различными.) Приведя равенство (3.4) к общему знаменателю, получим

$$Y(z) \sum_{i=0}^N b_i z^{-i} = X(z) \sum_{i=0}^N a_i z^{-i},$$

или

$$\sum_{i=0}^N b_i z^{-i} Y(z) = \sum_{i=0}^N a_i z^{-i} X(z).$$

Если рассматривать члены вида $z^{-k}Y(z)$ как прямые z -преобразования последовательностей $y(n-k)$, то, выполнив обратные z -преобразования обеих частей последнего равенства, можно получить искомое разностное уравнение:

$$\sum_{i=0}^N b_i y(n-i) = \sum_{i=0}^N a_i x(n-i).$$

Поскольку $b_0=1$, это уравнение можно разрешить относительно неизвестной функции $y(n)$:

$$y(n) = \sum_{i=0}^N a_i x(n-i) - \sum_{i=1}^N b_i y(n-i).$$

Простая структура реализации данного разностного уравнения имеет вид, показанный на рис. 3.9. Как упоминалось выше, она носит название прямой формы I. В ней для образования цепей, соответствующих числителю и знаменателю формулы (3.4), используются отдельные элементы задержки. Характерными чертами этой структуры являются ее простота и непосредственная связь с разностным уравнением.

Однако, как будет показано ниже, если полюсы $H(z)$ расположены близко друг от друга или от единичной окружности, как это имеет место для частотно-избирательных фильтров, то при использовании цепей данной структуры возникает трудноразрешимая проблема чувствительности характеристик фильтра к погрешности коэффициентов. По этой причине в большинстве практических случаев прямую форму I стараются не применять.

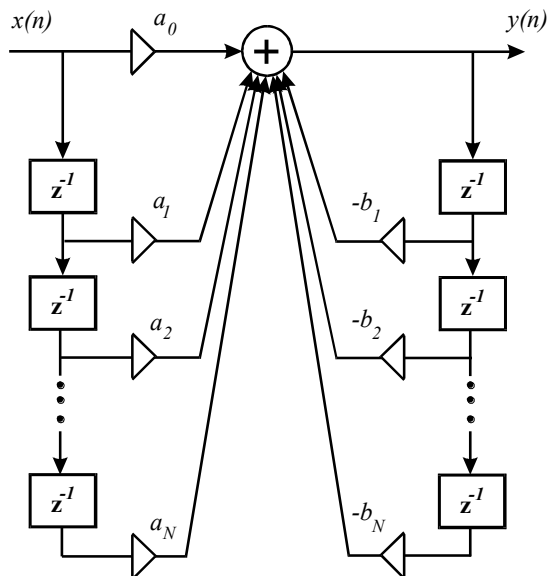


Рис. 3.9. Прямая форма I

Если записать формулу (3.4) в несколько ином виде, а именно

$$H(z) = \frac{Y(z)}{X(z)} = \underbrace{\left(\frac{1}{\sum_{i=0}^N b_i z^{-i}} \right)}_{H_1(z)} \underbrace{\left(\sum_{i=0}^N a_i z^{-i} \right)}_{H_2(z)},$$

то можно получить другую структуру цифровой цепи. Цифровая цепь, соответствующая последней формуле, состоит из двух последовательно соединенных цепей с коэффициентами передачи соответственно $H_1(z)$ и $H_2(z)$.

Первая из цепей имеет только полюсы, а вторая – только нули.

Если записать

$$H_1(z) = \frac{W(z)}{X(z)} = \frac{1}{\sum_{i=0}^N b_i z^{-i}},$$

и

$$H_2(z) = \frac{Y(z)}{W(z)} = \sum_{i=0}^N a_i z^{-i},$$

то получается пара разностных уравнений (в предположении, что $b_0=1$):

$$w(n) = x(n) - \sum_{i=1}^N b_i w(n-i),$$

$$y(n) = \sum_{i=0}^N a_i w(n-i),$$

которые можно реализовать следующим образом (рис. 3.10).
Такую структуру называют прямой формой II (неканонической).

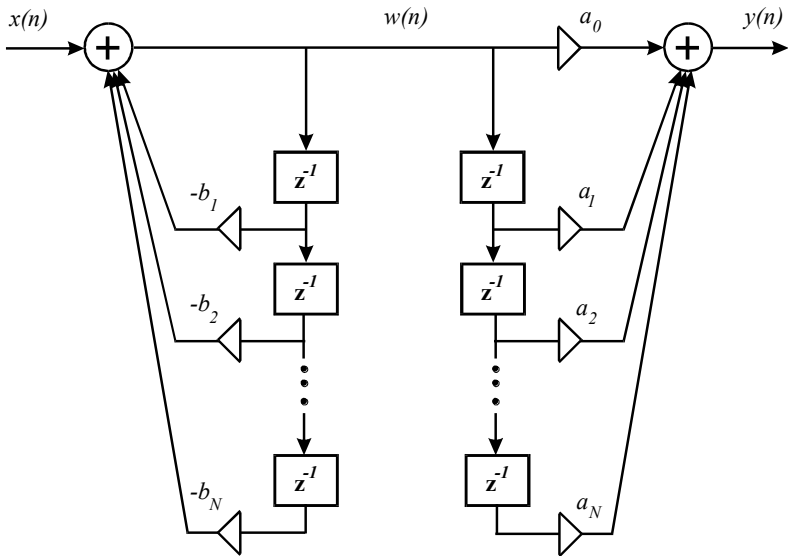


Рис. 3.10. Прямая форма II (неканоническая)

Поскольку в цепях соответствующих $H_1(z)$ и $H_2(z)$ сигнал $w(n)$ задерживается одинаково, то для построения цепи достаточно использовать один набор элементов задержки (рис. 3.11). Такую структуру называют прямой формой II или канонической структурой, т. к. в ней используется минимальное количество сумматоров и элементов задержки. (Некоторые другие схемы также обладают этим свойством, поэтому называть структуру рис. 3.11 канонической не рекомендуется.) Прямая форма II имеет такие же достоинства и недостатки, как и прямая форма I.

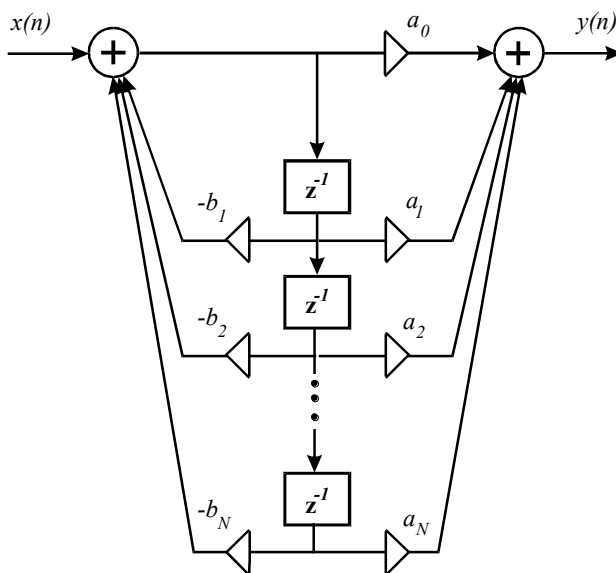


Рис. 3.11. Прямая форма II (каноническая)

Записав формулу (3.4) в виде

$$H(z) = \frac{Y(z)}{X(z)} = a_0 \prod_{k=1}^K H_k(z), \quad (3.5)$$

получим третью структуру построения цифровой цепи. Множители $H_k(z)$ соответствуют либо блокам второго порядка, т. е.

$$H_k(z) = \frac{1 + a_{1k}z^{-1} + a_{2k}z^{-2}}{1 + b_{1k}z^{-1} + b_{2k}z^{-2}}, \quad (3.6)$$

либо блокам первого порядка, т. е.

$$H_k(z) = \frac{1 + a_{1k}z^{-1}}{1 + b_{1k}z^{-1}}, \quad (3.7)$$

а K равно целой части числа $(N+1)/2$.

Схему, реализующую формулу (3.5), называют каскадной (или последовательной) формой (рис. 3.12):

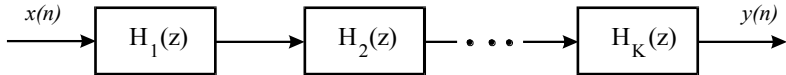


Рис. 3.12. Каскадная (последовательная) форма

Каждый из блоков второго порядка, образующих последовательную форму, можно реализовать в прямой форме I или в прямой форме II.

Использование блоков второго порядка (и, возможно, одного блока первого порядка) при построении цепи определяется тем, что для получения комплексного полюса или нуля цепь с действительными коэффициентами должна включать блок второго порядка.

В предположении неограниченной точности представления всех переменных порядок блоков и способ группирования нулей с полюсами не имеют значения. Однако для реальных устройств, с учетом эффектов квантования, эти вопросы имеют большое значение. Подробнее они будут рассмотрены ниже.

Еще одна трудность, связанная с особенностями последовательной формы, состоит в необходимости введения масштабирующих множителей между отдельными блоками. Эти множители не позволяют переменным фильтра принимать слишком большие или слишком малые значения.

Четвертую структурную схему цифровой цепи можно получить, разложив правую часть формулы (3.4) на простые дроби:

$$H(z) = C + \sum_{k=1}^K H_k(z),$$

где слагаемые $H_k(z)$ соответствуют или блокам второго порядка:

$$H_k(z) = \frac{a_{0k} + a_{1k}z^{-1}}{1 + b_{1k}z^{-1} + b_{2k}z^{-2}},$$

или блокам первого порядка:

$$H_k(z) = \frac{a_{0k}}{1 + b_{1k}z^{-1}},$$

причем K равно целой части числа $(N+1)/2$ и, как следует из формулы (3.4), $C = a_N/b_N$.

Покажем структурную схему, реализующую это соотношение (рис. 3.13).

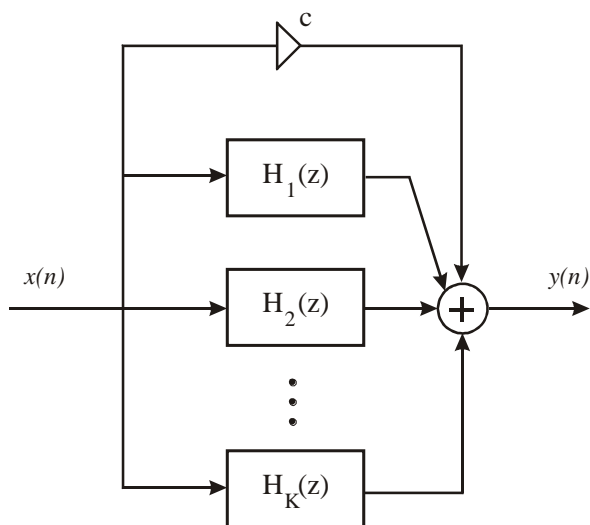


Рис. 3.13. Параллельная форма

Ее называют параллельной формой. Блоки первого и второго порядков строятся по схеме одной из рассмотренных выше прямых форм. Хотя рассмотренные структурные схемы цепей не исчерпывают всех возможных структур, при моделировании на ЭВМ и аппаратурной реализации они применяются наиболее часто.

Другие структуры можно получить множеством способов. Так, например, можно построить параллельно-последовательную структуру, в которой часть передаточной функции реализуется в параллельной форме, а остальная часть – в последовательной.

Выбор наилучшей из этих многочисленных структурных схем, как при аппаратурной реализации, так и при моделировании на ЭВМ, определяется экономическими соображениями. Последние, в свою очередь, зависят от свойств структур при ограниченной точности представления переменных и коэффициентов цепей (фильтров).

3.5. Структурные схемы цепей без полюсов

В важном частном случае знаменатель дроби (3.4) может быть постоянным (для простоты приравняем его к единице). При этом разностное уравнение, описывающее систему, становится рекурсивным, т. е. текущее значение отклика $y(n)$ зависит только от текущего и конечного числа предшествующих значений входной последовательности.

В этом случае правую часть (3.4) обычно преобразуют таким образом, чтобы выразить $H(z)$ непосредственно через ИХ цепи (фильтра):

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{n=0}^{N-1} h(n)z^{-n}. \quad (3.8)$$

Здесь верхний предел суммирования заменен на $N-1$, чтобы уравнение описывало физически реализуемый фильтр, длина ИХ которого равна N отсчетам.

Разностное уравнение, соответствующее выражению (3.8), имеет вид

$$y(n) = h(0)x(n) + h(1)x(n-1) + \dots + h(N-1)x(n-N+1),$$

т. е. является нерекурсивным уравнением.

Для построения цепей (фильтров) с КИХ рассматриваемого типа обычно применяют несколько структурных схем. Чаще всего используют прямую форму, рассмотренную в п. 3.4.

Для данного частного случая существует только одна прямая форма (рис. 3.14). Из-за сходства этой структуры с линией задержки с отводами ее часто называют цепью с многоотводной линией задержки (или иногда трансверсальным фильтром).

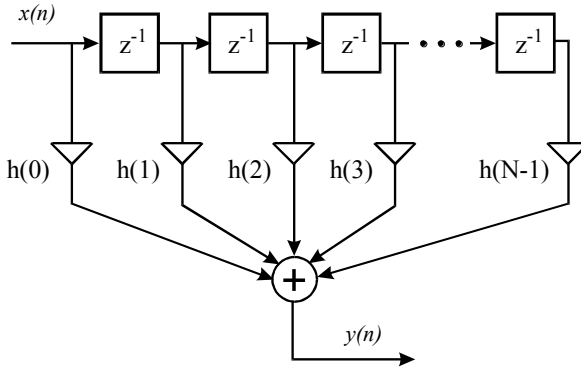


Рис. 3.14. Структурная схема КИХ-цепи

При построении цепей, не имеющих полюсов, удобна оказывается и последовательная структура. В этом случае z -преобразование ИХ цепи (3.8) представляется в виде произведения z -преобразований, соответствующих цепям первого и второго порядка, т. е.

$$H(z) = \prod_{k=1}^K H_k(z),$$

где

$$H_k(z) = a_{0k} + a_{1k}z^{-1} + a_{2k}z^{-2} \text{ — (цепь второго порядка),}$$

или

$$H_k(z) = a_{0k} + a_{1k}z^{-1} - \text{(цепь первого порядка)},$$

причем K равно целой части $(N+1)/2$.

При построении цепей без полюсов довольно часто применяют еще несколько структур, которые не имеют аналогов с цепями общего вида, содержащими и нули, и полюсы. Наиболее распространенная из них основана на так называемом методе быстрой свертки, когда свертка вычисляется с помощью обратного преобразования Фурье от произведения преобразования Фурье входной последовательности и ИХ системы.

3.6. Нерекурсивная цепь второго порядка

Структурная схема цепи при $a_0=1$ изображена на рис. 3.15. Выходной сигнал цепи описывается разностным уравнением

$$y(n) = x(n) + a_1x(n-1) + a_2x(n-2). \quad (3.9)$$

Цепь устойчива при любых значениях коэффициентов a_1, a_2 .

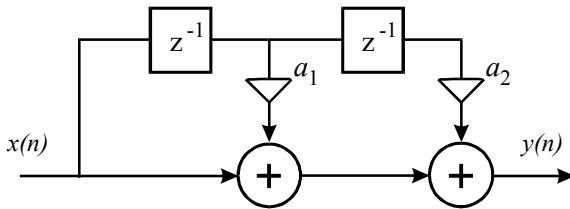


Рис. 3.15. Структурная схема нерекурсивной цепи второго порядка

3.6.1. Частотные свойства

Выполнив одностороннее z -преобразование обеих частей уравнения (3.9), полагая $x(-1)=0$ и $x(-2)=0$, получим

$$Y(z) = X(z) + a_1z^{-1}X(z) + a_2z^{-2}X(z).$$

Передаточная функция цепи выражается зависимостями

$$H(z) = \frac{Y(z)}{X(z)} = 1 + a_1 z^{-1} + a_2 z^{-2} = \frac{z^2 + a_1 z + a_2}{z^2}. \quad (3.10)$$

Эта функция имеет двукратный полюс в начале координат и два нуля в точках

$$z_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_2}.$$

В зависимости от соотношения между величинами a_1 и a_2 возможны четыре варианта.

Вариант I. $0 < a_2 < \frac{a_1^2}{4}, a_1 > 0.$

При этом нули вещественные, отрицательные, разные. Диаграмма нулей и полюсов показана на рис. 3.16а. Рассматриваемая цепь является фильтром нижних частот. В силу вещественности нулей и полюсов ее можно представить в виде последовательного соединения двух ФНЧ первого порядка, передаточные функции которых имеют нули z_1 и z_2 соответственно.

Вариант II. $0 < a_2 < \frac{a_1^2}{4}, a_1 < 0.$

Нули передаточной функции вещественные, положительные, разные. Диаграмма нулей и полюсов показана на рис. 3.16б. Цепь является ФВЧ. Она эквивалентна двум последовательно соединенным ФВЧ первого порядка, передаточные функции которых имеют нули z_1 и z_2 соответственно.

Вариант III. $a_2 < 0, a_1$ – произвольное.

Нули передаточной функции вещественные, с разными знаками. Диаграмма нулей и полюсов для $a_1 > 0$ приведена на рис. 3.16в. Цепь является полосовым фильтром (ПФ). Она эквивалентна последовательному соединению ФНЧ и ФВЧ первого порядка, передаточные функции которых имеют нули z_1 и z_2 соответственно.

Вариант IV. $a_2 > \frac{a_1^2}{4}$, a_1 – произвольное.

Нули передаточной функции комплексно сопряженные:

$$z_{1,2} = -\frac{a_1}{2} \pm j\sqrt{a_2 - \frac{a_1^2}{4}} = re^{\pm j\theta}.$$

Диаграмма нулей и полюсов изображена на рис. 3.16г. Цепь является режекторным фильтром (РФ).

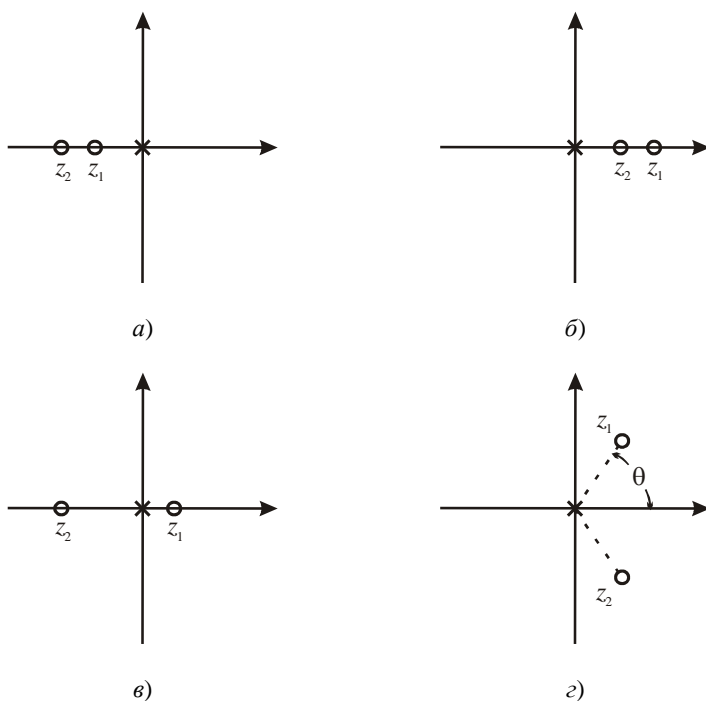


Рис. 3.16. Диаграммы нулей и полюсов
нерекурсивной цепи второго порядка:
а) вариант I; б) вариант II; в) вариант III; г) вариант IV

Покажем на рис. 3.17 разбиение плоскости параметров цепи (a_1, a_2) , соответствующее рассмотренным четырем вариантам.

Заменив в (3.10) $z=e^{j\omega}$, получим выражение для ЧХ нерекурсивной цепи второго порядка:

$$H(e^{j\omega})=1+a_1e^{-j\omega}+a_2e^{-j2\omega}.$$

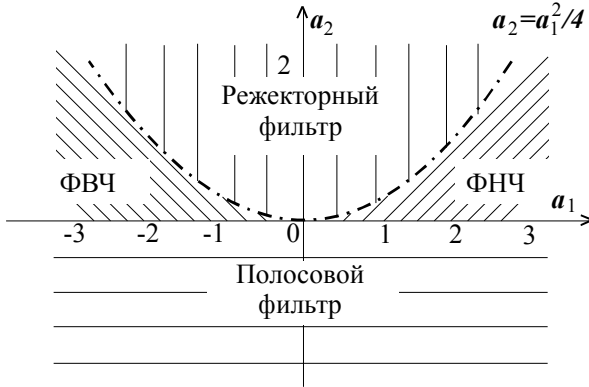


Рис. 3.17. Частотные свойства нерекурсивной цепи второго порядка

Ее АЧХ характеризуется зависимостью

$$H(\omega) = \sqrt{a_1^2 + (1-a_2)^2 + 2a_1(1+a_2)\cos\omega + 4a_2\cos^2\omega}, \quad (3.11)$$

а ФЧХ

$$\phi(\omega) = \arctg - \frac{a_1 \sin \omega + a_2 \sin 2\omega}{1 + a_1 \cos \omega + a_2 \cos 2\omega}. \quad (3.12)$$

Как и у любой цифровой цепи, АЧХ и ФЧХ нерекурсивной цепи второго порядка периодичны (период $\omega=2\pi$) и симметричны относительно $\omega=\pi$.

Выше указывалось, что нерекурсивные цепи обладают конечной ИХ и поэтому их обычно называют КИХ-цепями (КИХ-фильтрами).

В общем случае КИХ-цепь порядка $N-1$ (длины N) описывается передаточной функцией

$$H(z) = \sum_{n=0}^{N-1} a_n z^{-n},$$

которой соответствует ИХ:

$$h(n) = \sum_{i=0}^{N-1} a_i \delta(n-i).$$

Фильтр с КИХ обладает уникальным свойством – линейной ФЧХ, если его ИХ (т. е. коэффициенты a_n передаточной функции) симметрична, т. е.

$$h(n) = h(N-1-n), \quad a_n = a_{N-1-n},$$

или асимметрична, т. е.

$$h(n) = -h(N-1-n), \quad a_n = -a_{N-1-n}.$$

Отсюда следуют четыре возможных вида КИХ-фильтров с линейной ФЧХ в зависимости от симметричности или асимметричности ИХ и четности или нечетности N .

Фильтр вида 1: симметричная ИХ и нечетное N .

Фильтр вида 2: симметричная ИХ и четное N .

Фильтр вида 3: асимметричная ИХ и нечетное N .

Фильтр вида 4: асимметричная ИХ и четное N .

В рассматриваемой цепи $N=3$, $a_0=1$, поэтому могут быть реализованы фильтры:

вида 1: $a_0=1$, $a_2=1$, a_1 – произвольной величины, при этом цепь может являться ФНЧ, ФВЧ и РФ;

вида 3: $a_0=1$, $a_2=-1$, $a_1=0$, при этом цепь является ПФ.

КИХ-цепь, у которой все коэффициенты a_n одинаковы, называется однородным фильтром.

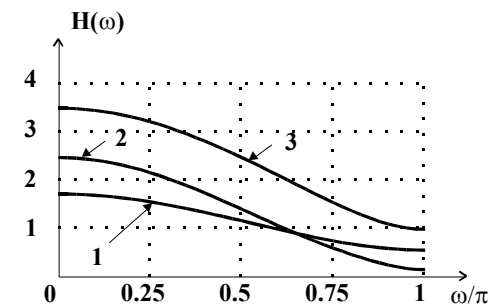
Заменой в (3.11) $\cos\omega = b$, $4a_2 = A$, $2a_1(1+a_2) = B$, $a_1^2 + (1-a_2)^2 = C$ выражение АЧХ преобразуется к виду

$$H(b) = \sqrt{Ab^2 + Bb + C}. \quad (3.13)$$

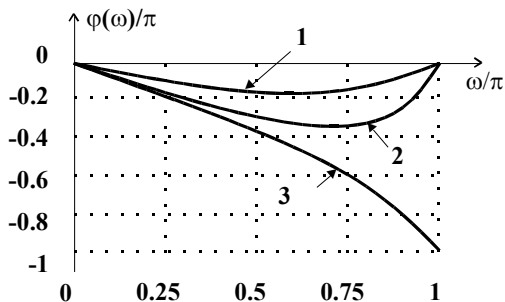
Рассмотрим каждый тип фильтра более подробно.

Фильтры нижних и верхних частот

Покажем АЧХ и ФЧХ фильтров нижних (рис. 3.18а, 3.18б) и верхних (рис. 3.19а, 3.19б) частот, вычисленные по формулам (3.11) и (3.12) для трех сочетаний параметров цепи.

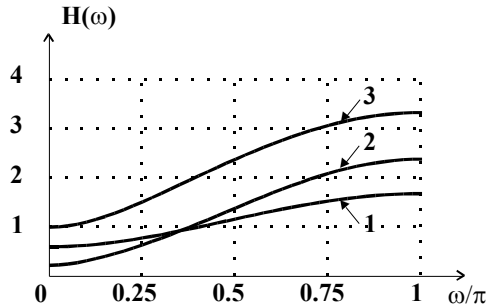


а)

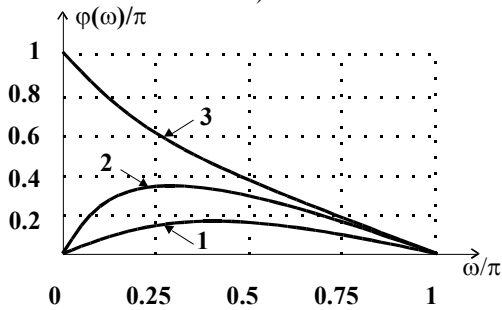


б)

Рис. 3.18. Характеристики фильтра нижних частот:
а) амплитудно-частотная; б) фазочастотная;
(кривая 1: $a_1=0.5$, $a_2=0.05$; кривая 2: $a_1=1$, $a_2=0.2$;
кривая 3: $a_1=2$, $a_2=0.08$)



a)



б)

Рис. 3.19. Характеристики фильтра верхних частот:
 а) амплитудно-частотная; б) фазочастотная;
 (кривая 1: $a_1 = -0.5$, $a_2 = 0.05$; кривая 2: $a_1 = -1$, $a_2 = 0.2$;
 кривая 3: $a_1 = -2$, $a_2 = 0.08$)

Нормированная АЧХ выражается зависимостью

$$\frac{H(b)}{H(b)_{\max}} = \sqrt{\frac{Ab^2 + Bb + C}{A + |B| + C}} = \frac{\sqrt{Ab^2 + Bb + C}}{a_2 + 1 + |a_1|}.$$

Воспользовавшись соотношением

$$\frac{Ab^2 + Bb + C}{(a_2 + 1 + |a_1|)^2} = \frac{1}{2},$$

получим выражение для частоты среза:

$$\omega_c = \arccos \frac{-B \pm \sqrt{B^2 - 2A [2C - (a_2 + 1 + |a_1|)^2]}}{2A}.$$

Полосовой и режекторный фильтры

Покажем АЧХ и ФЧХ для трех сочетаний параметров (a_1, a_2) ПФ на рис. 3.20а и 3.20б и РФ на рис. 3.21а и 3.21б соответственно.

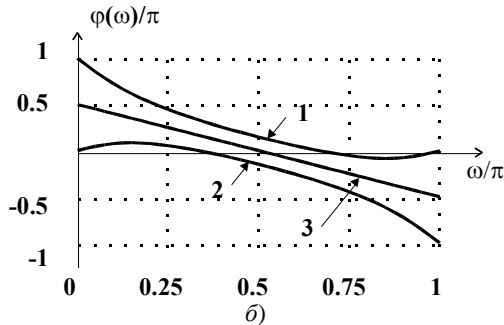
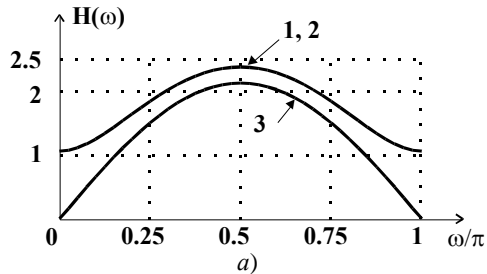


Рис. 3.20. Характеристики полосового фильтра:
 а) амплитудно-частотная; б) фазочастотная;
 (кривая 1: $a_1 = -1, a_2 = -1$; кривая 2: $a_1 = 1, a_2 = -1$;
 кривая 3: $a_1 = 0, a_2 = -1$)

Получаемая из (3.13) функция

$$H^2(b) = Ab^2 + Bb + C$$

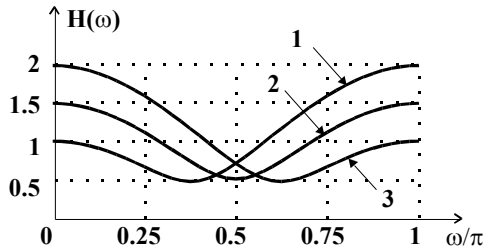
является квадратным трехчленом, который приводится к параболе

$$H^2(b) = A \left(b + \frac{B}{2A} \right)^2 - \frac{B^2 - 4AC}{4A}$$

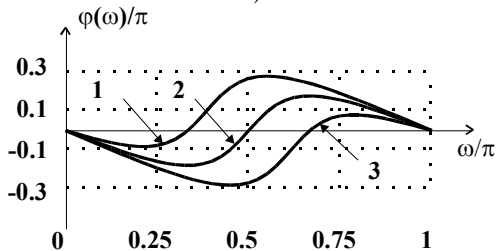
с координатами вершины $M(\alpha; \beta) = M\left(\frac{-B}{2A}; \frac{4AC - B^2}{4A}\right)$.

При $a_2 > 0$ (режекторный фильтр) ветви параболы направлены вверх, а при $a_2 < 0$ (полосовой фильтр) – вниз. Это означает, что резонансная частота полосового и режекторного фильтров выражается зависимостью

$$\omega_p = \arccos b_p = \arccos \frac{-B}{2A} = \arccos \frac{-a_1(1+a_2)}{4a_2}. \quad (3.14)$$



a)



б)

Рис. 3.21. Характеристики режекторного фильтра:
 а) амплитудно-частотная; б) фазочастотная;
 (кривая 1: $a_1 = -0.5, a_2 = 0.5$; кривая 2: $a_1 = 0, a_2 = 0.5$;
 кривая 3: $a_1 = 0.5, a_2 = 0.5$)

Покажем на рис. 3.22 графики зависимости $\omega_p(a_1)$ для РФ ($a_2=1$) и ПФ ($a_2=-0.5$).

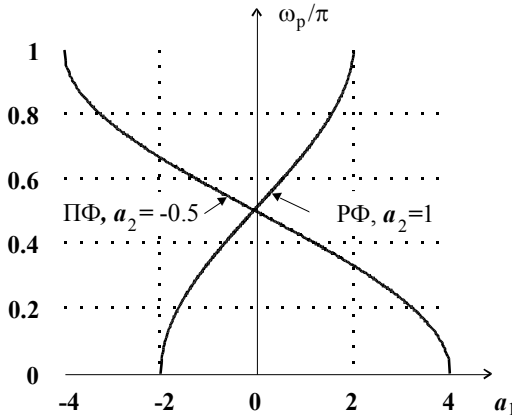


Рис. 3.22. Зависимость резонансной частоты фильтров от коэффициента a_1

Резонансное значение квадрата АЧХ равно

$$H^2(b_p) = \beta = \frac{4AC - B^2}{4A} = C - \frac{B^2}{4A^2} A = C - b_p^2 A. \quad (3.15)$$

Нормированная АЧХ выражается зависимостью

$$\frac{H(b)}{H(b_p)} = \sqrt{\frac{Ab^2 + Bb + C}{\beta}}. \quad (3.16)$$

Резонансные свойства РФ и ПФ принято оценивать с помощью добротности, вычисляемой по формуле

$$Q = \frac{\omega_p}{\omega_2 - \omega_1} = \frac{\arccos b_p}{\arccos b_2 - \arccos b_1}. \quad (3.17)$$

Здесь величины $\omega_{1,2}$, $b_{1,2}$ соответствуют границам: полосы пропускания (ПП) в ПФ по уровню $1/\sqrt{2}$ от максимального уровня,

полосы подавления (задерживания) (ПЗ) в РФ по уровню $\sqrt{2}$ от минимального уровня.

Воспользовавшись соотношением

$$\frac{Ab_{1,2}^2 + Bb_{1,2} + C}{\beta} = \frac{1}{2},$$

получим для ПФ

$$b_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}. \quad (3.18)$$

Из соотношения

$$\frac{Ab_{1,2}^2 + Bb_{1,2} + C}{\beta} = 2,$$

для РФ получим

$$b_{1,2} = \frac{-B \pm \sqrt{4AC - B^2}}{2A}. \quad (3.19)$$

Заметим, что в РФ при $a_2=1$ понятие добротности не существует, т. к. при этом $H(b_p)=0$. Покажем на рис. 3.23 график зависимости $Q(a_2)$ для РФ.

Резонансную кривую удобно представить как функцию относительной расстройки:

$$\gamma = \frac{\omega}{\omega_p} = \frac{\arccos b}{\arccos b_p}. \quad (3.20)$$

На границе полосы пропускания значения γ равны

$$\gamma_{1,2} = \frac{\omega_{1,2}}{\omega_p} = \frac{\arccos b_{1,2}}{\arccos b_p}.$$

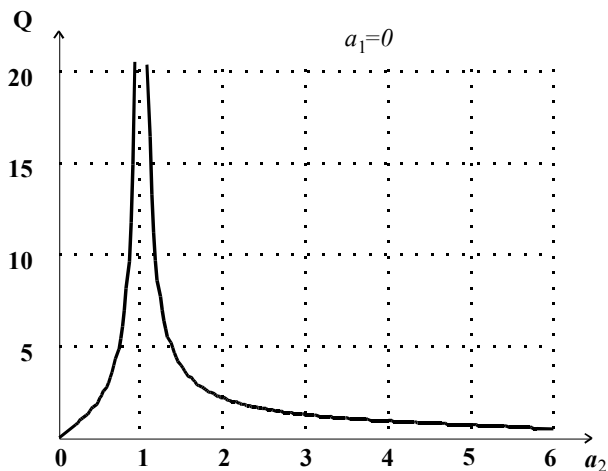


Рис. 3.23. Зависимость добротности режекторного фильтра от коэффициента a_2

Нормированная АЧХ через γ выражается следующим образом:

$$\frac{H(\gamma)}{H(1)} = \sqrt{\frac{A \cos^2(\omega_p \gamma) + B \cos(\omega_p \gamma) + C}{\beta}}.$$

Формула добротности как функция γ выглядит следующим образом:

$$Q = \frac{1}{\gamma_2 - \gamma_1}. \quad (3.21)$$

3.6.2. Временные характеристики

К временным характеристикам относятся: ИХ и ПХ, реакция на воздействие прямоугольного импульса (для ФНЧ и ФВЧ), реакция на воздействие прямоугольного радиоимпульса (для ПФ и РФ). Выражения для характеристик получаются прямой подстановкой входного воздействия в разностное уравнение (3.9).

Импульсная характеристика

Она определяется зависимостью

$$h(n) = \delta(n) + a_1 \delta(n-1) + a_2 \delta(n-2).$$

На рис. 3.24а–3.24г приведены графики ИХ соответственно ФНЧ с параметрами $a_1=1$, $a_2=0.2$, ФВЧ с параметрами $a_1=-1$, $a_2=0.2$, ПФ с параметрами $a_1=1$, $a_2=-1$, РФ с параметрами $a_1=1$, $a_2=1$.

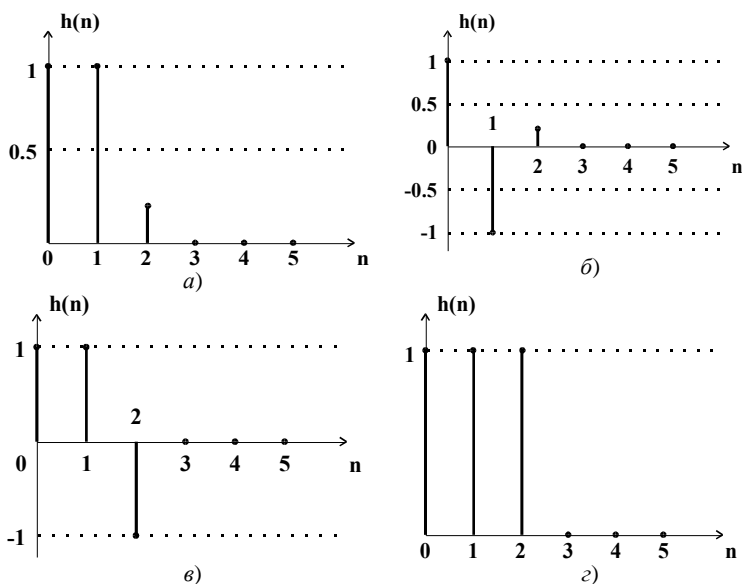


Рис. 3.24. Импульсные характеристики нерекурсивной цепи второго порядка: а) фильтра нижних частот; б) фильтра верхних частот; в) полосового фильтра; г) режекторного фильтра

Переходная характеристика

В данном случае она равна

$$g(n) = 1(n) + a_1 1(n-1) + a_2 1(n-2).$$

На рис. 3.25а–3.25г приведены графики ПХ соответственно ФНЧ, ФВЧ, ПФ и РФ с указанными параметрами.

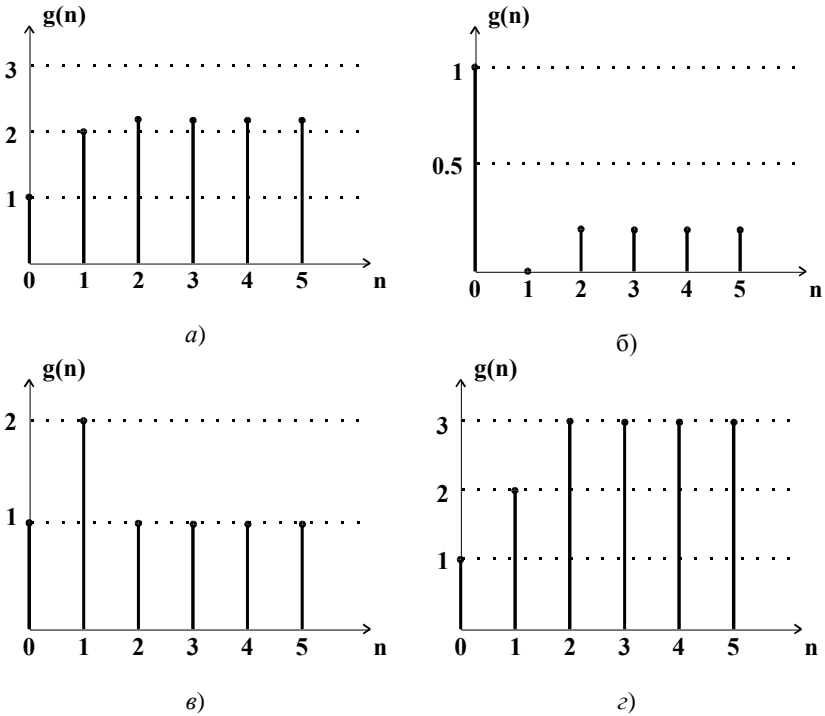


Рис. 3.25. Переходные характеристики нерекурсивной цепи второго порядка: а) фильтра нижних частот; б) фильтра верхних частот; в) полосового фильтра; г) режекторного фильтра

Реакция фильтров нижних и верхних частот на воздействие прямоугольного импульса

Входной сигнал в виде прямоугольного импульса выражается следующим образом:

$$x(n) = X[1(n) - 1(n - n_u)],$$

где n_u – длительность импульса, т. е. число отсчетов в импульсе, а X – амплитуда импульса. Выходной сигнал рассматриваемой цепи $y(n)$ находится прямой подстановкой этой функции в уравнение (3.9). На рис. 3.26а и рис. 3.26б приведены осциллограммы выходных сигналов соответственно ФНЧ и ФВЧ с указанными параметрами на воздействие прямоугольного импульса длительностью $n_u=4$ при $X=1$.

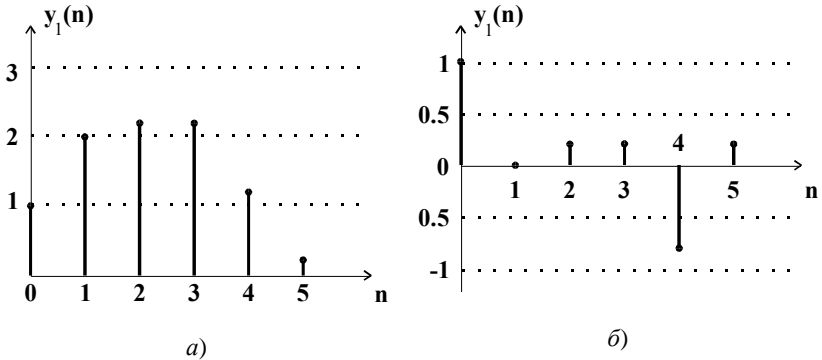


Рис. 3.26. Реакция на воздействие прямоугольного импульса:
а) фильтра нижних частот; б) фильтра верхних частот

Реакция полосового и режсекторного фильтров на воздействие прямоугольного радиоимпульса

Входной сигнал в виде прямоугольного радиоимпульса выражается следующим образом:

$$x(n) = X \sin \omega n [1(n) - 1(n - n_u)].$$

Реакции фильтра находится прямой подстановкой этой функции в уравнение (3.9). Осциллограмма модуля огибающей выходного сигнала ПФ с указанными параметрами на воздействие прямоугольного радиоимпульса длительностью $n_u=4$ при $X=1$ на резонансной частоте показана на рис. 3.27.

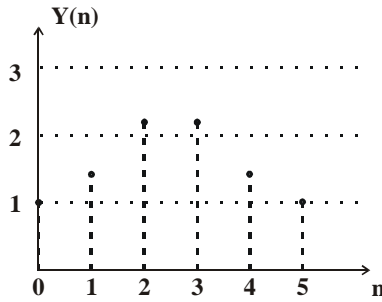


Рис. 3.27. Модуль огибающей реакции полосового фильтра на воздействие прямоугольного радиоимпульса

3.7. Рекурсивная цепь второго порядка

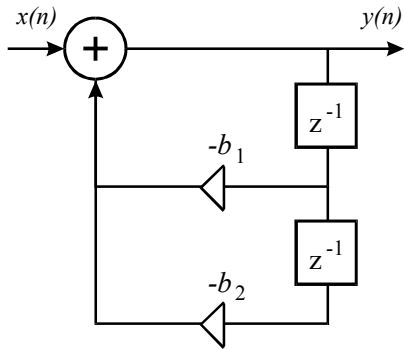


Рис. 3.28. Структурная схема рекурсивной цепи второго порядка

Структурная схема цепи изображена на рис. 3.28. Колебания в цепи под действием внешней силы описываются разностным уравнением

$$y(n) = x(n) - b_1 y(n-1) - b_2 y(n-2). \quad (3.22)$$

Цепь устойчива, если ее параметры находятся внутри треугольника, ограниченного отрезками прямых:

$$\begin{cases} b_2 = 1 \\ b_2 = b_1 - 1 \\ b_2 = -b_1 - 1. \end{cases}$$

Покажем этот треугольник на рис. 3.29. Данную фигуру называют треугольником устойчивости. Ниже полагается, что параметры цепи выбираются внутри этого треугольника.

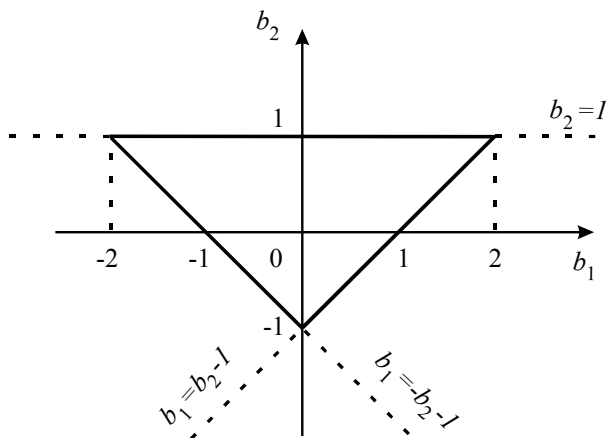


Рис. 3.29. Треугольник устойчивости рекурсивной цепи второго порядка

Выполнив одностороннее z -преобразование обеих частей уравнения (3.22), полагая $y(-1)=y(-2)=0$, получим

$$Y(z) = X(z) - b_1 z^{-1} Y(z) - b_2 z^{-2} Y(z).$$

Передаточная функция цепи выражается зависимостью

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 + b_1 z^{-1} + b_2 z^{-2}} = \frac{z^2}{z^2 + b_1 z + b_2}. \quad (3.23)$$

Она имеет двукратный нуль в начале координат $z_1=0$ и два полюса в точках

$$p_{1,2} = -\frac{b_1}{2} \pm \sqrt{\frac{b_1^2}{4} - b_2}. \quad (3.24)$$

3.7.1. Частотные свойства

В зависимости от соотношения между величинами b_1 и b_2 возможны четыре варианта.

Вариант I. $0 < b_2 < \frac{1}{4} b_1^2$, $b_1 < 0$.

При этом полюсы вещественные, положительные, разные. Диаграмма нулей и полюсов показана на рис. 3.30а. Рассматриваемая цепь является ФНЧ. В силу вещественности нулей и полюсов ее можно представить в виде последовательного соединения двух ФНЧ первого порядка, передаточные функции которых имеют полюсы p_1 и p_2 соответственно.

Вариант II. $0 < b_2 < \frac{1}{4}b_1^2, b_1 > 0$.

При этом полюсы вещественные, отрицательные, разные. Диаграмма нулей и полюсов показана на рис. 3.30б. Рассматриваемая цепь является ФВЧ. Она эквивалентна двум последовательно соединенным ФВЧ первого порядка, передаточные функции которых имеют полюсы p_1 и p_2 соответственно.

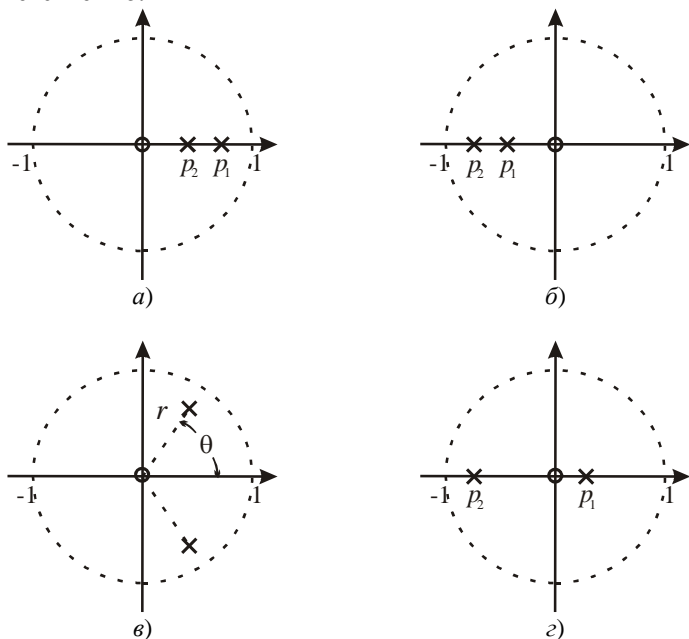


Рис. 3.30. Диаграммы нулей и полюсов рекурсивной цепи второго порядка:

а) вариант I; б) вариант II; в) вариант III; г) вариант IV

Вариант III. $b_2 > \frac{1}{4}b_1^2$, b_1 – произвольное.

Полюсы передаточной функции комплексно сопряженные:

$$p_{1,2} = -\frac{b_1}{2} \pm j\sqrt{b_2 - \frac{b_1^2}{4}} = re^{\pm j\theta}.$$

Диаграмма нулей и полюсов приведена на рис. 3.30в. Цепь является ПФ.

Вариант IV. $b_2 < 0$, b_1 – произвольное.

Полюсы передаточной функции вещественные, с разными знаками. Диаграмма нулей и полюсов для $b_1 > 0$ приведена на рис. 3.30г. Цепь является РФ.

Покажем на рис. 3.31 разбиение плоскости параметров цепи (b_1, b_2) , соответствующее рассмотренным четырем вариантам, с учетом треугольника устойчивости.

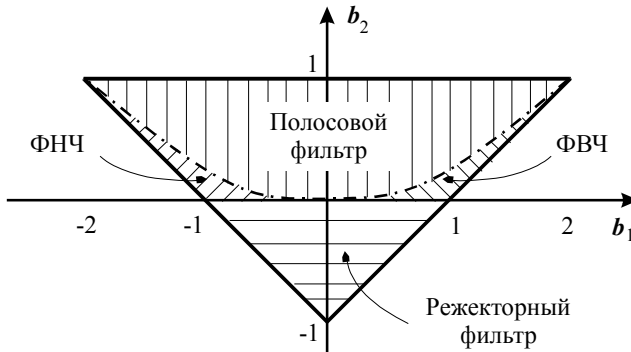


Рис. 3.31. Частотные свойства рекурсивной цепи второго порядка

Заменив в (3.23) $z=e^{j\omega}$, получим выражение для ЧХ рекурсивной цепи второго порядка:

$$H(e^{j\omega}) = \frac{1}{1 + b_1 e^{-j\omega} + b_2 e^{-j2\omega}}.$$

Ее АЧХ характеризуется зависимостью

$$H(\omega) = \frac{1}{\sqrt{b_1^2 + (1-b_2)^2 + 2b_1(1+b_2)\cos\omega + 4b_2\cos^2\omega}}, \quad (3.25)$$

а ФЧХ равна

$$\phi(\omega) = \arctg \frac{b_1 \sin \omega + b_2 \sin 2\omega}{1 + b_1 \cos \omega + b_2 \cos 2\omega}. \quad (3.26)$$

Сравнивая эти выражения с (3.11) и (3.12) для нерекурсивной цепи второго порядка, видим, что АЧХ и ФЧХ рекурсивной цепи второго порядка также периодичны (период $\omega=2\pi$) и симметричны относительно $\omega=\pi$.

Заменой $\cos \omega = a$, $4b_2 = A$, $2b_1(1+b_2) = B$, $b_1^2 + (1-b_2)^2 = C$ выражение для АЧХ преобразуется к виду

$$H(a) = \frac{1}{\sqrt{Aa^2 + Ba + C}}. \quad (3.27)$$

Рассмотрим каждый тип фильтра более подробно.

Фильтры нижних и верхних частот

Покажем на рис. 3.32а и 3.32б АЧХ и ФЧХ соответственно ФНЧ, а на рис. 3.33а и 3.33б – ФВЧ, вычисленные по формулам (3.25) и (3.26) для двух сочетаний параметров цепи.

Нормированная АЧХ выражается зависимостью

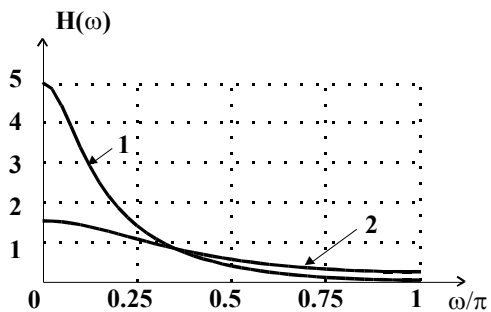
$$\frac{H(a)}{H(a)_{\max}} = \sqrt{\frac{A - |B| + C}{Aa^2 + Ba + C}} = \frac{b_2 + 1 - |b_1|}{\sqrt{Aa^2 + Ba + C}}.$$

Частоту среза найдем, воспользовавшись соотношением

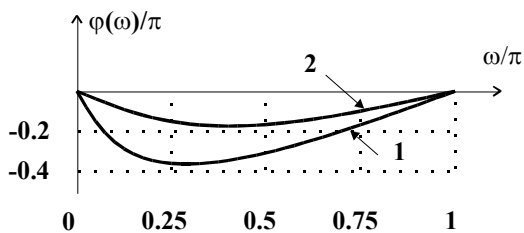
$$\frac{(b_2 + 1 - |b_1|)^2}{Aa_c^2 + Ba_c + C} = \frac{1}{2},$$

отсюда

$$\omega_c = \arccos \frac{-B \pm \sqrt{B^2 - 4A[c - 2(b_2 + 1 - |b_1|)^2]}}{2A}.$$



a)



б)

Рис. 3.32. Характеристики фильтра нижних частот:
 а) амплитудно-частотная; б) фазочастотная
 (кривая 1: $b_1=-1, b_2=0.2$; кривая 2: $b_1=-0.5, b_2=0.05$)

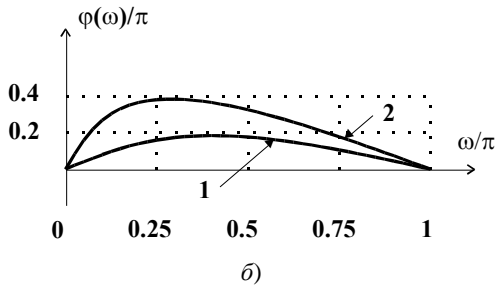
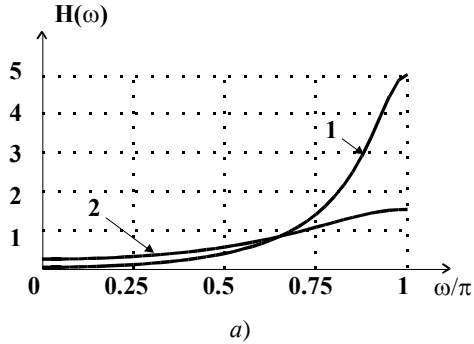


Рис. 3.33. Характеристики фильтра верхних частот:
 а) амплитудно-частотная; б) фазочастотная
 (кривая 1: $b_1=1, b_2=0.2$; кривая 2: $b_1=0.5, b_2=0.05$)

Полосовой и режекторный фильтры

Покажем на рис. 3.34 и 3.34б соответственно АЧХ и ФЧХ полосового фильтра, а на рис. 3.35а и 3.35б – РФ для двух сочетаний параметров b_1, b_2 .

Получаемая из (3.27) функция

$$f^2(a) = \frac{1}{H^2(a)} = Aa^2 + Va + C$$

является квадратичным трехчленом, который приводится к параболе

$$f^2(a) = A\left(a + \frac{B}{2A}\right)^2 - \frac{B^2 - 4AC}{4A}.$$

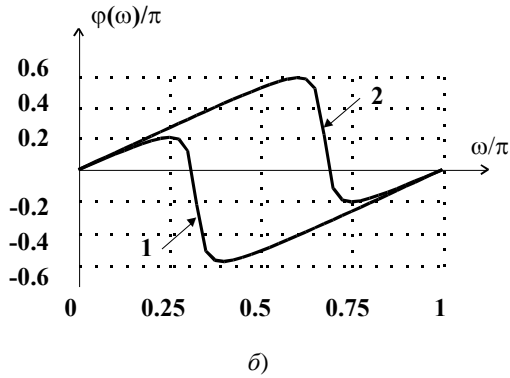
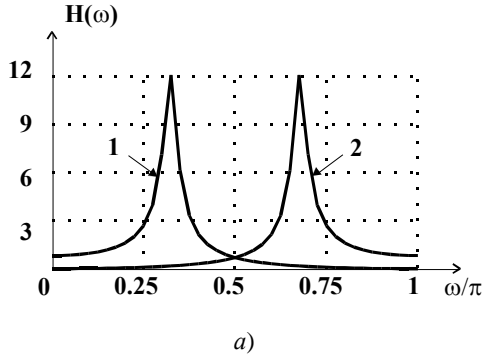


Рис. 3.34. Характеристики полосового фильтра:
 а) амплитудно-частотная; б) фазочастотная
 (кривая 1: $b_1=-1, b_2=0.9$; кривая 2: $b_1=1, b_2=0.9$)

Вершина параболы имеет координаты

$$M(\alpha, \beta) = M\left(-\frac{B}{2A}, \frac{4AC - B^2}{4A}\right),$$

которые совпадают с координатами вершины параболы $H^2(b)$ в нерекурсивной цепи второго порядка. При $b_2 > 0$ (ПФ) ветви параболы направлены вверх, а при $b_2 < 0$ (РФ) – вниз.

Резонансная частота ПФ и РФ выражается зависимостью

$$\omega_p = \arccos a_p = \arccos\left(-\frac{B}{2A}\right) = \arccos\left(-\frac{b_1(1+b_2)}{4b_2}\right).$$

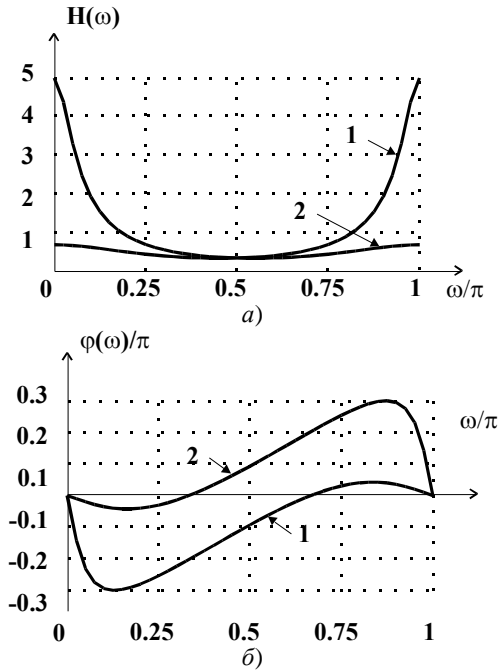


Рис. 3.35. Характеристики режекторного фильтра:
 а) амплитудно-частотная; б) фазочастотная
 (кривая 1: $b_1 = -0.4, b_2 = -0.4$; кривая 2: $b_1 = 0.4, b_2 = -0.4$)

Эта формула отличается от (3.14) для нерекурсивной цепи второго порядка только обозначением параметров.

Резонансное значение функции $f^2(a)$ равно

$$f^2(a_p) = \beta = \frac{4AC - B^2}{4A} = C - \frac{B^2}{4A} = C - \frac{B^2}{4A^2} A = C - a_p^2 A.$$

Оно отличается от (3.15) только обозначением переменной.

Нормированная АЧХ выражается зависимостью

$$\frac{H(a)}{H(a_p)} = \sqrt{\frac{\beta}{Aa^2 + Ba + C}}$$

(сравните с формулой (3.16)).

Резонансные свойства ПФ и РФ характеризуются добротностью, вычисляемой по формуле, аналогичной (3.17), т. е.

$$Q = \frac{\omega_p}{\omega_2 - \omega_1} = \frac{\arccos a_p}{\arccos a_2 - \arccos a_1}.$$

Как и в нерекурсивной цепи второго порядка, здесь величины $\omega_{1,2}$ соответствуют границам полосы:

пропускания в ПФ по уровню $1/\sqrt{2}$ от максимального уровня,

задерживания (подавления) в РФ по уровню $\sqrt{2}$ от минимального уровня.

Воспользовавшись соотношением

$$\frac{\beta}{Aa_{1,2}^2 + Ba_{1,2} + C} = \frac{1}{2},$$

получим для ПФ

$$a_{1,2} = \frac{-B \pm \sqrt{4AC - B^2}}{2A}.$$

Правая часть этого выражения совпала с правой частью формулы (3.19) для РФ, реализуемого на базе нерекурсивной цепи второго порядка.

Из соотношения

$$\frac{\beta}{Aa_{1,2}^2 + Ba_{1,2} + C} = 2$$

для РФ получим

$$a_{1,2} = \frac{-B \pm \sqrt{\frac{B^2 - 4AC}{2}}}{2A}.$$

Правая часть этого выражения совпала с правой частью формулы (3.18) для ПФ, реализуемого на базе нерекурсивной цепи второго порядка. Заметим, что при $b_2=1$ имеем $f^2(a_p) = \beta = 0$.

По аналогии с (3.20) введем относительную расстройку:

$$\gamma = \frac{\omega}{\omega_p} = \frac{\arccos a}{\arccos a_p}.$$

На границах полосы пропускания значения γ равны:

$$\gamma_{1,2} = \frac{\omega_{1,2}}{\omega_p} = \frac{\arccos a_{1,2}}{\arccos a_p}.$$

Нормированная АЧХ через γ выражается следующим образом:

$$\frac{H(\gamma)}{H(1)} = \sqrt{\frac{\beta}{A \cos^2(\omega_p \gamma) + B \cos(\omega_p \gamma) + C}}.$$

Величина добротности выражается через γ функцией (3.21).

3.7.2. Временные характеристики

К временным характеристикам относятся: ИХ и ПХ, реакция на воздействие прямоугольного импульса (для ФНЧ и ФВЧ), реакция на воздействие прямоугольного радиоимпульса (для полосового и режекторного фильтров). Временные характеристики находим с помощью обратного z -преобразования произведения изображения в z -плоскости входного воздействия и передаточной функции (3.23).

Импульсная характеристика

Определим ИХ с помощью обратного z -преобразования функции (3.23):

$$\begin{aligned} h(n) &= \sum_i \operatorname{res} \left. \frac{z^2 z^{n-1}}{z^2 + b_1 z + b_2} \right|_{p_i} = \sum_{i=1}^2 \operatorname{res} \left. \frac{z^{n+1}}{(z-p_1)(z-p_2)} \right|_{p_i} = \\ &= \frac{p_1^{n+1}}{p_1 - p_2} + \frac{p_2^{n+1}}{p_2 - p_1} = \frac{p_1^{n+1} - p_2^{n+1}}{p_1 - p_2}. \end{aligned}$$

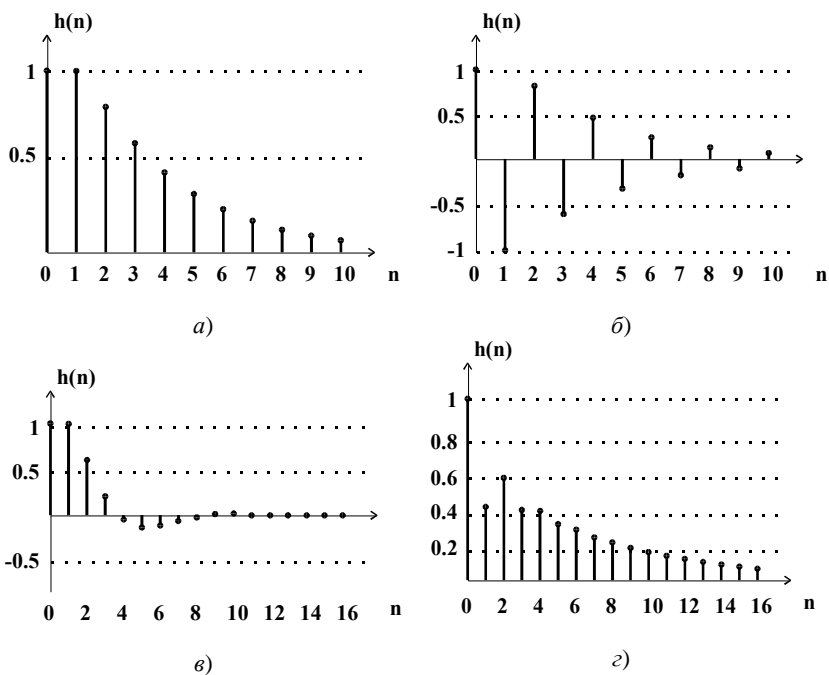


Рис. 3.36. Импульсные характеристики рекурсивной цепи второго порядка:
 а) фильтра нижних частот; б) фильтра верхних частот;
 в) полосового фильтра; з) режекторного фильтра

Здесь $p_i = p_{1,2}$ находятся из (3.24). Если $p_2 = p_1^*$, будем считать $p_{1,2} = -\alpha \pm j\omega_0$. Тогда

$$h(n) = 2\operatorname{Re} \frac{p_1^{n+1}}{p_1 - p_2} = 2\operatorname{Im} j \frac{p_1^{n+1}}{2j\omega_0} = \operatorname{Im} \frac{p_1^{n+1}}{\omega_0}.$$

На рис. 3.36а–3.36з приведены графики ИХ соответственно ФНЧ с параметрами $b_1 = -1$, $b_2 = 0.2$, ФВЧ с параметрами $b_1 = 1$, $b_2 = 0.2$, ПФ с параметрами $b_1 = -1$, $b_2 = 0.4$, РФ с параметрами $b_1 = -0.4$, $b_2 = -0.4$.

Переходная характеристика

Определим ПХ с помощью обратного z -преобразования произведения изображения единичного скачка и (3.23). В общем случае произвольных полюсов передаточной функции имеем

$$\begin{aligned}
 g(n) &= \sum_{i=1}^3 \operatorname{res} \frac{z \cdot z^2 \cdot z^{n-1}}{(z-p_3)(z-p_1)(z-p_2)} \Big|_{p_i} = \\
 &= \frac{1}{1+b_1+b_2} + \frac{p_1^{n+2}}{(p_1-p_3)(p_1-p_2)} + \frac{p_2^{n+2}}{(p_2-p_3)(p_2-p_1)} = \\
 &= \frac{1}{1+b_1+b_2} + \frac{1}{p_1-p_2} \left(\frac{p_1^{n+2}}{p_1-p_2} - \frac{p_2^{n+2}}{p_2-p_3} \right), \quad (3.28)
 \end{aligned}$$

где $p_3=1$. Если $p_2 = p_1^*$, то

$$\begin{aligned}
 g(n) &= \frac{1}{1+b_1+b_2} + 2 \operatorname{Re} \frac{p_1^{n+2}}{(p_1-1)(p_1-p_2)} = \\
 &= \frac{1}{1+b_1+b_2} + 2 \operatorname{Im} j \frac{p_1^{n+2}}{2j\omega_0(p_1-1)} = \frac{1}{1+b_1+b_2} + \operatorname{Im} \frac{p_1^{n+2}}{\omega_0(p_1-1)}.
 \end{aligned}$$

На рис. 3.37а–3.37г приведены графики ПХ соответственно ФНЧ, ФВЧ, ПФ и РФ с указанными параметрами.

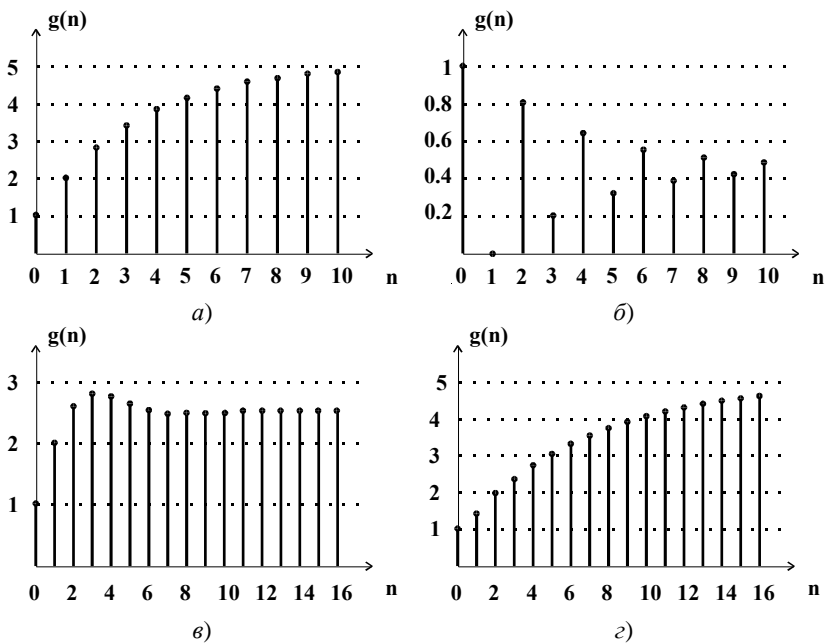


Рис. 3.37. Переходные характеристики рекурсивной цепи второго порядка:
a) фильтра нижних частот; *б)* фильтра верхних частот;
в) полосового фильтра; *г)* режекторного фильтра

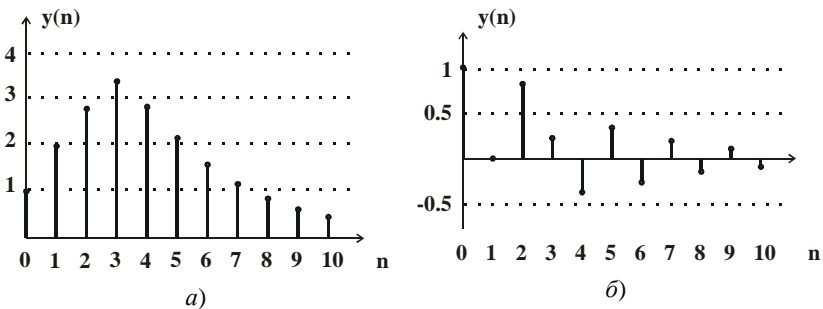


Рис. 3.38. Реакция на воздействие прямоугольного импульса:
a) фильтра нижних частот; *б)* фильтра верхних частот;
в) полосового фильтра; *г)* режекторного фильтра

Реакция фильтров нижних и верхних частот на воздействие прямоугольного импульса

Реакция ФНЧ и ФВЧ на воздействие прямоугольного импульса $x(n)=X[1(n)-1(n-n_u)]$ (при этом p_1 и p_2 вещественные) выражается следующим образом

$$\begin{aligned}
 y(n) &= X \sum_{i=1}^3 \operatorname{res}\left\{ Z[1(n)-1(n-n_u)] \frac{z^2 z^{n-1}}{(z-p_1)(z-p_2)} \right\} \Bigg|_{p_i} = \\
 &= X \sum_{i=1}^3 \operatorname{res}\left\{ Z[1(n)] \frac{z^{n+1}}{(z-p_1)(z-p_2)} \right\} \Bigg|_{p_i} - \\
 &\quad - X \sum_{i=1}^3 \operatorname{res}\left\{ Z[1(n-n_u)] \frac{z^{n+1}}{(z-p_1)(z-p_2)} \right\} \Bigg|_{p_i} = \\
 &= Xg(n)1(n) - Xg(n-n_u)1(n-n_u),
 \end{aligned}$$

где $g(n)$ находится из (3.34).

На рис. 3.38а и 3.38б приведены осциллограммы соответственно ФНЧ и ФВЧ с указанными параметрами на воздействие прямоугольного импульса длительностью $n_u=4$ при $X=1$.

Реакция полосового и режекторного фильтров на воздействие прямоугольного радиоимпульса

Прямоугольный радиоимпульс, воздействующий на ПФ и РФ, представляем в комплексной форме:

$$\begin{aligned}
 \dot{x}(n) &= X e^{j\omega n} [1(n) - 1(n-n_u)] = \\
 &= X e^{j\omega n} 1(n) - X e^{j\omega(n-n_u)} e^{j\omega n_u} 1(n-n_u) = \dot{x}_1(n) - \dot{x}_2(n).
 \end{aligned}$$

Здесь $\dot{x}_2(n) = e^{j\omega n_u} \dot{x}_1(n-n_u)$. Функция $x_1(n)$ имеет изображение

$$X_1(z) = X \sum_{n=0}^{\infty} e^{j\omega n} z^{-n} = X \frac{1}{1 - z^{-1} e^{j\omega}} = X \frac{z}{z - e^{j\omega}}$$

и полюс $p_3 = e^{j\omega}$.

Найдем реакцию $\dot{y}_1(n)$ на воздействие $\dot{x}_1(n)$ в общем случае:

$$\begin{aligned} \dot{y}_1(n) &= X \sum_{i=1}^3 \operatorname{res} \frac{z}{z - e^{j\omega}} \frac{z^2 z^{n-1}}{(z - p_1)(z - p_2)} \Big|_{p_i} = \\ &= X \left\{ \frac{e^{j\omega(n+2)}}{p_3^2 + b_1 p_3 + b_2} + \frac{p_1^{n+2}}{(p_1 - e^{j\omega})(p_1 - p_2)} + \right. \\ &\quad \left. + \frac{p_2^{n+2}}{(p_2 - e^{j\omega})(p_2 - p_1)} \right\} \mathbf{1}(n) = \\ &= X \left\{ \frac{e^{j\omega(n+2)}}{p_3^2 + b_1 p_3 + b_2} + \frac{1}{p_1 - p_2} \left[\frac{p_1^{n+2}}{p_1 - e^{j\omega}} - \frac{p_2^{n+2}}{p_2 - e^{j\omega}} \right] \right\} \mathbf{1}(n). \end{aligned}$$

Если $p_2 = p_1^*$ (это характерно для ПФ), то

$$\begin{aligned} \dot{y}_1(n) &= X \left\{ \frac{e^{j\omega(n+2)}}{p_3^2 + b_1 p_3 + b_2} + \frac{p_1^{n+2}}{(p_1 - e^{j\omega})(p_1 - p_1^*)} - \right. \\ &\quad \left. - \frac{p_1^{*(n+2)}}{(p_1^* - e^{j\omega})(p_1 - p_1^*)} \right\} \mathbf{1}(n) = \\ &= X \left\{ \frac{e^{j\omega(n+2)}}{p_3^2 + b_1 p_3 + b_2} + \frac{1}{2j\omega_0} \left[\frac{p_1^{n+2}}{p_1 - e^{j\omega}} - \frac{p_1^{*(n+2)}}{p_1^* - e^{j\omega}} \right] \right\} \mathbf{1}(n). \end{aligned}$$

Соответственно реакция $\dot{y}_2(n)$ на воздействие $\dot{x}_2(n)$ имеет вид

$$\dot{y}_2(n) = e^{j\omega n_u} \dot{y}_1(n - n_u).$$

Полная реакция цепи равна

$$\dot{y}(n) = \dot{y}_1(n) - \dot{y}_2(n),$$

а реальный сигнал

$$y(n) = \text{Im } \dot{y}(n).$$

Модуль его огибающей выражается следующим образом

$$Y(n) = |\dot{y}(n)|.$$

Осциллограмма модуля огибающей реакции ПФ с указанными параметрами на воздействие прямоугольного радиоимпульса длительностью $n_u=8$ при $X=1$ на резонансной частоте фильтра показана на рис. 3.39.

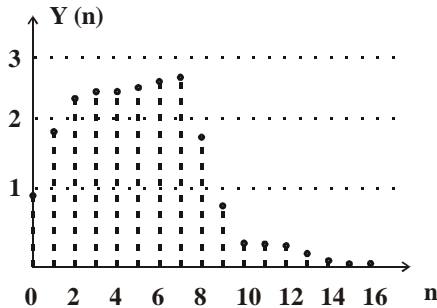


Рис. 3.39. Модуль огибающей реакции полосового фильтра на воздействие прямоугольного радиоимпульса

3.8. Контрольные вопросы

1. Дайте определение рекурсивной и нерекурсивной цепи.
2. Дайте определение устойчивости цифровой цепи.
3. Докажите условия устойчивости нерекурсивной, рекурсивной и рекурсивно-нерекурсивной цепей первого и второго порядков.
4. Дайте определение передаточной функции цифровой цепи.
5. Получите для передаточной функции выражение по известному разностному уравнению.
6. По передаточной функции запишите разностное уравнение.
7. С помощью нуль-полусных диаграмм объясните 4 типа фильтров, реализуемых на базе рекурсивной цепи второго порядка. Покажите соответствующую картину на плоскости параметров цепи (b_1, b_2).
8. Дайте определение и выведите формулы АЧХ рекурсивных и нерекурсивных цепей второго порядка. Приведите графики для нескольких случаев.
9. Дайте определение и выведите формулы ФЧХ рекурсивных и нерекурсивных цепей второго порядка. Приведите графики для нескольких случаев.
10. В чем отличие АЧХ рекурсивной и нерекурсивной цепей?
11. Каково влияние рекурсивной (нерекурсивной) части цепи на АЧХ?
12. В чем отличие ФЧХ рекурсивной и нерекурсивной цепей?
13. Каково влияние рекурсивной (нерекурсивной) части цепи на ФЧХ?
14. Изобразите нуль-полусную диаграмму рекурсивных и нерекурсивных цепей второго порядка, если известна передаточная функция (частотная характеристика).
15. Дайте определение резонансной частоты ПФ и РФ. Выведите формулу ω_r для рекурсивной цепи второго порядка. Покажите область существования ω_r . Как зависит ω_r от параметров цепи?
16. Дайте определение добротности ПФ и РФ.
17. Выведите формулу Q для рекурсивной цепи второго порядка.
18. Покажите область существования Q .

19. Как зависит Q от параметров цепи?
20. Сравните частотные свойства нерекursивных и recursивных цепей второго порядка.
21. Дайте определение ИХ цепи.
22. Вычислите несколько отсчетов по разностным уравнениям recursивных и нерекursивных цепей второго порядка. Постройте графики.
23. В чем принципиальное отличие ИХ recursивной и нерекursивной цепей?
24. Какова связь между ИХ и ЧХ цепи?
25. Дайте определение ПХ цепи. Как определить ее, если известна передаточная функция цепи?
26. В чем принципиальное отличие ПХ recursивной и нерекursивной цепей?
27. Какова связь между ИХ и ПХ цепи?
28. По одной из известных характеристик
 - разностное уравнение,
 - передаточная функция,
 - нуль-полусная диаграмма,
 - структурная схема
 получите все остальные.
29. Объясните зависимость искажений прямоугольного импульса в исследуемых цепях от параметров цепей.
30. Объясните связь между ЧХ, ИХ и ПХ исследуемых цепей и искажениями прямоугольного импульса.
31. Покажите области существования частоты среза recursивных и нерекursивных цепей второго порядка.
32. Изобразите графики зависимости частоты среза от параметра нерекursивной и recursивной цепей.

3.9. Литература

1. Брюханов, Ю. А. Цифровые цепи и сигналы: учеб. пособие / Ю. А. Брюханов. – Ярославль: ЯрГУ, 2005.
2. Солонина, А. И. Основы цифровой обработки сигналов: курс лекций / А. И. Солонина и др. – СПб.: БХВ-Петербург, 2005.

3. Рабинер, Л. Теория и применение цифровой обработки сигналов: пер. с англ. / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978.
4. Баскаков, С. И. Радиотехнические цепи и сигналы / С. И. Баскаков. – М.: Высшая школа, 2005.
5. Гоноровский, И. С. Радиотехнические цепи и сигналы / И. С. Гоноровский, М. П. Демин. – М.: Радио и связь, 1994.
6. Карташев, В. Г. Основы теории дискретных сигналов и цифровых фильтров / В. Г. Карташев. – М.: Высшая школа, 1982.
7. Гольденберг, Л. М. Цифровая обработка сигналов: учеб. пособие / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1990.

4. ЭФФЕКТЫ КВАНТОВАНИЯ В ЦИФРОВЫХ СИГНАЛАХ

В реальных условиях технической реализации значения цифровых сигналов квантуются по уровню. Это обусловлено конечным числом разрядов в представлении чисел. Рассмотрим вопросы представления чисел, квантования сигналов, аналогово-цифрового и цифроаналогового преобразования сигналов.

4.1. Представление чисел

Обычно применяются два представления – с фиксированной и плавающей запятой. При фиксированной запятой возможно несколько вариантов представления, отличающихся положением запятой (числа, выравненные справа и слева) и способом представления отрицательных чисел.

Когда используется форма с выравниванием справа, каждое число представляется в виде последовательности положительных степеней числа два. Для пояснения рассмотрим арифметическое представление с помощью четырех двоичных разрядов (3 – для величины, 1 – для знака). Любая из $8=2^3$ последовательностей трех двоичных разрядов, имеющих положительный знак (знаковый разряд равен нулю), представляет число, значение которого можно выразить формулой

$$X = \sum_{i=0}^{M-1} d_i 2^i,$$

где d_i – двоичная переменная, которая может принимать значения только 0 и 1, а $M=3$ – количество двоичных разрядов, т. е. число находится в интервале $[0, 2^M-1]$ с шагом 1.

При фиксированной запятой и выравнивании слева, когда предполагается, что запятая находится слева от старшего значащего разряда, каждая из последовательностей представляет число, значение которого равно

$$X = \sum_{i=1}^M d_i 2^{-i},$$

т. е. число находится в интервале $[0, 1-2^{-M}]$ с шагом 2^{-M} . Итак, можно сказать, что выравненные справа числа с фиксированной запятой являются целыми, а выравненные слева – дробными.

Основное различие между представлениями, отличающимися положением запятой, определяется выполнением операции умножения. Результат умножения двух дробей есть также дробь, и поэтому невозможно превысить длину регистров, если в процессе выполнения операции отбрасывать младшие значащие разряды. Очевидно, что переполнение при выполнении операции сложения может иметь место при обеих формах представления. Поэтому при реализации цифровых систем с фиксированной запятой вопросам переполнения должно уделяться особое внимание.

При представлении числа с плавающей запятой последовательность соответствующих двоичных разрядов делится на две последовательности. Первая, называемая мантиссой, является, как правило, выравненным слева числом с фиксированной запятой и знаком, которое представляет собой нормализованное значение числа. Вторая, называемая порядком, представляет собой показатель (со знаком) степени (основание степени – два), на который надо умножить мантиссу, чтобы получить значение числа. В этом случае число задается соотношением

$$X = \pm m 2^{\pm l}.$$

Представление с плавающей запятой более удобно, когда используется большой диапазон представления чисел, поскольку несколько дополнительных разрядов порядка можно использовать для согласования больших диапазонов представления чисел.

Другой вопрос, подлежащий обсуждению – представление отрицательных чисел. Более простым является представление в виде абсолютного значения со знаком (т. е. в прямом коде со знаком). В этом случае один из разрядов предназначается для

знака (обычно 0 – положительный знак, 1 – отрицательный), остальные же образуют абсолютное значение числа. Другой способ представления основан на принципе дополнения: представление в виде обратного и дополнительного кодов.

Рассмотрим представление числа в виде дополнительного кода как наиболее часто используемое. В этом случае положительные числа представляются так же, как и обычно, т. е. абсолютным значением со знаком (знак 0), тогда как отрицательные целые числа выражаются в виде

$$X_{omp} = 2^M - \sum_{i=0}^{M-1} d_i 2^i, \quad (4.1)$$

а дробные числа

$$X_{omp} = 1 - \sum_{i=1}^M d_i 2^{-i}, \quad (4.2)$$

где M – число двоичных разрядов системы.

Поэтому, например, в системе с четырьмя разрядами, рассмотренной выше, число 1 представляется в виде 0001, тогда как число -1 – как $2^M - 1$, т. е. в виде 1111. Можно отметить, что в этом случае отличительным признаком отрицательных чисел является знаковый разряд 1.

Можно показать, что такое представление является циклическим, т. е. если начнем с нуля и будем прибавлять 1 к каждому предшествующему числу, получим все положительные числа от 0 до $2^M - 1$. Прибавив еще 1, получим отрицательное число, поскольку знаковый разряд станет равным 1, а значение числа будет соответствовать -2^M . Продолжая прибавлять единицу, получаем все отрицательные числа вплоть до -1 , а затем получим 0, т. е. число, с которого была начата последовательность (для дробных чисел см. рис. 4.1).

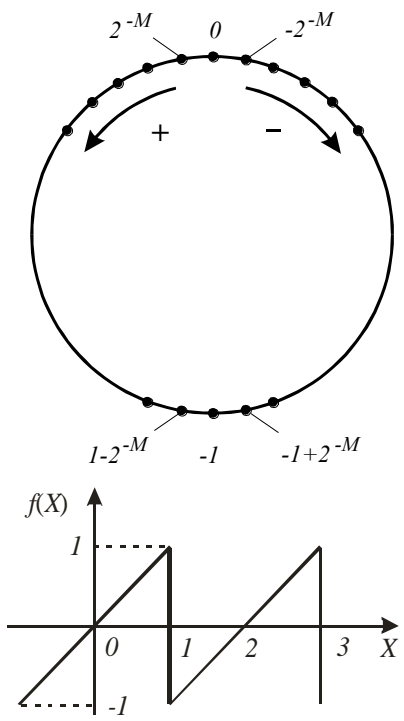


Рис. 4.1. Представление чисел в виде дополнительного кода

Следствием этого является важное свойство такой системы записи чисел, а именно если результат сложения более чем двух положительных и отрицательных чисел таков, что его можно представить имеющимся количеством двоичных разрядов, то этот результат является точным независимо от порядка выполнения операций сложения, даже если в процессе этих операций имело место переполнение.

Другим важным свойством представления в дополнительном коде является то, что результат вычитания можно получить путем сложения уменьшаемого с дополнительным кодом вычитаемого.

Действительно, рассмотрим, например, целые числа: сложение X_1 с дополнительным кодом X_2 , определенным в виде (4.1), соответствует получению числа $X_1 + 2^M - X_2$, а поскольку представление ведется по модулю 2^M , то это эквивалентно получению разности $X_1 - X_2$.

Таким образом, используя арифметическое представление в виде дополнительного кода, можно выполнять как сложение, так и вычитание с применением только сумматоров. Открытым остается пока вопрос о вычислении дополнительного кода числа. Из определений (4.1) и (4.2) следует, что для получения дополнительного кода числа необходимо выполнить вычитание.

Но существуют более эффективные методы. Например, дополнительный код числа можно получить, взяв обратные значения всех двоичных разрядов (включая знаковый) и прибавив единицу к полученному результату.

Из двух рассмотренных способов представления чисел способ с фиксированной запятой проще с точки зрения технической реализации. Просто реализуется сложение двух чисел, и можно получить умножение без переполнения, если для представления с фиксированной запятой используются дроби. Основные сложности применения арифметики такого типа при реализации цифровых устройств связаны с ограниченным динамическим диапазоном представления. Поэтому следует особое внимание уделять задаче масштабирования переменных, используемых в структуре фильтра, чтобы избежать переполнений.

Используя представление с плавающей запятой, можно избежать некоторых трудностей, связанных с ограничениями динамического диапазона, выделяя для этого, самое большее, некоторые дополнительные разряды в представлении. Однако реализация арифметики такого типа является более трудной задачей, и такая арифметика более медленная, чем с фиксированной запятой. Действительно, чтобы сложить два числа в системе с плавающей запятой, необходимо сначала нормализовать их, чтобы выравнять порядки, а затем выполнить сложение. Если необходимо, результат сложения следует еще раз нормализовать. Кроме того, умножение с плавающей запятой в действительности является объединением двух арифметических операций, а именно перемножения мантисс и сложения порядков.

Важно также отметить, что используя представление с фиксированной запятой, путем подходящего масштабирования можно избежать опасности переполнения в операциях сложения и ошибка будет вноситься лишь при умножении. Между тем, если используется представление с плавающей запятой, ошибка будет вноситься также в операциях сложения, поскольку в процессе нормализации некоторые разряды одного из чисел теряются.

Увеличение сложности арифметики с плавающей запятой часто приводит к выбору арифметики с фиксированной запятой

при разработке технических средств для выполнения операций цифровой фильтрации. В этом случае доступный выигрыш в объеме технических средств, емкости памяти и повышения быстродействия может оказаться более важным, чем усложнение работы по синтезу фильтра, которая в этом случае включает, во-первых, выбор наилучшей структуры фильтра, во-вторых, вычисление ограничений в случае переполнения и, в-третьих, исследование поведения ошибки для выбранной структуры.

4.2. Усечение и округление

При любом представлении чисел X' вносится ошибка $e = X - X'$, являющаяся следствием аппроксимации переменных конечным числом двоичных разрядов. Эта ошибка зависит от числа используемых при таком представлении разрядов M , способа представления (с фиксированной или плавающей запятой), способа представления отрицательных чисел и, наконец, метода, используемого для исключения разрядов, превышающих длину регистров.

Для выполнения последнего действия можно рассмотреть два различных метода. Первый состоит в простом отбрасывании лишних разрядов, второй – в аппроксимации на низшем уровне, а именно: число принимает то же значение, что и в первом методе, если отбрасываемая дробь меньше половины цены младшего разряда представления или же увеличивается на единицу младшего разряда в противном случае. Первый метод аппроксимации называется усечением, а второй – округлением.

Округление (см. рис. 4.2, где Q – шаг квантования, $Q = 2^{-M}$ при использовании дробных чисел, $Q = 1$ при использовании целых чисел), очевидно, является более точным способом, чем усечение, поскольку при округлении максимальная ошибка равна половине шага квантования. Поэтому, если число представляется в виде дроби из M двоичных разрядов (без знака), ошибка аппроксимации будет находиться в пределах

$$-Q/2 < e \leq Q/2. \quad (4.3)$$

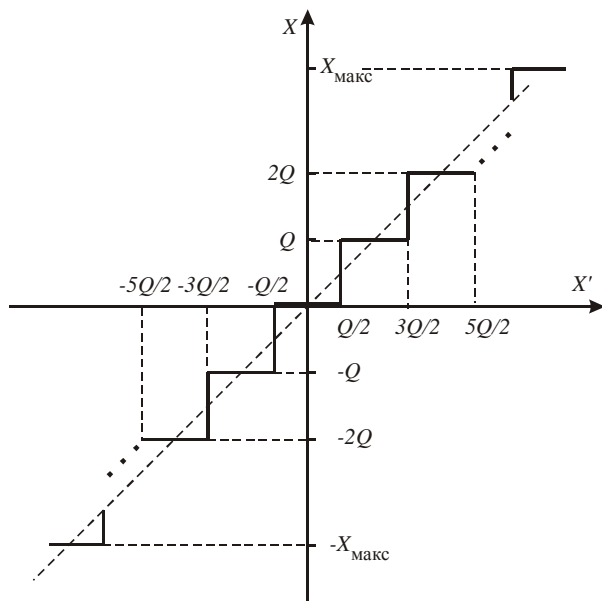


Рис. 4.2. Характеристика квантователя с округлением

При усечении максимальная ошибка больше и равна шагу квантования, а её значение (по знаку) зависит от представления отрицательных чисел. Если используется представление в прямом коде со знаком (рис. 4.3), то ошибка будет определяться как разность между полученным и истинным значениями и будет изменяться в пределах от 0 до $-Q$, если число положительно, и в пределах от Q до 0, если число отрицательно, т. е. знак ошибки всегда противоположен знаку числа. Если используется представление в дополнительном коде (см. рис. 4.4), ошибка не связана со знаком квантуемого числа и всегда отрицательна при изменении в данном диапазоне $[0, -Q]$.

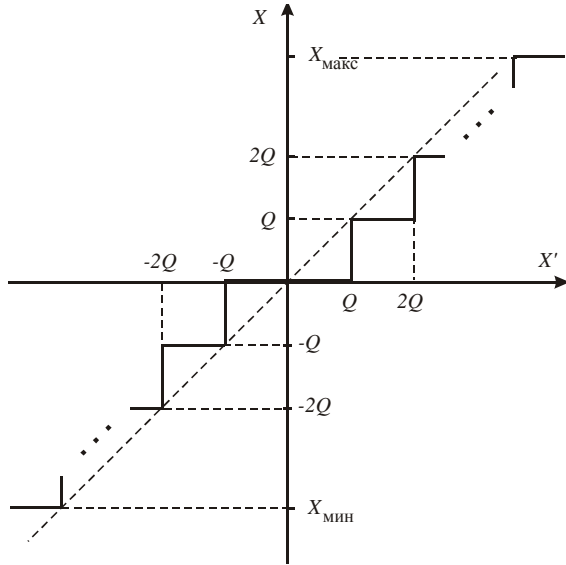


Рис. 4.3. Характеристика квантователя с усечением и представлением чисел в прямом и обратном кодах

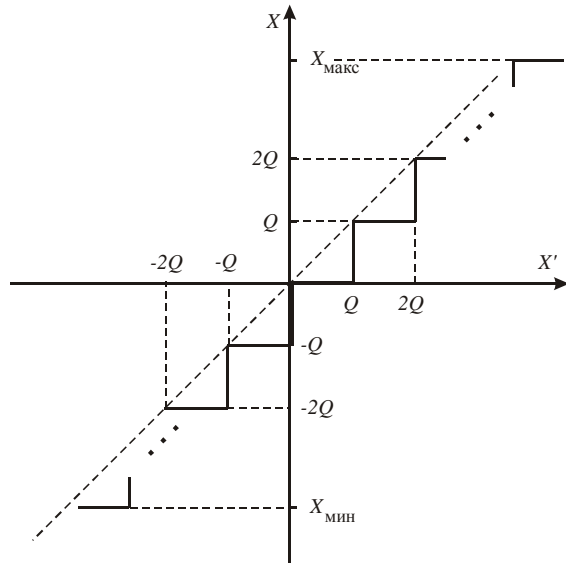


Рис. 4.4. Характеристика квантователя с усечением и представлением чисел в дополнительном коде

Если выбрано представление с плавающей запятой, ситуация усложняется. В этом случае аппроксимация (усечение или округление) выполняется только для мантиссы числа. Это означает, что абсолютное значение ошибки зависит от числа, т. е. в этом случае важна относительная ошибка, а не абсолютная. Следовательно, в случае с плавающей запятой ошибки, вносимые аппроксимацией, мультипликативны, а не аддитивны. Потому если α' – аппроксимируемое число, α – полученное число с плавающей запятой, ε – относительная ошибка, то

$$(\alpha - \alpha') / \alpha' = \varepsilon \quad (4.4)$$

и, следовательно,

$$\alpha = \alpha' + \varepsilon \alpha' = \alpha' (1 + \varepsilon).$$

Если используется, например, округление, ошибка в мантиссе будет определяться формулой (4.3) и поэтому

$$-2^l 2^{-M} / 2 < \alpha - \alpha' < 2^l 2^{-M} / 2,$$

где l – порядок. Используя (4.4) и неравенство $2^{l-1} < \alpha' < 2^l$, можно получить границы для ε :

$$-2^{-M} < \varepsilon \leq 2^{-M}.$$

4.3. Квантование сигнала

Операцию квантования сигнала можно рассматривать как представление его выборок M двоичными разрядами одним из рассмотренных в пункте 4.2 способов. Поэтому каждая квантованная выборка может быть представлена в виде

$$x(n) = x'(n) + e(n),$$

где $x'(n)$ – истинное значение выборки, а $e(n)$ – соответствующая ей ошибка.

Чтобы продолжить анализ, необходимо выдвинуть некоторые предположения относительно поведения этих ошибок. Принимается статистическая модель, основанная на следующих четырех предположениях: во-первых, ошибка $e(n)$ является последовательностью выборок стационарного случайного процесса; во-вторых, значения ошибки не коррелированы с точными значениями сигнала; в-третьих, значения ошибки не коррелированы между собой (представляют собой белый шум), и в-четвертых, распределение вероятностей ошибки равномерно во всем диапазоне ошибок квантования.

Разумеется, эти предположения, до некоторой степени, произвольны и легко можно построить примеры, для которых эти предположения неверны, например постоянный сигнал или синусоидальный сигнал, дискретизированный с частотой, рационально кратной частоте синусоиды. В первом случае все ошибки $e(n)$ будут одинаковы и равны, а во втором – они образуют периодическую последовательность, поэтому в обоих случаях выдвинутые предположения будут неверны. Однако для большинства сигналов, с которыми приходится встречаться, эти предположения практически применимы. Поэтому эта модель ошибки квантования имеет огромное практическое значение. Можно сказать, что модель можно использовать, когда поведение сигнала таково, что при переходе от одной выборки к другой его значение меняется на несколько уровней квантования и в то же время число уровней квантования, или, что то же самое, число разрядов в представлении числа не очень мало.

На рис. 4.5 показаны плотности распределения, полученные на основании четвертого предположения; на рис. 4.5а – для ошибки округления, на рис. 4.5б – для ошибки усечения с представлением отрицательных чисел в дополнительном коде, а на рис. 4.5в – для ошибки усечения с представлением отрицательных чисел в прямом и обратном кодах. Для этих распределений вероятности легко вычислить среднее значение и дисперсию ошибки для округления:

$$m_e=0, \quad (4.5)$$

$$\sigma_e^2 = \int_{-Q/2}^{Q/2} e^2 \frac{1}{Q} de = \frac{1}{3Q} \left(\frac{Q^3}{8} + \frac{Q^3}{8} \right) = \frac{Q^2}{12}, \quad (4.6)$$

для усечения с представлением чисел в дополнительном коде:

$$m_e = -Q/2, \quad (4.7)$$

$$\sigma_e^2 = Q^2/12, \quad (4.8)$$

для усечения с представлением чисел в прямом и обратном кодах:

$$m_e = 0, \quad (4.9)$$

$$\sigma_e^2 = Q^2/3. \quad (4.10)$$

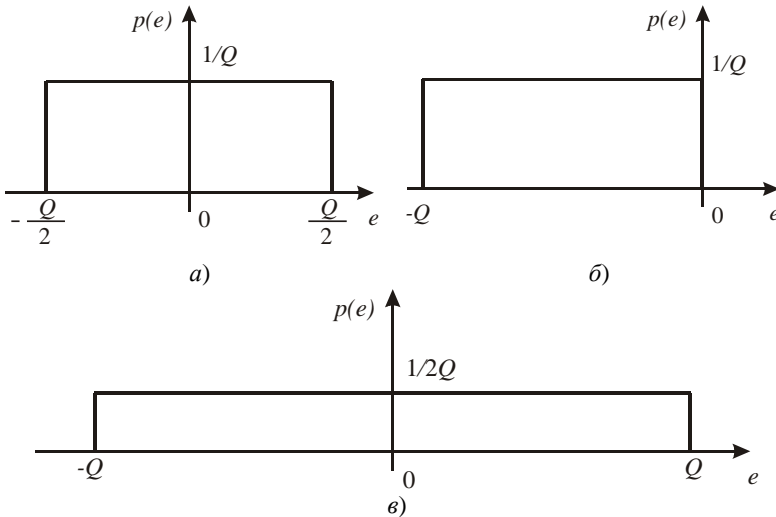


Рис. 4.5. Плотности распределения вероятностей ошибок квантования: а) при округлении; б) при усечении с представлением отрицательных чисел в дополнительном коде; в) при усечении с представлением чисел в прямом и обратном кодах

Из предшествующего следует, что процесс квантования можно рассматривать как наложение на сигнал $x(n)$, заданный с

бесконечной точностью, шума $e(n)$, имеющего средние значения и дисперсии, определяемые выражениями (4.5) – (4.10) (см. рис. 4.6).

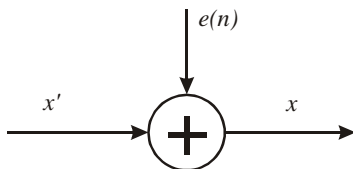


Рис. 4.6. Представление процесса квантования сигнала

Вклад этой ошибки можно условно измерять в виде отношения мощности сигнала к мощности шума (отношения сигнал/шум). Например, в случае округления дробных чисел можно написать

$$c / \omega = \sigma_x^2 / \sigma_e^2 = \sigma_x^2 / (2^{-2M} / 12) = 12(2^{2M} \sigma_x^2);$$

используя логарифмическую меру, получаем

$$c / \omega (\text{дБ}) = 10 \lg 12 + M 20 \lg 2 + 10 \lg \sigma_x^2.$$

Интересно отметить, что, поскольку $20 \lg 2 \approx 6$, добавление одного разряда в арифметике увеличивает отношение сигнал/шум приблизительно на 6 дБ.

Описанный процесс квантования нуждается в обсуждении выбора шага квантования, используемого в приложениях. Этот выбор определяется необходимой точностью представления, наличием шума в обрабатываемом сигнале и процедурой, которая применяется для обработки сигнала.

Содержащийся в сигнале шум определяет верхнюю границу числа уровней квантования. Очевидно, нет особого смысла использовать малый шаг квантования, когда в сигнале содержится большой шум, т. к. в этом случае будем квантовать шум, а не сигнал. Достаточно выбрать столько уровней, чтобы вклад шума квантования был мал по сравнению с шумом, содержащимся в сигнале.

В противном случае, когда шум очень мал, шаг квантования следует выбирать таким образом, чтобы получить желаемое

качество выходного сигнала. Фактически ухудшение качества входного сигнала может быть вызвано ошибками, возникающими в процессе обработки. Итак, квантование сигнала и структуру процесса обработки в части используемых алгоритмов, точности представления и т. д. следует выбирать так, чтобы они согласовались с требованиями к точности обработки сигналов.

4.4. Аналогово-цифровое преобразование

Одним из наиболее важных способов формирования цифрового сигнала является дискретизация и квантование непрерывного колебания. Устройство, предназначенное для преобразования непрерывного колебания в последовательность отсчетов, каждый из которых является аппроксимацией соответствующего отсчета входного колебания, называется аналогово-цифровым преобразователем (АЦП).

Приведем структурную схему АЦП (рис. 4.7).

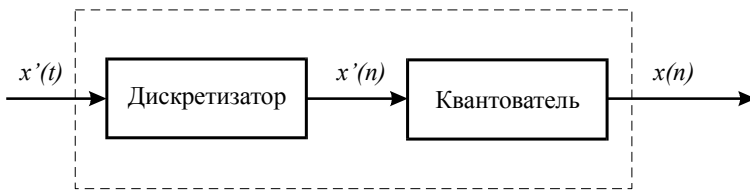


Рис. 4.7. Структурная схема аналогово-цифрового преобразователя

Его работу можно представить в виде двухэтапного процесса. На первом этапе формируется последовательность $x'(n) = x(t)|_{t=nT}$, в которой отсчеты $x'(n)$ представлены с неограниченной точностью. На втором этапе значение каждого отсчета $x'(n)$ представляется числом, состоящим из конечного числа двоичных разрядов. В результате получается новая последовательность $x(n)$ – цифровой сигнал.

В реальных АЦП обе операции выполняются совместно, т. е. имеется единый блок, на вход которого поступает колебание $x(t)$, а на выходе формируется сигнал $x(n)$. Разность $e(n) = x(n) - x'(n)$

называется шумом квантования или шумом аналогово-цифрового преобразования.

Как было отмечено в разделе 2, полоса входного колебания должна быть ограничена, т. к. иначе сигнал $x(n)$ не будет однозначно представлять $x(t)$. Поэтому АЦП предшествует аналоговый ФНЧ для видеосигналов или ПФ для радиосигналов.

4.5. Цифроаналоговое преобразование сигналов

Цифроаналоговое преобразование необходимо для преобразования цифрового сигнала $x(n)=x(nT)$ в аналоговый $x(t)$. Покажем теоретически возможность такой операции.

Из рис. 2.4б ясно видно, что если $\Omega \leq \pi/T$, т. е. если частота дискретизации, по крайней мере, вдвое больше наивысшей частоты спектра $X_H(j\Omega)$, то $X(e^{j\omega T})$ совпадает с $X_H(j\Omega)$ на интервале $-\pi/T \leq \omega \leq \pi/T$. Эта частота дискретизации иногда называется частотой Найквиста (Котельникова). В этом случае можно ожидать, что функция $x(t)$ может быть восстановлена по выборкам $x(nT)$ при помощи подходящей интерполяционной формулы.

Чтобы вывести интерполяционную формулу, предположим, что $\Omega \leq \pi/T$ (рис. 2.4б), тогда согласно (2.10) имеем

$$X(e^{j\Omega T}) = \frac{1}{T} X_H(j\Omega), \quad -\pi/T \leq \Omega \leq \pi/T.$$

В соответствии с преобразованием Фурье в непрерывном времени

$$x(t) = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} X_H(j\Omega) e^{j\Omega t} d\Omega.$$

Объединяя два последних выражения, можно записать

$$x(t) = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} TX(e^{j\Omega T}) e^{j\Omega t} d\Omega.$$

Так как согласно (1.2) (дискретный ряд Фурье) имеем

$$X(e^{j\Omega T}) = \sum_{n=-\infty}^{\infty} x(nT) e^{-j\Omega nT},$$

то

$$x(t) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \left(\sum_{n=-\infty}^{\infty} x(nT) e^{-j\Omega nT} \right) e^{j\Omega t} d\Omega$$

или, изменяя порядок суммирования и интегрирования,

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT) \left[\frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} e^{j\Omega(t-nT)} d\Omega \right].$$

Выполнив интегрирование, получим

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT) \left[\frac{\sin\left(\frac{\pi}{T}(t-nT)\right)}{\frac{\pi}{T}(t-nT)} \right]. \quad (4.11)$$

Выражение (4.11) дает интерполяционную формулу для восстановления аналогового сигнала $x(t)$ по его выборкам. Заметим, что это выражение описывает свертку входного сигнала и импульсной характеристики ФНЧ с прямоугольной АЧХ. Представление аналогового сигнала в виде (4.11) справедливо только для функций с ограниченным спектром при достаточно малом T , т. е. при отсутствии эффекта наложения.

Выражение (4.11) можно понимать как разложение сигнала с непрерывным временем в ряд вида

$$x(t) = \sum_{n=-\infty}^{\infty} C_n \Phi_n(t), \quad (4.12)$$

где коэффициенты C_n и функции $\Phi_n(t)$ определяются выражениями

$$C_n = x(nT) \quad (4.13a)$$

и

$$\Phi_n(t) = \frac{\sin\left(\frac{\pi}{T}(t - nT)\right)}{\frac{\pi}{T}(t - nT)}. \quad (4.13б)$$

Существует много классов функций $\Phi_n(t)$, которые можно использовать для представления функции с непрерывным временем в виде (4.12), включая синусоидальные функции, функции Лагерра и полиномы Лежандра. При любом представлении вида (4.12) последовательность коэффициентов C_n можно рассматривать как дискретный сигнал, представляющий аналоговый сигнал $x(t)$. Однако не все такие представления в равной степени полезны. Большим преимуществом выбора функции $\Phi_n(t)$ вида (4.13б) является то, что коэффициенты легко получаются путем дискретизации сигнала с непрерывным временем.

Цифроаналоговое преобразование поясняется на рис. 4.8 – 4.9. Входной сигнал $x(nT)$ через предварительный фильтр с ИХ $g(n)$ поступает на вход цифроаналогового преобразователя (ЦАП). Это устройство содержит интерполятор нулевого порядка, преобразующий цифровой сигнал $\hat{x}(nT)$ (см. рис. 4.9а) в некоторый аналоговый сигнал $\hat{x}(t)$ (см. рис. 4.9б), который после прохождения через аналоговый ФНЧ преобразуется в нужный сигнал $x(t)$.

Преобразование цифрового сигнала $\hat{x}(nT)$ в аналоговый $\hat{x}(t)$ в интерполяторе нулевого порядка сопряжено с искажением спектра. Частотная характеристика этого интерполятора выражается функцией

$$H(\omega) = \frac{2 \sin \frac{\omega}{2}}{\omega}.$$

Чтобы скомпенсировать эти искажения спектра, цифровой сигнал $x(nT)$ часто предварительно пропускают через ЦФ с ЧХ,

описывающейся функцией $G(e^{j\omega})=1/H(\omega)$. Таким образом, последовательное соединение устройств с ЧХ $G(e^{j\omega})$ и $H(\omega)$ в целом обеспечивает равномерную ЧХ.

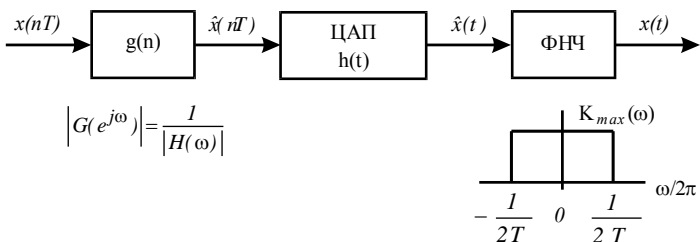


Рис. 4.8. Структурная схема цифроаналогового преобразования сигналов

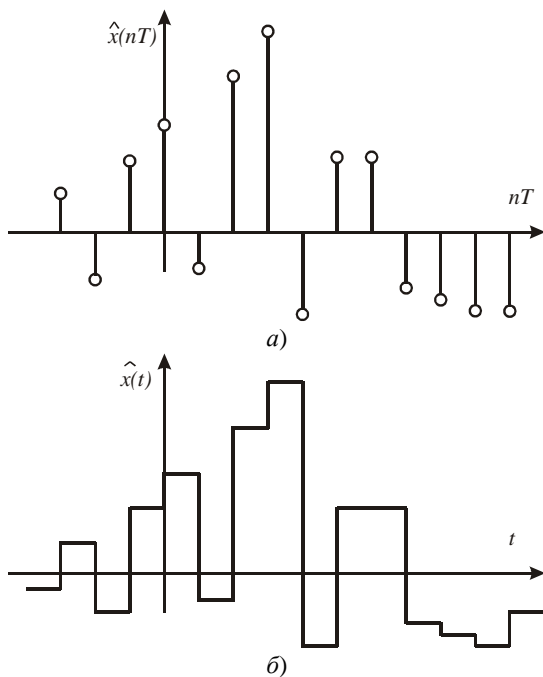


Рис. 4.9. Сигналы цифроаналогового преобразователя: *a)* входной; *б)* выходной

Следует иметь в виду, что цифровой компенсирующий фильтр, показанный на рис. 4.8, не обязательно выполнять в виде отдельного устройства; его можно включить в состав самой системы ЦОС.

4.6. Контрольные вопросы

1. Формы представления чисел в цифровых системах.
2. Преимущества и недостатки различных форм представления чисел.
3. Специфика арифметических операций с числами, представленными в разных формах.
4. Кодирование чисел при аналогово-цифровом преобразовании.
5. Особенности двоичных кодов (прямой, обратный и дополнительный).
6. Аппроксимация чисел способом округления. Характеристика квантователя. Плотность вероятности ошибки округления.
7. Аппроксимация чисел способом усечения. Характеристика квантователя для разных кодов. Плотность вероятности ошибки усечения.
8. Статистическая модель ошибки квантования сигнала.
9. Предположения, положенные в основу статистической модели.
10. Статистические характеристики шума квантования для различных кодов и способов квантования.

4.7. Литература

1. Брюханов, Ю. А. Цифровые цепи и сигналы: учеб. пособие / Ю. А. Брюханов. – Ярославль: ЯрГУ, 2005.
2. Солонина, А. И. Основы цифровой обработки сигналов: курс лекций / А. И. Солонина и др. – СПб.: БХВ-Петербург, 2005.
3. Рабинер, Л. Теория и применение цифровой обработки сигналов: пер. с англ. / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978.
4. Оппенгейм, А. В. Цифровая обработка сигналов: пер. с англ. / А. В. Оппенгейм, Р. В. Шафер. – М.: Связь, 1979.

5. Каппелини, В. Цифровые фильтры и их применение: пер. с англ. / В. Каппелини, А. Дж. Константиноидис, П. Эмилиани. – М.: Энергоатомиздат, 1983.
6. Карташев, В. Г. Основы теории дискретных сигналов и цифровых фильтров / В. Г. Карташев. – М.: Высшая школа, 1982.
7. Гольденберг, Л. М. Цифровая обработка сигналов: справочник / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1985.
8. Гольденберг, Л. М. Цифровая обработка сигналов: учеб. пособие / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1990.

5. ЭФФЕКТЫ КВАНТОВАНИЯ В ЦИФРОВЫХ ЦЕПЯХ

Цифровые цепи отличаются от дискретных конечной точностью представления (квантованием) коэффициентов и результатов арифметических операций, выполняемых в них. При реализации цифровых цепей точность коэффициентов (количество двоичных разрядов) должна быть согласована с длиной машинного слова или с разрядностью специальных вычислительных средств, используемых в операциях фильтрации.

Некоторое приближение появляется также в процессе выполнения операций фильтрации. Действительно, реализация алгоритма работы цифровой цепи состоит в последовательности умножений и сложений, и, когда два числа умножаются, результат имеет большую длину, чем операнды. Поскольку очевидно, что практически нецелесообразно конструировать аппаратные средства с увеличенной длиной регистров, необходимо вводить некоторую аппроксимацию результатов таких операций.

Задача исследования поведения ошибки в операциях цифровой фильтрации очень сложна и является существенно нелинейной, поскольку умножители, сумматоры, АЦП, выполняющие описанное выше квантование, представляют собой нелинейные устройства. Но если квантование не очень грубое, как для большинства сигналов, встречающихся в приложениях фильтрации, то можно использовать линейную модель поведения системы и статистическую модель ошибок.

В настоящем разделе рассматриваются только те нелинейные эффекты, которые связаны, во-первых, с явлением предельного цикла, которое может возникнуть, когда входной сигнал фильтра становится равным нулю, и, во-вторых, с пульсациями, вызванными переполнением.

5.1. Реакция идеальной цифровой цепи на шум квантования

В п. 4.3 описан процесс квантования и рассмотрена модель шума. Теперь рассмотрим, что произойдет, если шум квантования будет подан на вход цифровой цепи. В этом случае будем

предполагать, что арифметические операции в процедуре фильтрации выполняются с бесконечной точностью. Такой подход правомочен, т. к. рассмотренная в п. 4.3 модель линейна. Это означает, что различные источники ошибок в линейных системах можно изучать отдельно.

Влияние входного шума квантования на выход цепи легко определить путем вычисления реакции цепи на входной шумовой сигнал $e(n)$. Поэтому если $h(n)$ – ИХ линейной системы, а последовательность ошибки $e(n)$ является ее входом, то выход системы задается выражением

$$\varepsilon(n) = \sum_{m=0}^n h(m) \cdot e(n-m).$$

Дисперсию величины $\varepsilon(n)$ можно записать в следующем виде:

$$\begin{aligned} \sigma_0^2(n) &= E \cdot \left[\sum_{m=0}^n h(m) \cdot e(n-m) \sum_{l=0}^n h(l) \cdot e(n-l) \right] = \\ &= \sum_{m=0}^n \sum_{l=0}^n h(m) \cdot h(l) \cdot E[e(n-m) \cdot e(n-l)], \end{aligned}$$

где E – математическое ожидание. Поскольку $e(n)$ предполагается белым шумом, можно написать

$$\sigma_0^2(n) = \sum_{m=0}^n \sum_{l=0}^n h(m) \cdot h(l) \cdot \delta(l-m) \cdot \sigma_e^2 = \sigma_e^2 \cdot \sum_{m=0}^n h^2(m).$$

В случае БИХ-фильтра, если полюсы фильтра лежат в области устойчивости, $h(n) \rightarrow 0$ при $n \rightarrow \infty$ и дисперсию в установившемся режиме можно представить в виде

$$\sigma_0^2 = \sigma_e^2 \cdot \sum_{n=0}^{\infty} h^2(n), \quad (5.1)$$

тогда как в случае КИХ-фильтра в (5.1) следует суммировать лишь первые N ненулевых значений ИХ. Используя равенство Парсеваля, этот результат можно также записать в виде

$$\sigma_0^2 = \frac{\sigma_e^2}{\pi} \cdot \int_0^\pi |H(e^{j\omega})|^2 d\omega,$$

где $H(e^{j\omega})$ – ЧХ рассматриваемой системы.

Рассмотрим, например, рекурсивную систему первого порядка, описываемую разностным уравнением

$$y(n) = x(n) - b_1 y(n-1), \quad (5.2)$$

с ИХ

$$h(n) = (-b_1)^n.$$

Из (5.1) можно непосредственно вычислить величину дисперсии в установившемся режиме:

$$\sigma_0^2 = \frac{\sigma_e^2}{1 - b_1^2}. \quad (5.3)$$

Можно заметить, что, когда $b_1 \approx 1$, т. е. полюс цепи лежит вблизи единичной окружности, входной шум может значительно усиливаться системой.

Для рекурсивного звена второго порядка, описываемого уравнением

$$y(n) = x(n) - b_1 y(n-1) - b_2 y(n-2),$$

с двумя полюсами передаточной функции $p_{1,2} = r \cdot e^{\pm j\theta}$, где

$$r = \sqrt{b_2}, \text{ а } \cos \theta = \frac{b_1}{2\sqrt{b_2}}, \text{ имеем}$$

$$\sigma_0^2 = \frac{\sigma_e^2 (1 + r^2)}{(1 - r^2)(r^2 + 1 - 2r^2 \cos \theta)}.$$

Если $r \approx 1$, после замены $\rho = 1 - r$ можно написать

$$\sigma_0^2 = \frac{\sigma_e^2}{4\rho \sin^2 \theta}.$$

Из этого выражения следует, что выходной шум зависит от расстояния полюсов до единичной окружности и от их углового расположения, т. е. больше шума вносят полюса, близкие к 0 и π .

5.2. Квантование арифметических операций в БИХ-цепях

Как было отмечено выше, всегда бывает необходимо не только аппроксимировать входные данные и коэффициенты, но и учитывать квантование результатов арифметических операций в структуре цепи. Процедура фильтрации состоит из последовательности операций сложения и умножения.

Рассмотрим реализацию с фиксированной запятой и предположим, что коэффициенты масштабированы так, что в цепи не возникает переполнения. В этом случае следует выполнять аппроксимацию результатов умножения. Произведение двух чисел, состоящих соответственно из N и M двоичных разрядов, содержит самое большее $N+M$ разрядов. Итак, следует выполнить некоторую аппроксимацию, чтобы уместить результаты умножений в ячейки, длина которых меньше $N+M$ разрядов.

Теперь кратко рассмотрим вопрос о том, как можно решить задачу предсказания суммарной ошибки в выходном сигнале фильтра, являющейся следствием операции квантования. В такой общей постановке эта задача очень сложна и представляет собой по-прежнему нелинейную задачу. Однако, если длина чисел, которые используются в умножении, не очень мала, можно применять линейную статистическую модель. Предположим, что в структуре цепи присутствует несколько источников шума, по одному на каждый узел умножения структурной схемы.

При таком подходе можно изучить поведение ошибки, вычисляя суммарный шум, присутствующий в выходном сигнале, как результат квантования полученных произведений.

На рис. 5.1 показано, какой вид имеет узел умножения в структуре цепи; он состоит из добавленного источника белого шума, последовательно соединенного с неискаженными шумами результатом умножения. Этот источник представляет собой ошибку квантования, имеющую дисперсию $\sigma_e^2 = Q^2 / 12$, для округления и усечения в дополнительном коде. Отметим, что число разрядов M в представлении произведения может быть больше, чем число разрядов в представлении входного сигнала. Это означает, что при необходимости можно достичь меньших значений дисперсии ошибки, вносимой при аппроксимации произведений, по сравнению с ошибками квантования входного сигнала и делается это простым увеличением разрядности ячеек памяти.

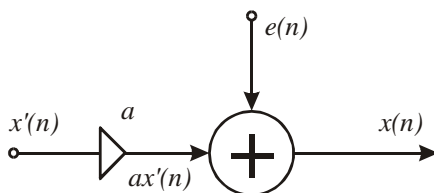


Рис. 5.1. Представление узла умножения с учетом ошибок квантования

Когда приведенное выше представление можно считать верным, дисперсию шума в выходном сигнале, соответствующую каждому источнику шума в структуре цепи, можно легко вычислить с помощью метода, рассмотренного в п. 5.1. Единственная разница в том, что в этом случае нужно определять ИХ отдельных частей цепи от точек, где включены источники шума, до выхода. Это значит, что если имеется L источников шума и через $h_k(n)$ обозначена ИХ части цепи от k -го источника шума до выхода, то дисперсия ошибки в выходном сигнале, соответствующая каждому источнику шума в установившемся режиме, равна

$$\sigma_{0k}^2 = \sigma_e^2 \sum_{n=0}^{\infty} h_k^2(n),$$

где σ_e^2 – дисперсия внесенного шума, и, следовательно, дисперсия полной ошибки в выходном сигнале будет равна

$$\sigma_0^2 = \sum_{k=1}^L \sigma_{0k}^2.$$

Здесь можно сделать несколько общих замечаний. Первое из них связано с тем, что поведение шума в рекурсивной структуре сильно зависит от особенности структуры, выбранной для реализации процедуры фильтрации. Действительно, ИХ части цепи от точки введения источника шума до выхода зависит от применяемой структуры. Это означает, что при выборе структуры цепи надо проявлять особую осторожность. Второе замечание, имеющее отношение к первому, заключается в том, что не все источники шума, введенные в структуру цепи, имеют одинаковую важность. Действительно, рассмотрим, например, звено второго порядка в каноническом виде.

На рис. 5.2 показано, что три источника ошибок: $e_1(n)$, $e_2(n)$, $e_3(n)$ – непосредственно добавляют ошибку на выходе цепи, тогда как источники $e_4(n)$, $e_5(n)$ введены в цепь обратной связи цепи и, как следствие, существует возможность усиления их шума цепью. Полнос передаточной функции цепи дает резонанс в ЧХ, что может привести к усилению шума, как это было показано в п. 5.1.

Теперь снова рассмотрим рекурсивные звенья первого и второго порядка, чтобы изучить их поведение с учетом этой новой ошибки.

На рис. 5.3 показана эквивалентная структура рекурсивного звена первого порядка с источником шума. Видно, что ошибка квантования произведения проходит по тому же пути, что и входной сигнал. Поэтому ИХ, соответствующая этому источнику шума, совпадает с ИХ всего звена и система ведет себя так, как будто входной шум $e(n)$ изменен и имеет вид

$$e_{ax}(n) = e(n) + e_1(n).$$

На рис. 5.4 показана эквивалентная структура с учетом шума для рекурсивного звена второго порядка. В этом случае также ошибки $e_1(n)$, $e_2(n)$ проходят до выхода по той же цепи, что и входной сигнал и для нахождения дисперсии шума квантования на выходе реальной цепи можно использовать результаты, полученные в пункте 5.1, с измененными значениями дисперсии.

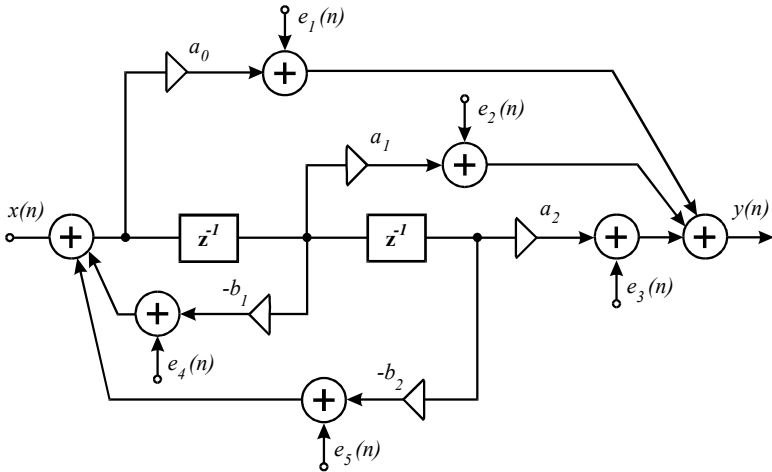


Рис. 5.2. Структурная схема звена второго порядка в каноническом виде с учетом источников ошибок квантования

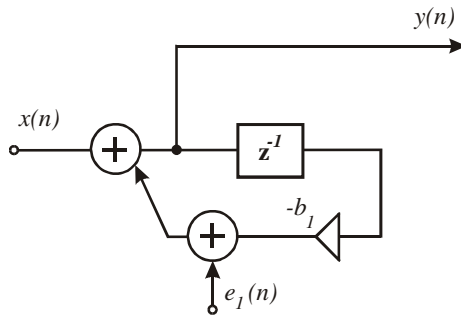


Рис. 5.3. Эквивалентная структурная схема рекурсивного звена первого порядка с источником шума квантования

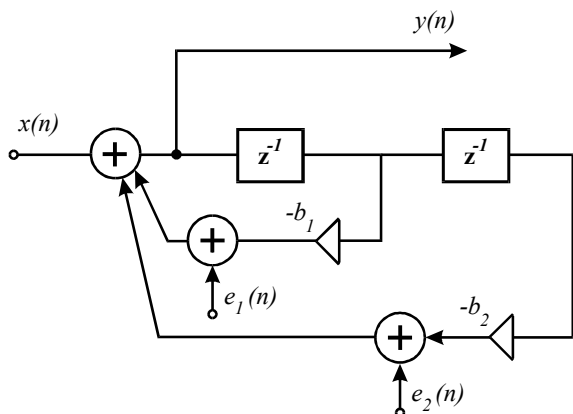


Рис. 5.4. Эквивалентная структурная схема рекурсивного звена второго порядка с источниками шума квантования

5.3. Ограничения, вызванные переполнением, и отношение сигнал/шум

Используя разработанную в пунктах 5.1 и 5.2 модель, можно вычислить дисперсию выходного шума в зависимости от дисперсии входного шума и шума, введенного в структуру цепи в результате аппроксимации арифметических операций. Однако чаще бывает нужно знать зависимость отношения сигнал/шум на выходе системы от входного отношения сигнал/шум. При вычислении этого отношения необходимо принять во внимание не только шум, генерируемый системой, но и ограничения, накладываемые системой на входной сигнал. Действительно, если имеются регистры длиной M разрядов, то, вообще говоря, невозможно использовать для входного сигнала весь динамический диапазон, обеспечиваемый такой разрядностью, так как следует избегать переполнений в процессе фильтрации.

Если задана система с ИХ $h(n)$, то максимально возможный уровень выходного сигнала

$$y_{\max} = x_{\max} \sum_{n=0}^{\infty} |h(n)|, \quad (5.4)$$

где x_{\max} – максимальное значение входного сигнала. Поэтому, если используется представление с фиксированной запятой, необходимо для каждого узла суммирования вычислить ИХ по каждому входу этого узла и нормализовать его входы так, чтобы избежать переполнения. Это, очевидно, приводит к уменьшению мощности входного сигнала и уменьшению отношения сигнал/шум при заданной длине машинного слова и при фиксированной дисперсии выходного шума.

К сожалению, ограничения, вызванные переполнением и, следовательно, максимально возможное отношение сигнал/шум на выходе системы зависят от вида структурной реализации цепи. Поэтому невозможно привести общие результаты для отношения сигнал/шум. Отметим, что указанные выше ограничения во многих приложениях являются очень жесткими. В ряде случаев можно сделать некоторые предположения о виде входного сигнала и использовать менее жесткие ограничения. Например, если входной сигнал является синусоидальным или имеет очень узкую полосу частот, максимальное значение сигнала на выходе системы при амплитуде входного сигнала, равной единице, может, самое большее, равняться усилению системы на рассматриваемой частоте.

Чтобы пояснить приведенные здесь рассуждения, рассмотрим звено первого порядка вида (5.2), на вход которого подается синусоидальный сигнал. Во избежание переполнения входной сигнал нужно нормализовать, выбрав амплитуду его равной $1-|b_1|$ (для устойчивости необходимо $|b_1| < 1$), чтобы амплитуда выходного сигнала не превышала единицу (полагаем, что используются числа, выравненные слева). Тогда дисперсия входного сигнала будет равна

$$\sigma_s^2 = (1 - |b_1|)^2 / 3.$$

Используя (5.3), получим дисперсию выходного сигнала:

$$\sigma_{os}^2 = (1 - |b_1|)^2 / 3(1 - b_1^2),$$

еще раз используя (5.3) для шума квантования (в случае округления и усечения в дополнительном коде), получим выходное отношение шум/сигнал, равное:

$$\sigma_{on}^2 / \sigma_{os}^2 = Q^2 / 4(1 - |b_1|)^2.$$

В итоге заметим, что в случае каскадных реализаций ограничения, вызванные переполнением, и поведение шума зависят также от порядка расположения звеньев и процедуры спаривания нулей и полюсов. В большинстве методов синтеза фильтров за основу структуры принимают биквадратные звенья, т. е. звенья, которые получают путем выбора двух комплексно-сопряженных, но в остальном произвольных пар нулей и полюсов. Очевидно, поведение шума различно при различном выборе пар нулей и полюсов.

5.4. Квантование арифметических операций в КИХ-цепях

Рассмотренные выше проблемы не так сложны для КИХ-цепей. Например, при синтезе структур в прямой форме шум, внесенный квантованием произведений, не проходит через всю структуру цепи, а просто прибавляется к ее выходу. Поэтому если в каждом узле умножения добавляется источник шума $e(n)$ с дисперсией σ_e^2 , то дисперсия шума в выходном сигнале равна $N\sigma_e^2$ ($N-1$ – порядок фильтра) и не зависит от коэффициентов и, следовательно, от типа цепи (например, низкочастотный, высокочастотный фильтр и т. д.), которую следует реализовать. Но значения коэффициентов влияют на выходное отношение сигнал/шум через ограничения, вызванные переполнением, которые можно выразить с помощью соотношения

$$y_{\max} = x_{\max} \sum_{n=0}^{N-1} |h(n)|.$$

Задача несколько осложняется, если рассматривается каскадная реализация КИХ-цепи. Действительно, шум на выходе каждого звена в этом случае будет последовательно проходить

через остальные звенья каскада. Если $h_k(n)$ – реакция на единичный импульс $\delta(n)$, прошедший от точки включения источника шума $e_k(n)$ в k -м каскаде до выхода, то для последовательного соединения звеньев второго порядка дисперсия выходного шума, соответствующего $e_k(n)$, выражается зависимостью

$$\sigma_{ok}^2 = 3\sigma_e^2 \sum_{n=0}^{N-3k} |h_k(n)|^2, \quad (5.5)$$

(здесь N – длина ИХ всей цепи). Тогда дисперсия полного шума

$$\sigma_0^2 = \sum_{k=1}^K \sigma_{ok}^2,$$

где K – число звеньев второго порядка. Из (5.5) следует, что выходной шум зависит от порядка расположения звеньев. Эти рассуждения, очевидно, верны для нерекурсивных реализаций цифровых КИХ-цепей. Если используется рекурсивная реализация, например структура с частотными выборками, необходимо исследовать поведение шума методами, изложенными в предыдущих пунктах.

5.5. Предельные циклы и пульсации, вызванные переполнением

Методы анализа, рассмотренные в предыдущих пунктах, основывались на предположении о взаимной некоррелированности ошибок между собой, а также по отношению к входной последовательности. Эти предположения верны, когда динамический диапазон сигналов таков, что изменение значения сигнала при переходе от одной выборки к другой составляет несколько уровней квантования. Однако в некоторых случаях это не так. Один из таких случаев имеет место, когда сигнал на входе цепи становится равным нулю. Тогда

ошибки, связанные с квантованием результатов арифметических операций внутри цепи, могут привести к постоянному поддержанию ненулевого выходного сигнала.

Для примера рассмотрим рекурсивную цепь, описываемую разностным уравнением первого порядка вида

$$y(n) = x(n) + 0.9y(n-1)$$

с нулевым сигналом на входе $x(n) = 0$ и начальными условиями $y(-1) = 10$. Полагая, что результаты арифметических операций округляются до ближайшего целого, получаем следующую последовательность выходных выборок: 9, 8, 7, 6, 5, 5, 5, 5, ... , и в дальнейшем выходной сигнал не сможет принять значений меньше 5.

Общее исследование явлений подобного рода представляется довольно сложным. Но некоторые полезные результаты можно получить путем простого качественного анализа. Рассмотрим разностное уравнение первого порядка вида

$$y(n) = x(n) + b_1 y(n-1),$$

где результат произведения $b_1 y(n-1)$ округляется до ближайшего целого. Если коэффициент b_1 равен 1, т. е. полюс системы первого порядка лежит на единичной окружности, система будет сохранять выходное значение $y(n-1)$, когда входной сигнал станет нулевым. Затем надо отыскать диапазон начальных значений $y(n-1)$, для которых коэффициент b_1 будет практически равен 1 с учетом округления. Этот интервал обычно называется мертвой зоной фильтра.

В рассматриваемом случае мертвую зону легко можно найти, заметив, что, для того чтобы выходной сигнал y не уменьшался, он должен удовлетворять неравенству

$$y \leq |b_1| y + 0.5, \quad (5.6)$$

так как в этом случае выходной сигнал после округления равен входному сигналу, т. е. правой части (5.6), и не может принять значение $Y+1$, поскольку $|b_1|$ должно быть меньше 1, чтобы

фильтр был устойчив. Из (5.6) видно, что мертвая зона определяется промежутком между $-Y$ и Y , где Y – наибольшее целое число, удовлетворяющее соотношению

$$Y \leq 0.5 / (1 - |b_1|).$$

Если значение b_1 отрицательно, предельный цикл имеет постоянные амплитуду и знак; если же положительно, предельный цикл имеет постоянную амплитуду и чередующийся знак. При $|b_1| < 0.5$ предельный цикл не может существовать.

Подобный анализ можно выполнить и для звена второго порядка вида

$$y(n) = x(n) - b_1 y(n-1) - b_2 y(n-2)$$

с использованием округления результатов. В этом случае также существует мертвая зона, которая равна промежутку между $-Y$ и Y , где Y – наибольшее целое число, удовлетворяющее соотношению

$$Y \leq 0.5 / (1 - b_2), \quad 0 \leq b_2 < 1,$$

которое является условием для того, чтобы полюсы практически лежали на единичной окружности. Коэффициент b_1 и округление соответствующих произведений влияет на вид предельного цикла, а именно на его частоту.

Еще один эффект нелинейных колебаний может иметь место, когда внутри цепи происходит переполнение. Когда используется представление чисел в дополнительном коде, из-за цикличности арифметики выход соответствующего узла цепи перескакивает с максимально положительного значения диапазона представления на минимальное отрицательное или наоборот. Это вызывает большие пульсации в выходном сигнале цепи. Для борьбы с этим явлением можно использовать два способа. Первый заключается в том, чтобы вычислить ограничения, вызванные переполнением в наиболее жесткой форме (5.4) и провести нормализацию, чтобы избежать возможности переполнения. Второй состоит в использовании арифметики с насыщением, а именно когда в

арифметических операциях наступает переполнение, в результат подставляют максимальное по абсолютному значению число, имеющее тот же знак, что и сумма (произведение).

5.6. Квантование коэффициентов

Отклонение коэффициентов фильтра от их номинальных значений вызывает отклонение характеристик передаточных функций. В действительности, это может вызвать довольно неприятные явления, когда эти коэффициенты связаны с полюсами передаточной функции. Поясним сказанное следующим примером. Рассмотрим передаточную функцию, заданную в виде

$$H(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}}.$$

Если полюсы этой функции комплексно сопряженные, обозначим их через $re^{\pm j\theta}$ (см. пункт 5.1). При малых изменениях b_1 и b_2 координаты полюсов изменятся на величины

$$\Delta r = \frac{\partial r}{\partial b_1} \Delta b_1 + \frac{\partial r}{\partial b_2} \Delta b_2, \quad \Delta \theta = \frac{\partial \theta}{\partial b_1} \Delta b_1 + \frac{\partial \theta}{\partial b_2} \Delta b_2.$$

Поскольку $r = \sqrt{b_2}$, $\cos \theta = \frac{b_1}{2\sqrt{b_2}}$, получим

$$\Delta r = \frac{\Delta b_2}{2r}, \quad \Delta \theta = -\frac{\Delta b_1}{2r \sin \theta} + \frac{\Delta b_2}{2r^2 \operatorname{tg} \theta}.$$

Можно заметить, что величина Δr резко изменяется при значениях r , а $\Delta \theta$ – при значениях θ , близких к нулю. Важно отметить, что изменения в расположении полюсов сильно зависят от структуры, выбранной для реализации заданной передаточной функции.

Квантование коэффициентов приводит к появлению ошибки АЧХ цепи:

$$\Delta H(\omega) = \tilde{H}(\omega) - H(\omega),$$

где $\tilde{H}(\omega)$ – АЧХ с квантованными коэффициентами, $H(\omega)$ – АЧХ с неквантованными коэффициентами. Величина $|\Delta H(\omega)|$ не должна превосходить допустимую величину $|\Delta H|_{\max}$, определяемую обычно из условия, чтобы отклонения реальной АЧХ от идеально заданной были в допустимых пределах.

Необходимое число разрядов в квантованных коэффициентах цепи можно определить путем вычисления $|\Delta H(\omega)|$ для последовательно возрастающего числа разрядов в кодах коэффициентов при выполнении условия $|\Delta H(\omega)| < |\Delta H|_{\max}$.

Более просто (с точки зрения вычислительной сложности) влияние квантования коэффициентов может быть учтено, если реальную цепь представить в виде параллельного соединения идеальной цепи $H(z)$ (с неквантованными коэффициентами a_i , b_i) и паразитной цепи $H_{\Delta}(z)$, коэффициенты которой зависят от погрешностей Δa_i и Δb_i квантования исходных коэффициентов, рассматриваемых как статистически независимые величины с равномерным распределением (рис. 5.5). При этом можно оценить среднеквадратическое отклонение АЧХ (или другой характеристики) реальной цепи от характеристики идеальной цепи и определить число разрядов, обеспечивающее допустимость упомянутого отклонения.

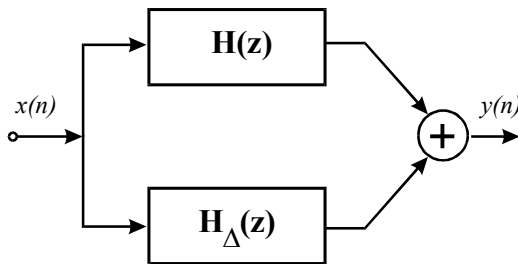


Рис. 5.5. Представление реальной цепи с учетом квантования коэффициентов

Возможны и практически применяются и другие методы, в частности методы, основанные на предварительном изучении чувствительности характеристик цепи к малым изменениям ее коэффициентов, а также методы, учитывающие конкретные структуры цифровых цепей.

5.7. Техническая реализация цифровых цепей

Реализация алгоритма работы цифровой цепи состоит из последовательности умножений выборок выходного и (или) входного сигналов и сложения полученных произведений. Поэтому основными стандартными модулями аппаратуры являются: запоминающее устройство; сумматоры; умножители; устройство, управляющее последовательностью операций.

На рис. 5.6 показана общая структурная схема технических средств для реализации цифровой цепи, которая включает следующие стандартные модули:

- а) память входных выборок;
- б) память выходных выборок, если выполняется процедура БИХ-фильтрации;
- в) память коэффициентов;
- г) арифметическое устройство, способное выполнять сложение и умножение;
- д) устройство управления последовательностью операций.

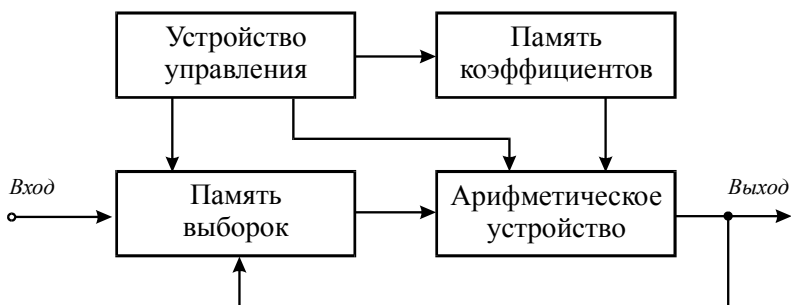


Рис. 5.6. Общая структурная схема технических средств для реализации цифровой цепи

Для реализации отдельных блоков этой структурной схемы в настоящее время имеются различные возможности, которые отличаются быстродействием, уровнем интеграции, потреблением мощности и т. д.

После расчета коэффициентов и выбора структуры цепи решаются вопросы выбора кодирования чисел (прямой или дополнительный код), способов их представления (с фиксированной или плавающей запятой) и выбора элементной базы. Выбор того или иного варианта аппаратной реализации определяется главным образом требованиями к быстродействию и объему аппаратных затрат.

Быстродействие цифровой цепи определяется наибольшей временной задержкой t_{\max} при прохождении сигнала за один период дискретизации, в течение которого вычисляется очередная выборка выходного сигнала. Величина t_{\max} находится непосредственно по структурной схеме цепи; при этом рассматриваются все возможные пути прохождения сигналов, и выбирается тот путь, по которому задержка сигнала наибольшая. Это время и принимается за t_{\max} .

Величина t_{\max} определяется в основном временем выполнения операции умножения двух операндов и числом операций умножения, которые необходимо выполнить для вычисления одного отсчета выходного сигнала. Поэтому число операций умножения часто рассматривается как основной критерий быстродействия и вообще качества реализации цифровой цепи.

Другим важным реализационным критерием является, как уже упомянуто, объем аппаратных затрат и, прежде всего, объем оперативной и постоянной памяти, необходимой для хранения коэффициентов (параметров) фильтра и результатов выполнения арифметических и логических операций, необходимых для реализации алгоритма цепи.

В качестве простейшего примера рассмотрим реализацию нерекурсивной цифровой цепи в прямой форме (рис. 5.7а), описываемой уравнением

$$y(n) = \sum_{i=0}^4 a_i x(n-i) = a_0 x(n) + a_1 x(n-1) + a_2 x(n-2) + a_3 x(n-3) + a_4 x(n-4)$$

и передаточной функцией

$$H(z) = \sum_{i=0}^4 a_i z^{-i}.$$

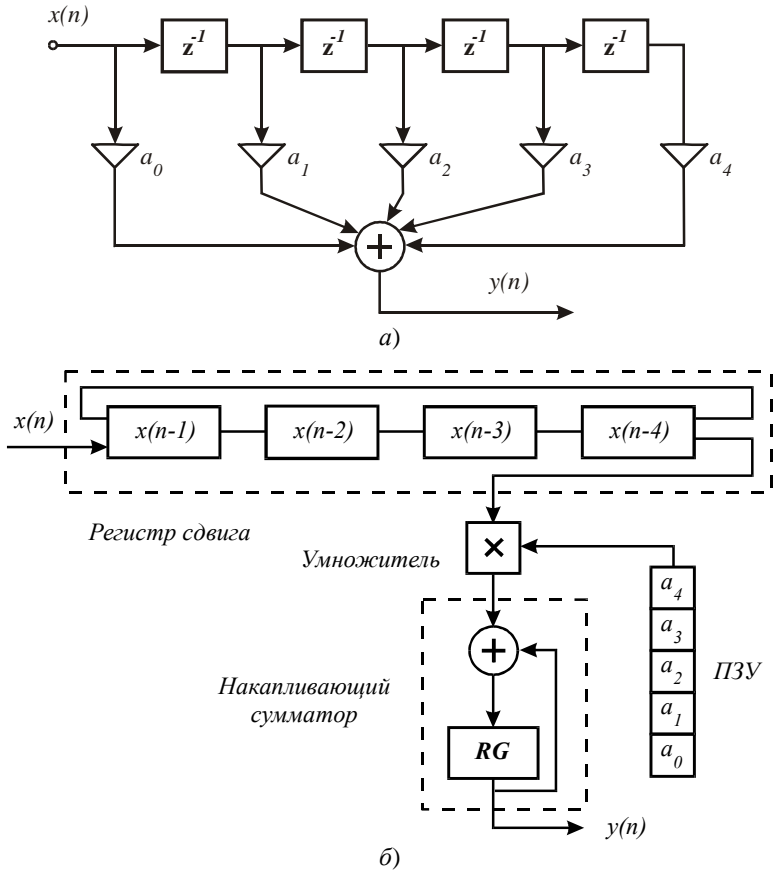


Рис. 5.7. Реализация нерекурсивной цифровой цепи в прямой форме: а) структурная схема цепи; б) простейшая схема специализированного устройства

Простейшая схема специализированного устройства приведена на рис. 5.7б. Она содержит регистр сдвига для получения единичных задержек, постоянное запоминающее устройство (ПЗУ) для хранения коэффициентов, умножитель для получения произведений и накапливающий сумматор, в выходном регистре **RG** которого и накапливается сумма произведений.

Последовательность работы следующая:

1-й такт: регистр **RG** очищается;

2-й такт: на вход регистра сдвига поступает входной сигнал $x(n)$, одновременно в этом регистре происходит сдвиг на M разрядов (M – число разрядов в выборке входного сигнала) и выталкивается число $x(n-4)$. Последнее умножается на коэффициент a_4 , и результат $a_4x(n-4)$ поступает в накапливающий сумматор;

3-й такт: в регистре сдвига происходит сдвиг и выталкивается значение $x(n-3)$, которое одновременно поступает на вход регистра сдвига и первый вход умножителя, вычисляется $a_3x(n-3)$, и в накапливающем сумматоре образуется $a_4x(n-4)+a_3x(n-3)$;

4-й такт: в регистре сдвига происходит сдвиг и выталкивается значение $x(n-2)$, которое одновременно поступает на вход регистра сдвига и первый вход умножителя, вычисляется $a_2x(n-2)$, и в регистре **RG** накапливающего сумматора будет $a_4x(n-4)+a_3x(n-3)+a_2x(n-2)$;

5-й такт: в регистре сдвига происходит сдвиг и выталкивается значение $x(n-1)$, которое одновременно поступает на вход регистра сдвига и первый вход умножителя, вычисляется $a_1x(n-1)$, и в регистре **RG** образуется $a_4x(n-4)+a_3x(n-3)+a_2x(n-2)+a_1x(n-1)$;

6-й такт: в регистре сдвига происходит сдвиг и выталкивается значение $x(n)$, которое одновременно поступает на вход регистра сдвига и первый вход умножителя, вычисляется $a_0x(n)$, и на выходе накапливающего сумматора **RG** образуется

$$y(n) = \sum_{i=0}^4 a_i x(n-i),$$
 а в регистре сдвига оказываются записанными коды $x(n), \dots, x(n-3)$.

Далее вычисления циклически повторяются для следующих моментов: $(n+1)$, $(n+2)$ и т. д. Таким образом, каждый выходной отсчет вычисляется за 6 тактов. Длительность такта τ должна быть не меньше времени, необходимого для реализации совокупности операций: сдвига, умножения, сложения.

На практике порядок фильтра N может достигать многих сотен, и максимальная скорость обработки сигналов f_{\max} может оказаться недостаточной для использования в системах, функционирующих в реальном масштабе времени.

Для различных областей применения требуются различные частоты дискретизации. Так, в системах связи $f_{\delta}=8$ кГц, в звукотехнике $f_{\delta}=40\text{--}48$ кГц, в системах обработки ТВ изображения $f_{\delta}=14$ МГц. Для обеспечения необходимой точности число разрядов в данных различно в различных областях применения, например, в технике связи, обычно, число разрядов в данных $M=8$, в цифровой звукотехнике $M \in [14, 16]$.

В некоторых случаях используются цифровые цепи на основе ПЗУ, в ячейки которого заносится весь возможный набор значений произведений или даже сумм произведений. С целью ускорения процессов ЦОС, наряду с применением современной быстродействующей элементной базы, используется ряд алгоритмических и архитектурных методов.

5.8. Контрольные вопросы

1. Укажите источники ошибок квантования в цифровых цепях.
2. Найдите реакцию идеальной цифровой цепи на шум квантования входного сигнала в общем случае.
3. Найдите реакцию идеальной рекурсивной цифровой цепи первого порядка на шум квантования входного сигнала. Проведите анализ полученного выражения.
4. Найдите реакцию идеальной рекурсивной цифровой цепи второго порядка на шум квантования входного сигнала. Проведите анализ полученного выражения.

5. Найдите дисперсию полной ошибки квантования арифметических операций в БИХ-цепи.
6. Как влияет структура цепи на выходной шум квантования?
7. Какие ограничения следует накладывать на входной сигнал цифровой цепи, чтобы избежать переполнения, и почему?
8. Укажите особенности квантования арифметических операций в КИХ-цепях.
9. Получите выражение для отношения сигнал/шум на выходе идеальной рекурсивной цепи первого порядка при воздействии квантованного гармонического сигнала.
10. Получите выражение для отношения сигнал/шум на выходе рекурсивной цепи первого порядка с учетом квантования результатов умножения при воздействии неквантованного гармонического сигнала.
11. Получите выражение для отношения сигнал/шум на выходе рекурсивной цепи первого порядка с учетом квантования результатов умножения при воздействии квантованного гармонического сигнала.
12. Получите выражение для отношения сигнал/шум на выходе идеальной рекурсивной цепи второго порядка при воздействии квантованного гармонического сигнала.
13. Получите выражение для отношения сигнал/шум на выходе рекурсивной цепи второго порядка с учетом квантования результатов умножения при воздействии неквантованного гармонического сигнала.
14. Получите выражение для отношения сигнал/шум на выходе рекурсивной цепи второго порядка с учетом квантования результатов умножения при воздействии квантованного гармонического сигнала.
15. Что такой предельный цикл? Как он образуется? Что понимают под «мертвой зоной»? Приведите примеры.
16. Что такое пульсации и как с ними бороться?

5.9. Литература

1. Брюханов, Ю. А. Цифровые цепи и сигналы: учеб. пособие / Ю. А. Брюханов. – Ярославль: ЯрГУ, 2005.
2. Солонина, А. И. Основы цифровой обработки сигналов: курс лекций / А. И. Солонина и др. – СПб.: БХВ-Петербург, 2005.
3. Куприянов, М. С. Цифровая обработка сигналов: процессоры, алгоритмы, средства программирования / М. С. Куприянов, Б. А. Матюшкин. – СПб.: Политехника, 1999.
4. Рабинер, Л. Теория и применение цифровой обработки сигналов: пер. с англ. / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978.
5. Оппенгейм, А. В. Цифровая обработка сигналов: пер. с англ. / А. В. Оппенгейм, Р. В. Шафер. – М.: Связь, 1979.
6. Каппелини, В. Цифровые фильтры и их применение: пер. с англ. / В. Каппелини, А. Дж. Константиnidис, П. Эмилиани. – М.: Энергоатомиздат, 1983.
7. Карташев, В. Г. Основы теории дискретных сигналов и цифровых фильтров / В. Г. Карташев. – М.: Высшая школа, 1982.
8. Гольденберг, Л. М. Цифровая обработка сигналов: справочник / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1985.
9. Гольденберг, Л. М. Цифровая обработка сигналов: учеб. пособие / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1990.

6. БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Ранее было показано, что дискретное преобразование Фурье может быть использовано для выполнения свертки. Здесь будут рассмотрены вычислительные требования ДПФ и будут получены несколько *быстрых алгоритмов* его вычисления. Эти алгоритмы известны как *быстрые преобразования Фурье* (БПФ). Рассмотрим алгоритм прореживания по основанию 2. Он был опубликован в 1965 году американскими учеными Кули и Тьюки. Отметим также, что существуют алгоритмы БПФ со смешанным основанием и БПФ простых чисел, но их рассмотрение выходит за рамки данной книги.

6.1. Алгоритмы БПФ по основанию 2

С использованием упрощенных обозначений N -точечное ДПФ от N -точечной последовательности $x(n)$ может быть найдено следующим образом:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk}, \quad (6.1)$$

где $W_N = e^{-j2\pi/N}$.

Поскольку $x(n)$ может быть и вещественным и комплексным, для оценки $X(k)$ требуется порядка N комплексных умножений и N комплексных сложений для каждого значения k . Таким образом, поскольку существует N значений $X(k)$, вычисление N -точечного ДПФ требует N^2 комплексных умножений и сложений.

Основная стратегия, которая используется в алгоритме БПФ, включает в себя разложения N -точечного ДПФ на последовательно меньшие ДПФ. Чтобы увидеть, как это работает, предположим, что длина $x(n)$ – чётное число (т. е. N делится на 2 без остатка). Если $x(n)$ разделить на две последовательности длины $N/2$, то на вычисление $N/2$ -точечного ДПФ каждой из этих двух последовательностей требуется примерно $(N/2)^2$ умножений и такое же количество сложений. Таким образом, два ДПФ

требуют всего $2(N/2)^2 = \frac{1}{2}N^2$ умножений и сложений.

Продолжая подобные разбиения, удаётся получить существенную экономию.

6.2. БПФ с прореживанием по времени

Алгоритм БПФ с прореживанием по времени основан на расщеплении (прореживании) $x(n)$ на более мелкие последовательности и нахождении $X(k)$ из ДПФ этих прореженных последовательностей. Рассмотрим, как это прореживание приводит к эффективному алгоритму, когда длина последовательности является степенью числа 2.

Пусть $x(n)$ – последовательность длины $N=2^v$. Предположим, что она разбивается (прореживается) на две подпоследовательности, длина каждой из которых $N/2$. Как показано на рис. 6.1, первая подпоследовательность $g(n)$ формируется из чётных членов последовательности:

$$g(n) = x(2n) \quad n = 0, 1, \dots, \frac{N}{2} - 1,$$

а вторая – $h(n)$ – из нечётных:

$$h(n) = x(2n+1) \quad n = 0, 1, \dots, \frac{N}{2} - 1.$$

В таком случае N -точечное ДПФ от $x(n)$ будет иметь вид

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n)W_N^{nk} = \sum_{n \text{ чётн}} x(n)W_N^{nk} + \sum_{n \text{ нечётн}} x(n)W_N^{nk} = \\ &= \sum_{l=0}^{\frac{N}{2}-1} g(l)W_N^{2lk} + \sum_{l=0}^{\frac{N}{2}-1} h(l)W_N^{(2l+1)k}. \end{aligned} \quad (6.2)$$

Так как $W_N^{2lk} = W_{N/2}^{lk}$, то выражение (6.2) может быть представлено в виде

$$X(k) = \sum_{l=0}^{\frac{N}{2}-1} g(l) W_{N/2}^{lk} + W_N^k \sum_{l=0}^{\frac{N}{2}-1} h(l) W_{N/2}^{lk} .$$

Заметьте, что первый член – это $N/2$ -точечное ДПФ от $g(n)$, а второй – $N/2$ -точечное ДПФ от $h(n)$:

$$X(k) = G(k) + W_N^k H(k) \quad k = 0, 1, \dots, N-1. \quad (6.3)$$

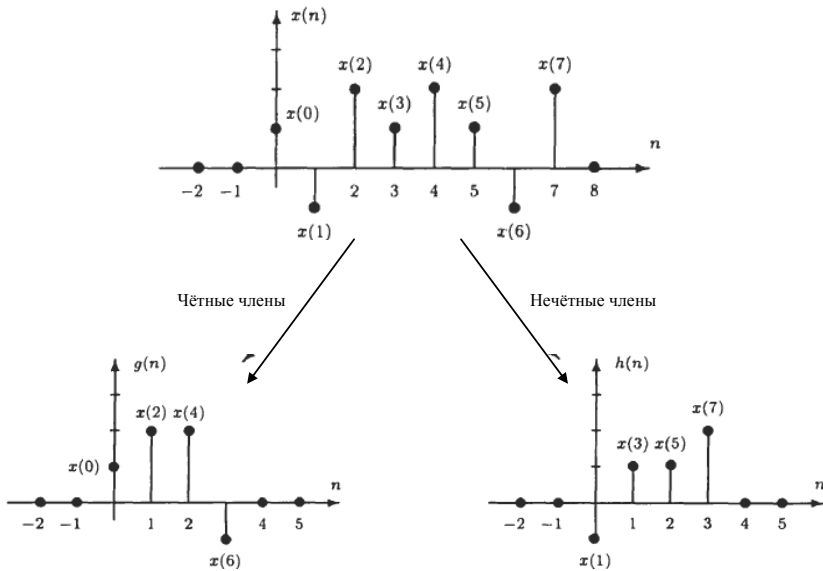


Рис. 6.1. Прореживание последовательности длины $N = 8$ по основанию 2

Хотя $N/2$ -точечные ДПФ от $g(n)$ и $h(n)$ являются последовательностями длины $N/2$, периодичность комплексных экспонент позволяет записать

$$G(k) = G\left(k + \frac{N}{2}\right) \quad H(k) = H\left(k + \frac{N}{2}\right).$$

Таким образом, $X(k)$ может быть вычислено из $N/2$ -точечных ДПФ от $G(k)$ и $H(k)$. Заметим, что

$$W_N^{k+N/2} = W_N^k W_N^{N/2} = -W_N^k,$$

тогда

$$W_N^{k+\frac{N}{2}} H\left(k + \frac{N}{2}\right) = -W_N^k H(k),$$

это является необходимым для формирования результирующих значений $W_N^k H(k)$, где $k=0, 1, \dots, N/2-1$. Комплексные экспоненты при $H(k)$ в выражении (6.3) называются *поворотными (поворачивающими) множителями*. Блок-схема вычислений, необходимых для первого этапа воздействия БПФ с прореживанием по времени на восьмиточечное ДПФ, показана на рис. 6.2.

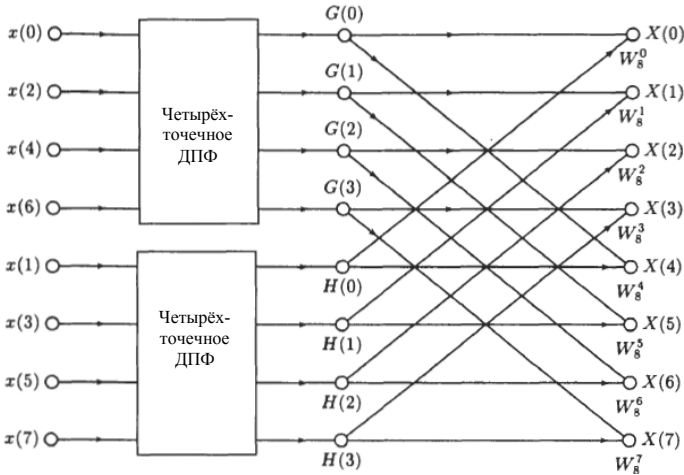


Рис. 6.2. Действие алгоритма БПФ с прореживанием по времени на восьмиточечное ДПФ после первого прореживания

Если $N/2$ четно, то $g(n)$ и $h(n)$ могут быть прорежены ещё раз. Например, $G(k)$ может быть представлено как

$$G(k) = \sum_{n=0}^{\frac{N}{2}-1} g(n) W_{N/2}^{nk} = \sum_{\substack{n=0 \\ \text{н чётн}}}^{\frac{N}{2}-1} g(n) W_{N/2}^{nk} + \sum_{\substack{n=0 \\ \text{н нечётн}}}^{\frac{N}{2}-1} g(n) W_{N/2}^{nk} .$$

Как и прежде, это приводит к тому, что

$$G(k) = \sum_{n=0}^{\frac{N}{4}-1} g(2n) W_{N/4}^{nk} + W_{N/2}^k \sum_{n=0}^{\frac{N}{4}-1} g(2n+1) W_{N/4}^{nk} ,$$

где первое слагаемое – $N/4$ -точечное ДПФ чётных членов $g(n)$, а второе – $N/4$ -точечное ДПФ нечётных членов. Блок-схема, иллюстрирующая это разложение, показана на рис. 6.3. Если N является степенью 2, прореживание может быть продолжено до тех пор, пока не останутся только двухточечные ДПФ вида, показанного на рис. 6.4.

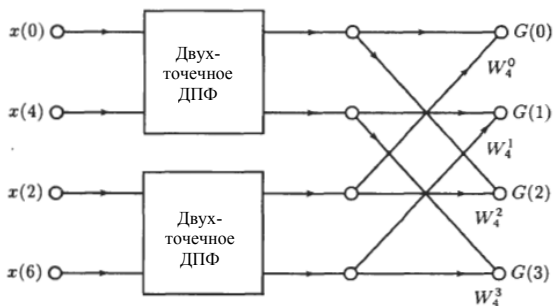


Рис. 6.3. Прореживание четырёхточечного ДПФ на два двухточечных ДПФ с помощью БПФ с прореживанием по времени

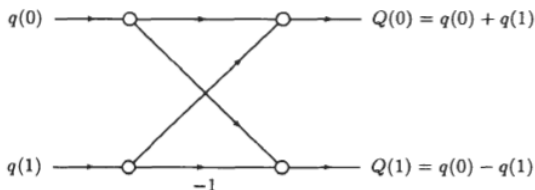


Рис. 6.4. Двухточечное ДПФ

Основная вычислительная единица БПФ показана на рис. 6.5а. Она называется *бабочка*. Эта структура может быть упрощена путем факторизации члена W_N^r с нижней ветви, как показано на рис. 6.5б. Оставшийся фактор $W_N^{N/2} = -1$. Полное прореживание восьмиточечного ДПФ с помощью БПФ с прореживанием по времени по основанию 2 показано на рис. 6.6.

Вычисление N -точечного ДПФ с использованием БПФ с прореживанием по времени по основанию 2 является гораздо более эффективным, чем вычисления ДПФ напрямую. Например, если $N=2^v$, то необходимо $\log_2 N = v$ этапов вычислений. Так как каждый этап требует $N/2$ комплексных операций умножения на поворотные множители W_N^r и N комплексных сложений, то всего необходимо $\frac{1}{2}N \log_2 N$ комплексных умножений и $N \log_2 N$ комплексных сложений.

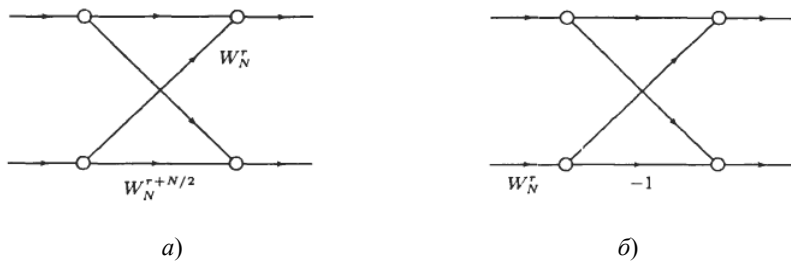


Рис. 6.5. Операции бабочка:

- а) бабочка – основной вычислительный элемент алгоритма БПФ;
- б) упрощённая бабочка, содержащая только одну операцию комплексного умножения

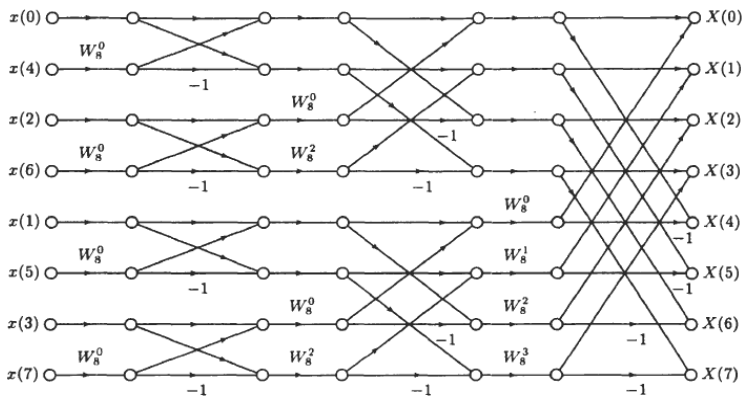


Рис. 6.6. Полное прореживание восьмиточечного ДПФ с помощью БПФ с прореживанием по времени по основанию 2

Обратите внимание, что как только в структуре алгоритма БПФ с прореживанием по времени была выполнена операция бабочки над парой комплексных чисел, исчезла необходимость сохранения исходной пары. Таким образом, выходная пара чисел может быть сохранена в том же регистре, что и входная пара. Следовательно, требуется только один массив размером N и вычисления в этом случае могут быть выполнены «на месте». Однако для этого входная последовательность $x(n)$ должна храниться (или поступать в обработку) в непоследовательном порядке, как показано на рис. 6.6. Такая уловка с входной последовательностью происходит из-за последовательного прореживания $x(n)$. Получающийся в результате порядок соответствует обратной индексации бит исходной последовательности. Другими словами, если индекс n записан в двоичной форме, порядок, в котором входная последовательность должна поступать в обработку, определяется чтением двоичного представления n в обратном порядке, как показано в табл. 6.1 для $N=8$.

Таблица 6.1

Обратная индексация бит исходной последовательности

n	Двоичное представление	Обратное двоичное представление	n'
0	000	000	0
1	001	100	4
2	010	010	2
3	011	110	6
4	100	001	1
5	101	101	5
6	110	011	3
7	111	111	7

Альтернативные формы алгоритмов БПФ могут быть получены из БПФ с прореживанием по времени, манипулированием блок-схемой и изменением порядка, в котором хранятся результаты каждого этапа вычислений. Например, узлы блок-схемы могут быть переставлены таким образом, чтобы входная последовательность $x(n)$ находилась в нормальном порядке. Однако с такой перестановкой утрачивается способность выполнять вычисления на месте.

6.3. БПФ с прореживанием по частоте

Другой класс алгоритмов БПФ может быть получен путем прореживания выходной последовательности $X(k)$ на все меньшие и меньшие подпоследовательности. Эти алгоритмы называют БПФ с прореживанием по частоте. Они могут быть получены следующим образом. Пусть N является степенью числа 2, $N=2^v$ и оценка чётных и нечётных членов $X(k)$ производится раздельно. Тогда чётные члены можно записать в виде

$$X(2k) = \sum_{n=0}^{N-1} x(n) W_N^{2nk} .$$

Разделяя эту сумму на первые $N/2$ точек и последние $N/2$ точек, а также используя равенство $W_N^{2nk} = W_{N/2}^{nk}$, выражение можно привести к виду

$$X(2k) = \sum_{n=0}^{\frac{N}{2}-1} x(n)W_{N/2}^{nk} + \sum_{n=N/2}^{N-1} x(n)W_{N/2}^{nk}.$$

Со сменой индексов второй суммы имеем

$$X(2k) = \sum_{n=0}^{\frac{N}{2}-1} x(n)W_{N/2}^{nk} + \sum_{n=0}^{\frac{N}{2}-1} x\left(n + \frac{N}{2}\right)W_{N/2}^{(n+\frac{N}{2})k}.$$

Наконец, так как $W_{N/2}^{(n+\frac{N}{2})k} = W_{N/2}^{nk}$,

$$X(2k) = \sum_{n=0}^{\frac{N}{2}-1} \left[x(n) + x\left(n + \frac{N}{2}\right) \right] W_{N/2}^{nk},$$

а это есть не что иное, как $N/2$ -точечное ДПФ последовательности, формируемой добавлением первых $N/2$ точек $x(n)$ к ее последним $N/2$ точкам.

Выполняя сходные операции над нечётными членами $X(k)$, получим

$$X(2k+1) = \sum_{n=0}^{\frac{N}{2}-1} W_N^n \left[x(n) - x\left(n + \frac{N}{2}\right) \right] W_{N/2}^{nk}. \quad (6.4)$$

Блок-схема, иллюстрирующая первый этап прореживания, показана на рис. 6.7. Как и в БПФ с прореживанием по времени, прореживание может быть продолжено, пока не останутся лишь двухточечные ДПФ. Полное прореживание восьмиточечного ДПФ с помощью БПФ с прореживанием по частоте по основанию 2 показано на рис. 6.8. Вычислительная сложность его такая же, как с прореживанием по времени, поэтому вычисления могут быть выполнены «на месте». Наконец, отметим, что, хотя входная последовательность $x(n)$ находится в нормальном порядке, члены частот $X(k)$ представлены с использованием обратной индексации бит.

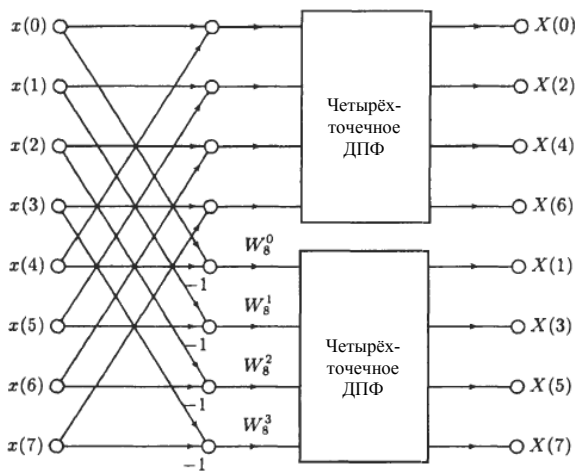


Рис. 6.7. Действие алгоритма БПФ с прореживанием по частоте на восьмиточечное ДПФ после первого прореживания

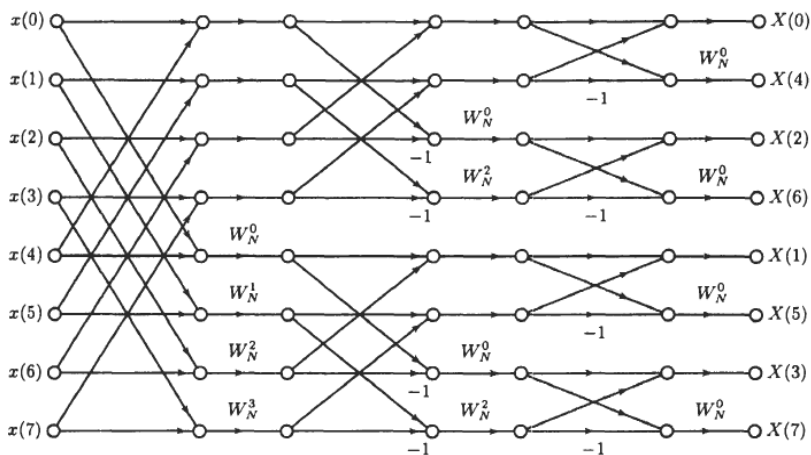


Рис. 6.8. Полное прореживание восьмиточечного ДПФ с помощью БПФ с прореживанием по частоте по основанию 2

6.4. Контрольные вопросы

1. Дайте определение ДПФ.
2. Дайте определение быстрого преобразования Фурье.
3. В чем заключается связь ДПФ и БПФ?
4. Оцените вычислительную сложность ДПФ.
5. Оцените вычислительную сложность БПФ.
6. Назовите способы применения БПФ.
7. Поясните, в чем заключается основная идея, используемая в алгоритме БПФ.
8. Назовите преимущества использования БПФ.
9. Назовите виды алгоритмов БПФ с прореживанием по основанию 2.
10. Расскажите про алгоритм БПФ с прореживанием по времени.
11. Что такое поворотные множители?
12. Изобразите блок-схему вычислений первого этапа БПФ с прореживанием по времени.
13. Что такое основная вычислительная единица БПФ?
14. Какие существуют способы упрощения структуры «бабочка»?
15. В чем преимущества использования БПФ с прореживанием по времени по основанию 2 по сравнению с вычислением ДПФ напрямую?
16. Расскажите о представлении данных в структуре алгоритма БПФ с прореживанием по времени.
17. Как могут быть получены альтернативные формы алгоритмов БПФ из БПФ с прореживанием по времени?
18. Расскажите про алгоритм БПФ с прореживанием по частоте.
19. Изобразите блок-схему вычислений, первого этапа БПФ с прореживанием по частоте.
20. Какие еще существуют виды алгоритмов БПФ?

6.5. Литература

1. Солонина, А. И. Основы цифровой обработки сигналов: курс лекций / А. И. Солонина и др. – СПб.: БХВ-Петербург, 2005.
2. Hayes, M. H. Digital signal processing / M. H. Hayes. – McGraw-Hill, 2011.
3. Брюханов, Ю. А. Цифровые цепи и сигналы: учеб. пособие / Ю. А. Брюханов. – Ярославль: ЯрГУ, 2005.
4. Рабинер, Л. Теория и применение цифровой обработки сигналов: пер. с англ. / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978.
5. Оппенгейм, А. В. Цифровая обработка сигналов: пер. с англ. / А. В. Оппенгейм, Р. В. Шафер. – М.: Связь, 1979.
6. Каппелини, В. Цифровые фильтры и их применение: пер. с англ. / В. Каппелини, А. Дж. Константиноидис, П. Эмилиани. – М.: Энергоатомиздат, 1983.
7. Карташев, В. Г. Основы теории дискретных сигналов и цифровых фильтров / В. Г. Карташев. – М.: Высшая школа, 1982.
8. Гольденберг, Л. М. Цифровая обработка сигналов: учеб. пособие / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М.: Радио и связь, 1990.

7. СИНТЕЗ ЦИФРОВЫХ КИХ-ФИЛЬТРОВ МЕТОДОМ ОКОН

7.1. Основные определения

Цифровой фильтр (ЦФ) в широком смысле – это любая цифровая система или цепь, которая согласно заданному алгоритму осуществляет извлечение цифрового сигнала либо его параметров из действующей на входе системы (цепи) смеси сигнала с помехой.

Цифровой фильтр в узком смысле – это частотно-избирательная цепь, которая обеспечивает селекцию цифровых сигналов по частоте. К таким фильтрам относятся рассмотренные ранее фильтры нижних и верхних частот, полосовой и режекторный фильтры.

Цифровые фильтры, как и любая цифровая система, могут быть реализованы аппаратно, программно или аппаратно-программно. Это определяется целью, назначением и местом ЦФ в предполагаемой системе.

Аппаратная реализация подразумевает использование разнообразных блоков (регистров, сумматоров, умножителей, устройств памяти, логических элементов и т. п.).

Программная реализация означает, что фильтр представлен в виде программы, написанной на определенном языке программирования. Так, для персонального компьютера это будет любой из языков высокого уровня, а для микропроцессора, микроконтроллера или цифрового сигнального процессора – язык соответствующего ассемблера.

Аппаратно-программная реализация предполагает, что часть функций фильтра выполняется аппаратно (аналого-цифровое и цифроаналоговое преобразования, умножение, прием/передача данных и т. д.), а другая часть функций выполняется программно.

Под *проектированием ЦФ* понимается процесс, в результате которого создается программа или цифровое устройство, отвечающее заданным требованиям и ограничениям.

Процесс проектирования ЦФ включает в себя следующие этапы.

1. *Синтез*, результатом которого является функциональная схема фильтра с коэффициентами. Процедуры синтеза фильтров с конечной и бесконечной импульсными характеристиками существенно различаются, однако имеют одинаковую последовательность: задание требований к фильтрам; решение задачи аппроксимации характеристик фильтра, в результате которой рассчитываются коэффициенты его передаточной функции (разностного уравнения); конструирование функциональной схемы ЦФ.

2. *Выбор или разработка рациональных алгоритмов вычислений* с учетом арифметики, используемой при заданном методе реализации – плавающая или фиксированная запятая (в англоязычной литературе – точка). Алгоритм зависит от разрядности регистров процессора, количества аккумуляторов, возможности распараллеливания операций, наличия устройств, выполняющих операцию умножения с накоплением, и других особенностей процессора. Конечной целью данного этапа является обеспечение функционирования фильтра в реальном времени при минимальных потерях качества обработки сигналов.

3. *Проверка моделированием* проектируемого фильтра в нереальном времени по стандартным сигналам с использованием программных эмуляторов. Задача проверки моделированием состоит в обнаружении и устранении возможных логических и прочих скрытых ошибок, испытании на соответствие сконструированного фильтра заданным характеристикам, включая частотные, временные, шумовые и др.

4. *Практическая реализация и отладка* в реальном времени с помощью отладочных модулей.

Результаты проверки моделированием и отладки могут оказать влияние на изменение ряда решений – от выбора структурной схемы ЦФ до задания новых требований.

7.2. Синтез цифровых фильтров

Каждый из классов ЦФ имеет свои принципиально отличные методы синтеза, которые рассматриваются отдельно для каждого класса, а сами методы синтеза могут быть как оптимальными

(«оптимальный» означает наилучший в некотором заранее определенном смысле, например по заданному критерию), так и не оптимальными. В связи с этим выделяют фильтры, в которых реализованы принципы оптимального синтеза, и фильтры, при синтезе которых принципы оптимальности не принимаются во внимание.

7.2.1. Задание требований к цифровым фильтрам

Требования (спецификации) к цифровым фильтрам могут формулироваться как во временной, так и в частотной областях, что определяется назначением фильтра и областью его описания.

Во временной области требования могут задаваться к импульсной $h(n)$ и переходной $g(n)$ характеристикам при широких допусках к частотным свойствам фильтра.

Частотно-избирательные фильтры обычно синтезируются в частотной области. При этом в процессе синтеза сами частоты могут задаваться как действительными (f [Гц], ω [рад/с]), так и нормированными на частоту дискретизации $\hat{\omega}$ [рад], что определяется удобствами и традициями.

В частотной области требования могут предъявляться:

- 1) только к амплитудно-частотным характеристикам без каких-либо ограничений на фазочастотные характеристики;
- 2) только к ФЧХ, когда важно сохранение фазовых, а потому и временных соотношений между гармоническими составляющими принимаемого сигнала;
- 3) одновременно и к АЧХ, и к ФЧХ.

Рассмотрим задание требований к частотно-избирательным ЦФ. Идеальные частотно-избирательные ЦФ являются физически нереализуемыми, поэтому необходимо аппроксимировать их АЧХ при заранее определенных допусках, зависящих от ряда практических ограничений. Эти ограничения связаны, прежде всего, с назначением синтезируемого фильтра. Все ограничения и допуски составляют требования, предъявляемые к фильтру.

Требования к частотно-избирательным фильтрам включают в себя:

- 1) задание частоты дискретизации $\hat{\omega}_\partial$ [рад] и типа избирательности;
- 2) задание требований к АЧХ $|H(e^{j\hat{\omega}})|$;
- 3) выбор метода аппроксимации АЧХ.

7.2.2. Типы частотно-избирательных фильтров и задание требований к ним

Диаграммы требований к АЧХ синтезируемых частотно-избирательных КИХ-фильтров представлены на рис. 7.1.

Фильтр нижних частот

Реальный ФНЧ имеет три частотных полосы (рис. 7.1а): полосу пропускания, переходную полосу и полосу задерживания.

Полоса пропускания ограничивается *граничной частотой* в ПП $\hat{\omega}_n$; ширина ПП $\Delta\hat{\omega}_{\text{ПП}} = \hat{\omega}_n$; δ_n – отклонение АЧХ от 1 в ПП, часто выражаемое в дБ с использованием величины α_n , получившей название *максимально допустимого затухания* в ПП:

$$\alpha_n = -20 \lg(1 - \delta_n).$$

Переходная полоса располагается между полосами пропускания и задерживания; ширина переходной полосы $\Delta\hat{\omega} = \hat{\omega}_\partial - \hat{\omega}_n$. Поскольку в этой полосе требования не задаются, *удовлетворительным окажется любое решение, если оно соответствует требованиям в полосах пропускания и задерживания.*

Полоса задерживания (подавления, ослабления) лежит в пределах от *граничной частоты* в ПЗ $\hat{\omega}_\partial$ до половины частоты дискретизации $\hat{\omega}_\partial/2$; ширина ПЗ $\Delta\hat{\omega}_{\text{ПЗ}} = \hat{\omega}_\partial/2 - \hat{\omega}_\partial$; δ_∂ – отклонение АЧХ от 0 в ПЗ, часто выражаемое в дБ с использованием величины α_∂ , получившей название *минимально допустимого затухания* в ПЗ:

$$\alpha_\partial = -20 \lg(\delta_\partial).$$

Фильтр верхних частот

Реальный ФВЧ имеет, как и ФНЧ, три частотных полосы, которые расположены в обратном, относительно ФНЧ, порядке (рис. 7.1б).

Полоса задерживания лежит в пределах от 0 до $\hat{\omega}_3$; ширина ПЗ $\Delta\hat{\omega}_{ПЗ} = \hat{\omega}_3$.

Переходная полоса лежит в пределах от $\hat{\omega}_3$ до $\hat{\omega}_n$; ширина переходной полосы $\Delta\hat{\omega} = \hat{\omega}_n - \hat{\omega}_3$.

Полоса пропускания лежит в пределах от $\hat{\omega}_n$ до $\hat{\omega}_\delta/2$; ширина ПП $\Delta\hat{\omega}_{ПП} = \hat{\omega}_\delta/2 - \hat{\omega}_n$.

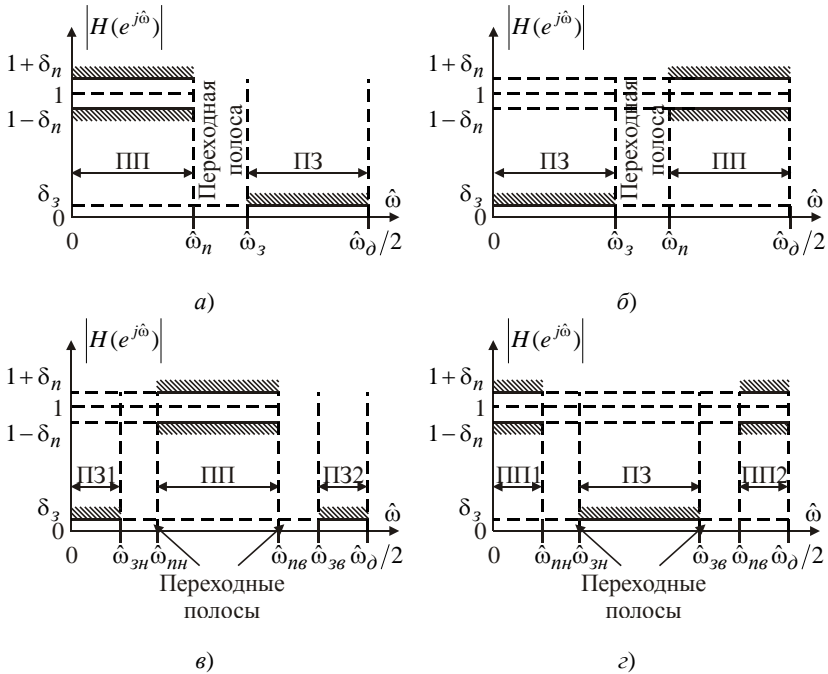


Рис. 7.1. Диаграммы требований к АЧХ частотно-избирательных КИХ-фильтров: а) фильтр нижних частот; б) фильтр верхних частот; в) полосовой фильтр; з) режекторный фильтр

Полосовой фильтр

Реальный ПФ, иначе полоснопропускающий, характеризуется пятью частотными полосами (рис. 7.1в), из которых одна – центральная – полоса пропускания, две полосы задерживания и две переходные полосы. Обозначения, приведенные на рис. 7.1в, следующие:

$\hat{\omega}_{3н}$ – граничная частота первой полосы задерживания (ПЗ1);
ширина ПЗ1 $\Delta\hat{\omega}_{ПЗ1} = \hat{\omega}_{3н}$;

$\hat{\omega}_{нн}$ – нижняя граничная частота ПП; $\hat{\omega}_{нв}$ – верхняя граничная частота ПП; ширина ПП $\Delta\hat{\omega}_{ПП} = \hat{\omega}_{нв} - \hat{\omega}_{нн}$;

$\hat{\omega}_{3в}$ – граничная частота второй полосы задерживания (ПЗ2);
ширина ПЗ2 $\Delta\hat{\omega}_{ПЗ2} = \hat{\omega}_{\partial}/2 - \hat{\omega}_{3в}$.

Переходные полосы 1 и 2 имеют ширину $\Delta\hat{\omega}_1 = \hat{\omega}_{нн} - \hat{\omega}_{3н}$ и $\Delta\hat{\omega}_2 = \hat{\omega}_{3в} - \hat{\omega}_{нв}$.

Режекторный фильтр

Реальный РФ, иначе полоснозаграждающий, подобно ПФ, характеризуется пятью полосами (рис. 7.1г), из которых две полосы пропускания, одна полоса задерживания и две переходные полосы. Обозначения, приведенные на рис. 7.1г, следующие:

$\hat{\omega}_{нн}$ – граничная частота первой полосы пропускания (ПП1);
ширина ПП1 $\Delta\hat{\omega}_{ПП1} = \hat{\omega}_{нн}$;

$\hat{\omega}_{3н}$ – нижняя граничная частота ПЗ; $\hat{\omega}_{3в}$ – верхняя граничная частота ПЗ; ширина ПЗ $\Delta\hat{\omega}_{ПЗ} = \hat{\omega}_{3в} - \hat{\omega}_{3н}$;

$\hat{\omega}_{нв}$ – граничная частота второй полосы пропускания (ПП2);
ширина ПП2 $\Delta\hat{\omega}_{ПП2} = \hat{\omega}_{\partial}/2 - \hat{\omega}_{нв}$.

Переходные полосы 1 и 2 имеют ширину $\Delta\hat{\omega}_1 = \hat{\omega}_{нн} - \hat{\omega}_{3н}$ и $\Delta\hat{\omega}_2 = \hat{\omega}_{3в} - \hat{\omega}_{нв}$.

Замечание. При синтезе ЦФ вместо частот $\hat{\omega}_3$ и $\hat{\omega}_н$, $\hat{\omega}_{3н}$ и $\hat{\omega}_{нн}$, $\hat{\omega}_{3в}$ и $\hat{\omega}_{нв}$ могут задаваться частоты среза $\hat{\omega}_с$, $\hat{\omega}_{сн}$, $\hat{\omega}_{св}$ соответственно. В случае синтеза КИХ-фильтров методом окон, при заданных частотах $\hat{\omega}_3$ и $\hat{\omega}_н$, $\hat{\omega}_{3н}$ и $\hat{\omega}_{нн}$, $\hat{\omega}_{3в}$ и $\hat{\omega}_{нв}$, частоты

$\hat{\omega}_c, \hat{\omega}_{сн}, \hat{\omega}_{св}$ могут быть найдены с использованием следующих математических выражений:

$$\hat{\omega}_c = \frac{\hat{\omega}_з + \hat{\omega}_п}{2}, \hat{\omega}_{сн} = \frac{\hat{\omega}_{зн} + \hat{\omega}_{пн}}{2}, \hat{\omega}_{св} = \frac{\hat{\omega}_{зв} + \hat{\omega}_{пв}}{2}.$$

Значение величины АЧХ на частотах среза $\hat{\omega}_c, \hat{\omega}_{сн}, \hat{\omega}_{св}$ у фильтров, синтезированных с использованием метода окон, составляет в логарифмическом масштабе -6 дБ.

7.3. Синтез КИХ-фильтров методом окон

Среди достаточного разнообразия методов синтеза КИХ-фильтров наибольшее применение в различных приложениях получили два «классических» – метод окон и метод оптимального (по Чебышеву) синтеза.

7.3.1. Общая характеристика задачи синтеза КИХ-фильтров методом окон

Без потери общности зададимся целью рассчитать коэффициенты одного из частотно-избирательных фильтров, представленных на рис. 7.2, например ФНЧ. Для нерекурсивной фильтрации коэффициенты фильтра являются и отсчетами его ИХ.

Рассматриваемый фильтр имеет идеальную АЧХ, а потому его ЧХ $H_u(e^{j\hat{\omega}})$ может быть представлена только бесконечным рядом Фурье:

$$H_u(e^{j\hat{\omega}}) = \sum_{n=-\infty}^{\infty} h_u(n) e^{-j\hat{\omega}n}. \quad (7.1)$$

Тогда идеальная ИХ может быть рассчитана как коэффициенты ряда Фурье:

$$h_u(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_u(e^{j\hat{\omega}}) e^{j\hat{\omega}n} d\hat{\omega}. \quad (7.2)$$

Ясно, что это – физически нереализуемый БИХ-фильтр, поскольку ИХ бесконечна и начинается в $-\infty$, т. е. реакция

предшествует воздействию. Простейший путь конструирования физически реализуемой передаточной функции фильтра состоит в исключении всех членов ряда (7.1), имеющих отрицательный индекс $n < 0$. В результате получается ИХ, соответствующая БИХ-фильтру.

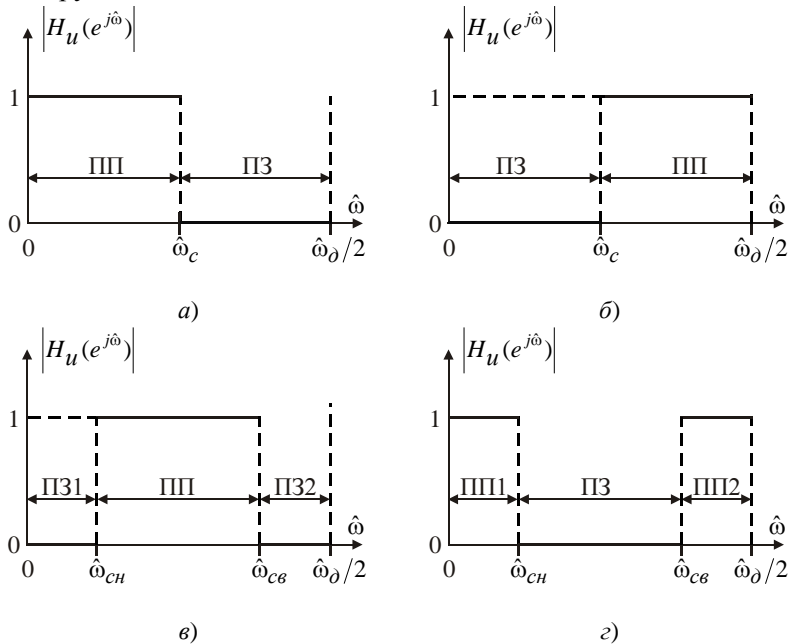


Рис. 7.2. Идеальные АЧХ частотно-избирательных фильтров:
 а) фильтр нижних частот; б) фильтр верхних частот;
 в) полосовой фильтр; з) режекторный фильтр

Для получения импульсной характеристики КИХ-фильтра необходимо ограничить ряд (7.1) сверху до $N + 1$ членов. С точки зрения ИХ $h_u(n)$ рассматриваемая операция представляет собой ее *усечение* в пределах от $0 \leq n \leq N$, т. е. умножение ИХ на последовательность $w_n(n)$ вида

$$w_n(n) = \begin{cases} 1, & 0 \leq n \leq N, \\ 0, & \text{при других } n. \end{cases} \quad (7.3)$$

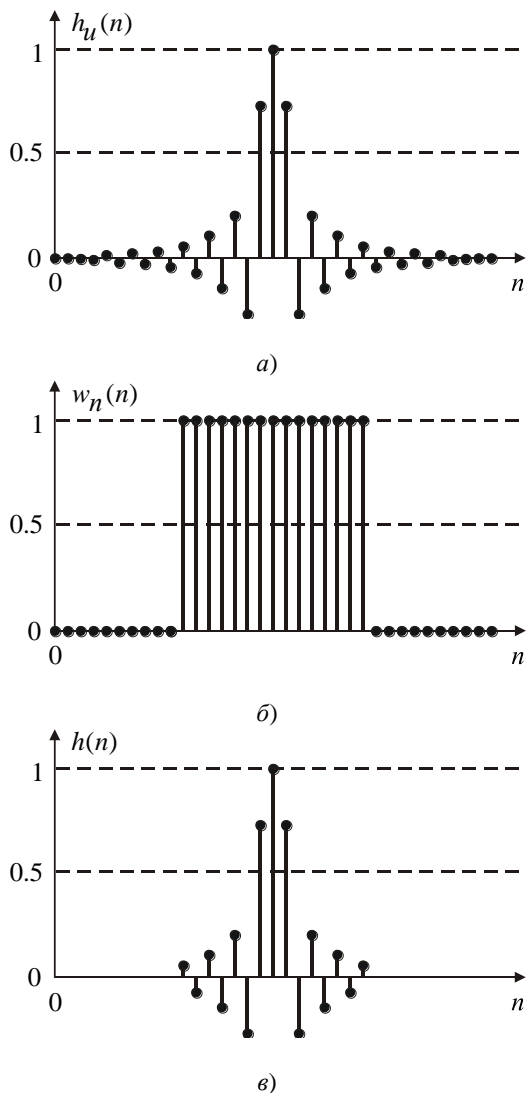


Рис. 7.3. Процедура усечения импульсной характеристики: а) «идеальная» ИХ; б) последовательность (7.3); в) результат усечения

Процедура усечения с использованием (7.3), в результате которой образуется частотная характеристика КИХ-фильтра

$$H(e^{j\hat{\omega}}) = \sum_{n=0}^N h(n) e^{-j\hat{\omega}n}, \quad (7.4)$$

демонстрируется на рис. 7.3. Из (7.4) после замены $e^{j\hat{\omega}} = z$ имеем передаточную функцию:

$$H(z) = \sum_{n=0}^N h(n) z^{-n}, \quad (7.5)$$

где реальная ИХ (т. е. коэффициенты $H(z)$) равна произведению идеальной ИХ (7.2) на функцию (7.3):

$$h(n) = w_n(n) h_u(n). \quad (7.6)$$

Казалось бы, решение найдено. Действительно, подбирая значения N и контролируя поведение АЧХ, за несколько итераций можно найти такое N , при котором требования к заданному фильтру будут выполнены.

Однако усечение ряда Фурье вследствие разрывности АЧХ приводит к существенным ошибкам.

7.3.2. Явление Гиббса

Появляющиеся ошибки связаны с характером сходимости ряда Фурье в точке разрыва первого рода, каковой и является точка $\hat{\omega}_c$ (рис. 7.2а).

Во-первых, в точке разрыва $\hat{\omega}_c$ первого рода *ряд Фурье сходится к среднему предельных значений функции слева и справа*. Предел слева равняется 1, а справа – 0, поэтому $\left| H(e^{j\hat{\omega}_c}) \right| = 0.5$.

Во-вторых, в точке разрыва *сходимость ряда Фурье не является равномерной* и носит особый характер, который выражается в появлении пульсаций вблизи точки разрыва, максимум которых слева и справа составляет 9% от АЧХ и остается таковым вне зависимости от N . Этот феномен получил название *явления Гиббса*.

В результате формируются пульсации как в полосе задерживания, так и в полосе пропускания фильтра; кроме того,

образуется переходная полоса, ширина которой тем меньше, чем больше значение N .

Рассмотрим явление Гиббса более подробно, для чего обратимся к выражению (7.6). Здесь и далее ограничимся четными значениями N .

Выражение (7.6) представляет собой умножение двух функций во временной области. Для перехода в частотную область воспользуемся тем свойством, что умножение двух функций во временной области соответствует свертке их преобразований Фурье. Преобразование Фурье последовательности $h_u(n)$ соответствует ЧХ $H_u(e^{j\hat{\omega}})$ идеального фильтра, а преобразование Фурье весовой функции $w_n(n)$ соответствует $W_n(e^{j\hat{\omega}})$:

$$W_n(e^{j\hat{\omega}}) = \sum_{n=0}^N w_n(n) e^{-j\hat{\omega}n} = e^{-jN\hat{\omega}/2} \frac{\sin\left(\frac{\hat{\omega}(N+1)}{2}\right)}{\sin\left(\frac{\hat{\omega}}{2}\right)}. \quad (7.7)$$

Функция (7.7) известна под названием *ядра Дирихле*. Если не учитывать множитель, характеризующий линейный фазовый сдвиг, то график функции

$$\hat{W}_n(e^{j\hat{\omega}}) = \sin\left(\frac{\hat{\omega}(N+1)}{2}\right) / \sin\left(\frac{\hat{\omega}}{2}\right) \quad (7.8)$$

будет иметь форму, показанную на рис. 7.4а, а ее модуль – на рис. 7.4б. Это быстроколеблющаяся функция с максимальным значением N при $\hat{\omega} = 0$, резко спадающая по амплитуде. *Область с максимальной амплитудой называется главным лепестком, а остальные области – боковыми лепестками.*

Функция $W_n(e^{j\hat{\omega}})$ принимает нулевые значения на частотах

$$\hat{\omega} = k \frac{2\pi}{N+1}, \quad k = 1, 2, \dots, N,$$

а ширина всех лепестков, включая главный, одинакова и равна

$$\Delta\hat{\omega} = \frac{2\pi}{N+1}, \quad (7.9)$$

причем с увеличением N ширина главного лепестка уменьшается и увеличивается число пульсаций.

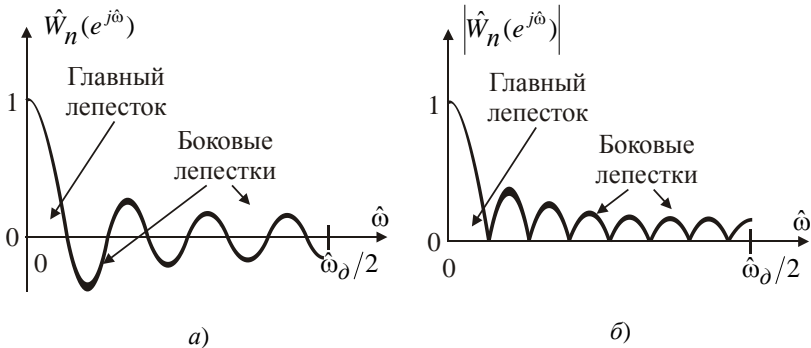


Рис. 7.4. Примеры графиков: а) функции (7.7) при $N = 14$;
б) функции (7.8) при $N = 14$

Замечание. В некоторых источниках ширина главного лепестка выбирается вдвое больше той, которая определяется с использованием выражения (7.9). Это связано с тем, что главный лепесток модуля частотной характеристики окна математически определен как в положительной, так и в отрицательной областях частот.

Таким образом, усечение ряда Фурье (7.1) до N членов (7.4) эквивалентно свертке ЧХ идеального ФНЧ $H_u(e^{j\hat{\omega}})$, АЧХ которой имеет форму прямоугольника (рис. 7.2а), с ядром Дирихле $W_n(e^{j\hat{\omega}})$. Это означает, что колебания функции $W_n(e^{j\hat{\omega}})$ попадают в область частотного прямоугольника, в которой свертка точно воспроизводит эти колебания. В результате получается АЧХ реального фильтра, у которой вблизи точки разрыва наблюдается два эффекта:

1) возникают ошибки аппроксимации в виде пульсаций ЧХ $H(e^{j\hat{\omega}})$, которые обусловлены боковыми лепестками функции $W_n(e^{j\hat{\omega}})$;

2) образуется сглаживающий разрыв – переходная полоса, ширина которой зависит от главного лепестка функции $W_n(e^{j\hat{\omega}})$ и приблизительно равняется его ширине.

7.3.3. Окна и их основные параметры

Явление Гиббса объясняется, как уже отмечалось, неравномерной сходимостью ряда Фурье в точке разрыва. Управлять сходимостью ряда Фурье (7.1) можно с помощью весовой последовательности конечной длины $w(n)$, называемой *окном* или *весовой функцией*. Метод состоит в том, что коэффициенты ряда Фурье, т. е. ИХ $h_u(n)$ идеального фильтра, умножаются на $w(n)$. В результате получается ИХ реального фильтра

$$h(n) = w(n)h_u(n) \quad (7.10)$$

и передаточная функция

$$H(z) = \sum_{n=0}^N h(n)z^{-n} = \sum_{n=0}^N w(n)h_u(n)z^{-n}, \quad (7.11)$$

где $N+1$ – длина функции $w(n)$.

Ограниченная на интервале $0 \leq n \leq N$ и равная нулю вне этого интервала положительная симметричная весовая функция называется окном:

$$w(n) = \begin{cases} w(N-n), & 0 \leq n \leq N, \\ 0, & \text{при других } n. \end{cases} \quad (7.12)$$

Метод конструирования передаточной функции с помощью ограничения ряда (7.1) окном (7.12) часто называется *методом окон*, или *методом взвешивания*.

Известно большое количество окон, применение которых определяется характером поставленной задачи синтеза фильтров. Ниже рассматриваются три из них: прямоугольное окно, обобщенное окно Хэмминга и окно Кайзера. Данные оконные функции обладают свойствами всех возможных видов окон и

позволяют достаточно хорошо понять преимущества и недостатки метода взвешивания.

Прямоугольное окно (окно Дирихле)

$$w_n(n) = \begin{cases} 1, & 0 \leq n \leq N, \\ 0, & \text{при других } n. \end{cases}$$

Именно прямоугольное окно (рис. 7.3б, 7.4) было использовано при простом усечении ряда Фурье и изучении явления Гиббса. Ранее была получена ЧХ этого окна:

$$W_n(e^{j\hat{\omega}}) = e^{-jN\hat{\omega}/2} \frac{\sin\left(\frac{\hat{\omega}(N+1)}{2}\right)}{\sin\left(\frac{\hat{\omega}}{2}\right)}$$

и исследовано ее поведение. Установлено, что чем больше N , тем уже главный лепесток и переходная полоса синтезируемого фильтра, а также меньше амплитуда пульсаций в области полос пропускания и задерживания при неизменной величине амплитуды пульсаций $\approx 9\%$ вблизи частоты среза $\hat{\omega}_c$.

Таким образом, прямоугольное окно позволяет сделать вывод о том, что «хорошее» окно должно обладать двумя свойствами:

- 1) ширина главного лепестка ЧХ окна должна быть малой;
- 2) энергия боковых лепестков ЧХ окна должна быстро уменьшаться с увеличением частоты $\hat{\omega}$.

Эти два требования несовместимы, поэтому на практике необходим компромиссный вариант.

Обобщенное окно Хэмминга

Второе из рассматриваемых окон, называемое обобщенным окном Хэмминга, имеет следующий вид:

$$w_x(n) = \begin{cases} \alpha - (1-\alpha) \cos\left(\frac{2\pi n}{N}\right), & 0 \leq n \leq N, \\ 0, & \text{при других } n, \end{cases}$$

где α лежит в пределах $[0, 1]$. Случай $\alpha = 0.5$ соответствует окну Хэннинга, а случай $\alpha = 0.54$ – окну Хэмминга.

Частотную характеристику рассматриваемого окна можно получить, если представить его во временной области в виде произведения

$$w_x(n) = w_n(n) \left[\alpha - (1-\alpha) \cos\left(\frac{2\pi n}{N}\right) \right],$$

где $w_n(n)$ – прямоугольное окно, длина которого $N + 1$ равна длине окна $w_x(n)$.

Для перехода к ЧХ заменим косинусы экспоненциальными функциями, используя формулу Эйлера, и учтем, что умножение последовательности $x(n)$ на $e^{j\hat{\omega}_0 n}$ дает в частотной области сдвиг на $\hat{\omega}_0$: $X(e^{j(\hat{\omega} - \hat{\omega}_0)})$. Окончательно ЧХ обобщенного окна Хэмминга имеет следующий вид:

$$W_x(e^{j\hat{\omega}}) = \alpha W_n(e^{j\hat{\omega}}) + \frac{1-\alpha}{2} W_n\left(e^{j\left(\hat{\omega} - \frac{2\pi}{N}\right)}\right) + \frac{1-\alpha}{2} W_n\left(e^{j\left(\hat{\omega} + \frac{2\pi}{N}\right)}\right). \quad (7.13)$$

Видно, что ЧХ обобщенного окна Хэмминга представляет собой сумму сдвинутых по частоте характеристик прямоугольного окна $W_n(e^{j\hat{\omega}})$. При этом вторая и третья компоненты существенно уменьшают амплитуду бокового лепестка вблизи главного лепестка (так как боковые лепестки второй и третьей компонент (7.13) находятся в противофазе с боковыми лепестками первой компоненты). Следовательно, уровни боковых лепестков (7.13) оказываются значительно ниже, чем у прямоугольного окна.

Однако ширина главного лепестка обобщенного окна Хэмминга оказывается в два раза больше, чем ширина главного лепестка прямоугольного окна.

На рис. 7.5 приведен пример временной и модуля частотной характеристик окна Хэмминга ($\alpha = 0.54$).

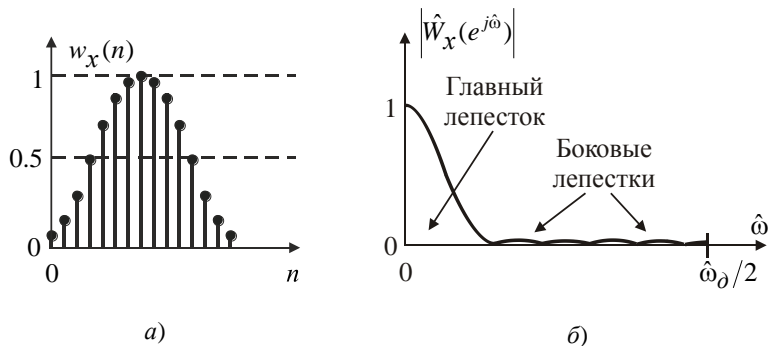


Рис. 7.5. Примеры характеристик окна Хэмминга:
 а) временной характеристики при $N = 14$;
 б) модуля частотной характеристики при $N = 14$

Замечание. Применение параметра $\alpha = 0.54$ в выражении (7.13) позволяет сконцентрировать 99.96% общей энергии спектра в главном лепестке, при этом максимальный уровень боковых лепестков на 40 дБ ниже главного максимума.

Окно Кайзера

Установлено, что задача расчета «хороших» окон фактически сводится к математической задаче отыскания ограниченных во времени функций, преобразования Фурье которых наилучшим образом аппроксимируют функции, ограниченные по частоте, т. е. имеют минимальную энергию за пределами заданного интервала частот. При решении данной задачи в замкнутой форме для непрерывных функций времени был введен класс так называемых вытянутых сфероидальных функций, которые имеют достаточно сложный вид.

Поэтому Кайзер в качестве «наилучшего» окна предложил относительно простую аппроксимацию этих функций, получившую название *окна Кайзера* и имеющую следующий вид:

$$w_{\kappa}(n) = \frac{I_0[\beta(1 - [(n - \alpha) / \alpha]^2)^{1/2}]}{I_0(\beta)}, \quad 0 \leq n \leq N,$$

где $\alpha = N/2$, а $I_0(\cdot)$ – модифицированная функция Бесселя первого рода нулевого порядка, которая может быть легко получена из следующего выражения:

$$I_0(x) = 1 + \sum_{k=1}^{\infty} \left[\frac{(x/2)^k}{k!} \right]^2.$$

Параметр β определяет форму окна и компромисс между максимальным уровнем боковых лепестков и шириной главного лепестка ЧХ окна. *Окно Кайзера является по существу оптимальным окном, в том смысле, что оно имеет максимум энергии в главном лепестке для данного уровня боковых лепестков.*

Значения β и N вычисляются по полученным Кайзером эмпирическим формулам при условии, что отклонение от единицы в ПП и отклонение от нуля в ПЗ фильтра одинаковы $\delta_n = \delta_z = \delta$.

В приведенные ниже математические выражения входят затухание в ПЗ $\alpha_3 = -20 \lg(\delta_3)$ и ширина переходной полосы $\Delta\hat{\omega}$, причем параметр β зависит от требований к α_3 :

$$\beta = \begin{cases} 0.1102(\alpha_3 - 8.7), & \alpha_3 > 50, \\ 0.5842(\alpha_3 - 21)^{0.4} + 0.07886(\alpha_3 - 21), & 21 \leq \alpha_3 \leq 50, \\ 0, & \alpha_3 < 21. \end{cases}$$

Окно Кайзера обладает одним замечательным свойством: для фиксированного отклонения δ произведение порядка КИХ-фильтра N на ширину его переходной полосы $\Delta\hat{\omega}$ является примерно постоянной величиной, которая называется *D-фактором*:

$$D = N \cdot \Delta \hat{\omega} \Big|_{\delta = \text{const}} \cong \text{const}.$$

D -фактор можно вычислить по формуле

$$D = \begin{cases} \frac{\alpha_3 - 7.95}{2.2855}, & \alpha_3 \geq 21, \\ 5.7944, & \alpha_3 < 21. \end{cases}$$

Частотная характеристика окна Кайзера в замкнутой форме не получена, однако с очень хорошим приближением ее можно вычислить с помощью следующего выражения:

$$W_{\kappa}(\hat{f}) = \frac{N \cdot sh \left[\sqrt{\beta^2 - (N\hat{f}/2)^2} \right]}{I_0(\beta) \sqrt{\beta^2 - (N\hat{f}/2)^2}},$$

где $\hat{f} = \hat{\omega}/(2 \cdot \pi)$, а $sh(x)$ – гиперболический синус:

$$sh(x) = \frac{e^x - e^{-x}}{2}.$$

7.3.4. Методика синтеза КИХ-фильтров на основе оконных функций

Как видно из выражения (7.5), отсчеты импульсной характеристики КИХ-фильтров одновременно являются и коэффициентами его передаточной функции (разностного уравнения), поэтому задача синтеза в данном случае сводится к получению импульсной характеристики.

Методика синтеза включает в себя:

- 1) задание требований к фильтру (см. п. 7.2);
- 2) вычисление импульсной характеристики идеального фильтра $h_u(n)$ (см. п. 7.3.1);
- 3) выбор окна и длины фильтра (см. п. 7.3.2, п. 7.3.3);

4) расчет импульсной характеристики реального фильтра (см. п. 7.3.1);

5) проверку выполнений заданных требований. Для проверки выполнения заданных требований рассчитывается АЧХ: если требования выполняются, то на этом процедуру конструирования передаточной функции можно завершить; если требования не выполняются, необходимо увеличить N и повторить п. 1–5.

Может случиться и так, что требования выполняются с большим запасом – тогда следует проверить, нельзя ли уменьшить N . Или может оказаться, что при выбранном окне длина фильтра оказывается слишком большой, а фильтр по каким-либо соображениям (обеспечение работы в реальном масштабе времени, элементная база и т. п.) не может быть реализован. Тогда выбирается другое окно и процедура повторяется.

7.3.5. Особенности использования метода окон

Метод взвешивания является весьма удобным для проектирования КИХ-фильтров, однако он обладает некоторыми особенностями, которые ограничивают его использование.

1. Импульсная характеристика идеального фильтра $h_u(n)$ должна представлять замкнутое математическое выражение. Не имея его, для вычисления взвешенных коэффициентов, трудно вообще говорить об использовании метода окон.

2. Для частотно-избирательных фильтров метод взвешивания одинаково аппроксимирует переходные полосы и пульсации в ПП и ПЗ.

3. Метод окон не является оптимальным в том смысле, что он не позволяет получить наименьшую по амплитуде пульсацию для заданного порядка фильтра N и заданного множества частот среза.

7.4. Контрольные вопросы

1. Дайте определение цифрового фильтра в широком и узком смысле.
2. Дайте определение КИХ-фильтра.
3. Какие способы реализации цифровых фильтров существуют?
4. Что собой представляет процесс проектирования цифровых фильтров?
5. Какие существуют способы задания требований к цифровым фильтрам?
6. Какие требования предъявляются при синтезе КИХ-фильтров нижних частот?
7. Какие требования предъявляются при синтезе КИХ-фильтров верхних частот?
8. Какие требования предъявляются при синтезе полосовых КИХ-фильтров?
9. Какие требования предъявляются при синтезе режекторных КИХ-фильтров?
10. Что такое «окно»? Приведите примеры «окон».
11. Перечислите параметры окон, которые оказывают влияние на синтез КИХ-фильтров.
12. Как связаны между собой ширина главного лепестка и уровень боковых лепестков с шириной переходной полосы и пульсациями частотной характеристики в ПП и ПЗ синтезируемого КИХ-фильтра.
13. Дайте определение метода окон (взвешивания). На примере частотно-избирательных фильтров (ФНЧ, ФВЧ, ПФ, РФ) перечислите этапы синтеза КИХ-фильтров с использованием метода взвешивания.
14. Какие ошибки возникают при использовании метода взвешивания?
15. Каковы основные особенности использования метода окон при синтезе КИХ-фильтров?

7.5. Литература

1. Брюханов, Ю. А. Цифровые фильтры: учеб. пособие / Ю. А. Брюханов, А. Л. Приоров. – Ярославль: ЯрГУ, 2002.
2. Приоров, А. Л. Основы цифровой обработки сигналов: лабораторный практикум / А. Л. Приоров, В. А. Волохов, А. М. Лаврентьев. – ВУНЦ ВВС «ВВА им. профессора Н. Е. Жуковского и Ю. А. Гагарина» (филиал, г. Ярославль), 2012.
3. Айфичер, Э. Цифровая обработка сигналов: практический подход. – 2-е изд. / Э. Айфичер, Б. Джервис. – Вильямс, 2004.
4. Сергиенко, А. Б. Цифровая обработка сигналов. – 3-е изд. / А. Б. Сергиенко. – СПб.: БХВ-Петербург, 2011.
5. Дьяконов, В. П. MATLAB 6.0/6.1/6.5/6.5+SP1 + Simulink 4/5. Обработка сигналов и изображений / В. П. Дьяконов. – М.: СОЛОН-Пресс, 2005.
6. Хрящев, В. В. Основы теории цепей: сборник задач / В. В. Хрящев, А. Л. Приоров, В. А. Волохов – Ярославль: ЯрГУ, 2008.
7. Солонина, А. И. Цифровая обработка сигналов. Моделирование в MATLAB / А. И. Солонина, С. М. Арбузов. – СПб.: БХВ-Петербург, 2008.

8. СИНТЕЗ ЦИФРОВЫХ БИХ-ФИЛЬТРОВ МЕТОДОМ БИЛИНЕЙНОГО Z-ПРЕОБРАЗОВАНИЯ

8.1. Основные определения

Синтез ЦФ по заданным требованиям к его характеристикам сводится к задаче аппроксимации, которая решается чисто математическими методами в области, определяемой назначением фильтра. Если аппроксимация производится в p -области, т. е. в области преобразования Лапласа, получим аналоговый фильтр, а если в z -области, т. е. в области z -преобразования, – цифровой фильтр. В обоих случаях математический аппарат должен соответствовать особенностям выбранных областей синтеза. Однако при синтезе стандартных линейных частотно-избирательных БИХ-фильтров (ФНЧ, ФВЧ, ПФ, РФ) удобнее всего воспользоваться хорошо развитым аппаратом расчета аналоговых фильтров и методами отображения p -области в z -область, т. е. методами преобразования аналоговых фильтров в цифровые. Такой синтез включает в себя:

- 1) выбор метода отображения p -области в z -область;
- 2) расчет аналогового фильтра по требованиям, заданным к ЦФ;
- 3) применение к аналоговому фильтру выбранного метода отображения p -области в z -область.

Рассчитываемый по требованиям, заданным к ЦФ, аналоговый фильтр называется *фильтром-прототипом*, или просто *прототипом*.

Основными ограничениями для методов синтеза ЦФ по фильтрам-прототипам являются:

- 1) *сохранение существенных свойств АЧХ фильтра-прототипа в АЧХ ЦФ*, что означает необходимость отображения мнимой оси $j\Omega$ p -области на единичную окружность z -области;
- 2) *преобразование устойчивого фильтра-прототипа в устойчивый ЦФ*, что означает необходимость отображения полюсов устойчивого прототипа из левой p -полуплоскости внутрь единичного круга z -области;

3) отображение p -области в z -область должно представлять рациональную функцию, с использованием которой рациональная передаточная функция $H_a(p)$ аналогового фильтра-прототипа отображается в рациональную передаточную функцию $H(z)$ цифрового фильтра:

$$H(z) = K \frac{(z - z_1)(z - z_2) \dots (z - z_M)}{(z - p_1)(z - p_2) \dots (z - p_N)},$$

где z_i – i -й нуль, p_i – i -й полюс, K – коэффициент усиления, а $M \leq N$, для того чтобы фильтр с передаточной функцией $H(z)$ был физически реализуемым.

Из ряда известных методов преобразования аналоговых фильтров-прототипов в ЦФ можно выделить следующие:

- 1) инвариантного преобразования импульсной характеристики;
- 2) согласованного z -преобразования;
- 3) отображения дифференциалов;
- 4) билинейного z -преобразования.

Необходимо отметить, что синтез ЦФ с использованием аналогового фильтра-прототипа начинается с построения аналогового ФНЧ. Поэтому в следующем разделе описываются классические аналоговые ФНЧ, рассматриваемые в работе, и их основные характеристики.

Замечание. При дальнейшем изложении материала для обозначения характеристик аналоговых фильтров вводится в рассмотрение аналоговая частота Ω [рад/с], а для обозначения характеристик цифровых фильтров – цифровая частота ω [рад/с].

8.2. Аналоговые фильтры-прототипы нижних частот

Исторически сложилось так, что большинство методов аппроксимации аналоговых фильтров являются развитыми для синтеза пассивных систем, т. е. систем, состоящих из пассивных компонентов, таких как, например, конденсаторы и резисторы. Пассивные системы не требуют источника энергии для своего

функционирования, и, в отличие от активных систем, в них не происходит усиления сигнала по мощности. Поэтому, если обозначить через $H_a(j\Omega)$ ЧХ аналогового ФНЧ, спецификации на его синтез могут быть записанными в следующем виде:

$$\frac{1}{\sqrt{1+\varepsilon^2}} \leq |H_a(j\Omega)| \leq 1, \quad |\Omega| \leq \Omega_n,$$

$$0 \leq |H_a(j\Omega)| \leq \frac{1}{A}, \quad |\Omega| \geq \Omega_3,$$

где ε – параметр, определяющий пульсации в ПП, A – параметр, определяющий пульсации в ПЗ, Ω_n и Ω_3 – граничные частоты в ПП и ПЗ соответственно, измеряемые в $[rad/c]$. Диаграмма требований к АЧХ аналогового ФНЧ в терминах ε и A представлена на рис. 8.1а. Дополнительно необходимо отметить, что

$$|H_a(j\Omega_n)| = \frac{1}{\sqrt{1+\varepsilon^2}}, \quad |H_a(j\Omega_3)| = \frac{1}{A}.$$

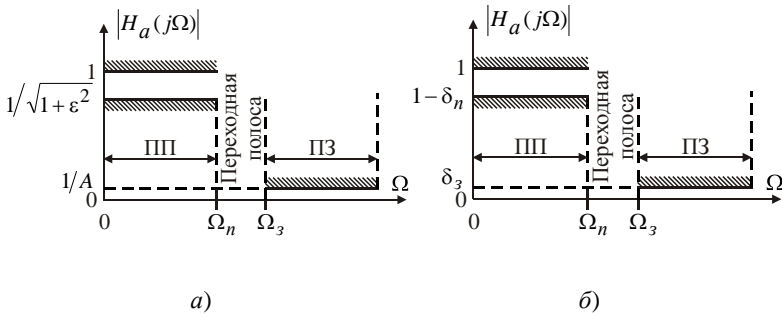


Рис. 8.1. Диаграммы требований к АЧХ аналогового ФНЧ:
а) в терминах ε и A ; б) в терминах δ_n и δ_3

На рис. 8.1б приведен другой способ задания требований на синтез аналогового ФНЧ в терминах отклонений АЧХ в ПП (δ_n) и ПЗ (δ_s). Взаимосвязи между параметрами δ_n , δ_s и ε , A определяются следующими выражениями:

$$1 - \delta_n = \frac{1}{\sqrt{1 + \varepsilon^2}}, \quad \delta_s = \frac{1}{A}.$$

Известны четыре стандартных вида классических аналоговых ФНЧ, получивших свое название по виду аппроксимации: Баттерворта, Чебышёва типа I и типа II, эллиптические, или Золотарева-Кауэра.

Предварительно отметим, что в приводимых ниже методах аппроксимации используется квадрат модуля ЧХ аналогового фильтра $|H_a(j\Omega)|^2$, т. е. квадрат АЧХ. Это объясняется тем, что функция $|H_a(j\Omega)|^2$, являясь вещественной функцией вещественного аргумента, существенно упрощает решение задачи наилучшей аппроксимации.

8.2.1. Фильтры Баттерворта

Функция квадрата АЧХ порядка N аналогового ФНЧ в данном случае имеет вид

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \left(\frac{\Omega}{\Omega_c}\right)^{2N}},$$

где Ω_c – частота среза. Пример квадрата АЧХ аналогового ФНЧ Баттерворта приведен на рис. 8.2.

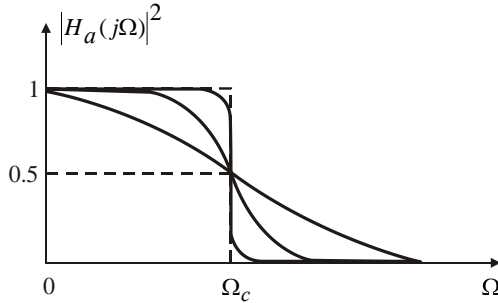


Рис. 8.2. Квадрат АЧХ аналогового ФНЧ Баттерворта

Основные свойства квадрата модуля ЧХ $|H_a(j\Omega)|^2$ аналогового ФНЧ Баттерворта следующие:

1. $|H_a(j0)|^2 = 1$ для всех значений N ;
2. $|H_a(j\Omega_c)|^2 = \frac{1}{2}$ для всех значений N ;
3. $|H_a(j\Omega)|^2$ является монотонно убывающей функцией частоты Ω ;
4. $|H_a(j\Omega)|^2$ стремится к идеальному ФНЧ при $N \rightarrow \infty$;
5. $|H_a(j\Omega)|^2$ является максимально плоской на частотах $\Omega = 0$ и $\Omega = \infty$, так как в этих точках функция квадрата АЧХ аналогового ФНЧ Баттерворта N -го порядка имеет равными нулю первые N производных. Поэтому фильтр Баттерворта называют *оптимальным по критерию максимальной плоскости функции* $|H_a(j\Omega)|^2$ в точках частот $\Omega = 0$ и $\Omega = \infty$.

Для дальнейшего анализа введем в рассмотрение передаточную функцию аналогового ФНЧ Баттерворта $H_a(p)$, с использованием которой можно записать следующее выражение:

$$H_a(p)H_a(-p) = \left| H_a(j\Omega) \right|_{\Omega=p/j}^2 = \frac{1}{1 + \left(\frac{p}{j\Omega_c} \right)^{2N}} = \frac{(j\Omega_c)^{2N}}{p^{2N} + (j\Omega_c)^{2N}}.$$

Корни полинома знаменателя, или полюсы $H_a(p)H_a(-p)$, имеют следующий вид:

$$p_k = (-1)^{2N} (j\Omega_c) = \Omega_c e^{j\frac{\pi}{2N}(2k+N+1)}, \quad k = 0, 1, \dots, 2N-1. \quad (8.1)$$

Интерпретация выражения (8.1) является следующей:

1) функция $H_a(p)H_a(-p)$ имеет $2N$ полюсов, которые являются равномерно распределенными на окружности радиуса Ω_c с угловым шагом π/N [рад];

2) для нечетного значения N выражение (8.1) принимает вид

$$p_k = \Omega_c e^{j\frac{\pi k}{N}}, \quad k = 0, 1, \dots, 2N-1;$$

3) для четного значения N выражение (8.1) принимает вид

$$p_k = \Omega_c e^{j\left(\frac{\pi}{2N} + \frac{\pi k}{N}\right)}, \quad k = 0, 1, \dots, 2N-1;$$

4) полюсы функции $H_a(p)H_a(-p)$ являются симметрично расположенными относительно мнимой оси $j\Omega$;

5) полюсы функции $H_a(p)H_a(-p)$ при любых значениях N не лежат на мнимой оси $j\Omega$, а в случае нечетного значения N лежат на действительной оси p -плоскости.

Пример расположения полюсов функции $H_a(p)H_a(-p)$ для нечетного и четного порядка фильтров Баттерворта представлен на рис. 8.3.

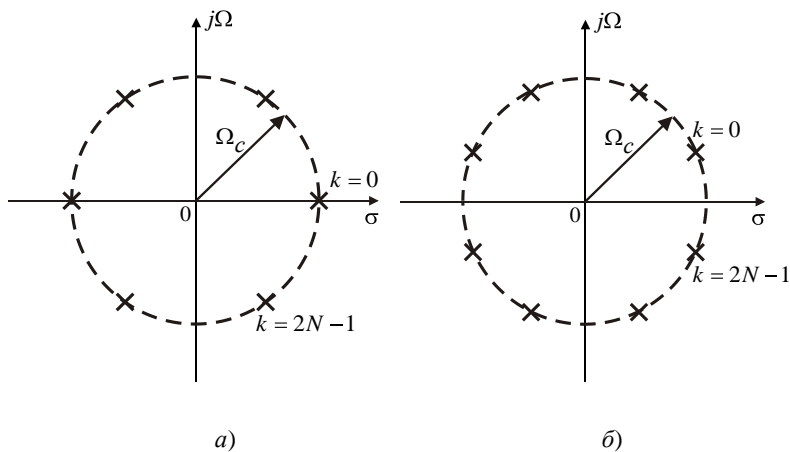


Рис. 8.3. Пример расположения полюсов функции $H_a(p)H_a(-p)$ для фильтров Баттерворта: а) нечетный порядок N фильтра; б) четный порядок N фильтра

Передаточная функция $H_a(p)$ устойчивого, физически реализуемого аналогового ФНЧ Баттерворта состоит из N полюсов функции $H_a(p)H_a(-p)$, лежащих в левой p -полуплоскости:

$$H_a(p) = \frac{\Omega_c^N}{\prod_{k=0}^{N-1} (p - p_k)}, \quad (8.2)$$

где $p_k = \Omega_c e^{j\frac{\pi}{2N}(2k+N+1)}$, $k = 0, 1, \dots, N-1$.

Основные этапы синтеза аналогового ФНЧ Баттерворта при заданных значениях $\Omega_{П}$, Ω_3 и $\delta_{П}$, δ_3 (ε , A) включают в себя:

1) определение порядка фильтра Баттерворта, позволяющего выполнить заданные спецификации на синтез фильтра:

$$N \geq \frac{\lg \left(\frac{(1-\delta_n)^{-2} - 1}{\delta_3^{-2} - 1} \right)^{1/2}}{\lg \left(\frac{\Omega_n}{\Omega_3} \right)} = \frac{\lg \left(\frac{\varepsilon}{\sqrt{A^2 - 1}} \right)}{\lg \left(\frac{\Omega_n}{\Omega_3} \right)};$$

2) определение частоты среза Ω_c из множества частот:

$$\Omega_n \left[(1-\delta_n)^{-2} - 1 \right]^{-1/2N} \leq \Omega_c \leq \Omega_3 \left[\delta_3^{-2} - 1 \right]^{-1/2N};$$

3) синтез передаточной функции фильтра с использованием выражения (8.2).

Замечание. Достоинством фильтров Баттерворта является их близкая к линейной ФЧХ в ПП. С этой точки зрения они имеют существенное преимущество перед другими, рассматриваемыми далее, классическими аналоговыми фильтрами.

Основное свойство ФНЧ Баттерворта – равенство АЧХ величине $1/\sqrt{2}$ на частоте среза Ω_c – говорит о том, что процедура аппроксимации сосредоточивается на двух частотах – $\Omega = 0$ и $\Omega = \infty$. Поэтому единственным параметром для задания функции $|H_a(j\Omega)|^2$ ФНЧ Баттерворта является порядок фильтра N , определяющий степень плоскости АЧХ на указанных крайних частотах.

Фильтры Баттерворта применяются при необходимости сохранения соотношений составляющих сигнала по амплитуде и фазе.

8.2.2. Фильтры Чебышёва

Удовлетворение требований к фильтру может быть обеспечено при меньших порядках, если ошибку аппроксимации равномерно распределить по ПП или по ПЗ или, что еще лучше, – по обеим указанным полосам. Тогда получим равноволновую аппроксимацию и порядок фильтра, существенно меньший порядка фильтра Баттерворта. Такими свойствами обладают

фильтры Чебышёва. В зависимости от области, в которой минимизируется ошибка аппроксимации, различают фильтры Чебышёва типа I и типа II:

1) фильтры Чебышёва типа I в ПП имеют равноволновый характер аппроксимации, а в ПЗ АЧХ монотонно убывает;

2) фильтры Чебышёва типа II в ПЗ имеют равноволновый характер аппроксимации, а в ПП АЧХ является максимально плоской (подобно фильтрам Баттерворта); т. е. вид аппроксимации фильтров Чебышёва типа II является обратным к виду аппроксимации фильтров Чебышева типа I, поэтому фильтры Чебышева типа II часто называют *инверсными*.

Фильтры Чебышёва типа I

Как уже отмечалось, фильтры Чебышёва типа I обладают равноволновой АЧХ в ПП и монотонной АЧХ – в ПЗ. Функция квадрата АЧХ N -го порядка ФНЧ Чебышёва типа I имеет вид

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \varepsilon^2 T_N^2\left(\frac{\Omega}{\Omega_c}\right)},$$

где ε – параметр, управляющий величиной пульсаций в ПП, $T_N(x)$ – полином Чебышёва N -го порядка, который определяется следующим выражением:

$$T_N(x) = \begin{cases} \cos(N \arccos(x)), & 0 \leq x \leq 1, \\ \cosh(N \cosh^{-1}(x)), & 1 < x < \infty, \end{cases}$$

где $x = \frac{\Omega}{\Omega_c}$, $\cosh(x) = \frac{e^x + e^{-x}}{2}$ и $\cosh^{-1}(x) = \lg[x + (x^2 - 1)^{1/2}]$.

Необходимо отметить, что полиномы Чебышёва могут быть получены рекурсивно:

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \quad k \geq 1, \quad T_0(x) = 1, \quad T_1(x) = x.$$

Основные свойства полиномов Чебышёва приведены ниже.

1. Если $|x| \leq 1$, то $|T_N(x)| \leq 1$. Если $|x| > 1$, то $T_N(x)$ монотонно увеличивается с ростом x .

2. $T_N(1) = 1$ для любых значений N .

3. $T_N(0) = \pm 1$ для четных N и $T_N(0) = 0$ для нечетных N .

4. Функция $T_N(x)$ обращается в 0 только при $|x| \leq 1$.

Примеры функции $|H_a(j\Omega)|^2$ для фильтров Чебышёва типа I приведены на рис. 8.4.

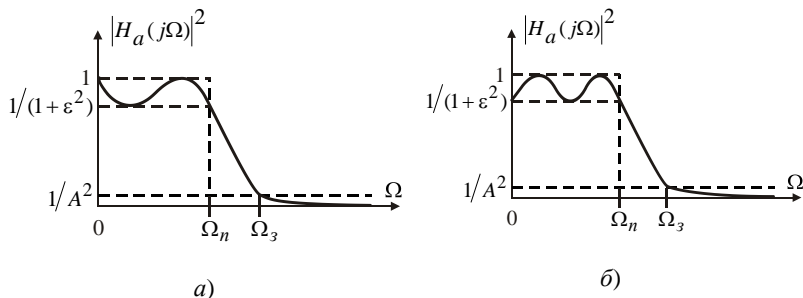


Рис. 8.4. Примеры квадратов АЧХ фильтров Чебышёва типа I: а) нечетный порядок N фильтра; б) четный порядок N фильтра

Основные свойства квадрата модуля частотной характеристики $|H_a(j\Omega)|^2$ аналогового ФНЧ Чебышёва типа I следующие:

$$1. |H_a(j0)|^2 = 1 \text{ для нечетных значений } N, |H_a(j0)|^2 = \frac{1}{1 + \varepsilon^2}$$

для четных значений N ;

$$2. |H_a(j\Omega_c)|^2 = \frac{1}{1 + \varepsilon^2} \text{ для всех значений } N;$$

$$3. \frac{1}{1 + \varepsilon^2} \leq |H_a(j\Omega)|^2 \leq 1 \text{ при } 0 \leq \Omega \leq \Omega_c;$$

$$4. |H_a(j\Omega)|^2 \text{ монотонно уменьшается до } 0 \text{ при } \Omega > \Omega_c;$$

$$5. |H_a(j\Omega_3)|^2 = \frac{1}{A^2}.$$

Для определения передаточной функции $H_a(p)$ устойчивого, физически реализуемого аналогового ФНЧ Чебышёва типа I необходимо найти полюсы функции $H_a(p)H_a(-p)$ и выбрать из них те, которые лежат в левой p -полуплоскости. Полюсы $H_a(p)H_a(-p)$ находятся путем вычисления корней следующей функции:

$$1 + \varepsilon^2 T_N^2 \left(\frac{p}{j\Omega_c} \right). \quad (8.3)$$

Можно показать, что если $p_k = \sigma_k + j\Omega_k$, $k = 0, \dots, N-1$ являются корнями функции (8.3), лежащими в левой полуплоскости, то

$$\sigma_k = a\Omega_c \cos \left(\frac{\pi}{2} + \frac{(2k+1)\pi}{2N} \right), \quad \Omega_k = b\Omega_c \sin \left(\frac{\pi}{2} + \frac{(2k+1)\pi}{2N} \right),$$

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

где $a = \frac{1}{2} \left(\sqrt[N]{\alpha} - \sqrt[N]{1/\alpha} \right)$, $b = \frac{1}{2} \left(\sqrt[N]{\alpha} + \sqrt[N]{1/\alpha} \right)$ и $\alpha = \frac{1}{\varepsilon} + \sqrt{1 + \frac{1}{\varepsilon^2}}$.

Эти корни лежат на эллипсе с большой осью $b\Omega_c$ и малой осью $a\Omega_c$. С учетом сказанного выше передаточная функция $H_a(p)$ устойчивого, физически реализуемого аналогового ФНЧ Чебышёва типа I имеет вид

$$H_a(p) = H_a(j0) \prod_{k=0}^{N-1} \frac{-p_k}{p - p_k}, \quad (8.4)$$

$$H_a(j0) = \begin{cases} 1, & \text{при нечетном } N, \\ \frac{1}{\sqrt{1 + \varepsilon^2}}, & \text{при четном } N. \end{cases}$$

Основные этапы синтеза аналогового ФНЧ Чебышёва типа I при заданных значениях $\Omega_n = \Omega_c$, Ω_3 и δ_n , δ_3 (ε , A) включают в себя:

1) определение порядка фильтра Чебышёва типа I, позволяющего выполнить заданные спецификации на синтез фильтра:

$$N \geq \frac{\cosh^{-1}\left(\frac{\delta_s^{-2}-1}{(1-\delta_n)^{-2}-1}\right)^{1/2}}{\cosh^{-1}\left(\frac{\Omega_s}{\Omega_n}\right)} = \frac{\cosh^{-1}\left(\frac{\sqrt{A^2-1}}{\varepsilon}\right)}{\cosh^{-1}\left(\frac{\Omega_s}{\Omega_n}\right)};$$

2) синтез передаточной функции фильтра с использованием выражения (8.4).

Замечание. Фильтры Чебышёва типа I наиболее широко применяются в радиосистемах для подавления помех от соседних радиостанций.

Фильтры Чебышёва типа II

Фильтры Чебышёва типа II обладают монотонной АЧХ в ПП и равноволновой – в ПЗ. Функция квадрата АЧХ N -го порядка ФНЧ Чебышёва типа II имеет вид

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \left[\varepsilon^2 T_N^2\left(\frac{\Omega_c}{\Omega}\right) \right]^{-1}},$$

откуда следует, что фильтры Чебышева типа II обладают не только полюсами, но и нулями. Дополнительно необходимо отметить, что выражение для квадрата АЧХ фильтра Чебышёва типа II фактически получено из аналогичного выражения для фильтра Чебышёва типа I путем замены в последнем компонента

$$\varepsilon^2 T_N^2\left(\frac{\Omega}{\Omega_c}\right) \text{ на } \left[\varepsilon^2 T_N^2\left(\frac{\Omega_c}{\Omega}\right) \right]^{-1}.$$

Уравнения синтеза для фильтров Чебышёва типа II являются аналогичными тем, что были введены для фильтров Чебышёва типа I, за исключением того, что $\Omega_s = \Omega_c$, так как пульсации АЧХ

для данного типа фильтров наблюдаются в ПЗ. Примеры квадратов АЧХ фильтров Чебышёва типа II приведены на рис. 8.5.

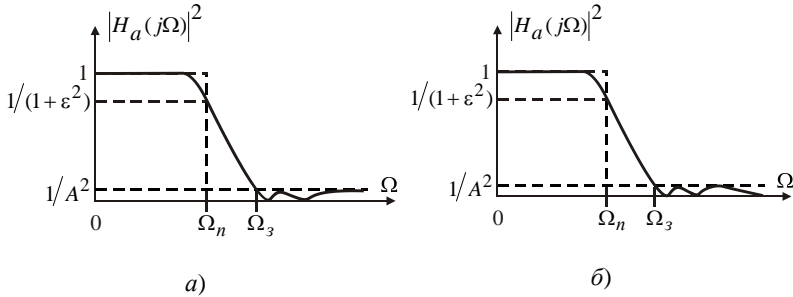


Рис. 8.5. Примеры квадратов АЧХ фильтров Чебышёва типа II: а) нечетный порядок N фильтра; б) четный порядок N фильтра

Замечание. Фильтры Чебышёва типа II применяются значительно реже, чем фильтры Чебышёва типа I, поскольку они требуют реализации нулей. Это не только усложняет программную и аппаратную реализацию, но и увеличивает собственный шум фильтра за счет введения умножений в числитель передаточной функции.

8.2.3. Эллиптические фильтры (фильтры Золотарева-Кауэра)

Эллиптические фильтры характеризуются равноволновой АЧХ как в ПП, так и в ПЗ. Они являются оптимальными в том смысле, что среди всех других фильтров имеют наименьшую ширину переходной полосы для заданного порядка N , $\Omega_n = \Omega_c$ и пульсаций δ_n , δ_3 (ε , A) в ПП и ПЗ соответственно.

Функция квадрата АЧХ N -го порядка ФНЧ Золотарева-Кауэра имеет вид

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \varepsilon^2 U_N^2 \left(\frac{\Omega}{\Omega_c} \right)},$$

где ε – параметр, управляющий величиной пульсаций в ПП, $U_N(x)$ – эллиптическая функция Якоби N -го порядка.

Примеры квадратов АЧХ эллиптических фильтров приведены на рис. 8.6.

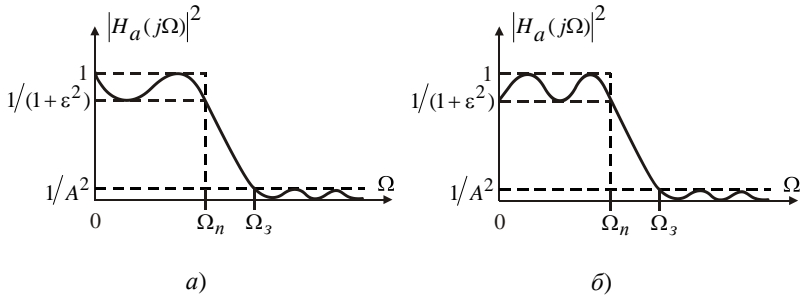


Рис. 8.6. Примеры квадратов АЧХ эллиптических фильтров: а) нечетный порядок N фильтра; б) четный порядок N фильтра

Теория эллиптических функций очень сложна и здесь не рассматривается. Заметим только, что при вычислении полюсов фильтров Золотарева-Кауэра необходимо обращаться к эллиптическим интегралам, для вычисления которых в программах синтеза фильтров используются численные методы.

Замечание. Фильтры Золотарева-Кауэра применяются в задачах, требующих высокой избирательности и некритичных к виду ФЧХ. Дополнительно необходимо отметить, что эллиптические фильтры являются очень трудными для анализа и, следовательно, для синтеза. Однако порядок фильтров Золотарева-Кауэра при заданных спецификациях на фильтр можно оценить с использованием сравнительно простого выражения:

$$N \geq \frac{\lg \left(16 \frac{\delta_s^{-2} - 1}{(1 - \delta_n)^2 - 1} \right)}{\lg \left(\frac{1}{q_0 + 2q_0^5 + 15q_0^9 + 150q_0^{13}} \right)} = \frac{\lg \left(16 \frac{A^2 - 1}{\varepsilon^2} \right)}{\lg \left(\frac{1}{q_0 + 2q_0^5 + 15q_0^9 + 150q_0^{13}} \right)},$$

$$\text{где } q_0 = \frac{1}{2} \cdot \frac{1 - (1 - k^2)^{1/4}}{1 + (1 - k^2)^{1/4}}, \quad k = \frac{\Omega_n}{\Omega_z}.$$

8.3. Билинейное z -преобразование

Билинейное z -преобразование представляет метод синтеза цифровых БИХ-фильтров по аналоговому прототипу. Метод основан на представлении комплексного оператора p как функции комплексной переменной z . Такое представление оператора p , имеющее вид дробно-рациональной функции от z^{-1} первого порядка, и называется билинейным z -преобразованием.

8.3.1. Определение билинейного z -преобразования

Для перехода от p -области, которая описывает аналоговый фильтр-прототип, в z -область, которая описывает цифровой фильтр, можно воспользоваться следующим соотношением между комплексными переменными p и z :

$$z = e^{pT_\Delta} \quad \text{или} \quad p = \frac{1}{T_\Delta} \ln(z), \quad (8.5)$$

где T_Δ – период дискретизации. Формально с использованием выражения (8.5), которое получается из прямой взаимосвязи между преобразованием Лапласа и z -преобразованием, можно осуществить отображение передаточной функции $H_a(p)$ аналогового фильтра-прототипа в передаточную функцию $H(z)$ цифрового фильтра. Однако подстановка выражения $p = \frac{1}{T_\Delta} \ln(z)$ в рациональную функцию $H_a(p)$ приводит к получению физически нереализуемой передаточной функции, так как последняя не представляет отношения двух полиномов от z .

В связи с этим необходимо вводить в рассмотрение другие отображения p -области в z -область, которые должны удовлетворять трем основным свойствам, перечисленным в п. 8.1.

Одно из таких отображений может быть получено, если выражение (8.5) представить в следующем виде:

$$z = e^{pT_0} = \frac{e^{\frac{pT_0}{2}}}{e^{-\frac{pT_0}{2}}} \approx \frac{1 + \frac{pT_0}{2}}{1 - \frac{pT_0}{2}}. \quad (8.6)$$

Из выражения (8.6) можно получить следующую взаимосвязь между комплексными переменными p и z :

$$p = \frac{2}{T_0} \frac{1 - z^{-1}}{1 + z^{-1}}. \quad (8.7)$$

Передаточная функция $H(z)$ цифрового фильтра может быть получена из передаточной функции $H_a(p)$ аналогового фильтра путем замены переменной p на переменную z , согласно выражению (8.7), т. е.

$$H(z) = H_a(p) \Big|_{p = \frac{2}{T_0} \frac{1 - z^{-1}}{1 + z^{-1}}}. \quad (8.8)$$

Преобразование (8.7), представляющее собой конформное отображение p -области в z -область и являющееся рациональной функцией первого порядка от z^{-1} , называется билинейным z -преобразованием. Этот тип преобразования обеспечивает однозначное отображение p -плоскости на z -плоскость и позволяет реализовать эффективный метод синтеза ЦФ, который удовлетворяет трем основным свойствам, перечисленным в п. 8.1.

8.3.2. Свойства билинейного z -преобразования

1. *Билинейное z -преобразование обеспечивает однозначное отображение p -плоскости на z -плоскость (рис. 8.7).*

Частотная ось $j\Omega$ однозначно отображается в единичную окружность. Действительно, единичная окружность на комплексной плоскости описывается равенством $z = e^{j\omega T_0}$, а оператор $p = \sigma + j\Omega$ для частотной оси имеет вид $p = j\Omega$, поскольку $\sigma = 0$. Тогда из (8.7) получаем следующее выражение:

$$z = \frac{2/T_0 + j\Omega}{2/T_0 - j\Omega} = r e^{j\phi(\Omega)} = \underbrace{\sqrt{\frac{(2/T_0)^2 + \Omega^2}{(2/T_0)^2 + \Omega^2}}}_{r=1} e^{j\phi(\Omega)}. \quad (8.9)$$

Поскольку радиус окружности $r = 1$, становится ясным, что выражение (8.9) описывает единичную окружность на комплексной плоскости.

Рассмотрим, каким образом отображаются положительная и отрицательная области частотной оси $j\Omega$ на z -плоскость, для чего определим аргумент z из выражения (8.9):

$$\arg(z) = \phi(\Omega) = \arctg\left(\frac{\Omega T_0}{2}\right) - \arctg\left(-\frac{\Omega T_0}{2}\right) = 2\arctg\left(\frac{\Omega T_0}{2}\right).$$

Рассмотрим отображение некоторых точек $p = j\Omega$ на z -плоскость:

$$\Omega = 0; \phi(\Omega) = 0; z = e^{j\phi(\Omega)} = 1;$$

$$\Omega = \pm\infty; \phi(\Omega) = \pm 2\arctg(\infty) = \pm 2\frac{\pi}{2} = \pm\pi; z = e^{j\phi(\Omega)} = e^{\pm j\pi} = -1;$$

$$\Omega = \pm\frac{2}{T_0}; \phi(\Omega) = \pm 2\arctg(1) = \pm 2\frac{\pi}{4} = \pm\frac{\pi}{2}; z = e^{j\phi(\Omega)} = e^{\pm j\pi/2} = \pm j.$$

Полученные соответствия указаны в табл. 8.1.

Отображение оси $j\Omega$ на z -плоскость

Частота Ω	Значение $\varphi(\Omega)$	Значение z
0	0	1
$\pm \infty$	$\pm \pi$	-1
$\pm 2/T_\theta$	$\pm \pi/2$	$\pm j$

Продолжая подобные вычисления, можно убедиться (рис. 8.7), что ось $j\Omega$ однозначно отображается в единичную окружность на z -плоскости, причем положительная область частотной оси $j\Omega$ отображается на верхнюю, а отрицательная – на нижнюю полуокружность.

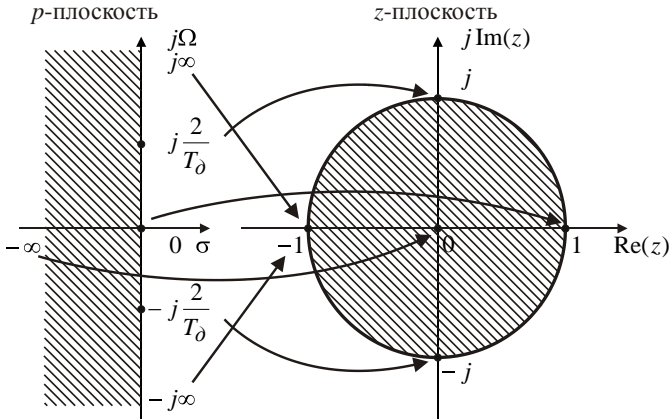


Рис. 8.7. Отображение p -плоскости на z -плоскость при билинейном z -преобразовании

Левая p -полуплоскость однозначно отображается внутрь единичного круга, а правая p -полуплоскость – вне его.

Для доказательства этого утверждения достаточно знать, в каких пределах лежат значения радиуса r . Воспользуемся выражением (8.7), в которое подставим оператор $p = \sigma + j\Omega$, и исследуем $r = |z|$ при положительных и отрицательных значениях σ :

$$|z| = \left| \frac{2/T_\delta + \sigma + j\Omega}{2/T_\delta - \sigma - j\Omega} \right| = \sqrt{\frac{(2/T_\delta + \sigma)^2 + \Omega^2}{(2/T_\delta - \sigma)^2 + \Omega^2}}.$$

Числитель подкоренного выражения при $\sigma < 0$, что соответствует левой p -полуплоскости, оказывается меньше знаменателя, поэтому $r < 1$ и левая p -полуплоскость отображается внутрь единичного круга z -плоскости. При $\sigma > 0$, что соответствует правой p -полуплоскости, числитель оказывается больше знаменателя, поэтому $r > 1$ и правая p -полуплоскость отображается на всю z -плоскость вне единичного круга.

2. ЦФ устойчив, если устойчив его аналоговый фильтр-прототип, поскольку полюсы последнего лежат в левой p -полуплоскости, отображаемой внутрь единичного круга z -плоскости.

3. Соотношение между аналоговыми Ω и цифровыми ω частотами принципиально нелинейно, т. е. шкала частот деформируется.

Найдем это соотношение из (8.7) при $p = j\Omega$ и $z = e^{j\omega T_\delta}$, поскольку рассматриваются только частотная ось и единичная окружность:

$$j\Omega = \frac{2}{T_\delta} \frac{1 - e^{-j\omega T_\delta}}{1 + e^{-j\omega T_\delta}} = j \frac{2}{T_\delta} \frac{(e^{j\omega T_\delta/2} - e^{-j\omega T_\delta/2})}{(e^{j\omega T_\delta/2} + e^{-j\omega T_\delta/2})},$$

откуда следует

$$\Omega = \frac{2}{T_\delta} \operatorname{tg} \left(\frac{\omega T_\delta}{2} \right), \quad (8.10)$$

$$\omega = \frac{2}{T_\delta} \operatorname{arctg} \left(\frac{\Omega T_\delta}{2} \right). \quad (8.11)$$

При $\omega < 0.3/T_\delta$ получается $\Omega \approx \omega$. Следовательно, приблизительно для одних и тех же частот цифровой и аналоговый фильтры имеют одинаковые АЧХ. Однако для более

высоких частот соотношение между Ω и ω становится нелинейным и, как показано на рис. 8.8, в шкалу частот ЦФ вводятся искажения по отношению к частотной шкале аналогового фильтра. Этот эффект известен как *деформация шкалы частот*.

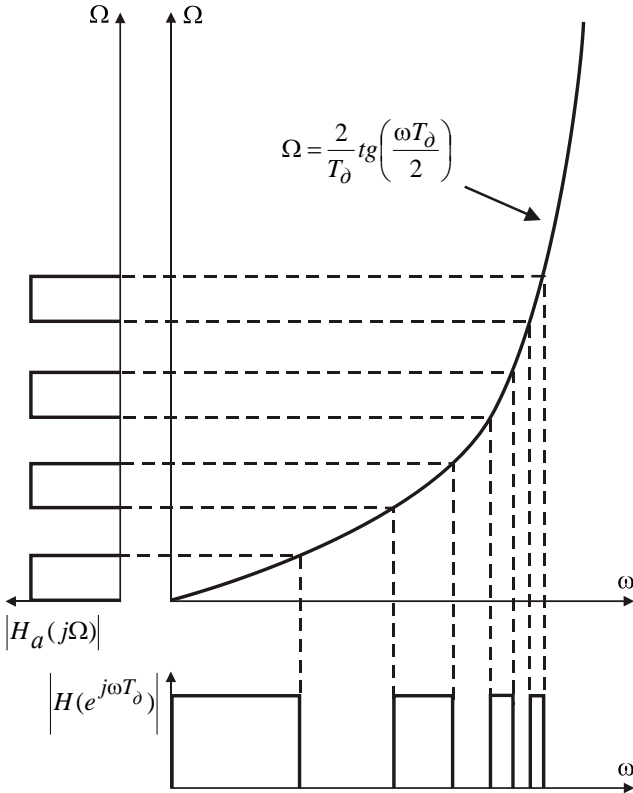


Рис. 8.8. Влияние деформации шкалы частот на АЧХ ЦФ

Влияние деформации шкалы частот можно продемонстрировать при рассмотрении аналогового фильтра, имеющего несколько полос пропускания, равномерно распределенных на частотной шкале (см. рис. 8.8). Получаемый ЦФ будет иметь такое же число полос пропускания, однако в

высокочастотной части шкалы ω полосы пропускания будут иметь тенденцию к непропорциональному уменьшению.

Если интересоваться только АЧХ, то деформация шкалы частот может быть скомпенсирована во всех практически важных случаях с помощью введения предыскажений в аналоговый фильтр. Пусть $\Omega_1, \Omega_2, \dots, \Omega_n$ – граничные частоты (частоты среза) полос пропускания и задерживания аналогового фильтра. Значения граничных частот соответствующих полос ЦФ определяются выражением

$$\omega_i = \frac{2}{T_d} \arctg \left(\frac{\Omega_i T_d}{2} \right), \quad i = 1, 2, \dots, n. \quad (8.12)$$

Таким образом, если у ЦФ должны быть получены заданные значения граничных частот $\tilde{\omega}_1, \tilde{\omega}_2, \dots, \tilde{\omega}_n$ полос пропускания и задерживания, то до применения билинейного z -преобразования в шкалу частот исходного аналогового фильтра должны быть введены предыскажения, обеспечивающие выполнение условий:

$$\Omega_i = \frac{2}{T_d} \operatorname{tg} \left(\frac{\tilde{\omega}_i T_d}{2} \right), \quad i = 1, 2, \dots, n. \quad (8.13)$$

При выполнении этих условий в соответствии с (8.12) и (8.13) получаем требуемые соотношения $\omega_i = \tilde{\omega}_i$. Влияние деформации шкалы частот на ФЧХ можно продемонстрировать на примере анализа аналогового фильтра с линейной ФЧХ $\arg(H_a(j\Omega))$. Как показано на рис. 8.9, ФЧХ $\arg(H(e^{j\omega T_d}))$ рассчитываемого цифрового фильтра нелинейна. Таким образом, если в ЦФ требуется сохранить линейность ФЧХ, то необходимо использовать другие методы расчета.

4. В цифровой области сохраняются свойства оптимальности АЧХ аналогового прототипа вследствие однозначности отображения частотной оси в единичную окружность, что исключает эффект наложения АЧХ.

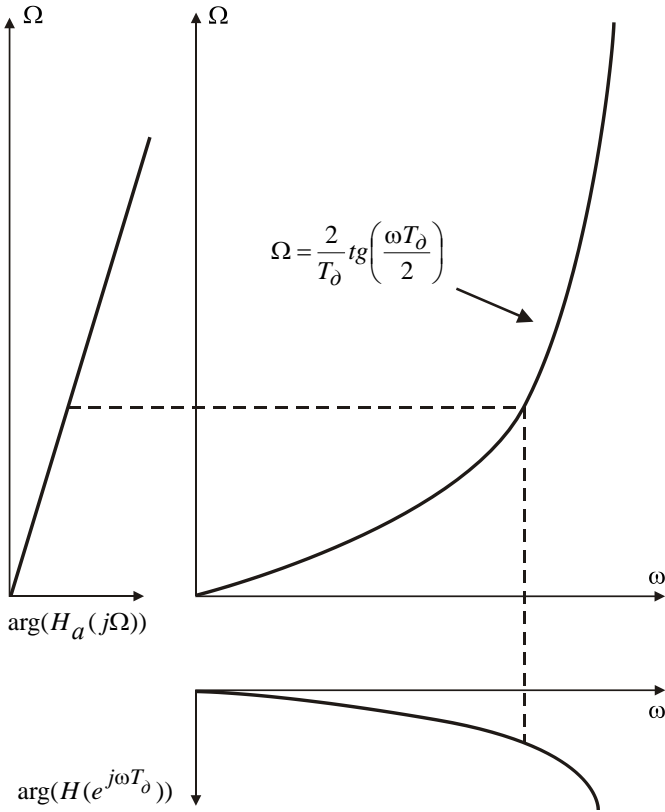


Рис. 8.9. Влияние деформации шкалы частот на ФЧХ ЦФ

8.3.3. Частотные преобразования

При расчете цифровых ФНЧ, ФВЧ, ПФ, РФ используются два подхода, представленные на рис. 8.10.

Различие между ними заключается в следующем. В первом случае аналоговый фильтр-прототип нижних частот с частотой среза Ω_c предварительно преобразуется в другой аналоговый фильтр (ФНЧ, ФВЧ, ПФ, РФ), из которого затем путем дискретизации, осуществляемой, например, с использованием

билинейного z -преобразования, рассчитывается ЦФ с заданными характеристиками.

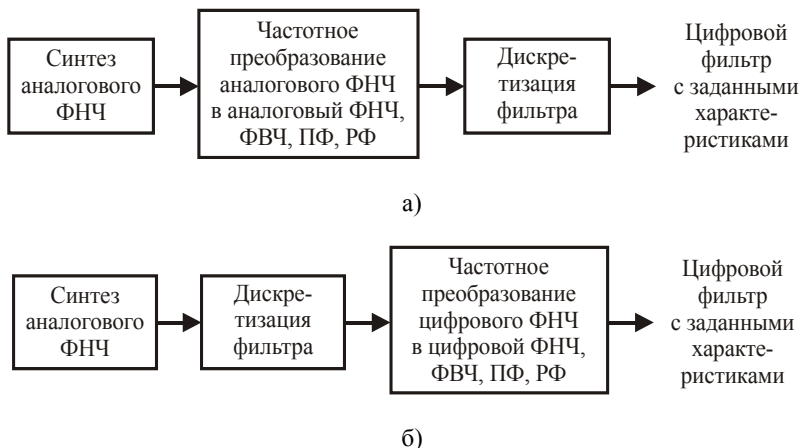


Рис. 8.10. Частотные преобразования ФНЧ

Во втором случае аналоговый фильтр-прототип нижних частот с частотой среза Ω_c дискретизируется сразу же, а затем путем преобразования его полосы формируется ЦФ (ФНЧ, ФВЧ, ПФ, РФ) с заданными характеристиками. Рассмотрим различные методы преобразования для аналоговых и цифровых фильтров.

Преобразование полосы частот для аналоговых фильтров

Наиболее простые преобразования аналогового фильтра-прототипа нижних частот с частотой среза Ω_c в другой аналоговый фильтр (ФНЧ, ФВЧ, ПФ, РФ) имеют следующий вид.

1. Преобразование ФНЧ с частотой среза Ω_c в ФНЧ с частотой среза Ω'_c : $p \rightarrow \frac{\Omega_c}{\Omega'_c} p$.

2. Преобразование ФНЧ с частотой среза Ω_c в ФВЧ с частотой среза Ω'_c : $p \rightarrow \frac{\Omega_c \Omega'_c}{p}$.

3. Преобразование ФНЧ с частотой среза Ω_c в ПФ с нижней $\Omega'_{сн}$ и верхней $\Omega'_{св}$ частотой среза: $p \rightarrow \Omega_c \frac{p^2 + \Omega'_{св} \Omega'_{сн}}{p(\Omega'_{св} - \Omega'_{сн})}$.

4. Преобразование ФНЧ с частотой среза Ω_c в РФ с нижней $\Omega'_{сн}$ и верхней $\Omega'_{св}$ частотой среза: $p \rightarrow \Omega_c \frac{p(\Omega'_{св} - \Omega'_{сн})}{p^2 + \Omega'_{св} \Omega'_{сн}}$.

Преобразование полосы частот для цифровых фильтров

По аналогии с фильтрами непрерывного времени существует несколько простых, предложенных Константином Константинидисом преобразований цифрового ФНЧ с частотой среза ω_c в ФНЧ, ФВЧ, ПФ, РФ. Приведем формулы для этих преобразований.

1. Преобразование ФНЧ с частотой среза ω_c в ФНЧ с частотой среза ω'_c : $z^{-1} \rightarrow \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}}$, где $\alpha = \frac{\sin[(\omega_c - \omega'_c)T_\delta / 2]}{\sin[(\omega_c + \omega'_c)T_\delta / 2]}$, где T_δ – период дискретизации.

2. Преобразование ФНЧ с частотой среза ω_c в ФВЧ с частотой среза ω'_c : $z^{-1} \rightarrow -\frac{z^{-1} - \alpha}{1 - \alpha z^{-1}}$, где $\alpha = \frac{\cos[(\omega_c - \omega'_c)T_\delta / 2]}{\cos[(\omega_c + \omega'_c)T_\delta / 2]}$, где T_δ – период дискретизации.

3. Преобразование ФНЧ с частотой среза ω_c в ПФ с нижней $\omega'_{сн}$ и верхней $\omega'_{св}$ частотой среза: $z^{-1} \rightarrow -\frac{z^{-2} - \alpha_1 z^{-1} + \alpha_2}{\alpha_2 z^{-2} - \alpha_1 z^{-1} + 1}$, где

$$\alpha_1 = \frac{2\beta\gamma}{\gamma+1}, \quad \alpha_2 = \frac{\gamma-1}{\gamma+1}, \quad \beta = \frac{\cos[(\omega'_{св} + \omega'_{сн})T_\delta / 2]}{\cos[(\omega'_{св} - \omega'_{сн})T_\delta / 2]},$$

$\gamma = ctg\left(\frac{(\omega'_{св} - \omega'_{сн})T_\delta}{2}\right)tg\left(\frac{\omega_c T_\delta}{2}\right)$, где T_δ – период дискретизации.

4. Преобразование ФНЧ с частотой среза ω_c в РФ с нижней

$\omega'_{сн}$ и верхней $\omega'_{св}$ частотой среза: $z^{-1} \rightarrow \frac{z^{-2} - \alpha_1 z^{-1} + \alpha_2}{\alpha_2 z^{-2} - \alpha_1 z^{-1} + 1}$, где

$$\alpha_1 = \frac{2\beta}{1+\gamma}, \quad \alpha_2 = \frac{1-\gamma}{1+\gamma}, \quad \beta = \frac{\cos[(\omega'_{св} + \omega'_{сн})T_\Delta / 2]}{\cos[(\omega'_{св} - \omega'_{сн})T_\Delta / 2]},$$

$\gamma = \operatorname{tg}\left(\frac{(\omega'_{св} - \omega'_{сн})T_\Delta}{2}\right) \operatorname{tg}\left(\frac{\omega_c T_\Delta}{2}\right)$, где T_Δ – период дискретизации.

Замечание. В пакете Matlab в качестве функций преобразования полосы частот используются функции, представленные в п. 8.3.3. Применительно к билинейному z -преобразованию результат, получаемый при синтезе цифрового фильтра с использованием частотных преобразований, изложенных в п. 8.3.3, является идентичным при одинаково заданных требованиях на синтез фильтра. Однако последнее является несправедливым, например, для синтеза БИХ-фильтров с использованием метода инвариантного преобразования импульсной характеристики.

8.3.4. Методика синтеза цифровых фильтров на основе билинейного z -преобразования

С использованием основных положений, сформулированных в п. 8.3.3, опишем два метода синтеза ЦФ на основе билинейного z -преобразования.

Методика синтеза ЦФ на основе билинейного z -преобразования с использованием трансформации полосы частот для аналоговых фильтров

1. Определить и деформировать граничные частоты (частоты среза) нужного фильтра с использованием выражения (8.13).

2. На основе спецификаций ЦФ и набора граничных частот, полученных на шаге 1, определить подходящий аналоговый фильтр-прототип нижних частот с передаточной функцией $H_a(p)$.

3. С использованием подхода, приведенного в п. 8.3.3, осуществить преобразование аналогового фильтра-прототипа нижних частот с частотой среза Ω_c в другой аналоговый фильтр (ФНЧ, ФВЧ, ПФ, РФ).

4. Применить билинейное z -преобразование с использованием выражения (8.8) к передаточной функции аналогового фильтра, полученной на шаге 3, и получить передаточную функцию $H(z)$ требуемого ЦФ.

5. Проверить выполнение заданных требований.

Методика синтеза ЦФ на основе билинейного z -преобразования с использованием трансформации полосы частот для цифровых фильтров

1. Определить и деформировать граничные частоты (частоты среза) нужного фильтра с использованием выражения (8.13).

2. На основе спецификаций ЦФ и набора граничных частот, полученных на шаге 1, определить подходящий аналоговый фильтр-прототип нижних частот с передаточной функцией $H_a(p)$.

3. Применить билинейное z -преобразование к функции $H_a(p)$ с использованием выражения (8.8) и получить передаточную функцию ЦФ.

4. С использованием подхода, приведенного в п. 8.3.3, осуществить преобразование цифрового ФНЧ с частотой среза ω_c , полученного на шаге 3, в требуемый ЦФ (ФНЧ, ФВЧ, ПФ, РФ).

5. Проверить выполнение заданных требований.

8.3.5. Особенности использования метода билинейного z -преобразования

Основным фактором, создающим ограничение на использование билинейного z -преобразования, является деформация шкалы частот. Последнее означает следующее.

1. Билинейное z -преобразование пригодно лишь в тех случаях, когда требования к АЧХ фильтра-прототипа задаются ступенчатой функцией частоты, что свойственно частотно-избирательным фильтрам (ФНЧ, ФВЧ, ПФ, РФ).

2. Билинейное z -преобразование не пригодно для синтеза корректирующих фильтров (фильтров с произвольными АЧХ), фильтров с линейной ФЧХ, а также для сохранения импульсной характеристики фильтра-прототипа. Иначе говоря, ни ФЧХ, ни ИХ аналогового прототипа не сохраняются.

8.4. Контрольные вопросы

1. Что такое аналоговый фильтр-прототип?
2. Дайте определение БИХ-фильтра.
3. Какие основные ограничения накладываются на методы синтеза ЦФ по аналоговым фильтрам-прототипам?
4. Приведите примеры методов синтеза БИХ-фильтров.
5. Какие существуют способы задания требований к цифровым БИХ-фильтрам?
6. Какую роль играют фильтры непрерывного времени при проектировании ЦФ?
7. Приведите примеры аналоговых фильтров-прототипов, перечислите их основные параметры.
8. Объясните сущность метода билинейного z -преобразования.
9. Перечислите и поясните основные свойства билинейного z -преобразования.
10. Что такое эффект деформации полосы частот? Приведите примеры.
11. Поясните два подхода, применяемые в задаче синтеза цифровых БИХ-фильтров по аналоговым фильтрам-прототипам.
12. Каковы основные особенности использования метода билинейного z -преобразования?

8.5. Литература

1. Брюханов, Ю. А. Цифровые фильтры: учеб. пособие / Ю. А. Брюханов, А. Л. Приоров. – Ярославль: ЯрГУ, 2002.
2. Приоров, А. Л. Основы цифровой обработки сигналов: лабораторный практикум / А. Л. Приоров, В. А. Волохов, А. М. Лаврентьев. – ВУНЦ ВВС «ВВА им. профессора Н. Е. Жуковского и Ю. А. Гагарина» (филиал, г. Ярославль), 2012.
3. Айфичер, Э. Цифровая обработка сигналов: практический подход. – 2-е изд. / Э. Айфичер, Б. Джервис. – Вильямс, 2004.
4. Сергиенко, А. Б. Цифровая обработка сигналов. – 3-е изд. / А. Б. Сергиенко. – СПб.: БХВ-Петербург, 2011.
5. Дьяконов, В. П. MATLAB 6.0/6.1/6.5/6.5+SP1 + Simulink 4/5. Обработка сигналов и изображений / В. П. Дьяконов. – М.: СОЛОН-Пресс, 2005.
6. Хрящев, В. В. Основы теории цепей: сборник задач / В. В. Хрящев, А. Л. Приоров, В. А. Волохов – Ярославль: ЯрГУ, 2008.
7. Солонина, А. И. Цифровая обработка сигналов. Моделирование в MATLAB / А. И. Солонина, С. М. Арбузов. – СПб.: БХВ-Петербург, 2008.

9. СИНТЕЗ ОПТИМАЛЬНЫХ ПО ЧЕБЫШЁВУ ЦИФРОВЫХ КИХ-ФИЛЬТРОВ

9.1. Основные определения

Простота расчета коэффициентов передаточной функции КИХ-фильтров методом окон сопровождается получением фильтров, которые обладают рядом недостатков:

1) расположение граничных частот синтезированного фильтра не соответствует заданным на него спецификациям;

2) пульсации в различных полосах фильтра аппроксимируются одинаково;

3) ошибка аппроксимации, т. е. разница между идеальной и реальной ЧХ фильтра, распределена неравномерно по различным частотным полосам. Причем пульсации АЧХ фильтра возрастают при приближении к границам соответствующих полос. Необходимо отметить, что равномерное распределение пульсаций АЧХ в различных полосах фильтра позволяет значительно снизить его порядок.

В теории ЦОС разработана методика, позволяющая устранить отмеченные выше недостатки синтезируемых ЦФ. Однако она является достаточно сложной для понимания и требует программной реализации.

В классе КИХ-фильтров с линейной ФЧХ существует возможность получить множество условий, для которых доказывается, что решение задачи синтеза ЦФ является оптимальным по Чебышёву, т. е. оптимальным в смысле минимума максимальной ошибки аппроксимации (иногда называемой минимаксной ошибкой, или ошибкой Чебышёва). Фильтры, обладающие свойством оптимальности, отмеченным выше, называются фильтрами с равномерными (равноволновыми) пульсациями, поскольку ошибка аппроксимации данного типа фильтров является равномерно распределенной по его частотным полосам. Как следствие последнего, ЦФ, оптимальные в смысле минимума максимальной ошибки аппроксимации, будут иметь более низкий порядок, чем ЦФ, синтезированные с использованием метода окон, при одинаково заданных спецификациях на синтез фильтров.

Далее будет сформулирована задача синтеза оптимальных фильтров в смысле минимума максимальной ошибки аппроксимации и проанализировано общее число максимумов и минимумов амплитудной характеристики КИХ-фильтров с линейной ФЧХ. С использованием этой формулировки может быть рассмотрен обобщенный алгоритм синтеза КИХ-фильтров с равномерными пульсациями, использующий полиномиальную интерполяцию при нахождении оптимального решения. Этот алгоритм получил название алгоритма Паркса-Макклеллана, включающего в себя алгоритм замены Ремеза для полиномиальных решений.

9.2. КИХ-фильтры с линейной ФЧХ

Фильтры с линейной ФЧХ могут быть классифицированы на четыре типа в зависимости от симметрии или антисимметрии ИХ $h(n)$, а также четности или нечетности N . Каждый из этих типов имеет собственные ограничения на расположение нулей передаточной функции $H(z)$, которая, в свою очередь, накладывает ограничения на характер ЧХ

$$H(e^{j\hat{\omega}}) = e^{j\beta} e^{-jN\hat{\omega}/2} A(\hat{\omega}),$$

где β – константа, зависящая от типа КИХ-фильтра с линейной ФЧХ, $A(\hat{\omega})$ – амплитудная характеристика, представляющая собой действительную функцию частоты $\hat{\omega}$ [рад] и принимающая как положительные, так и отрицательные значения. Модуль амплитудной характеристики есть АЧХ КИХ-фильтра с линейной ФЧХ.

9.2.1. Фильтры типа I – симметричная импульсная характеристика, N – четное

Центр симметрии ИХ проходит в точке $\alpha = N/2$, а ЧХ фильтра типа I с линейной ФЧХ выражается следующим образом:

$$H(e^{j\hat{\omega}}) = e^{-jN\hat{\omega}/2} \sum_{k=0}^{N/2} a(k) \cos(k\hat{\omega}),$$

где $a(k) = 2h\left(\frac{N}{2} - k\right)$, $k = 1, 2, \dots, \frac{N}{2}$, $a(0) = h\left(\frac{N}{2}\right)$.

Амплитудная характеристика имеет следующий вид:

$$A(\hat{\omega}) = \sum_{k=0}^{N/2} a(k) \cos(k\hat{\omega}).$$

9.2.2. Фильтры типа II – симметричная импульсная характеристика, N – нечетное

Центр симметрии ИХ проходит в точке $\alpha = N/2$, а ЧХ фильтра типа II с линейной ФЧХ может быть представлена в виде

$$H(e^{j\hat{\omega}}) = e^{-jN\hat{\omega}/2} \sum_{k=1}^{(N+1)/2} b(k) \cos\left[\left(k - \frac{1}{2}\right)\hat{\omega}\right],$$

где $b(k) = 2h\left(\frac{N+1}{2} - k\right)$, $k = 1, 2, \dots, \frac{N+1}{2}$.

Амплитудная характеристика имеет следующий вид:

$$A(\hat{\omega}) = \sum_{k=1}^{(N+1)/2} b(k) \cos\left[\left(k - \frac{1}{2}\right)\hat{\omega}\right].$$

Последнее выражение может быть представлено в виде

$$A(\hat{\omega}) = \cos\left(\frac{\hat{\omega}}{2}\right) \sum_{k=0}^{(N-1)/2} \tilde{b}(k) \cos(k\hat{\omega}),$$

где коэффициенты $\tilde{b}(k)$ являются линейно связанными с коэффициентами $b(k)$. Можно показать, что эта взаимосвязь имеет следующий вид:

$$\tilde{b}(0) = \frac{1}{2}b(1),$$

$$\tilde{b}(k) = 2b(k) - \tilde{b}(k-1), \quad k = 1, 2, \dots, \frac{N-3}{2},$$

$$\tilde{b}\left(\frac{N-1}{2}\right) = \frac{1}{2}b\left(\frac{N+1}{2}\right).$$

9.2.3. Фильтры типа III – антисимметричная импульсная характеристика, N – четное

Центр антисимметрии ИХ проходит в точке $\alpha = N/2$, а ЧХ фильтра типа III с линейной ФЧХ может быть представлена в виде

$$H(e^{j\hat{\omega}}) = je^{-jN\hat{\omega}/2} \sum_{k=1}^{N/2} c(k) \sin(k\hat{\omega}),$$

где $c(k) = 2h\left(\frac{N}{2} - k\right)$, $k = 1, 2, \dots, \frac{N}{2}$.

Амплитудная характеристика имеет следующий вид:

$$A(\hat{\omega}) = \sum_{k=1}^{N/2} c(k) \sin(k\hat{\omega}).$$

Последнее выражение может быть представлено в виде

$$A(\hat{\omega}) = \sin(\hat{\omega}) \sum_{k=0}^{(N-2)/2} \tilde{c}(k) \cos(k\hat{\omega}),$$

где коэффициенты $\tilde{c}(k)$ являются линейно связанными с коэффициентами $c(k)$. Можно показать, что эта взаимосвязь имеет следующий вид:

$$\tilde{c}\left(\frac{N-2}{2}\right) = c\left(\frac{N}{2}\right),$$

$$\tilde{c}\left(\frac{N-4}{2}\right) = c\left(\frac{N-2}{2}\right),$$

$$\tilde{c}(k-1) - \tilde{c}(k+1) = 2c(k), \quad 2 \leq k \leq \frac{N-4}{2},$$

$$\tilde{c}(0) - \frac{1}{2}\tilde{c}(2) = c(1).$$

9.2.4. Фильтры типа IV – антисимметричная импульсная характеристика, N – нечетное

Центр антисимметрии ИХ проходит в точке $\alpha = N/2$, при этом α не является целым числом, а ЧХ фильтра типа IV с линейной ФЧХ может быть представлена в виде

$$H(e^{j\hat{\omega}}) = je^{-jN\hat{\omega}/2} \sum_{k=1}^{(N+1)/2} d(k) \sin\left[\left(k - \frac{1}{2}\right)\hat{\omega}\right],$$

где $d(k) = 2h\left(\frac{N+1}{2} - k\right)$, $k = 1, 2, \dots, \frac{N+1}{2}$.

Амплитудная характеристика имеет следующий вид:

$$A(\hat{\omega}) = \sum_{k=1}^{(N+1)/2} d(k) \sin\left[\left(k - \frac{1}{2}\right)\hat{\omega}\right].$$

Последнее выражение может быть представлено в виде

$$A(\hat{\omega}) = \sin\left(\frac{\hat{\omega}}{2}\right) \sum_{k=0}^{(N-1)/2} \tilde{d}(k) \cos(k\hat{\omega}),$$

где коэффициенты $\tilde{d}(k)$ являются линейно связанными с коэффициентами $d(k)$. Можно показать, что эта взаимосвязь имеет следующий вид:

$$\tilde{d}\left(\frac{N-1}{2}\right) = 2d\left(\frac{N+1}{2}\right),$$

$$\tilde{d}(k-1) - \tilde{d}(k) = 2d(k), \quad 2 \leq k \leq \frac{N-1}{2},$$

$$\tilde{d}(0) - \frac{1}{2}\tilde{d}(1) = d(1).$$

В заключение рассмотренного выше необходимо отметить, что общее выражение для представления амплитудных характеристик КИХ-фильтров с линейной ФЧХ может быть записано в виде

$$A(\hat{\omega}) = Q(\hat{\omega})P(\hat{\omega}),$$

где выражения для $Q(\hat{\omega})$ и $P(\hat{\omega})$ приведены в табл. 9.1.

Таблица 9.1

$Q(\omega)$, L , $P(\omega)$ для КИХ-фильтров с линейной ФЧХ

Тип КИХ-фильтра с линейной ФЧХ	$Q(\hat{\omega})$	L	$P(\hat{\omega})$
тип 1	1	$\frac{N}{2}$	$\sum_{k=0}^L a(n) \cos(k\hat{\omega})$
тип 2	$\cos\left(\frac{\hat{\omega}}{2}\right)$	$\frac{N-1}{2}$	$\sum_{k=0}^L \tilde{b}(n) \cos(k\hat{\omega})$
тип 3	$\sin(\hat{\omega})$	$\frac{N-2}{2}$	$\sum_{k=0}^L \tilde{c}(n) \cos(k\hat{\omega})$
тип 4	$\sin\left(\frac{\hat{\omega}}{2}\right)$	$\frac{N-1}{2}$	$\sum_{k=0}^L \tilde{d}(n) \cos(k\hat{\omega})$

9.3. Теорема чередования

Для дальнейшего рассмотрения материала помимо функции, описывающей амплитудную характеристику $A(\hat{\omega})$ реального фильтра, введем функцию, описывающую амплитудную характеристику $A_u(\hat{\omega})$ идеального фильтра и весовую функцию $W(\hat{\omega})$, регулирующую относительный размер ошибки

аппроксимации в различных частотных полосах синтезируемого фильтра. Амплитудная характеристика $A_u(\hat{\omega})$ идеального фильтра принимает значение 1 в ПП фильтра и значение 0 в ПЗ для класса частотно-избирательных фильтров (ФНЧ, ФВЧ, ПФ, РФ). Весовую функцию $W(\hat{\omega})$ необходимо выбирать так, чтобы существовала возможность независимого контроля ошибки аппроксимации в различных частотных полосах фильтра. Например, для ФНЧ весовая функция $W(\hat{\omega})$ может быть выбрана в следующем виде:

$$W(\hat{\omega}) = \begin{cases} \frac{\delta_z}{\delta_n}, & \text{для ПП,} \\ 1, & \text{для ПЗ,} \end{cases} \quad (9.1)$$

где δ_n – отклонение АЧХ от 1 в ПП, δ_z – отклонение АЧХ от 0 в ПЗ.

Определив амплитудные характеристики реального и идеального фильтров, а также весовую функцию, введем в рассмотрение взвешенную ошибку аппроксимации:

$$\begin{aligned} E(\hat{\omega}) &= W(\hat{\omega})[A_u(\hat{\omega}) - A(\hat{\omega})] = \\ &= W(\hat{\omega})[A_u(\hat{\omega}) - Q(\hat{\omega})P(\hat{\omega})] = \\ &= W(\hat{\omega})Q(\hat{\omega}) \left[\frac{A_u(\hat{\omega})}{Q(\hat{\omega})} - P(\hat{\omega}) \right]. \end{aligned}$$

Для удобства введем модифицированную весовую функцию $\hat{W}(\hat{\omega})$ и модифицированную амплитудную характеристику идеального фильтра $\hat{A}_u(\hat{\omega})$ в виде

$$\begin{aligned} \hat{W}(\hat{\omega}) &= W(\hat{\omega})Q(\hat{\omega}), \\ \hat{A}_u(\hat{\omega}) &= \frac{A_u(\hat{\omega})}{Q(\hat{\omega})}. \end{aligned}$$

Тогда взвешенная ошибка аппроксимации для четырех типов КИХ-фильтров с линейной ФЧХ может быть представлена в виде

$$E(\hat{\omega}) = \hat{W}(\hat{\omega})[\hat{A}_u(\hat{\omega}) - P(\hat{\omega})].$$

В результате процедура расчета ЦФ заключается в нахождении коэффициентов фильтра $\alpha(k)$, которые минимизируют максимум абсолютной ошибки аппроксимации $E(\hat{\omega})$:

$$\min_{\alpha(k)} \left[\max_{\hat{\omega} \in S} |E(\hat{\omega})| \right] = \min_{\alpha(k)} \left[\max_{\hat{\omega} \in S} \left| \hat{W}(\hat{\omega})[\hat{A}_u(\hat{\omega}) - \sum_{k=0}^L \alpha(k) \cos(k\hat{\omega})] \right| \right], \quad (9.2)$$

где S – совокупность интервалов аппроксимации. По существу S содержит множество полос пропускания и полос задерживания фильтра.

Применительно к рассмотренной выше проблеме синтеза цифровых фильтров Паркс и Макклеллан переформулировали теорему теории аппроксимации Чебышёва, получив в результате следующую формулировку.

Теорема чередования. Пусть S – любое замкнутое подмножество закрытого интервала $0 \leq \hat{\omega} \leq \pi$. Для того чтобы функция $P(\hat{\omega})$, представленная в виде

$$P(\hat{\omega}) = \sum_{k=0}^L \alpha(k) \cos(k\hat{\omega}), \quad (9.2)$$

была единственной и наилучшей аппроксимацией на подмножестве S для функции $\hat{A}_u(\hat{\omega})$, необходимо и достаточно, чтобы функция ошибки $E(\hat{\omega})$ имела на S , по крайней мере, $L+2$ чередований. Т.е. существует, по крайней мере, $L+2$ экстремальных частот $\hat{\omega}_i$ на подмножестве S , таких что $\hat{\omega}_1 < \hat{\omega}_2 < \dots < \hat{\omega}_{L+2}$, $E(\hat{\omega}_i) = -E(\hat{\omega}_{i+1})$ и $|E(\hat{\omega}_i)| = \max_{\hat{\omega} \in S} |E(\hat{\omega})|$,

$i = 1, 2, \dots, L+2$.

Для конкретизации теоремы чередования рассмотрим синтез ФНЧ с ПП $0 \leq \hat{\omega} \leq \hat{\omega}_n$ и ПЗ $\hat{\omega}_3 \leq \hat{\omega} \leq \pi$. Так как амплитудная характеристика идеального фильтра $A_u(\hat{\omega})$ и весовая функция $W(\hat{\omega})$ являются кусочно-постоянными, можно записать

$$\begin{aligned} \frac{dE(\hat{\omega})}{d\hat{\omega}} &= \frac{d}{d\hat{\omega}} \{W(\hat{\omega})[A_u(\hat{\omega}) - A(\hat{\omega})]\} = \\ &= -W(\hat{\omega}) \frac{dA(\hat{\omega})}{d\hat{\omega}} = 0. \end{aligned}$$

Поэтому частоты $\hat{\omega}_i$, соответствующие пикам $E(\hat{\omega})$, также соответствуют пикам, в которых $A(\hat{\omega})$ удовлетворяет допустимой ошибке. Так как $A(\hat{\omega})$, например для КИХ-фильтров типа 1 с линейной ФЧХ, представляет собой тригонометрический полином степени L

$$\begin{aligned} A(\omega) &= \sum_{k=0}^L \alpha(k) \cos(k\hat{\omega}) = \\ &= \sum_{k=0}^L \alpha(k) \left[\sum_{n=0}^k \beta_{nk} (\cos(\hat{\omega}))^n \right] = \\ &= \sum_{k=0}^L \tilde{\alpha}(k) (\cos(\hat{\omega}))^k, \end{aligned} \quad (9.3)$$

то она имеет не больше чем $L-1$ локальных максимумов и минимумов на открытом интервале $0 < \hat{\omega} < \pi$. Дополнительно можно отметить, что частоты $\hat{\omega} = 0$ и $\hat{\omega} = \pi$ обычно являются экстремумами амплитудной характеристики $A(\hat{\omega})$, а также взвешенной ошибки аппроксимации $E(\hat{\omega})$. Следовательно, функция $A(\hat{\omega})$ имеет не больше чем $L+1$ экстремальных частот. Кроме того, граничные частоты в ПП и ПЗ, $\hat{\omega}_n$ и $\hat{\omega}_z$ также являются экстремумами функции $E(\hat{\omega})$, так как $|E(\hat{\omega})|$ является максимальным на этих частотах. Следовательно, должно существовать $L+3$ экстремальных частот функции $E(\hat{\omega})$ для единственной и наилучшей аппроксимации идеального ФНЧ.

С другой стороны, теорема чередования устанавливает существование, по крайней мере, $L+2$ экстремальных частот $E(\hat{\omega})$. Таким образом, единственная и наилучшая аппроксимация для идеальной амплитудной характеристики фильтра $A_u(\hat{\omega})$ имеет либо $L+2$, либо $L+3$ чередований взвешенной ошибки

аппроксимации $E(\hat{\omega})$. Синтезированные фильтры, содержащие $L+3$ чередований, были названы Парксом и Макклелланом фильтрами с дополнительной пульсацией. Последнее связано с тем фактом, что данного типа фильтры обладали большим, чем минимальное число $L+2$ чередований функции ошибки, требуемым для оптимальности.

Теорема чередования гарантирует единственность решения для достижения задачи оптимизации, представленной выражением (9.2). Из теоремы чередования следует, что для заданного множества экстремальных частот $\hat{\omega}_i$ можно записать

$$\hat{W}(\hat{\omega}_i)[\hat{A}_u(\hat{\omega}_i) - P(\hat{\omega}_i)] = (-1)^i \delta, \quad i = 0, 1, \dots, L+1, \quad (9.4)$$

где δ – максимальная величина взвешенной ошибки аппроксимации $E(\hat{\omega})$. Фактически, если весовая функция $W(\hat{\omega})$ выбрана так, как представлено в выражении (9.1), то $\delta = \delta_3$.

Множество линейных уравнений, представленных выражением (9.4), можно записать в виде

$$P(\hat{\omega}_i) + \frac{(-1)^i \delta}{\hat{W}(\hat{\omega}_i)} = \hat{A}_u(\hat{\omega}_i), \quad i = 0, 1, \dots, L+1,$$

или

$$\sum_{k=0}^L \alpha(k) \cos(k\hat{\omega}_i) + \frac{(-1)^i \delta}{\hat{W}(\hat{\omega}_i)} = \hat{A}_u(\hat{\omega}_i), \quad i = 0, 1, \dots, L+1. \quad (9.5)$$

Выражение (9.5) может быть записано в матричной форме относительно неизвестных коэффициентов фильтра $\alpha(k)$ и максимальной величины δ взвешенной ошибки аппроксимации $E(\hat{\omega})$:

$$\begin{bmatrix} 1 & \cos(\hat{\omega}_0) & \cdots & \cos(L\hat{\omega}_0) & \frac{1}{\hat{W}(\hat{\omega}_0)} \\ 1 & \cos(\hat{\omega}_1) & \cdots & \cos(L\hat{\omega}_1) & \frac{1}{\hat{W}(\hat{\omega}_1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cos(\hat{\omega}_{L+1}) & \cdots & \cos(L\hat{\omega}_{L+1}) & \frac{1}{\hat{W}(\hat{\omega}_{L+1})} \end{bmatrix} \begin{bmatrix} \alpha(0) \\ \alpha(1) \\ \vdots \\ \alpha(L) \\ \delta \end{bmatrix} = \begin{bmatrix} \hat{A}_u(\hat{\omega}_0) \\ \hat{A}_u(\hat{\omega}_1) \\ \vdots \\ \hat{A}_u(\hat{\omega}_{L+1}) \end{bmatrix}. \quad (9.6)$$

При заданных экстремальных частотах выражение (9.6) можно решить относительно неизвестных коэффициентов фильтра $\alpha(k)$ и максимальной величины δ взвешенной ошибки аппроксимации $E(\hat{\omega})$.

9.4. Алгоритм замены Ремеза

Рассматривая теорему чередования, необходимо сказать, что она гарантирует для минимаксной аппроксимации, представленной выражением (9.2), существование и единственность решения, однако она не дает представления об алгоритме, с использованием которого оптимальное решение может быть получено. Паркс и Макклеллан предложили итерационную процедуру, основанную на алгоритме замены Ремеза и позволяющую найти оптимальное решение задачи минимаксной аппроксимации амплитудной характеристики фильтра. Данная процедура предполагает, что порядок фильтра N , граничные частоты $\hat{\omega}_n$, $\hat{\omega}_z$ и отношение δ_z/δ_n являются известными. Однако во время процедуры синтеза фильтров более естественно точно задавать граничные частоты $\hat{\omega}_n$, $\hat{\omega}_z$ и пульсации δ_n , δ_z , зная которые далее определять порядок фильтра N , удовлетворяющий заданным спецификациям. Сравнительно простая формула, применяемая для оценки порядка N ФНЧ с равномерными пульсациями по заданным спецификациям $\hat{\omega}_n$, $\hat{\omega}_z$, δ_n и δ_z , была предложена Кайзером и имеет следующий вид:

$$\hat{N} = \frac{-10 \log_{10}(\delta_n \delta_3) - 13}{14.6 \Delta f}, \quad (9.7)$$

где Δf – ширина переходной полосы, определяемая как $\Delta f = \frac{\hat{\omega}_3 - \hat{\omega}_n}{2\pi}$. Более точная формула была предложена Херрманном:

$$\hat{N} = \frac{D_\infty(\delta_n, \delta_3) - f(\delta_n, \delta_3)(\Delta f)^2}{\Delta f}, \quad (9.8)$$

где $D_\infty(\delta_n, \delta_3)$ и $f(\delta_n, \delta_3)$ определяются следующим образом:

$$\begin{aligned} D_\infty(\delta_n, \delta_3) = & [0.005309(\log_{10}(\delta_n))^2 + 0.07114(\log_{10}(\delta_n)) - \\ & - 0.4761](\log_{10}(\delta_3)) - \\ & - [0.00266(\log_{10}(\delta_n))^2 + 0.5941(\log_{10}(\delta_n)) + 0.4278], \end{aligned}$$

$$f(\delta_n, \delta_3) = 11.022 + 0.51244(\log_{10}(\delta_n) - \log_{10}(\delta_3)).$$

Ниже приведены основные этапы синтеза цифрового фильтра с использованием алгоритма замены Ремеза.

1. Приблизительная оценка $L+2$ экстремальных частот $\hat{\omega}_j$.

2. Вычисление максимальной величины δ взвешенной ошибки аппроксимации $E(\hat{\omega})$ с использованием следующего выражения:

$$\delta = \frac{\sum_{k=0}^{L+1} \gamma(k) \hat{A}_u(\hat{\omega}_k)}{\sum_{k=0}^{L+1} \frac{(-1)^k \gamma(k)}{\hat{W}(\hat{\omega}_k)}}, \quad (9.9)$$

где $\gamma(k) = \prod_{n=0, n \neq k}^{L+1} \frac{1}{\cos(\hat{\omega}_k) - \cos(\hat{\omega}_n)}$. Формула (9.9) следует из выражения (9.6). Таким образом, задавая начальные значения экстремальных частот на шаге 1, вычисляем δ с использованием выражения (9.9).

3. Вычисление взвешенной ошибки аппроксимации $E(\hat{\omega})$ на густой сетке частот.

Поскольку функция $P(\hat{\omega})$ представляет собой тригонометрический полином вида

$$P(\hat{\omega}) = \sum_{k=0}^L \tilde{\alpha}(k) x^k, \quad x = \cos(\hat{\omega}),$$

и поскольку в точках $x_i = \cos(\hat{\omega}_i)$, $i = 0, 1, \dots, L+1$ известны значения функции $P(\hat{\omega})$, принимающие вид

$$P(\hat{\omega}_i) = \hat{A}_u(\hat{\omega}_i) - \frac{(-1)^i \delta}{\hat{W}(\hat{\omega}_i)}, \quad i = 0, 1, \dots, L+1, \quad (9.10)$$

можно осуществить интерполяцию Лагранжа функции $P(\hat{\omega})$ по значениям $P(\hat{\omega}_i)$ с использованием следующего математического выражения:

$$P(\hat{\omega}) = \frac{\sum_{i=0}^L P(\hat{\omega}_i) [\beta(i)/(x - x_i)]}{\sum_{i=0}^L [\beta(i)/(x - x_i)]}, \quad (9.11)$$

где $P(\hat{\omega}_i)$ определяются выражением (9.10), $x = \cos(\hat{\omega})$, $x_i = \cos(\hat{\omega}_i)$ и $\beta(i) = \prod_{i=0, i \neq k}^L \frac{1}{x_k - x_i}$.

Зная функцию $P(\hat{\omega})$, определяемую выражением (9.11), можно вычислить взвешенную ошибку аппроксимации $E(\hat{\omega})$:

$$E(\hat{\omega}) = \hat{W}(\hat{\omega}) [\hat{A}_u(\hat{\omega}) - P(\hat{\omega})]$$

на густой сетке частот. Обычно число точек для вычисления ошибки аппроксимации выбирается равным значению $16(N + 1)$, где N – порядок фильтра.

4. Если среди значений взвешенной ошибки аппроксимации, вычисленных на шаге 3, найдутся такие, что $|E(\hat{\omega})| \geq \delta$, то необходимо выбрать новое множество $L + 2$ экстремальных частот $\hat{\omega}_i$, которые соответствуют $L + 2$ наибольшим значениям $|E(\hat{\omega})|$, и затем повторить процедуру нахождения единственного и наилучшего решения, начиная с шага 2. Если $|E(\hat{\omega})| \leq \delta$ для всех точек на густой сетке частот, то процедура нахождения единственного и наилучшего решения задачи аппроксимации заканчивается. Тем самым коэффициенты функции $P(\hat{\omega})$ являются найденными и, как следствие, является полученной частотная характеристика реального фильтра $H(e^{j\hat{\omega}})$.

Блок-схема описанного выше алгоритма, приведенная на рис. 9.1, впервые была представлена Ремезом.

В заключение необходимо отметить, что приведенный выше оптимальный метод синтеза – это простой и эффективный путь вычисления коэффициентов КИХ-фильтра. Для большинства приложений оптимальный метод синтеза дает хорошие АЧХ фильтра при его разумном порядке N . Однако, хотя метод и позволяет полностью контролировать спецификации, задаваемые на фильтр, необходимым требованием для его выполнения является наличие программной реализации разрабатываемого фильтра.

9.5. Контрольные вопросы

1. Какие существуют методы синтеза КИХ-фильтров?
2. Что такое оптимальные методы синтеза цифровых фильтров?
3. Сформулируйте и докажите условия линейности ФЧХ КИХ-фильтра.
4. Назовите причины синтеза КИХ-фильтров с линейной ФЧХ.
5. Запишите общее математическое выражение, с использованием которого можно представить все типы фильтров с линейной ФЧХ.

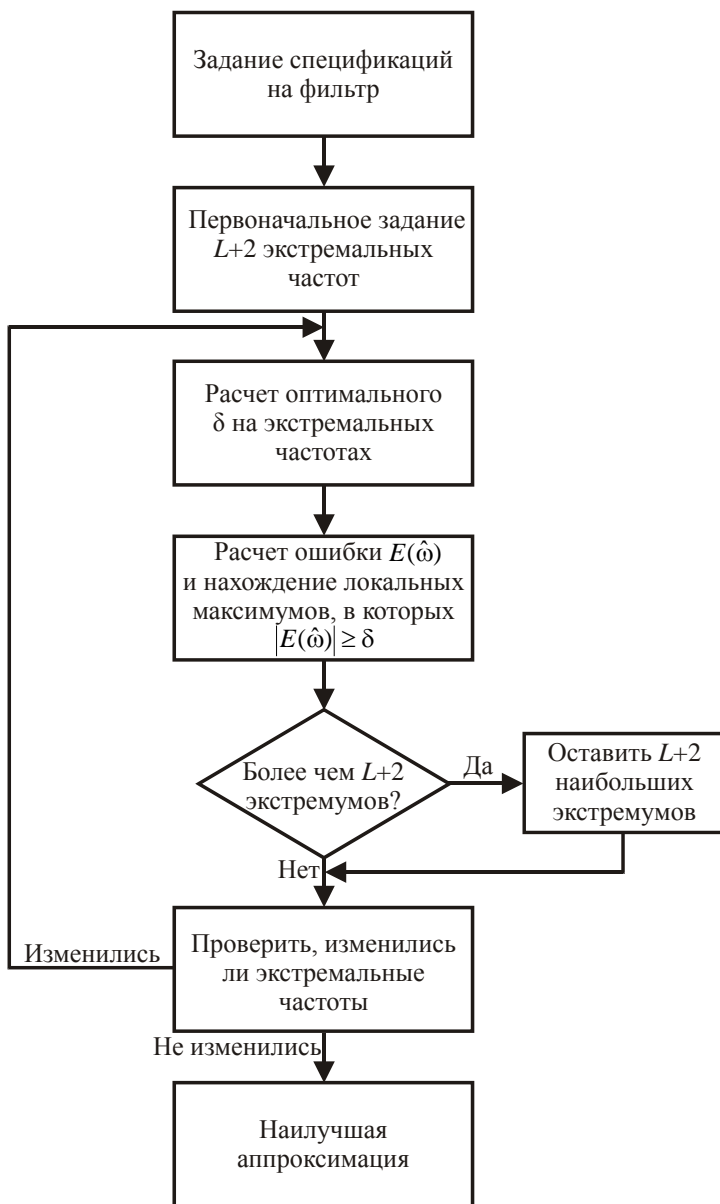


Рис. 9.1. Блок-схема алгоритма замены Ремеза

6. Что такое взвешенная ошибка аппроксимации?
7. Что означает оптимальность КИХ-фильтров по минимаксному критерию?
8. Сформулируйте теорему чередования.
9. Поясните основные этапы алгоритма Паркса-Макклеллана.
10. Поясните основные достоинства и недостатки алгоритма Паркса-Макклеллана.
11. Произведите сравнение алгоритмов синтеза КИХ-фильтров на основе метода окон и метода Паркса-Макклеллана.

9.6. Литература

1. Брюханов, Ю. А. Цифровые фильтры: учеб. пособие / Ю. А. Брюханов, А. Л. Приоров. – Ярославль: ЯрГУ, 2002.
2. Приоров, А. Л. Основы цифровой обработки сигналов: лабораторный практикум / А. Л. Приоров, В. А. Волохов, А. М. Лаврентьев. – ВУНЦ ВВС «ВВА им. профессора Н. Е. Жуковского и Ю. А. Гагарина» (филиал, г. Ярославль), 2012.
3. Айфичер, Э. Цифровая обработка сигналов: практический подход. – 2-е изд. / Э. Айфичер, Б. Джервис. – Вильямс, 2004.
4. Сергиенко, А. Б. Цифровая обработка сигналов. – 3-е изд. / А. Б. Сергиенко. – СПб.: БХВ-Петербург, 2011.
5. Дьяконов, В. П. MATLAB 6.0/6.1/6.5/6.5+SP1 + Simulink 4/5. Обработка сигналов и изображений / В. П. Дьяконов. – М.: СОЛОН-Пресс, 2005.
6. Хрящев, В. В. Основы теории цепей: сборник задач / В. В. Хрящев, А. Л. Приоров, В. А. Волохов. – Ярославль: ЯрГУ, 2008.
7. Солонина, А. И. Цифровая обработка сигналов. Моделирование в MATLAB / А. И. Солонина, С. М. Арбузов. – СПб.: БХВ-Петербург, 2008.
8. Proakis, J. G. Digital Signal Processing / J. G. Proakis, D. G. Manolakis, Prentice-Hall International, Inc., 1996.

10. МНОГОСКОРОСТНАЯ ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

10.1. Основные операторы изменения частоты дискретизации

Двумя основными видами операторов изменения частоты дискретизации являются дециматор и интерpolator, соответственно понижающий и повышающий частоту дискретизации. Анализ их ЧХ покажет, почему данные устройства обычно используются совместно с дополнительными фильтрами. Кроме того, частотный анализ поможет понять работу других многоуровневых систем, рассматриваемых ниже.

10.1.1. Временные соотношения

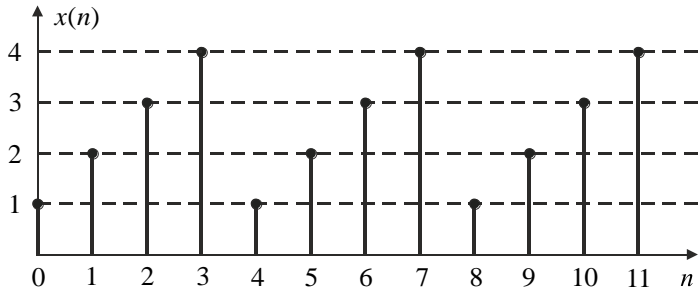
Увеличение частоты дискретизации

Рассмотрим временные характеристики двух основных устройств преобразования цифрового сигнала. Устройство увеличения частоты дискретизации в L раз, где L – положительное целое, имеет на выходе последовательность $x_u(n)$ с шагом дискретизации в L раз меньше, чем входной сигнал $x(n)$. Операция увеличения частоты дискретизации задается введением $L - 1$ нулевых отсчетов во входную последовательность и выражается формулой

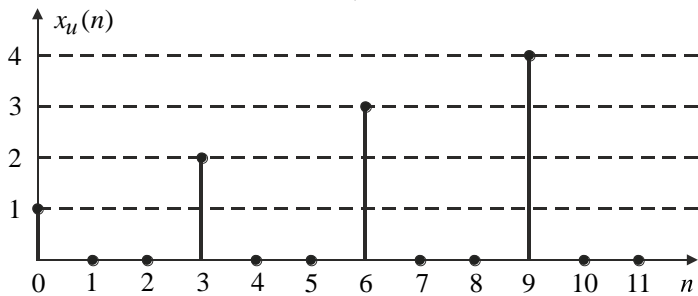
$$x_u(n) = \begin{cases} x(n/L), & n = 0, \pm L, \pm 2L, \dots, \\ 0, & \text{иначе.} \end{cases} \quad (10.1)$$

Пример выполнения операции увеличения частоты дискретизации в 3 раза ($L = 3$) для произвольного периодического сигнала (рис. 10.1а), представлен на рис. 10.1б. На рис. 10.1в представлена блок-схема операции увеличения частоты дискретизации в L раз.

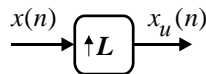
На практике обычно с помощью определенного процесса фильтрации вместо нулей, не несущих информационной нагрузки, вставляются ненулевые отсчеты. Этот процесс называется интерполяцией и будет рассмотрен далее.



a)



б)



в)

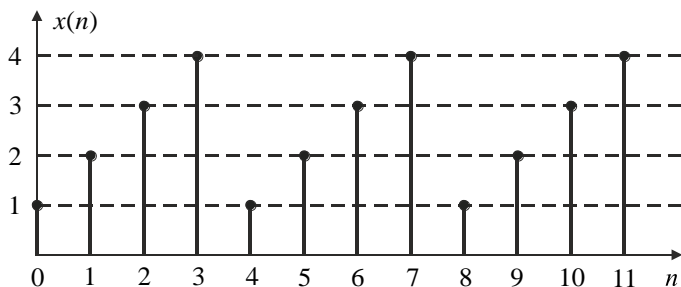
Рис. 10.1. Увеличение частоты дискретизации в 3 раза:
 а) входной сигнал; б) выходной сигнал при увеличении частоты дискретизации в 3 раза; в) блок-схема увеличения частоты дискретизации в L раз

Уменьшение частоты дискретизации

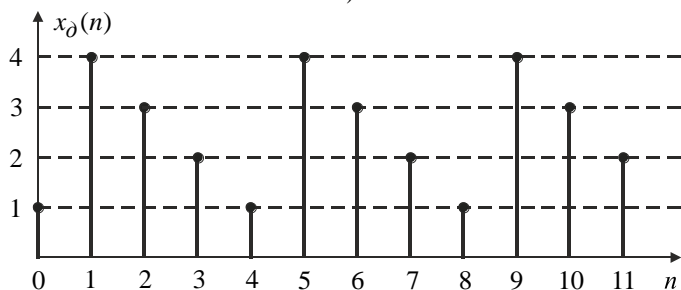
При уменьшении частоты дискретизации в M раз, где M – положительное целое, получаем выходную последовательность $x_{\partial}(n)$ с шагом дискретизации в M раз больше, чем у входного сигнала $x(n)$. Понижение частоты дискретизации получается

путем сохранения одного отсчета входной последовательности и удалением $M - 1$ последующих отсчетов сигнала, что выражается следующей формулой:

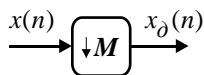
$$x_{\partial}(n) = x(nM). \quad (10.2)$$



a)



б)



в)

Рис. 10.2. Уменьшение частоты дискретизации в 3 раза:
 а) входной сигнал; б) выходной сигнал при уменьшении частоты дискретизации в 3 раза; в) блок-схема уменьшения частоты дискретизации в M раз

В результате шаг дискретизации выходной последовательности возрастает в M раз. Пример выполнения операции уменьшения частоты дискретизации в 3 раза ($M = 3$) для произвольного периодического сигнала (рис. 10.2а) представлен на рис. 10.2б. На рис. 10.2в представлена блок-схема операции уменьшения частоты дискретизации в M раз.

10.1.2. Соотношения в частотной области

Увеличение частоты дискретизации

Определим соотношения спектров входного и выходного сигналов при увеличении частоты дискретизации в L раз. Для этого вычислим z -преобразование последовательности $x_u(n)$, описываемой формулой (1),

$$X_u(z) = \sum_{n=-\infty}^{\infty} x_u(n)z^{-n} = \sum_{\substack{n=-\infty \\ n - \text{кратно } L}}^{\infty} x(n/L)z^{-n} = \sum_{m=-\infty}^{\infty} x(m)z^{-mL} = X(z^L), \quad (10.3)$$

где $X(z)$ – z -преобразование входной последовательности $x(n)$. Выполняя в выражении (10.3) замену $z = e^{j\hat{\omega}}$, получим соотношение для связи преобразований Фурье (спектров) сигналов на входе и выходе блока увеличения частоты дискретизации (рис. 10.1в):

$$X_u(e^{j\hat{\omega}}) = X(e^{j\hat{\omega}L}). \quad (10.4)$$

Здесь частота $\hat{\omega}$ является нормированной на частоту дискретизации, измеряемую в Γ_{ω} , и измеряется в радианах. Выражение (10.4) фактически показывает, что для получения преобразования Фурье сигнала на выходе блока повышения частоты дискретизации необходимо выполнить операцию масштабирования по частоте преобразования Фурье сигнала на входе данного блока.

На рис. 10.3 показан пример преобразования функции $|X(e^{j\hat{\omega}})|$ в функцию $|X_u(e^{j\hat{\omega}})|$ для случая увеличения частоты

дискретизации в 2 раза ($L = 2$). Функции $|X(e^{j\hat{\omega}})|$ и $|X_u(e^{j\hat{\omega}})|$ были получены для последовательностей $x(n)$ и $x_u(n)$ соответственно, являющихся комплексными. Асимметричное (нестандартное) решение было выбрано для того, чтобы более детально показать результат уменьшения шага дискретизации.

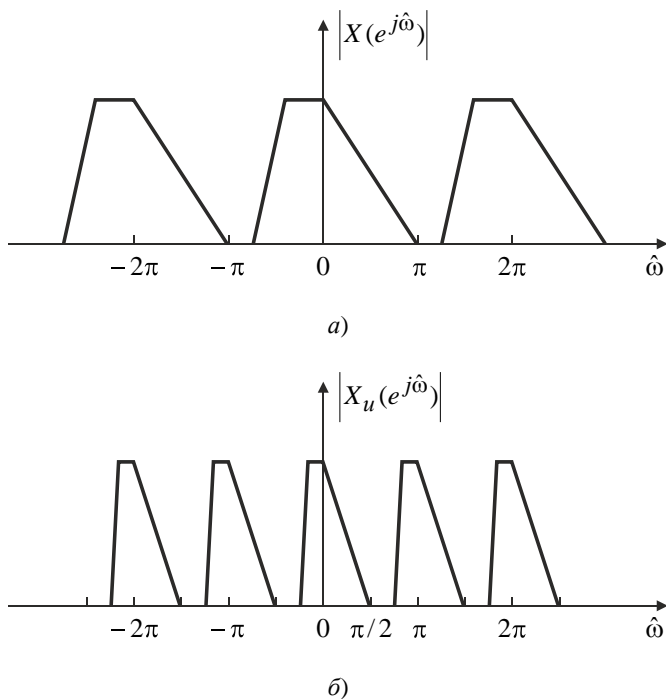


Рис. 10.3. Результат увеличения частоты дискретизации в спектральной области в 2 раза:
а) для входного сигнала; б) для выходного сигнала

Как показано на рис. 10.3, при $L = 2$ модуль преобразования Фурье сигнала $x_u(n)$ сжимается в 2 раза. Этот процесс называется «формированием изображения», так как в результате получается дополнительное изображение входного спектра. В случае с коэффициентом интерполяции, равным L , получим

дополнительно $L-1$ «изображений» входного спектра сигнала. Данные «изображения», не несущие дополнительной информации, можно устранить с помощью низкочастотной цифровой фильтрации.

Уменьшение частоты дискретизации

Рассчитаем соотношение спектров входного и выходного сигналов при уменьшении частоты дискретизации. Для этого вычислим z -преобразование последовательности $x_\partial(n)$, представленной в (10.2):

$$X_\partial(z) = \sum_{n=-\infty}^{\infty} x_\partial(n)z^{-n} = \sum_{n=-\infty}^{\infty} x(nM)z^{-n}. \quad (10.5)$$

Введем промежуточную последовательность $x_n(n)$, которую можно представить через отсчеты сигнала $x(n)$ на входе блока уменьшения частоты дискретизации в виде

$$x_n(n) = \begin{cases} x(n), & n = 0, \pm M, \pm 2M, \dots, \\ 0, & \text{иначе.} \end{cases} \quad (10.6)$$

Далее получим

$$\begin{aligned} X_\partial(z) &= \sum_{n=-\infty}^{\infty} x(nM)z^{-n} = \sum_{n=-\infty}^{\infty} x_n(nM)z^{-n} = \\ &= \sum_{k=-\infty}^{\infty} x_n(k)z^{-k/M} = X_n(z^{1/M}). \end{aligned} \quad (10.7)$$

Здесь $X_n(z)$ – z -преобразование последовательности $x_n(n)$.

Последовательность $x(n)$ может быть приведена к последовательности $x_n(n)$ с использованием выражения $x_n(n) = c(n)x(n)$, где $c(n)$ определяется как

$$c(n) = \begin{cases} 1, & n = 0, \pm M, \pm 2M, \dots, \\ 0, & \text{иначе.} \end{cases} \quad (10.8)$$

Более удобное представление последовательности $c(n)$ имеет вид

$$c(n) = \frac{1}{M} \sum_{k=0}^{M-1} e^{j \frac{2\pi}{M} kn}. \quad (10.9)$$

Используя замену $x_n(n) = c(n)x(n)$ и воспользовавшись уравнением (10.9) при z -преобразовании $x_n(n)$, получим

$$\begin{aligned} X_n(z) &= \sum_{n=-\infty}^{\infty} x_n(n)z^{-n} = \sum_{n=-\infty}^{\infty} c(n)x(n)z^{-n} = \\ &= \frac{1}{M} \sum_{n=-\infty}^{\infty} \left(\sum_{k=0}^{M-1} e^{j \frac{2\pi}{M} kn} \right) x(n)z^{-n} = \\ &= \frac{1}{M} \sum_{k=0}^{M-1} \left(\sum_{n=-\infty}^{\infty} x(n) e^{j \frac{2\pi}{M} kn} z^{-n} \right) = \frac{1}{M} \sum_{k=0}^{M-1} X \left(e^{-j \frac{2\pi}{M} k} z \right). \end{aligned} \quad (10.10)$$

Взаимосвязь между z -преобразованиями сигналов на входе и выходе блока уменьшения частоты дискретизации (рис. 10.2б), с учетом подстановки выражения (10.10) в выражение (10.7), имеет вид:

$$X_o(z) = \frac{1}{M} \sum_{k=0}^{M-1} X \left(e^{-j \frac{2\pi}{M} k} z^{1/M} \right). \quad (10.11)$$

Выполняя в выражении (10.11) замену $z = e^{j\hat{\omega}}$, получим соотношение для связи преобразований Фурье сигналов на входе и выходе блока уменьшения частоты дискретизации (рис. 10.2б):

$$X_o(e^{j\hat{\omega}}) = \frac{1}{M} \sum_{k=0}^{M-1} X(e^{j(\hat{\omega} - 2\pi k)/M}). \quad (10.12)$$

Выражение (10.12) подразумевает, что $X_o(e^{j\hat{\omega}})$ является суммой M слагаемых смещенных функций $X(e^{j\hat{\omega}})$ с увеличенным периодом, которая дополнительно умножается на

$1/M$. В силу уменьшения частоты дискретизации возможно возникновение эффекта наложения спектров. Этого можно избежать, если при уменьшении частоты дискретизации в M раз сигнал $x(n)$ ограничен полосой частот до π/M .

Возможный вид функции $|X_\partial(e^{j\hat{\omega}})|$ при отсутствии и наличии эффекта наложения спектров как следствия выполнения операции понижения частоты дискретизации представлен на рис. 10.4 и рис. 10.5 соответственно.

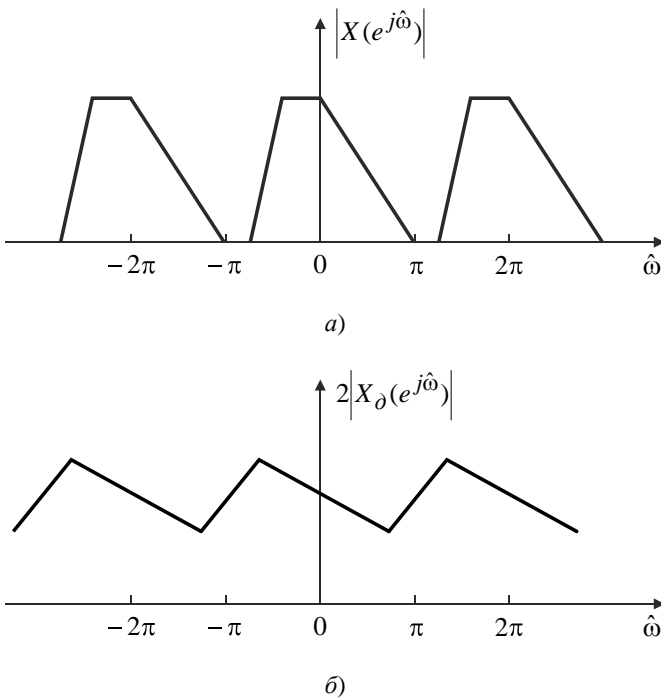


Рис. 10.4. Эффект наложения в частотной области, вызванный понижением частоты дискретизации в 2 раза:
 а) спектр исходного сигнала; б) суммарный спектр выходного сигнала после понижения частоты дискретизации в 2 раза

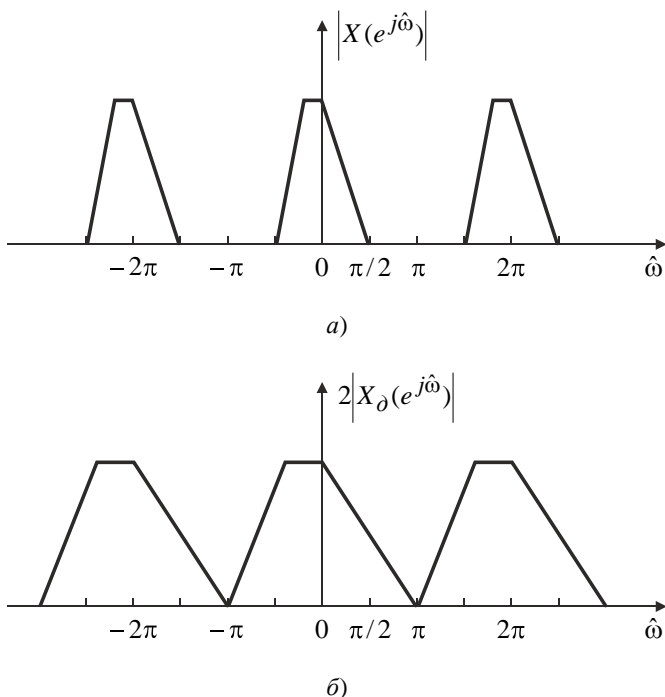


Рис. 10.5. Отсутствие наложения при понижении частоты дискретизации в 2 раза: а) спектр исходного сигнала; б) суммарный спектр выходного сигнала после понижения частоты дискретизации в 2 раза

10.1.3. Многоскоростные тождества

1. *Коммутативность изменения частоты дискретизации.* Схемы, представленные на рис. 10.6а, эквивалентны, если коэффициенты повышения L и понижения M частоты дискретизации являются взаимно простыми числами, т. е. с учетом того, что L и M являются положительными целыми числами, они не имеют общих делителей, больших 1.

2. *Чередование операции фильтрации и повышения частоты дискретизации.* Каскадное соединение фильтра с передаточной

функцией $H(z)$ и блока повышения частоты дискретизации в L раз является эквивалентным каскадному соединению блока повышения частоты дискретизации в L раз и фильтра с передаточной функцией $H(z^L)$ (рис. 10.6б).

3. *Чередование операции фильтрации и понижения частоты дискретизации.* Каскадное соединение блока понижения частоты дискретизации в M раз и фильтра с передаточной функцией $H(z)$ является эквивалентным каскадному соединению фильтра с передаточной функцией $H(z^M)$ и блока понижения частоты дискретизации в M раз (рис. 10.6в).

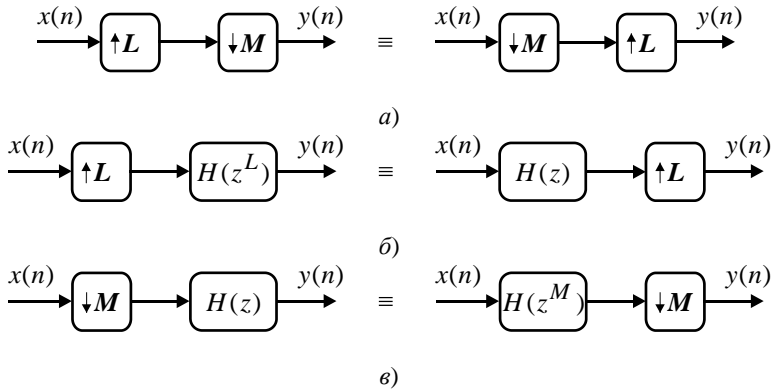


Рис. 10.6. Многоскоростные тождества: а) коммутативность изменения частоты дискретизации выполняется, если целые числа L и M являются взаимно простыми; б) чередование операции фильтрации и повышения частоты дискретизации; в) чередование операции фильтрации и понижения частоты дискретизации

10.1.4. Многоскоростные системы для преобразования частоты дискретизации

Процесс уменьшения частоты дискретизации сигналов более известен как «децимация», а многоскоростные устройства, используемые для децимации, называются дециматорами. Обратный процесс повышения частоты дискретизации сигналов

называется «интерполяцией», а устройства, используемые для интерполяции, называются интерполяторами. Выше было показано, что нельзя произвольно снижать частоту дискретизации сигнала так, чтобы при этом отсутствовал эффект наложения. Поэтому для предельно дискретизированного сигнала в первую очередь необходимо преобразовать его с помощью ФНЧ, а только затем уменьшать частоту дискретизации. Подобным образом после вставки нулевых отсчетов в последовательность при увеличении частоты дискретизации сигнал должен быть пропущен через ФНЧ.

В связи с тем, что преобразователи частоты дискретизации с коэффициентами дискретизации, являющимися дробными числами, могут быть реализованы с помощью каскадов интерполяции и децимации, в многоскоростных системах необходимо использовать цифровые фильтры. Рассмотрим порядок использования таких фильтров.

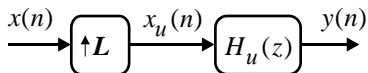
В первую очередь определим соотношения между входной и выходной последовательностями многоскоростной системы, в которой используются преобразователи частоты. Затем определим частотные характеристики понижающих фильтров, используемых в многоскоростных структурах. Далее рассмотрим децимацию и интерполяцию последовательностей и определим выходные характеристики понижающих цифровых фильтров.

10.1.5. Основные структуры

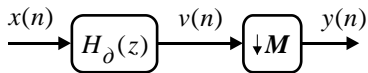
При повышении частоты дискретизации в L раз происходит повторение сигналов основного спектра (рис. 10.3), поэтому интерполятор для целых значений изменения частоты дискретизации содержит ФНЧ с передаточной функцией $H_u(z)$, имеющего частоту среза π/L , как показано на рис. 10.7а. Данный ФНЧ, называемый фильтром интерполяции, перемещает $L-1$ не несущих информационной нагрузки нулевых отсчетов $x_u(n)$ в спектр сигнала с повышенной частотой.

С другой стороны, устройство понижения частоты с индексом M , равным целому числу, может привести к эффекту наложения. Поэтому в основной структуре дециматора

применяется ФНЧ с передаточной функцией $H_\delta(z)$, имеющий частоту среза π/M , который расположен перед блоком понижения частоты дискретизации и формирует некоторый сигнал $v(n)$, как показано на рис. 10.7б. Данный ФНЧ называется фильтром децимации. Он ограничивает полосу частот входного сигнала $x(n)$ до $|\hat{\omega}| \leq \pi/M$ перед началом процесса уменьшения частоты дискретизации, чтобы не допустить эффекта наложения спектров.

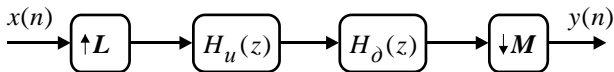


а)

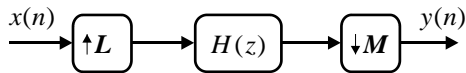


б)

Рис. 10.7. Применение ФНЧ: а) для устранения $L-1$ копий сжатого спектра сигнала; б) для устранения наложения



а)



б)

Рис. 10.8. Эквивалентное преобразование при изменении частоты дискретизации в дробное число раз:
а) исходная структура; б) структура после преобразования

Изменение частоты дискретизации в дробное число раз может быть получено совместным использованием интерполятора с коэффициентов L и дециматора с коэффициентом M , где L и

M – положительные целые числа. Интерполяция должна предшествовать децимации, как показано на рис. 10.8а. Оба фильтра – интерполяции $H_u(z)$ и децимации $H_\partial(z)$ – могут быть объединены в один фильтр $H(z)$. Тем самым можно избежать эффекта наложения, который может возникнуть при понижении частоты дискретизации и повлиять на исчезновение «лишних» составляющих сигнала при увеличении частоты дискретизации. Результат такого объединения фильтров показан на рис. 10.8б.

10.1.6. Связь входного и выходного сигналов

Опуская математические выкладки, определим соотношения вход-выход для систем, представленных на рис. 10.7 и рис. 10.8.

Рассмотрим соотношение вход-выход для интерполятора, изображенного на рис. 10.7а. Пусть $h_u(n)$ обозначает ИХ фильтра интерполяции с передаточной функцией $H_u(z)$. Во временной области желаемая связь вход-выход может быть описана следующим образом:

$$y(n) = \sum_{m=-\infty}^{\infty} h_u(n - Lm)x(m), \quad (10.13)$$

а в z -области

$$Y(z) = H(z)X(z^L). \quad (10.14)$$

Рассмотрим соотношение вход-выход для дециматора, изображенного на рис. 10.7б. Пусть $h_\partial(n)$ обозначает ИХ фильтра децимации с передаточной функцией $H_\partial(z)$. Во временной области желаемая связь вход-выход выглядит следующим образом:

$$y(n) = \sum_{m=-\infty}^{\infty} h_\partial(Mn - m)x(m), \quad (10.15)$$

а в z -области

$$Y(z) = \frac{1}{M} \sum_{k=0}^{M-1} H \left(e^{-j\frac{2\pi}{M}k} z^{1/M} \right) X \left(e^{-j\frac{2\pi}{M}k} z^{1/M} \right). \quad (10.16)$$

Рассмотрим соотношение вход-выход для системы изменения частоты дискретизации в дробное число раз, изображенной на рис. 10.8б. Пусть $h(n)$ обозначает ИХ фильтра с передаточной функцией $H(z)$. Во временной области желаемая связь вход-выход описывается следующим образом:

$$y(n) = \sum_{m=-\infty}^{\infty} h(Mn - mL)x(m), \quad (10.17)$$

а в z -области

$$Y(z) = \frac{1}{M} \sum_{k=0}^{M-1} H \left(e^{-j\frac{2\pi}{M}k} z^{1/M} \right) X \left(e^{-j\frac{2\pi}{M}kL} z^{L/M} \right). \quad (10.18)$$

10.1.7. Требования к цифровым фильтрам

Рассмотрим особенности ФНЧ, показанных на рис. 10.7 и рис. 10.8. В первую очередь рассмотрим интерполяционные фильтры. Пусть $x(n)$ – последовательность, полученная в результате дискретизации с частотой Найквиста ограниченного во времени сигнала $x_a(t)$. Если $X_a(j\Omega)$ и $X(e^{j\hat{\omega}})$ являются результатом преобразования Фурье $x_a(t)$ и $x(n)$ соответственно, то получим, что эти преобразованные функции связаны между собой следующим выражением:

$$X(e^{j\hat{\omega}}) = \frac{1}{T_\delta} \sum_{k=-\infty}^{\infty} X_a \left(\frac{j\hat{\omega} - j2\pi k}{T_0} \right), \quad (10.19)$$

где T_δ – период дискретизации. Если вместо дискретной функции $x(n)$ подставить функцию $y(n)$ с меньшим периодом T_δ/L , преобразование Фурье функции $Y(e^{j\hat{\omega}})$ может быть выражено в виде

$$Y(e^{j\hat{\omega}}) = \frac{L}{T_\delta} \sum_{k=-\infty}^{\infty} X_a \left(\frac{j\hat{\omega} - j2\pi k}{T_\delta / L} \right). \quad (10.20)$$

С другой стороны, если сигнал $x(n)$ будет пропущен через интерполятор с коэффициентом увеличения частоты дискретизации L и ФНЧ с частотой среза π/L , и усилением L , выходной сигнал интерполятора будет точно соответствовать сигналу $y(n)$, преобразование Фурье которого описывается выражением (10.20). В результате описание идеального интерполяционного ФНЧ будет следующим:

$$|H_u(e^{j\hat{\omega}})| = \begin{cases} L, & |\hat{\omega}| \leq \pi/L, \\ 0, & \text{иначе.} \end{cases} \quad (10.21)$$

Аналогично задаются спецификации для идеального децимационного ФНЧ, который будет задан выражением

$$|H_\delta(e^{j\hat{\omega}})| = \begin{cases} 1, & |\hat{\omega}| \leq \pi/M, \\ 0, & \text{иначе.} \end{cases} \quad (10.22)$$

В случае преобразователя частоты дискретизации с дробным множителем, как показано на рис. 10.8б, спецификации на идеальный ФНЧ могут быть заданы в виде

$$|H(e^{j\hat{\omega}})| = \begin{cases} L, & |\hat{\omega}| \leq \min\left(\frac{\pi}{L}, \frac{\pi}{M}\right), \\ 0, & \text{иначе.} \end{cases} \quad (10.23)$$

Тем самым подавляются «изображения», вызванные интерполятором, и одновременно обеспечивается отсутствие наложения, вызванное дециматором. Синтез фильтров (10.21) – (10.23) – стандартная задача разработки ФНЧ с конечной или бесконечной ИХ.

10.1.8. Банки цифровых фильтров

В ЦОС существует ряд приложений, в которых необходимо разделять сигнал на множество поддиапазонных сигналов, занимающих обычно неперекрывающиеся части исходной полосы частот. В других приложениях необходимо объединять много поддиапазонных сигналов в один составной сигнал. Таким образом, в данных практических приложениях банки цифровых фильтров играют важную роль.

Основные определения

Банк цифровых фильтров является набором цифровых полосовых фильтров с общим входом или суммирующим выходом, как показано на рис. 10.9. Устройство, показанное на рис. 10.9а, называется M -канальным банком фильтров анализа с фильтрами $H_k(z)$, известными как фильтры анализа. Банк фильтров анализа необходим для разложения входного сигнала $x(n)$ в набор M поддиапазонных сигналов $v_k(n)$, где каждый поддиапазонный сигнал занимает часть исходной полосы частот. Тем самым сигнал будет «проанализирован», когда будет разделен в частотной области на набор узких спектральных полос.

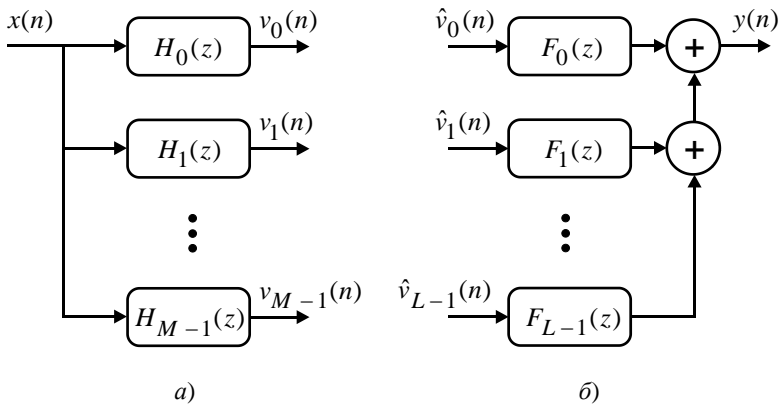


Рис. 10.9. Банки фильтров: а) анализа; б) синтеза

Двойственной операцией анализа является синтез, посредством которого набор поддиапазонных сигналов $v_k(n)$ объединяется в один сигнал $y(n)$. На рис. 10.9б показан L -диапазонный синтезирующий банк фильтров, где каждый фильтр $F_k(z)$ называется синтезирующим фильтром.

Равномерные банки фильтров

Ниже описывается техника создания класса банков фильтров с равной шириной полосы пропускания. Допустим, что $H_0(z)$ представляет собой физически реализуемый цифровой ФНЧ с действительной ИХ $h_0(n)$:

$$H_0(z) = \sum_{n=0}^{\infty} h_0(n)z^{-n}, \quad (10.24)$$

где допускается, что фильтр может иметь бесконечную импульсную характеристику. Пусть $H_0(z)$ имеет граничные частоты в полосе пропускания $\hat{\omega}_n$ и полосе задерживания $\hat{\omega}_z$, сосредоточенные вокруг частоты π/M (рис. 10.10а), где M – некоторое произвольное целое число. Теперь рассмотрим физически реализуемый импульсный отклик $h_k(n)$, $0 \leq k \leq M-1$, полученный модулированием $h_0(n)$ экспоненциальной последовательностью

$$h_k(n) = h_0(n)e^{j\frac{2\pi}{M}kn}. \quad (10.25)$$

Соответствующая передаточная функция равна

$$H_k(z) = \sum_{n=0}^{\infty} h_k(n)z^{-n} = \sum_{n=0}^{\infty} h_0(n) \left(e^{-j\frac{2\pi}{M}k} z \right)^{-n} = H_0 \left(e^{-j\frac{2\pi}{M}k} z \right), \quad (10.26)$$

где $0 \leq k \leq M-1$. При этом ЧХ может быть представлена в виде

$$H_k(e^{j\hat{\omega}}) = H_0(e^{j(\hat{\omega}-2\pi k/M)}), \quad 0 \leq k \leq M-1. \quad (10.27)$$

Таким образом, ЧХ $H_k(e^{j\hat{\omega}})$ получена смещением $H_0(e^{j\hat{\omega}})$ на $2\pi k/M$ (рис. 10.10б). Соответствующие ИХ $h_k(n)$ в общем

случае являются комплексными, и, следовательно, функция $|H_k(e^{j\hat{\omega}})|$ необязательно является симметричной относительно нулевой частоты. На рис. 10.10б представлены отклики $M-1$ фильтров, которые получены сдвигом низкочастотного фильтра-прототипа.

Отметим, что M фильтров $H_k(z)$, определенных в выражении (10.26), могут быть использованы как анализирующие фильтры в банке фильтров анализа (рис. 10.9а) или как фильтры банка синтеза $F_k(z)$ (рис. 10.9б). Полученный банк цифровых фильтров в данном случае называют равномерным.

10.2. Контрольные вопросы

1. Что собой представляет процесс увеличения частоты дискретизации во временной области? Приведите примеры.
2. Что собой представляет процесс уменьшения частоты дискретизации во временной области? Приведите примеры.
3. Опишите процесс увеличения частоты дискретизации во временной области. Приведите примеры.
4. Выполните вывод формулы (10.4).
5. Опишите процесс уменьшения частоты дискретизации во временной области. Приведите примеры.
6. Выполните вывод формулы (10.12).
7. Что такое многоскоростные тождества?
8. Как устроен интерполятор?
9. Для каких целей в структуре интерполятора используется ФНЧ?
10. Какие требования задаются на ФНЧ, используемый в структуре интерполятора?
11. Как устроен дециматор?
12. Для каких целей в структуре дециматора используется ФНЧ?
13. Какие требования задаются на ФНЧ, используемый в структуре дециматора?
14. Опишите процесс изменения частоты дискретизации в дробное число раз.

15. Какие требования задаются на ФНЧ, используемый в структуре схемы изменения частоты дискретизации в дробное число раз?
16. Определите соотношения вход-выход во временной области и в области z-преобразования для структур, изображенных на рис. 10.7 и 10.8.
17. Что такое банк фильтров?
18. Какие функции выполняет банк анализа?
19. Какие функции выполняет банк синтеза?
20. Что такое равномерный банк фильтров?
21. Как построить равномерный банк фильтров?

10.3. Литература

1. Брюханов, Ю. А. Цифровые фильтры: учеб. пособие / Ю. А. Брюханов, А. Л. Приоров. – Ярославль: ЯрГУ, 2002.
2. Приоров, А. Л. Основы цифровой обработки сигналов: лабораторный практикум / А. Л. Приоров, В. А. Волохов, А. М. Лаврентьев. – ВУНЦ ВВС «ВВА им. профессора Н. Е. Жуковского и Ю. А. Гагарина» (филиал, г. Ярославль), 2012.
3. Чобану, М. К. Цифровые многоскоростные системы обработки сигналов: учеб. пособие / М. К. Чобану. – М.: Издательский дом МЭИ, 2009.
4. Айфичер, Э. Цифровая обработка сигналов: практический подход. – 2-е изд. / Э. Айфичер, Б. Джервис. – Вильямс, 2004.
5. Сергиенко, А. Б. Цифровая обработка сигналов. – 3-е изд. / А. Б. Сергиенко. – СПб.: БХВ-Петербург, 2011.
6. Дьяконов, В. П. MATLAB 6.0/6.1/6.5/6.5+SP1 + Simulink 4/5. Обработка сигналов и изображений / В. П. Дьяконов. – М.: СОЛОН-Пресс, 2005.
7. Солонина, А. И. Цифровая обработка сигналов. Моделирование в MATLAB / А. И. Солонина, С. М. Арбузов. – СПб.: БХВ-Петербург, 2008.
8. Ingle, V. K. Digital Signal Processing using Matlab V.4 / V. K. Ingle, J. G. Proakis, PWS Publishing Company, 1997.

11. ОБРАБОТКА СИГНАЛОВ НА ОСНОВЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

11.1. Основные определения

Непрерывное преобразование Фурье (НПФ) является важнейшим средством анализа стационарных непрерывных сигналов, которые широко используются в радиотехнике и инфокоммуникациях. При этом сигнал раскладывается в базис синусов и косинусов различных частот. Количество этих функций бесконечно большое, а коэффициенты преобразования находятся путем вычисления скалярного произведения сигнала с комплексными экспонентами:

$$F(\Omega) = \int_{-\infty}^{\infty} f(t)e^{-j\Omega t} dt, \quad (11.1)$$

где $f(t)$ – сигнал, а $F(\Omega)$ – его преобразование Фурье.

С позиции точного представления произвольных сигналов и функций НПФ имеет ряд недостатков.

1. Даже для одной заданной частоты требуется знание сигнала не только в прошлом, но и в будущем, что является математической абстракцией.

2. В условиях практически неизбежного ограничения числа гармоник или спектра колебаний точное восстановление сигнала после прямого и обратного преобразований Фурье теоретически (и тем более практически) невозможно, в частности, из-за появления эффекта Гиббса.

3. Базисной функцией при разложении в ряд Фурье является гармоническое колебание, которое математически определено в интервале времени от $-\infty$ до ∞ и имеет неизменные во времени параметры.

4. Отдельные особенности сигнала (разрывы, пики) вызывают незначительные изменения частотного образа сигнала во всем интервале частот, которые «размазываются» по всей частотной оси, что делает их обнаружение по спектру практически невозможным.

5. Такая плавная базисная функция, как синусоида, в принципе не может представлять перепады сигналов с

бесконечной крутизной (прямоугольные импульсы и др.), хотя подобные сигналы применяются весьма широко.

6. Для нестационарных сигналов трудности НПФ многократно возрастают.

7. Единственным приспособлением к представлению быстрых изменений сигналов, таких как пики или перепады, является резкое увеличение числа гармоник, которые оказывают влияние на форму сигнала и за пределами локальных особенностей сигнала.

8. По составу высших составляющих спектра практически невозможно оценить местоположение особенностей на временной зависимости сигнала и их характер.

Проблемы спектрального анализа и синтеза сигналов, ограниченных во времени, частично решаются переходом к так называемому оконному преобразованию Фурье (ОПФ):

$$F(\Omega, b) = \int_{-\infty}^{\infty} f(t)w(t-b)e^{-j\Omega t} dt, \quad (11.2)$$

которое перед использованием НПФ, выражаемого формулой (11.1), применяет операцию умножения сигнала на некоторую весовую функцию или окно. Окном $w(t-b)$ называется локальная функция, которая перемещается вдоль временной оси для вычисления преобразования Фурье в нескольких позициях b . Таким образом, преобразование становится зависимым от времени, и в результате получается частотно-временное описание сигнала.

Недостатком ОПФ является то, что при его вычислении используется фиксированное окно, которое не может быть адаптировано к локальным свойствам сигнала. Вейвлет-преобразование, рассматриваемое далее, решает эту и некоторые другие проблемы, связанные с обработкой сигналов.

11.2. Введение в теорию вейвлет-преобразования

За последние два-три десятилетия в мире возникло и оформилось новое научное направление, связанное с так

называемым вейвлет-преобразованием. Слово «wavelet», являющееся переводом французского «ondelette», означает небольшие волны, следующие друг за другом. В узком смысле вейвлеты – это семейство функций, получающихся путем масштабирования и сдвигов одной материнской функции. В широком смысле вейвлеты – это функции, обладающие хорошей частотной локализацией, чье среднее значение равно нулю.

Вейвлет-преобразование одномерного сигнала – это его представление в виде обобщенного ряда или интеграла Фурье по системе базисных функций (в настоящей работе ортогональных)

$$\Psi_{ab}(t) = |a|^{-1/2} \Psi\left(\frac{t-b}{a}\right), \quad (11.3)$$

сконструированных из *материнского* (порождающего) вейвлета $\psi(t)$, обладающего определенными свойствами, за счет операций сдвига во времени – b и изменения временного масштаба – a . Множитель $|a|^{-1/2}$ обеспечивает независимость нормы этих функций от масштабирующего числа a . Таким образом, для заданных значений параметров a и b функция $\Psi_{ab}(t)$ и есть вейвлет, порождаемый материнским вейвлетом $\psi(t)$.

11.2.1. Непрерывное вейвлет-преобразование (НВП)

Сконструлируем базис $\Psi_{ab}(t)$ с помощью непрерывных масштабных преобразований и переносов материнского вейвлета $\psi(t)$ с произвольными значениями базисных параметров a и b в формуле (11.3). Тогда по определению прямое вейвлет-преобразование сигнала $f(t)$ будет представлено в виде

$$W_f(a,b) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \Psi^*\left(\frac{t-b}{a}\right) dt. \quad (11.4)$$

Если для порождающего вейвлета $\psi(t)$ выполняется равенство $C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\Omega)|^2}{\Omega} d\Omega < \infty$, то возможно обратное преобразование:

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_f(a, b) \psi_{ab}(t) \frac{dadb}{a^2}. \quad (11.5)$$

Здесь $\Psi(\Omega)$ – преобразование Фурье функции $\psi(t)$.

Из (11.4) следует, что вейвлет-спектр $W_f(a, b)$, в отличие от спектра Фурье, является функцией двух аргументов: первый аргумент a (временной масштаб) аналогичен периоду осцилляций, т. е. является обратной частоте, а второй аргумент b аналогичен смещению сигнала по оси времени. Следует отметить, что $W_f(a_0, b)$ характеризует временную зависимость при фиксированном значении $a = a_0$, тогда как зависимости $W_f(a, b_0)$ можно поставить в соответствие частотную зависимость при фиксированном значении $b = b_0$.

При непрерывном изменении параметров a и b для расчета вейвлет-спектра необходимы большие вычислительные затраты. Множество функций $\psi_{ab}(t)$ является избыточным. Поэтому необходима дискретизация параметров a и b при сохранении возможности восстановления сигнала из его трансформант. Дискретизация, как правило, осуществляется через степени двойки:

$$a = 2^j, b = k2^j, \psi_{jk}(t) = a^{-1/2} \psi\left(\frac{t-b}{a}\right) = 2^{-j/2} \psi(2^{-j}t - k),$$

где j и k – целые числа, а j называется параметром масштаба. В отличие от НВП в данном случае рассматриваются не все сдвиги и растяжения базисной функции, а только взятые на некоторой дискретной сетке (обычно логарифмической). Здесь необходимо отметить, что если сигнал остается непрерывным, то называть это преобразование дискретным неверно. В литературе, посвященной

вейвлет-анализу, его называют диадным вейвлет-преобразованием, прямой и обратный вид которого представлен в виде:

$$d_{jk} = \int_{-\infty}^{\infty} f(t)\psi_{jk}^*(t)dt, \quad (11.6)$$

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{jk}\psi_{jk}(t). \quad (11.7)$$

Если же сигнал дискретный, то аналогичное преобразование правильно называть дискретным вейвлет-преобразованием (ДВП).

11.2.2. Признаки вейвлета

1. *Локализация.* Вейвлет-преобразование, в отличие от преобразования Фурье, использует локализованные базисные функции. Вейвлет должен быть локализован как во временной области, так и в частотной. Для этого достаточно, чтобы выполнялись условия

$$|\psi(t)| \leq C(1 + |t|)^{-1-\varepsilon}, \quad |\Psi(\Omega)| \leq C(1 + |\Omega|)^{-1-\varepsilon}, \quad \text{при } \varepsilon > 0.$$

Например, дельта-функция и гармоническая функция не удовлетворяют необходимому условию одновременной локализации во временной и частотной областях.

2. *Нулевое среднее.* График вейвлет-функции должен осциллировать (быть знакопеременным) вокруг нуля на оси времени и иметь нулевую площадь:

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

Из этого условия становится понятным значение слова «wavelet» – небольшая волна. Часто для приложений необходимо, чтобы не только нулевой, но и все первые n моментов были равны нулю:

$$\int_{-\infty}^{\infty} t^n \psi(t) dt = 0.$$

Такой вейвлет называется вейвлетом n -го порядка. Обладающие большим числом нулевых моментов вейвлеты позволяют, игнорируя наиболее регулярные полиномиальные составляющие сигнала, анализировать мелкомасштабные флуктуации и особенности высокого порядка.

3. *Ограниченность.* Квадрат нормы функции должен быть конечным:

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty.$$

4. *Автомодельность базиса.* Характерным признаком базиса вейвлет-преобразования является его самоподобие. Все вейвлеты данного семейства $\Psi_{ab}(t)$ имеют то же число осцилляций, что и базисный вейвлет $\psi(t)$, поскольку получены из него посредством масштабных преобразований и сдвигов.

11.2.3. Примеры вейвлетобразующих функций (материнских вейвлетов)

Вейвлет-преобразование есть скалярное произведение анализирующего вейвлета на заданном масштабе и анализируемого сигнала. Вейвлет-коэффициенты $W_f(a, b)$ содержат комбинированную информацию об анализирующем вейвлете и анализируемом сигнале (как и коэффициенты преобразования Фурье, которые содержат информацию о сигнале и синусоидальной волне).

Выбор анализирующего вейвлета, как правило, определяется тем, какую информацию необходимо извлечь из сигнала. Каждый вейвлет имеет характерные особенности во временной и в частотной области, поэтому иногда с помощью разных вейвлетов можно полнее выявить и подчеркнуть те или иные свойства анализируемого сигнала. Примеры наиболее распространенных вейвлет-функций приведены в табл. 11.1.

Таблица 11.1

Примеры вейвлет-функций $\psi(t)$

Вейвлеты	Аналитическая запись $\psi(t)$	Спектральная плотность $\Psi(\Omega)$
Вещественные непрерывные базисы		
<i>Гауссовы:</i> первого порядка или WAVE- вейвлет	$-te^{-t^2/2}$	$(i\Omega)\sqrt{2\pi}e^{-\Omega^2/2}$
второго порядка или МНАТ- вейвлет («мексиканская шляпа» – mexican hat),	$(1-t^2)e^{-t^2/2}$	$(i\Omega)^2\sqrt{2\pi}e^{-\Omega^2/2}$
<i>n</i> -го порядка	$(-1)^n \frac{d^n}{dt^n} \left(e^{-t^2/2} \right)$	$(-1)^n (i\Omega)^n \sqrt{2\pi} e^{-\Omega^2/2}$
<i>DOG – difference of Gaussians</i>	$e^{-t^2/2} - 0.5e^{-t^2/8}$	$\sqrt{2\pi} \left(e^{-\Omega^2/2} - 0.5e^{-2\Omega^2} \right)$
<i>LP-Littlewood & Paley</i>	$(\pi t)^{-1} (\sin(2\pi t) - \sin(\pi t))$	$\begin{cases} (2\pi)^{-1/2}, & \pi \leq \Omega \leq 2\pi, \\ 0, & \text{иначе.} \end{cases}$
Вещественные дискретные базисы		
<i>Вейвлет Хаара (Haar)</i>	$\begin{cases} 1, & 0 \leq t < 1/2, \\ -1, & 1/2 \leq t < 1, \\ 0, & \text{иначе.} \end{cases}$	$\Psi(\Omega) = i \frac{\sin^2(\Omega/4)}{\Omega/4} e^{i\Omega/2}$
Комплексные базисы		
<i>Морле (Morlet)</i>	$e^{i\Omega_0 t} e^{-t^2/2}$	$\sigma(\Omega)\sqrt{2\pi}e^{-(\Omega-\Omega_0)^2/2}$
<i>Пауля (Paul)</i>	$\Gamma(n+1) \frac{i^n}{(1-n)^{n+1}}$	$\sigma(\Omega)\sqrt{2\pi}\Omega^n e^{-\Omega}$

Наиболее распространенные вещественные базисы конструируются на основе производных функции Гаусса:

$$\psi(t) = e^{-t^2/2}.$$

Это обусловлено тем обстоятельством, что функция Гаусса имеет наилучшие показатели локализации как во временной, так и в частотной областях. При $n=1$ получаем вейвлет первого порядка, называемый WAVE-вейвлетом, с равным нулю нулевым моментом. При $n=2$ получаем МНАТ-вейвлет, называемый «мексиканская шляпа» (mexican hat – похож на сомбреро). У него нулевой и первый моменты равны нулю. Он имеет лучшее разрешение, чем WAVE-вейвлет.

Наиболее простой пример *дискретного* вейвлета – это вейвлет Хаара (Haar). Недостатком его являются несимметричность формы и негладкость – резкие границы во временной области, вследствие чего возникает бесконечное чередование «лепестков» в частотной области, хотя и убывающих как $1/\omega$.

Среди *комплексных* вейвлетов наиболее часто используется базис, основанный на хорошо локализованном и во временной, и в частотной областях вейвлете Морле (Morlet). Характерный параметр ω_0 позволяет изменять избирательность базиса. Вещественная и мнимая части $\psi(t)$ – это амплитудно-модулированные колебания.

Выше был представлен небольшой перечень вейвлетов, описываемых аналитически в явном виде. Однако большинство вейвлет-функций не имеют аналитического описания в виде одной формулы, а задаются итерационными выражениями, легко вычисляемыми компьютерами. Примером таких вейвлетов являются функции Добеши (Daubechies).

Выбор конкретного материнского вейвлета (будь то непрерывный или дискретный) целиком зависит от характера поставленной задачи и от конкретного анализируемого сигнала. Разные сигналы удастся анализировать тем или иным способом, и критерием успеха обычно служит простота получаемого разложения. При этом решающим фактором оказываются интуиция и практический опыт исследователя.

11.2.4. Свойства вейвлет-анализа

Как уже отмечалось, прямое вейвлет-преобразование (11.4) содержит комбинированную информацию об анализируемом сигнале и анализирующем вейвлете. Несмотря на это, вейвлет-преобразование позволяет получить объективную информацию о сигнале, потому что некоторые свойства вейвлет-преобразования не зависят от выбора анализирующего вейвлета. Независимость от вейвлета делает эти простые свойства очень важными.

1. *Линейность*. Она следует из скалярного произведения (11.4):

$$W(\alpha f_1(t) + \beta f_2(t)) = \alpha W_{f_1}(a, b) + \beta W_{f_2}(a, b).$$

2. *Сдвиг*. Смещение сигнала во временной области на b_0 ведет к сдвигу вейвлет-образа также на b_0 :

$$W(f(t - b_0)) = W_f(a, b - b_0).$$

3. *Масштабирование*. Растяжение (сжатие) сигнала приводит также к растяжению (сжатию) его вейвлет-образа $W_f(a, b)$:

$$W(f(t/a_0)) = \frac{1}{a_0} W_f\left[\frac{a}{a_0}, \frac{b}{a_0}\right].$$

4. *Дифференцирование*:

$$W(d_t^n f(t)) = (-1)^n \int_{-\infty}^{\infty} f(t) d_t^n [\psi_{a,b}(t)] dt,$$

где $d_t^n f(t) = d^n f(t) / dt^n$, $n \geq 1$. Из этого свойства следует, что проигнорировать, например, крупномасштабные составляющие и проанализировать особенности высокого порядка или мелкомасштабные вариации сигнала $f(t)$ можно дифференцированием нужное число раз либо вейвлета, либо самого сигнала.

5. *Масштабно-временная локализация.* Она обусловлена тем, что элементы базиса вейвлет-преобразования хорошо локализованы и обладают подвижным частотно-временным окном.

За счет изменения масштаба (увеличение a приводит к сужению спектра Фурье функции $\psi_{ab}(t)$) вейвлеты способны выявлять различие в характеристиках на разных шкалах (частотах), а за счет сдвига – проанализировать свойства сигнала в разных точках на всем исследуемом интервале. Поэтому при анализе *нестационарных сигналов*, за счет свойства локальности вейвлетов, удастся получить существенное преимущество перед преобразованием Фурье, которое дает только глобальные сведения о частотах (масштабах) анализируемого сигнала, так как используемая при этом система функций (комплексная экспонента или синусы и косинусы) определена на бесконечном интервале.

По этой причине многие исследователи называют вейвлет-анализ *«математическим микроскопом»*. Это название хорошо отражает замечательные свойства метода сохранять хорошее разрешение на разных масштабах. Параметр сдвига b фиксирует точку фокусировки микроскопа, масштабный коэффициент a – увеличение, и, наконец, выбором материнского вейвлета $\psi(t)$ определяют его оптические качества.

11.2.5. Дискретное вейвлет-преобразование

Статьи, касающиеся практического использования вейвлет-преобразования, содержат в основной своей массе результаты компьютерных расчетов, в которых использовано ДВП, которое по аналогии с ДПФ широко используется в ЦОС. При этом не только параметры a и b , но и сигналы также дискретизируются во времени.

Подобно тому, как коэффициенты ДПФ обычно вычисляются при помощи БПФ, вейвлет-коэффициенты ДВП можно вычислить с помощью итерационной процедуры, известной под названием *быстрого вейвлет-преобразования* (БВП). Алгоритм БВП приведен далее. При этом, если необходимо, можно *сжать* полученные данные, отбросив

некоторую несущественную часть закодированной таким образом информации. Осуществляется это *квантованием*, в процессе которого приписываются разные весовые множители различным вейвлет-коэффициентам. Аккуратно проведенная процедура позволяет не только удалить некоторые статистические флуктуации и повысить роль динамических характеристик сигнала, но и существенно сократить компьютерную память и требования к передаче информации и, следовательно, снизить расходы.

11.2.6. Быстрое вейвлет-преобразование

При исследовании сигналов полезно их представление в виде совокупности последовательных приближений грубой (аппроксимирующей) $A_{j_0}(t)$ и уточненной (детализирующей) $D_j(t)$ составляющих

$$f(t) = A_{j_0}(t) + \sum_{j=1}^{j_0} D_j(t), \quad (11.8)$$

с последующим их уточнением итерационным методом. Каждый шаг уточнения соответствует определенному масштабу, т. е. уровню j_0 анализа (декомпозиции) и синтеза (реконструкции, восстановления) сигнала. Такое представление каждой составляющей сигнала вейвлетами можно рассматривать как во временной, так и в частотной областях. В этом суть кратномасштабного анализа (КМА).

Как уже отмечалось, на практике в большинстве случаев приходится сталкиваться с использованием ДВП. Однако формулы для вейвлет-преобразования дискретных сигналов не могут быть получены простой дискретизацией формул (11.6–11.7) диадного вейвлет-преобразования для непрерывного сигнала. Получим их из предпосылок КМА.

Пусть имеется непрерывный сигнал $f(t) \in V_0$. Дискретный сигнал интерпретируем как последовательность коэффициентов a_{0k} , полученную в ходе КМА сигнала $f(t)$ при масштабирующих функциях φ_{0k} :

$$f(t) = A_0(t) = \sum_k a_{0k} \phi_{0k}(t), \quad (11.9)$$

$$a_{0k} = \langle f(t), \phi_{0k}(t) \rangle = \int f(t) \phi_{0k}(t) dt,$$

где a_{0k} – коэффициенты аппроксимации на уровне $j_0 = 0$.

По концепции КМА сигнал $f(t)$ разлагается на две составляющие (принадлежащие подпространствам V_1 и W_1):

$$f(t) = A_1(t) + D_1(t) = \sum_k a_{1k} \phi_{1k}(t) + \sum_k d_{1k} \psi_{1k}(t). \quad (11.10)$$

Следовательно, получены две новые последовательности a_{1k} и d_{1k} . Отметим, что последовательности a_{1k} и d_{1k} имеют половинную длину по сравнению с a_{0k} . Далее процесс декомпозиции может быть продолжен по $A_1(t)$ (подпространства V_2 и W_2). Сигнал $f(t)$ на уровне разложения j_0 будет представлен совокупностью коэффициентов a_{j_0k} и d_{j_0k} .

Однако вычисления a_{j_0k} и d_{j_0k} по-прежнему зависят от непрерывных базисных функций $\phi(t)$ и $\psi(t)$. Эти функции однозначно определяются коэффициентами h_l :

$$\phi(t) = 2 \sum_l h_l \phi(2t-l), \quad (11.11)$$

$$\psi(t) = 2 \sum_l g_l \phi(2t-l), \quad (11.12)$$

$$h_l = \langle \phi(t), \phi(2t-l) \rangle = \int \phi(t) \phi(2t-l) dt,$$

$$g_l = (-1)^l h_{2n-l-1},$$

где $l = 0, 1, \dots, 2n-1$, n – порядок вейвлета. Вейвлеты n -го порядка существуют только на интервале длиной $2n-1$ и имеют $2n$ отличающихся от нуля коэффициентов h_l .

Из (11.11) и (11.12) можно получить следующие соотношения:

$$a_{j_0 k} = \langle f(t), \phi_{j_0 k}(t) \rangle = \sum_l h_{l-2k} \langle \phi(t), \phi_{j_0-1, l}(t) \rangle = \sum_l h_{l-2k} a_{j_0-1, l},$$

$$d_{j_0 k} = \langle f(t), \psi_{j_0 k}(t) \rangle = \sum_l g_{l-2k} \langle \phi(t), \phi_{j_0-1, l}(t) \rangle = \sum_l g_{l-2k} a_{j_0-1, l}.$$

Итерационная процедура быстрого вейвлет-анализа получила название анализа от «тонкого» к «грубому» масштабу.

На практике наименьший возможный масштаб (наибольший уровень разрешения) определяется числом N дискретных значений сигнала. На самом «тонком» значении масштаба $j_0 = 0$ за аппроксимирующие коэффициенты a_{0k} принимаются сами отсчеты дискретного сигнала. При переходе от текущего масштаба j_0 к следующему $j_0 + 1$ число вейвлет-коэффициентов уменьшается в два раза и они определяются по рекуррентным соотношениям:

$$a_{j_0+1, k} = \sum_l h_{l-2k} a_{j_0, k}, \quad (11.13)$$

$$d_{j_0+1, k} = \sum_l g_{l-2k} a_{j_0, k}. \quad (11.14)$$

При восстановлении (реконструкции) сигнала по его вейвлет-коэффициентам процесс идет от крупных масштабов к мелким и на каждом шаге описывается выражением

$$a_{j_0-1, k} = \sum_l (h_{k-2l} a_{j_0, l} + g_{k-2l} a_{j_0, l}), \quad (11.15)$$

которое получается из соотношений (11.11) и (11.12).

Таким образом, в практических приложениях с применением БВП используются только коэффициенты h_l , сами же вейвлеты не вычисляются и в расчетах не используются, что, как и в случае с БПФ, сокращает вычислительные затраты при обработке сигналов.

11.2.7. Частотный подход к вейвлет-преобразованию

До сих пор рассмотрение вейвлет-преобразования базировалось на временном подходе. Однако также плодотворна трактовка вейвлет-преобразования в частотной области на базе частотной фильтрации. В этом случае КМА сигнала рассматривается как поэтапная процедура фильтрации.

Обратимся к схеме, представленной на рис. 11.1. Сигнал $f(k)$ подается на низкочастотный (верхняя часть схемы) и высокочастотный (нижняя часть схемы) фильтры декомпозиции LoD и HiD соответственно, в которых вычисляется свертка (цифровая фильтрация) по формуле

$$y(k) = \sum_{l=0}^{2n-1} f(l)q(k-l), \quad (11.16)$$

где $2n$ – число отсчетов ИХ $q(k)$ фильтра.

В соответствии с (11.8) и (11.16) на выходе фильтров будут высокочастотная $y_H(k)$ и низкочастотная $y_L(k)$ компоненты сигнала.

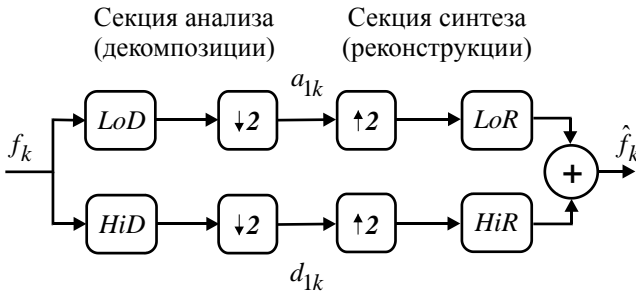


Рис. 11.1. Выполнение вейвлет-преобразования на базе частотной фильтрации

Из сопоставления (11.13), (11.14) и (11.16) следует, что для вычисления коэффициентов $a_{j_0 k}$ и $d_{j_0 k}$ (на первом этапе $j_0 = 1$) аргументы весовых коэффициентов фильтров должны быть взяты с обратным порядком следования, т. е. $h_{-l} = h(-l)$ и $g_{-l} = g(-l)$. Такие фильтры называются транспонированными.

Во вторых блоках схемы выполняется *децимация* с фактором 2, т. е. прореживание сигналов на выходе фильтров в два раза (из-за множителя 2 при аргументе k в формулах (11.13) и (11.14)).

Правая часть схемы рис. 11.1 осуществляет вейвлет-реконструкцию сигнала. Эта процедура использует операции интерполяции и фильтрации фильтрами реконструкции *LoR* и *HiR*. Операция *интерполяции* с фактором 2, обратная децимации с фактором 2, осуществляется путем увеличения в два раза числа составляющих добавлением нулевых компонентов вперемежку с имеющимися. При сложении сигналов, полученных на выходе фильтров *LoR* и *HiR*, получим сигнал $\hat{f}(k)$, близкий к исходному $f(k)$, т. е. произойдет его реконструкция на начальном уровне.

Для последующей итерации $j_0 = 2$ используются значения a_{1k} с предыдущей итерации и т. д.

Схема многошаговой итерационной процедуры анализа-синтеза показана на рис. 11.2, где представлена структура многошагового алгоритма декомпозиции и реконструкции сигнала, называемого алгоритмом Малла (Mallat).

Таким образом, быстрое вейвлет-преобразование во временной и частотной областях – это две стороны единой многошаговой структуры, позволяющей быстро осуществить как декомпозицию, так и реконструкцию цифрового сигнала.

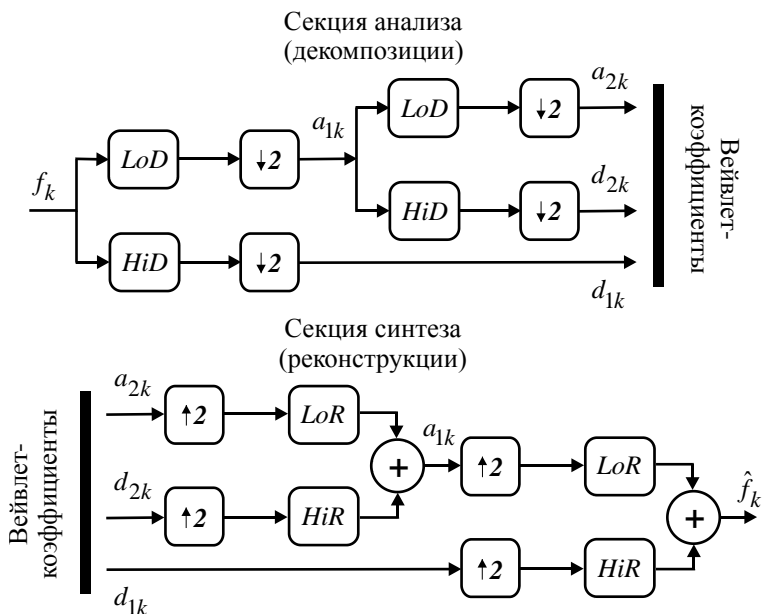


Рис. 11.2. Выполнение вейвлет-преобразования на базе частотной фильтрации (многошаговая итерационная процедура)

11.3. Удаление шумов из сигнала с использованием вейвлет-преобразования

Основу данной методики составляет использование пороговых функций различной формы, на основе которых происходит ограничение уровня детализирующих коэффициентов. Задав определенный порог для их уровня и «отсекая» коэффициенты ниже этого порога, можно значительно снизить уровень шума и сжать сигнал. Примеры пороговых функций, рассматриваемых в данной работе и зачастую используемых в современных алгоритмах фильтрации, основанных на кратномасштабных схемах преобразования, представлены на рис. 11.3.

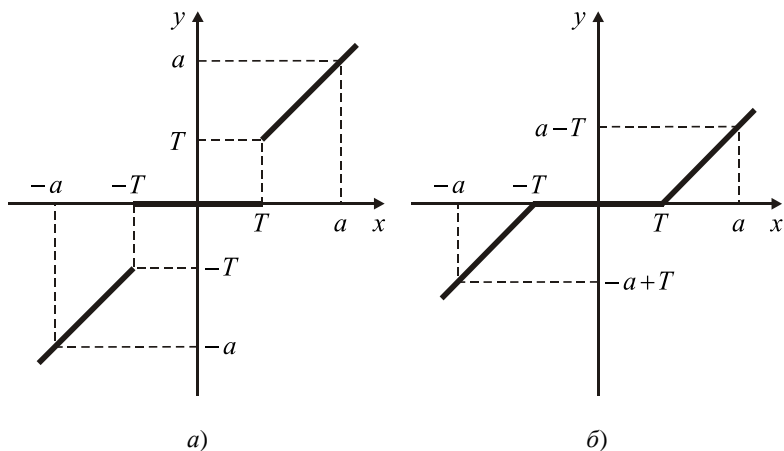


Рис. 11.3. Пороговые функции обработки коэффициентов преобразования:

а) жесткая пороговая функция; б) мягкая пороговая функция

На рис. 11.3а представлена *жесткая пороговая функция* (*жесткая пороговая оценка*), описываемая следующим выражением:

$$y(x) = \begin{cases} x, & \text{если } |x| \geq T, \\ 0, & \text{если } |x| < T. \end{cases} \quad (11.17)$$

Здесь в качестве величины T выступает некоторое *пороговое значение*, о возможности определения которого будет сказано ниже, в качестве x и y выступают входное и выходное значения коэффициентов преобразования.

На рис. 11.3б представлена *мягкая пороговая функция* (*мягкая пороговая оценка*), описываемая следующим выражением:

$$y(x) = \begin{cases} \text{sign}(x)(|x| - T), & \text{если } |x| \geq T, \\ 0, & \text{если } |x| < T. \end{cases} \quad (11.18)$$

Обозначения, представленные в выражение (11.18), те же, что и описанные выше. В качестве функции $sign(x)$, определяющей знак коэффициента x , выступает функция вида

$$sign(x) = \begin{cases} 1, & \text{если } x > 0, \\ 0, & \text{если } x = 0, \\ -1, & \text{если } x < 0. \end{cases} \quad (11.19)$$

Основное отличие мягкой пороговой функции от жесткой заключается в том, что первая из них не содержит разрыва в точке, определяемой пороговым значением T . Иными словами, мягкая пороговая функция является непрерывной. Последнее обстоятельство, в случае мягкой пороговой обработки, позволяет лучшим образом обрабатывать зашумленный сигнал (изображение) в окрестности точки его разрыва.

Нужно отметить, что понижение значения коэффициентов разложения на величину порогового значения, в случае мягкой пороговой обработки, в целом для большого количества сигналов оказывает неблагоприятное воздействие на итоговую оценку качества восстановленного сигнала. Поэтому в целом, как показывают эксперименты, лучшей с точки зрения численной оценки качества восстановленного сигнала является жесткая пороговая оценка.

Для того чтобы в случае мягкой пороговой обработки численная оценка качества восстановленного сигнала приближалась к последней в случае жесткой пороговой обработки, необходимо выбрать значение порога T , равное половине значения порога жесткой обработки сигнала.

Основной проблематикой использования пороговых методов обработки является выбор порогового значения. В большом количестве научных работ выбор порога осуществлялся с использованием следующего выражения:

$$T = \sigma \sqrt{2 \ln(N)}. \quad (11.20)$$

Формула (11.20) записана для случая одномерного сигнала. В ней использованы следующие обозначения: N – общее

количество отсчетов обрабатываемого сигнала, σ – средне-квадратическое отклонение шума, оцениваемое с использованием следующего выражения:

$$\sigma = \frac{M_X}{0.6745}, \quad (11.21)$$

где M_X – медиана абсолютного значения множества коэффициентов $\{X\}$. В качестве $\{X\}$ в данном случае выступают детализирующие коэффициенты, лежащие на самом низком уровне разложения.

Замечание. *Выражение (11.21) хорошо подходит для оценки среднеквадратического отклонения (СКО) белого гауссовского шума (БГШ) с нулевым математическим ожиданием, который является аддитивно подмешанным в некоторый полезный сигнал.*

11.4. Контрольные вопросы

1. В чем заключаются основные проблемы практического использования преобразования Фурье?
2. Что такое оконное преобразование Фурье?
3. Перечислите основные недостатки оконного преобразования Фурье.
4. Что такое вейвлет?
5. Приведите примеры вейвлет-функций.
6. Как осуществляется построение вейвлет-базиса?
7. Что такое непрерывное вейвлет-преобразование?
8. Что такое диадное вейвлет-преобразование?
9. Каковы основные признаки вейвлета?
10. Перечислите основные свойства вейвлет-анализа.
11. Что такое дискретное вейвлет-преобразование?
12. Как ДВП связано с кратномасштабным анализом?
13. Как получить аппроксимирующие и детализирующие коэффициенты при выполнении кратномасштабного анализа.
14. Что такое масштабирующая функция?
15. Что такое вейвлет-функция?

16. Что такое быстрое вейвлет-преобразование?
17. В чем заключается частотный подход к выполнению вейвлет-преобразования?
18. Что такое белый гауссовский шум?
19. Как построить алгоритм фильтрации сигналов с использованием вейвлет-преобразования?
20. В чем состоит отличие обработки вейвлет-коэффициентов на основе мягкой и жесткой пороговых функций?
21. Как оценить значение порога, используемого при обработке вейвлет-коэффициентов, на основе мягкой и жесткой пороговых функций?

11.5. Литература

1. Яковлев, А. Н. Введение в вейвлет-преобразования / А. Н. Яковлев. – Новосибирск: НГТУ, 2003.
2. Солонина, А. И. Цифровая обработка сигналов. Моделирование в MATLAB / А. И. Солонина, С. М. Арбузов. – СПб.: БХВ-Петербург, 2008.
3. Приоров, А. Л. Основы цифровой обработки сигналов: лабораторный практикум / А. Л. Приоров, В. А. Волохов, А. М. Лаврентьев. – ВУНЦ ВВС «ВВА им. профессора Н. Е. Жуковского и Ю. А. Гагарина» (филиал, г. Ярославль), 2012.
4. Дьяконов, В. П. MATLAB 6.0/6.1/6.5/6.5+SP1 + Simulink 4/5. Обработка сигналов и изображений / В. П. Дьяконов. – М.: СОЛОН-Пресс, 2005.
5. Смоленцев, Н. К. Основы теории вейвлетов. Вейвлеты в MATLAB / Н. К. Смоленцев, – М.: ДМК, 2005.
6. Малла, С. Вейвлеты в обработке сигналов / С. Малла. – М.: Мир, 2005.
7. Добеши, И. Десять лекций по вейвлетам / И. Добеши. – Москва-Ижевск: НИЦ «Регулярная и хаотическая динамика», 2004.
8. Блаттер, К. Вейвлет-анализ. Основы теории / К. Блаттер. – М.: Техносфера, 2004.

9. Короновский, А. Непрерывный вейвлетный анализ и его приложения / А. Короновский, А. Храмов. – М.: Физматлит., 2003.
10. Нефедов, В. И. Основы радиоэлектроники и связи: учеб. пособие / В. И. Нефедов, А. С. Сигов. – М.: Высшая школа, 2009.
11. Чуи, К. Введение в вейвлеты: пер. с англ. / К. Чуи. – М.: Мир, 2001.
12. Баскаков, С. И. Радиотехнические цепи и сигналы / С. И. Баскаков. – М.: Высшая школа, 2005.

12. ОСНОВЫ АДАПТИВНОЙ ОБРАБОТКИ СИГНАЛОВ

12.1. Вводные замечания

Сегодня адаптивная обработка сигналов является одним из направлений цифровой обработки сигналов, широко востребованным в современных радиотехнических системах. Часто под термином «адаптация» понимается процесс приспособления живых организмов или их органов к условиям окружающей среды. Однако адаптивными могут быть не только живые организмы, но и искусственные устройства и системы, которые в процессе своего функционирования могут самообучаться. В основе этих устройств обычно находится адаптивный фильтр (АФ).

Фильтрация сигналов – это обработка, целью которой является извлечение интересующей информации, например сигналов, обладающих определенными характеристиками.

Когда требуется использовать адаптивные фильтры? Известно, что свойства цифрового фильтра с фиксированными параметрами определяются его передаточной функцией. Если спецификацию к передаточной функции невозможно сформулировать заранее или она может меняться в процессе работы фильтра, то вместо фильтра с фиксированными параметрами необходимо использовать фильтр с изменяемыми параметрами, например адаптивный. Определение этих параметров во многом зависит от критерия работы адаптивного фильтра. Этим критерием часто является достижение минимума некоторой целевой функции, как правило, квадратичной функции ошибок между так называемым «требуемым» и выходным сигналами адаптивного фильтра.

Выходной сигнал АФ приближается к требуемому сигналу за счет изменения весовых коэффициентов (ВК), рассчитываемых на основе обработки требуемого и входного (входных в многоканальном случае) сигналов. В установившемся режиме значения ВК соответствуют минимуму целевой функции или находятся в небольшой окрестности этого минимума. Так как состояние АФ меняется в процессе настройки ВК, то считается, что он приспособляется или адаптируется к имеющимся

условиям функционирования. Поэтому такой фильтр и алгоритм вычисления его ВК называются адаптивными.

Одноканальный АФ представляет собой устройство, представленное на рис. 12.1. Здесь $x(k)$ – входной сигнал, $d(k)$ – требуемый сигнал (используемый в обозначении символ « d » происходит от английского слова «desired», т. е. требуемый), $y(k)$ – выходной сигнал, $\alpha(k) = d(k) - y(k)$ – сигнал ошибки, k – индекс дискретного времени или номер отсчета обрабатываемого сигнала. Эти отсчеты обычно равномерно распределены на оси времени как $t(k) = kT_s = k/F_s$, где T_s – период дискретизации, а F_s – частота дискретизации обрабатываемых сигналов.

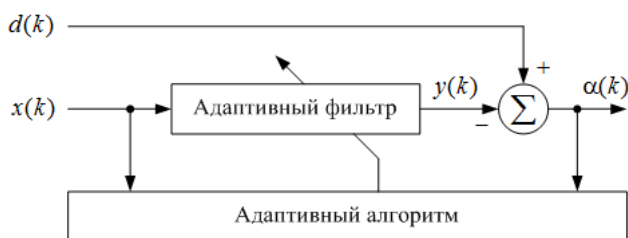


Рис. 12.1. Структурная схема адаптивного фильтра

Для работы АФ, как правило, недостаточно информации, получаемой только из входного сигнала. Источником недостающей информации обычно служит требуемый сигнал $d(k)$, физическая природа которого определяется конкретным приложением АФ. Приложение является одним из факторов, определяющих требования, предъявляемые к адаптивному фильтру на этапе его проектирования.

12.2. Требования к адаптивным фильтрам

Основные требования к адаптивному фильтру определяются такими составляющими, как приложение, структура и алгоритм.

1. Приложения. Тип приложения определяется сигналами, поступающими из окружающей среды в качестве входного $x(k)$ и требуемого $d(k)$ сигналов. Наиболее известными и широко используемыми приложениями АФ являются эхо-компенсаторы, эквалайзеры, адаптивные решетки и ряд других адаптивных устройств.

2. Структуры адаптивных фильтров. Подобно фильтрам с фиксированными ВК, существуют две основные структуры АФ. Это – фильтры с конечной импульсной характеристикой, или трансверсальные, и фильтры с бесконечной импульсной характеристикой, или рекурсивные.

Далее будут рассматриваться одноканальные адаптивные КИХ-фильтры с действительными ВК. Выходной сигнал $y(k)$ адаптивного КИХ-фильтра формируется как линейная комбинация задержанных отсчетов входного сигнала $x(k-n+1)$, взятых с весами $h_n(k-1)$, вычисляемыми на предыдущих итерациях $(k-1)$ по отношению к текущим итерациям k , т. е. как

$$y(k) = \sum_{n=1}^N h_n(k-1)x(k-n+1) = \mathbf{h}_N^T(k-1)\mathbf{x}_N(k), \quad (12.1)$$

где N – число ВК фильтра,

$\mathbf{h}_N(k-1) = [h_1(k-1), \dots, h_{N-1}(k-1), h_N(k-1)]^T$ – вектор ВК.

Используя выходной сигнал адаптивного фильтра (12.1), можно вычислить сигнал априорной ошибки моделирования требуемого сигнала:

$$\alpha(k) = d(k) - y(k) = d(k) - \mathbf{h}_N^T(k-1)\mathbf{x}_N(k). \quad (12.2)$$

Сигнал апостериорной ошибки определяется при ВК, равных $\mathbf{h}_N(k)$, как

$$e(k) = d(k) - y'(k) = d(k) - \mathbf{h}_N^T(k)\mathbf{x}_N(k). \quad (12.3)$$

Термины «априорный» и «апостериорный» связаны с используемыми значениями ВК адаптивного фильтра,

вычисляемыми соответственно на предыдущей и текущей итерациях алгоритма адаптивной фильтрации.

В большинстве приложений используются АФ с действительными ВК. Однако существуют АФ и с комплексными ВК, которые обычно применяются при построении адаптивных антенных решеток (ААР), эхо-компенсаторов и эквалайзеров в цифровых системах связи с квадратурной модуляцией, так как сигналы, обрабатываемые этими устройствами, являются комплексными. Это естественным образом приводит к необходимости использования АФ с комплексными ВК.

3. Алгоритмы. Адаптивный алгоритм (АА) – это процедура вычисления ВК, обеспечивающих минимизацию целевой функции адаптивного фильтра, т. е. выполнения критерия работы этого фильтра. Большинство вычислений в адаптивных алгоритмах представляют собой итерационные процедуры. При этом вычисляемые переменные обновляются на каждой итерации путем прибавления некоторых добавок к предыдущим значениям или вычитания этих добавок из предыдущих значений.

Например, в АА по критерию наименьшего среднего квадрата (Least Mean Square, LMS) ВК вычисляются как

$$\left\{ \begin{array}{l} \text{текущее} \\ \text{значение} \\ \text{вектора ВК} \end{array} \right\} = \left\{ \begin{array}{l} \text{предыдущее} \\ \text{значение} \\ \text{вектора ВК} \end{array} \right\} + \left\{ \begin{array}{l} \text{шаг} \\ \text{сходимости} \end{array} \right\} \left\{ \begin{array}{l} \text{вектор} \\ \text{входных} \\ \text{сигналов} \end{array} \right\} \{ \text{ошибка} \}.$$

В рекурсивных алгоритмах по критерию наименьших квадратов (Recursive Least Squares, RLS) ВК вычисляются как

$$\left\{ \begin{array}{l} \text{текущее} \\ \text{значение} \\ \text{вектора ВК} \end{array} \right\} = \left\{ \begin{array}{l} \text{предыдущее} \\ \text{значение} \\ \text{вектора ВК} \end{array} \right\} + \left\{ \begin{array}{l} \text{вектор} \\ \text{коэффициентов} \\ \text{Калмана} \end{array} \right\} \{ \text{ошибка} \}.$$

Выбор АА обычно обусловлен рядом требований, предъявляемых к АФ.

12.3. Критерии функционирования адаптивных фильтров

В процессе работы АА осуществляется поиск таких параметров АФ, при которых значение заданной целевой функции является минимальным. В качестве таких параметров обычно выступает вектор ВК $\mathbf{h}_N(k)$, а в качестве целевой функции некоторая функция входного, выходного и требуемого сигналов, т. е. $F = f\{x(k), y(k), d(k)\}$. Целевую функцию можно рассматривать и как функцию ошибок, т. е. $F = f\{e(k)\}$. Она должна быть действительной и неотрицательной, т. е. $F \geq 0$ для всех $x(k)$, $y(k)$ и $d(k)$. При работе АФ вектор его ВК $\mathbf{h}_N(k) \rightarrow \mathbf{h}_{N,o}$, где $\mathbf{h}_{N,o}$ – вектор координат минимума многомерной целевой функции в пространстве ВК. Достижение минимума целевой функции является критерием работы АФ.

Существуют разные способы определения целевой функции АФ, влияющие, в конечном счете, на сложность алгоритмов ее минимизации. Целевыми функциями, наиболее часто используемыми в адаптивной фильтрации, являются:

1) $F = f\{e(k)\} = E\{e^2(k)\}$ – среднеквадратическая ошибка (Mean Square Error, MSE),

2) $F = f\{e(k)\} = \sum_{i=1}^k e^2(i)$ – наименьшие квадраты (Least Squares, LS),

3) $F = f\{e(k)\} = \sum_{i=1}^k \lambda^{k-i} e^2(i)$ – взвешенные наименьшие квадраты (Weighted Least Squares, WLS),

4) $F = f\{e(k)\} = e^2(k)$ – квадрат мгновенной ошибки.

На практике используются и другие целевые функции. Однако выбор квадратичных целевых функций в АФ во многом обусловлен тем, что в результате их использования можно получать алгоритмы адаптивной фильтрации в виде итерационных вычислительных процедур, не содержащих логических операций, что отличает такие алгоритмы от вычислительных процедур оптимизации общего вида. Кроме того, эти функции являются

унимодальными, что гарантирует сходимость алгоритмов адаптивной фильтрации на их основе к единственному оптимальному решению.

Строго говоря, MSE-функция 1) на практике редко применяется, так как для ее вычисления требуется наличие бесконечного объема данных. В то же время эта функция широко используется при анализе АФ.

Целевые функции 2), 3) и 4) отличаются как сложностью реализации алгоритмов на их основе, так и характеристиками сходимости и остаточными ошибками этих алгоритмов в установившемся состоянии. Так, квадрат мгновенной ошибки является самой простой с точки зрения реализации алгоритмов целевой функцией. Однако алгоритмы на основе этой функции характеризуются медленной сходимостью, поскольку это сильно упрощенная целевая функция. Целевая LS-функция обычно используется при обработке стационарных сигналов, а WLS – при обработке медленно меняющихся сигналов.

12.4. Приложения адаптивных фильтров

Существует ряд известных приложений АФ. Наиболее простым из них является прямая идентификация неизвестной системы (рис. 12.2), базирующаяся на определении импульсного отклика этой системы. При решении задачи идентификации сигнал $x(k)$ является входным и для идентифицируемой системы, и для АФ. Поэтому минимизация некоторой функции ошибок $\alpha(k)$ между выходным сигналом неизвестной системы $d(k)$ и выходным сигналом АФ $y(k)$ означает, что адаптивный фильтр моделирует неизвестную систему, т. е. $\mathbf{h}_N(k) \rightarrow \mathbf{w}_N$, так как фильтр «воспроизводит» выходной сигнал системы $y(k)$.

Компенсация сигналов электрического или акустического эха является примером задачи идентификации неизвестной системы.

Другим примером использования АФ является задача обратной идентификации объекта. Решение этой задачи широко

используется для выравнивания характеристик электрических каналов связи (рис. 12.3) с помощью адаптивных эквалайзеров.



Рис. 12.2. Адаптивная идентификация неизвестной системы

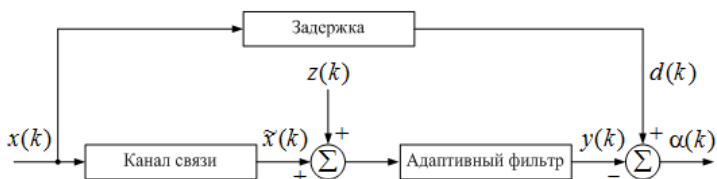


Рис. 12.3. Выравнивание электрического канала связи

Необходимость выравнивания канала связи возникает из-за того, что при передаче цифровых данных по каналам связи с ограниченной полосой пропускания возникает явление, называемое межсимвольной интерференцией (МСИ). Наряду с шумом $z(k)$ (рис. 12.3) МСИ сказывается на правильности решения о соответствии принятого информационного символа $\tilde{x}(k)$ символу $x(k)$, переданному по каналу связи. Если бы канал связи представлял собой просто линию задержки на время t_0 , то принимаемый сигнал (информационные символы) можно было бы легко оценивать (различать). Однако большинство каналов связи с математической точки зрения представляют собой КИХ-фильтр с импульсной характеристикой \mathbf{w} .

В каждый отсчет времени k , совпадающий с моментом смены информационных символов, следующих с периодом T , на

выходе канала образуется взвешенная сумма передаваемых подряд символов:

$$\tilde{x}(t_0 + kT) = w(t_0)x(t_0 + kT) + \sum_{n \neq k} w(t_0 + kT - nT)x(t_0 - nT). \quad (12.4)$$

В этой сумме только один символ несет текущую информацию, а остальные являются шумом (межсимвольной интерференцией), который не позволяет правильно различать символы на приемной стороне. Для уменьшения МСИ используют эквалайзеры.

Эквалайзер функционирует таким образом, что его АЧХ становится близкой к обратной АЧХ канала связи. Поэтому совместная АЧХ канала связи и эквалайзера становится близкой к равномерной. Это означает, что свертка импульсной характеристики канала связи \mathbf{w} и эквалайзера \mathbf{h} в моменты времени kT близка к дельта-функции, в результате чего МСИ уменьшается. В этом случае выходной сигнал эквалайзера определяется как

$$y(t_0 + kT) = \delta(t_0)x(t_0 + kT) + 0, \quad (12.5)$$

т. е. может быть правильно обнаружен (оценен и декодирован).

Если сравнить рис. 12.2 и 12.3, то можно заметить, что в последнем случае также решается задача идентификации линейной системы, определяющей задержку приема передаваемого сигнала. Известно, что ИХ линии задержки равна дельта-функции (т. е. содержит один ненулевой ВК), а АЧХ, как следствие, является равномерной. Это означает, что свертка фиксированной ИХ канала связи с изменяемым импульсным откликом АФ приближается в процессе адаптации к дельта-функции, а АЧХ становится равномерной.

Еще одним широко используемым приложением АФ является ААР (рис. 12.4), которая представляет собой многоканальный АФ с одним ВК в каждом из каналов.

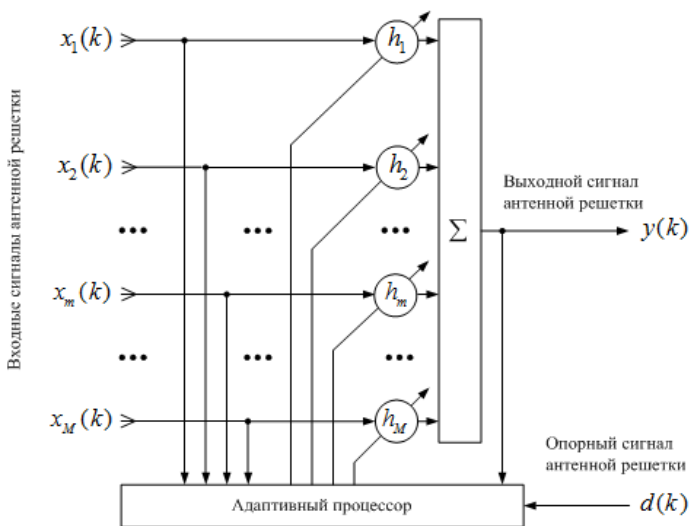


Рис. 12.4. Адаптивная антенная решетка

Антенная решетка характеризуется диаграммой направленности (ДН), т. е. амплитудно-угловой характеристикой $F(\theta)$, которая с помощью ВК может менять свою форму, обеспечивая требуемые значения усиления в направлениях на источники принимаемых сигналов, например единичное в направлении на источник полезного сигнала и нулевое или «очень маленькое» в направлениях на источники помех.

Широко используемым приложением АФ также является шумоочистка (компенсация шумов, шумоподавление). В задаче шумоочистки (рис. 12.5) в качестве сигнала $d(k)$ используется зашумленный полезный сигнал $x(k) + z_1(k)$, а в качестве входного сигнала – сигнал $z_2(k)$, коррелированный с сигналом $z_1(k)$ и некоррелированный с $x(k)$. Адаптивный фильтр из сигнала $z_2(k)$ выделяет сигнал $z_1(k)$, т. е. $y(k) \approx z_1(k)$, поэтому сигнал ошибки $\alpha(k) = d(k) - y(k) \approx x(k)$.

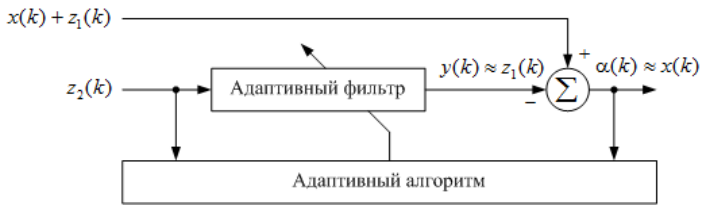


Рис. 12.5. Адаптивная шумоочистка сигналов

Адаптивные фильтры используются и для линейного предсказания наблюдаемых сигналов. При линейном предсказании в качестве требуемого сигнала АФ используется наблюдаемый сигнал, а в качестве входного сигнала – его задержанная копия. По окончании переходного процесса АФ представляет собой модель источника сигнала, а потому может быть использован для предсказания этого сигнала. Линейное предсказание, в частности, используется при кодировании речи (в вокодерах), при построении быстрых АФ, а также для выделения узкополосных сигналов $x(k)$ (спектральных линий, «line enhancement») на фоне широкополосных помех $z(k)$ (рис. 12.6).

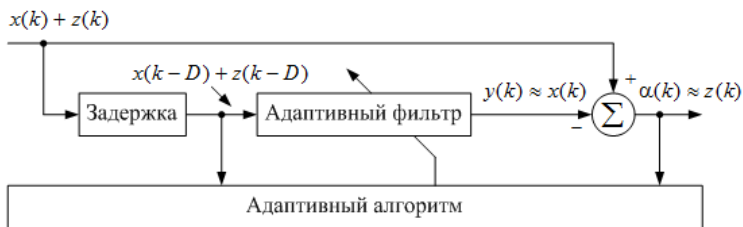


Рис. 12.6. Адаптивная фильтрация узкополосных сигналов

Принцип работы этой схемы базируется на различии корреляционных свойств широкополосного и узкополосного сигналов. Здесь величина задержки D в отсчетах обрабатываемых сигналов выбирается такой, что $k_z < D < k_x$, где k_z – ширина

основного лепестка автокорреляционной функции (АКФ) широкополосного сигнала $z(k)$ и k_x – ширина основного лепестка АКФ узкополосного сигнала $x(k)$. Будучи задержанным на D отсчетов, сигнал $x(k - D)$ остается коррелированным с сигналом $x(k)$, поэтому в процессе адаптации на выходе АФ формируется сигнал, близкий к узкополосному сигналу $x(k)$, а на выходе сигнала ошибки – сигнал, близкий к широкополосному сигналу $z(k)$. В зависимости от того, какой из обрабатываемых сигналов является «полезным» – узкополосный или широкополосный, в качестве «полезного» выходного сигнала используется выходной сигнал АФ или сигнал ошибки.

Среди других применений АФ можно отметить широкополосные акустические решетки, компенсаторы боковых лепестков направленных антенн, активные компенсаторы шумов, многоканальные эквалайзеры акустических каналов в аппаратуре высококачественного воспроизведения звука, одноканальные и многоканальные компенсаторы сигналов акустического эха в аппаратуре озвучивания помещений, полиномиальные (нелинейные) АФ и ряд других устройств. Их описание можно найти в современной научно-технической литературе.

12.5. Корреляционная матрица

Адаптивная фильтрация сигналов базируется на ряде принципов, вытекающих из оптимальной винеровской фильтрации. Решение задачи оптимальной винеровской фильтрации (вектор ВК фильтра) представляет собой решение задачи минимизации MSE-функции дискретных или цифровых фильтров. Это решение зависит от корреляционной матрицы входных сигналов сумматора и вектора взаимной корреляции входных и требуемого сигналов.

В АФ используются входные сигналы линейного сумматора (рис. 12.7), измеренные в дискретные моменты времени k .

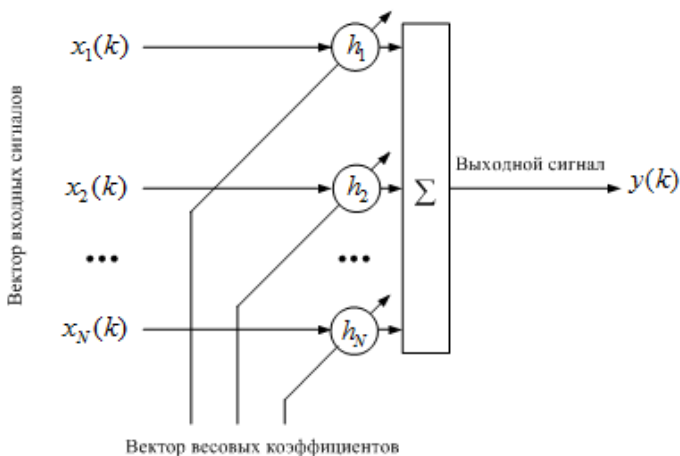


Рис. 12.7. Линейный сумматор

Результаты этих измерений образуют вектор отсчетов входных сигналов:

$$\mathbf{x}_N(k) = [x_1(k), x_2(k), \dots, x_N(k)]^T. \quad (12.6)$$

Если АФ по своей структуре является одноканальным КИХ-фильтром, то вектор

$$\mathbf{x}_N(k) = [x(k), x(k-1), \dots, x(k-N+1)]^T \quad (12.7)$$

представляет собой задержанные отсчеты входного сигнала этого фильтра, т. е. дискретный стохастический процесс.

Корреляционная матрица сигналов линейного сумматора определяется как

$$\mathbf{R}_N = E\{\mathbf{x}_N(k)\mathbf{x}_N^T(k)\} =$$

$$= \begin{bmatrix} E\{x_1^2(k)\} & E\{x_1(k)x_2(k)\} & \cdots & E\{x_1(k)x_N(k)\} \\ E\{x_2(k)x_1(k)\} & E\{x_2^2(k)\} & \cdots & E\{x_2(k)x_N(k)\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{x_N(k)x_1(k)\} & E\{x_N(k)x_2(k)\} & \cdots & E\{x_N^2(k)\} \end{bmatrix},$$

(12.8)

где $E\{\bullet\}$ – операция усреднения по ансамблю реализаций вектора $\mathbf{x}_N(k)$.

Корреляционная матрица является положительно полуопределенной. Если для произвольных комплексных векторов $\mathbf{x}_N(k)$ и $\mathbf{h}_N \neq \mathbf{0}_N$ определить скалярную переменную $y(k)$ как скалярное произведение этих векторов, т. е.

$$y(k) = \mathbf{h}_N^T \mathbf{x}_N(k), \quad (12.9)$$

то среднеквадратическое значение этой переменной – неотрицательная величина

$$E\{y^2(k)\} = E\{\mathbf{h}_N^T \mathbf{x}_N(k) \mathbf{x}_N^T(k) \mathbf{h}_N\} =$$

$$= \mathbf{h}_N^T E\{\mathbf{x}_N(k) \mathbf{x}_N^T(k)\} \mathbf{h}_N = \mathbf{h}_N^T \mathbf{R}_N \mathbf{h}_N \geq 0. \quad (12.10)$$

Неравенство (12.10) означает, что матрица \mathbf{R}_N положительно полуопределенная. Если $\mathbf{h}_N^H \mathbf{R}_N \mathbf{h}_N > 0$, то матрица \mathbf{R}_N положительно определенная.

12.6. Собственные числа и собственные векторы корреляционной матрицы

Одно из важных свойств корреляционной матрицы состоит в том, что эта матрица может быть разложена по собственным числам и соответствующим этим числам собственным векторам.

Число λ называется собственным числом матрицы \mathbf{R}_N , если выполняется условие

$$\mathbf{R}_N \mathbf{q}_N = \lambda \mathbf{q}_N, \quad (12.11)$$

где $\mathbf{q}_N \neq \mathbf{0}_N$ – собственный вектор, соответствующий этому собственному числу.

Уравнение (12.11) можно представить как

$$\mathbf{R}_N \mathbf{q}_N - \lambda \mathbf{q}_N = (\mathbf{R}_N - \lambda \mathbf{I}_N) \mathbf{q}_N = \mathbf{0}_N. \quad (12.12)$$

Уравнение (12.12) имеет нетривиальное решение $\mathbf{q}_N \neq \mathbf{0}_N$, только если определитель

$$\det \{ \mathbf{R}_N - \lambda \mathbf{I}_N \} = 0. \quad (12.13)$$

Уравнение (12.13) называется характеристическим уравнением матрицы \mathbf{R}_N и представляет собой полином N -й степени:

$$p(\lambda) = \lambda^N + a_{N-1} \lambda^{N-1} + \dots + a_1 \lambda + a_0 = 0. \quad (12.14)$$

Уравнение (12.14) имеет N корней (необязательно разных) $\lambda_1, \lambda_2, \dots, \lambda_N$, называемых собственными числами матрицы \mathbf{R}_N .

Каждый вектор $\mathbf{q}_{N,n} \neq \mathbf{0}_N$, где $n = 1, 2, \dots, N$, удовлетворяющий уравнению

$$\mathbf{R}_N \mathbf{q}_{N,n} = \lambda_n \mathbf{q}_{N,n}, \quad (12.15)$$

называется собственным вектором, соответствующим собственному числу λ_n .

Собственные числа λ_n и собственные векторы $\mathbf{q}_{N,n}$ корреляционной матрицы \mathbf{R}_N стационарного дискретного стохастического процесса характеризуются рядом свойств, основные из которых приводятся ниже.

Свойство 1. Если $\lambda_1, \lambda_2, \dots, \lambda_N$ – собственные числа корреляционной матрицы \mathbf{R}_N , то для любого целого числа $k > 0$

собственные числа матрицы $\mathbf{R}_N^k = \underbrace{\mathbf{R}_N \mathbf{R}_N \cdots \mathbf{R}_N}_k$ равны $\lambda_1^k, \lambda_2^k, \dots, \lambda_N^k$.

Свойство 2. Ненулевые собственные векторы $\mathbf{q}_{N,1}, \mathbf{q}_{N,2}, \dots, \mathbf{q}_{N,N}$ корреляционной матрицы \mathbf{R}_N , соответствующие различным собственным числам $\lambda_1, \lambda_2, \dots, \lambda_N$, являются линейно-независимыми.

Векторы $\mathbf{q}_{N,n}$ линейно-зависимые, если существуют такие $a_n \neq 0$, при которых

$$a_1 \mathbf{q}_{N,1} + a_2 \mathbf{q}_{N,2} + \dots + a_N \mathbf{q}_{N,N} = \sum_{n=1}^N a_n \mathbf{q}_{N,n} = \mathbf{0}_N. \quad (12.16)$$

Если равенство (12.16) выполняется только при всех $a_n = 0$, то векторы $\mathbf{q}_{N,n}$ являются линейно-независимыми.

Свойство 3. Собственные числа $\lambda_1, \lambda_2, \dots, \lambda_N$ корреляционной матрицы \mathbf{R}_N являются неотрицательными действительными числами.

Свойство 4. Собственные векторы $\mathbf{q}_{N,1}, \mathbf{q}_{N,2}, \dots, \mathbf{q}_{N,N}$ корреляционной матрицы \mathbf{R}_N , соответствующие различным собственным числам $\lambda_1, \lambda_2, \dots, \lambda_N$, являются ортогональными друг другу.

Ортогональность означает, что

$$\mathbf{q}_{N,n}^H \mathbf{q}_{N,m} = 0, \quad (12.17)$$

если $n \neq m$.

Так как каждому собственному числу λ_n соответствует бесконечное множество собственных векторов, то без потери общности в дальнейшем будем рассматривать только нормированные собственные векторы, т. е. такие, что

$$\mathbf{q}_{N,n}^T \mathbf{q}_{N,m} = \sum_{i=1}^N q_{in} q_{im} = \begin{cases} 1, & \text{при } n = m \\ 0, & \text{при } n \neq m \end{cases} \quad (12.18)$$

для всех $n=1, \dots, N$ и $m=1, \dots, N$. Векторы, удовлетворяющие условию (12.18), называются ортонормированными.

Свойство 5. Если собственные векторы $\mathbf{q}_{N,1}, \mathbf{q}_{N,2}, \dots, \mathbf{q}_{N,N}$ корреляционной матрицы \mathbf{R}_N , соответствующие различным собственным числам $\lambda_1, \lambda_2, \dots, \lambda_N$, представить матрицей $\mathbf{Q}_N = [\mathbf{q}_{N,1}, \mathbf{q}_{N,2}, \dots, \mathbf{q}_{N,N}]$ ортонормированных векторов (12.18), а собственные числа – диагональной матрицей $\Lambda_N = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$, то матрица \mathbf{R}_N может быть приведена к диагональному виду:

$$\mathbf{Q}_N^T \mathbf{R}_N \mathbf{Q}_N = \Lambda_N. \quad (12.19)$$

Из этого свойства также следует соотношение

$$\mathbf{Q}_N^T \mathbf{Q}_N = \mathbf{Q}_N \mathbf{Q}_N^T = \mathbf{I}_N, \quad (12.20)$$

откуда

$$\mathbf{Q}_N^T = \mathbf{Q}_N^{-1}. \quad (12.21)$$

Матрица, обладающая свойством (12.21), называется ортонормированной.

Используя (12.20) и (12.21), можно показать, что

$$\mathbf{R}_N = \mathbf{Q}_N \Lambda_N \mathbf{Q}_N^T \quad (12.22)$$

и

$$\mathbf{R}_N^{-1} = \mathbf{Q}_N \Lambda_N^{-1} \mathbf{Q}_N^T. \quad (12.23)$$

Свойство 6. Сумма собственных чисел $\lambda_1, \lambda_2, \dots, \lambda_N$ корреляционной матрицы \mathbf{R}_N равна следу этой матрицы:

$$\text{tr}\{\mathbf{R}_N\} = \sum_{n=1}^N \lambda_n. \quad (12.24)$$

Свойство 7. Произведение собственных чисел $\lambda_1, \lambda_2, \dots, \lambda_N$ корреляционной матрицы \mathbf{R}_N равно определителю этой матрицы, т. е.

$$\det \{ \mathbf{R}_N \} = \prod_{n=1}^N \lambda_n . \quad (12.25)$$

Рассмотренные свойства корреляционной матрицы, ее собственных чисел и собственных векторов используются при анализе процессов адаптивной фильтрации сигналов.

12.7. Винеровская фильтрация

Одной из целевых функций, используемых в теории адаптивной фильтрации, является MSE-функция, или среднеквадратическая ошибка, определяемая как

$$\xi(k) = E \{ |e(k)|^2 \} .$$

Для линейного сумматора (рис. 12.8) MSE-функция определяется следующим образом:

$$\xi = E \{ e^2(k) \} = \sigma_d^2 - \mathbf{h}_N^T \mathbf{r}_N - \mathbf{r}_N^T \mathbf{h}_N + \mathbf{h}_N^T \mathbf{R}_N \mathbf{h}_N , \quad (12.26)$$

где $e(k) = d(k) - y(k)$, $y(k) = \mathbf{h}_N^H \mathbf{x}_N(k)$, $\sigma_d^2 = E \{ d^2(k) \}$ – дисперсия (средняя мощность) требуемого сигнала, $\mathbf{r}_N = E \{ \mathbf{x}_N(k) d(k) \}$ – вектор взаимной корреляции между входными и требуемым сигналами, $\mathbf{R}_N = E \{ \mathbf{x}_N(k) \mathbf{x}_N^T(k) \}$ – корреляционная матрица входных сигналов (12.8). Вектор \mathbf{r}_N определяется как

$$\mathbf{r}_N = E \left\{ \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_N(k) \end{bmatrix} d(k) \right\} = \begin{bmatrix} E \{ x_1(k) d(k) \} \\ E \{ x_2(k) d(k) \} \\ \vdots \\ E \{ x_N(k) d(k) \} \end{bmatrix} . \quad (12.27)$$

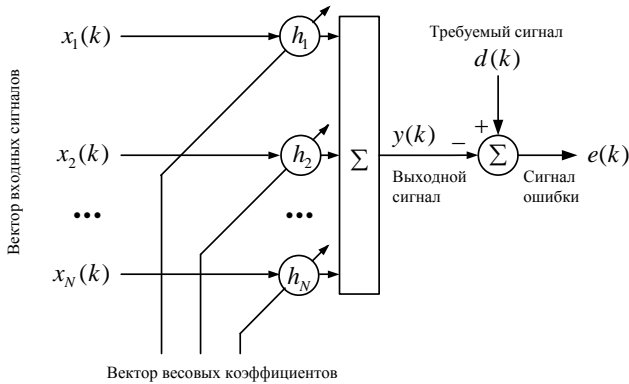


Рис. 12.8. Линейный сумматор: формирование сигнала ошибки

Вектор градиента MSE-функции (12.26) определяется как

$$\nabla_{\mathbf{h}_N} \xi = \frac{\partial \xi}{\partial \mathbf{h}_N} = \left[\frac{\partial \xi}{\partial h_1}, \frac{\partial \xi}{\partial h_2}, \dots, \frac{\partial \xi}{\partial h_N} \right]^T = -2\mathbf{r}_N + 2\mathbf{R}_N \mathbf{h}_N. \quad (12.28)$$

Полагая, что матрица \mathbf{R}_N является обратимой, условие

$$\nabla_{\mathbf{h}_N} \xi = -2\mathbf{r}_N + 2\mathbf{R}_N \mathbf{h}_N = \mathbf{0}_N \quad (12.29)$$

позволяет найти оптимальный вектор ВК

$$\mathbf{h}_{N,o} = \mathbf{R}_N^{-1} \mathbf{r}_N, \quad (12.30)$$

соответствующий минимуму целевой функции (12.26). Уравнение (12.29) называется уравнением Винера-Хопфа, вектор ВК (12.30) – оптимальным винеровским решением, а фильтр (рис. 12.8), ВК которого определяются уравнением (12.30), – оптимальным винеровским фильтром.

Используя (12.30) в уравнении (12.26), можно определить минимальное значение MSE:

$$\begin{aligned}\xi_{\min} &= E\{e^2(k)\} = \sigma_d^2 - \mathbf{h}_{N,o}^T \mathbf{r}_N - \mathbf{r}_N^T \mathbf{h}_{N,o} + \mathbf{h}_{N,o}^T \mathbf{R}_N \mathbf{h}_{N,o} = \\ &= \sigma_d^2 - \mathbf{r}_N^T \mathbf{h}_{N,o} = \sigma_d^2 - \mathbf{r}_N^T \mathbf{R}_N^{-1} \mathbf{r}_N.\end{aligned}\quad (12.31)$$

Вектор градиента функции (12.26) можно выразить и как

$$\nabla_{\mathbf{h}_N} \xi = \frac{\partial \xi}{\partial \mathbf{h}_N} = \frac{\partial E\{e^2(k)\}}{\partial \mathbf{h}_N} = -2E\{\mathbf{x}_N(k)e(k)\}, \quad (12.32)$$

Если требуемый сигнал определяется как

$$d(k) = \mathbf{h}_{N,o}^T \mathbf{x}_N(k) + e_o(k), \quad (12.33)$$

т. е. в нем присутствует аддитивный шум $e_o(k)$, часто называемый измерительным, и число ВК N линейной системы с ВК $\mathbf{h}_{N,o}$, порождающей сигнал $d(k)$, совпадает числом ВК линейного сумматора (рис. 12.8) \mathbf{h}_N , то при $\mathbf{h}_N = \mathbf{h}_{N,o} = \mathbf{R}_N^{-1} \mathbf{r}_N$ ошибка $e(k)$ определяется как

$$e(k) = d(k) - y(k) = \mathbf{h}_{N,o}^H \mathbf{x}_N(k) + e_o(k) - \mathbf{h}_{N,o}^H \mathbf{x}_N(k) = e_o(k), \quad (12.34)$$

а, согласно (12.28) и (12.32), градиент

$$\nabla_{\mathbf{h}_N} \xi = -2E\{\mathbf{x}_N(k)e_o(k)\} = -2\mathbf{r}_N + 2\mathbf{R}_N \mathbf{h}_N = -2\mathbf{r}_N + 2\mathbf{R}_N \mathbf{R}_N^{-1} \mathbf{r}_N = \mathbf{0}_N. \quad (12.35)$$

Уравнение (12.35) означает, что при $\mathbf{h}_N = \mathbf{h}_{N,o}$ сигнал ошибки ортогонален компонентам вектора входного сигнала, т. е. $E\{x_N(k-n+1)e_o(k)\} = 0$, $n=1, 2, \dots, N$, а также выходному сигналу фильтра, так как

$$\mathbf{h}_{N,o}^T E\{\mathbf{x}_N(k)e_o(k)\} = E\{\mathbf{h}_{N,o}^T \mathbf{x}_N(k)e_o(k)\} = E\{y_o(k)e_o(k)\} = 0. \quad (12.36)$$

Уравнения (12.35) и (12.36) известны как принцип ортогональности, означающий, что при $\mathbf{h}_N = \mathbf{h}_{N,o}$ сигнал ошибки фильтра ортогонален и к его входным и к выходному сигналам, т. е. сигнал ошибки и эти сигналы являются некоррелированными.

Природа шума $e_o(k)$ определяется применением адаптивного фильтра. Например, в компенсаторах сигналов акустического эха – это шум окружающей среды, в компенсаторах электрического эха – шум линий связи. Шум $e_o(k)$ также может порождаться в случае, если число ВК адаптивного фильтра, т. е. линейного сумматора (рис. 12.8), не совпадает с числом ВК импульсного отклика идентифицируемой линейной системы.

12.8. Поверхность среднеквадратической ошибки

MSE-функция (12.26), т. е. среднеквадратическая ошибка, является квадратичной функцией ВК \mathbf{h}_N . Она образует в N -мерном пространстве ВК эллиптический параболоид. Его поверхность вогнута и содержит только положительные значения. Пример MSE-функции при $N = 2$ показан на рис. 12.9.

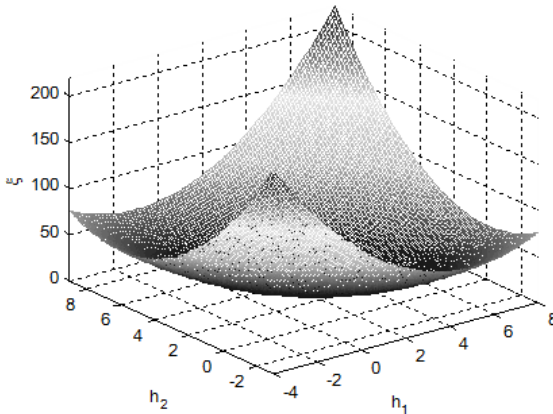


Рис. 12.9. Поверхность среднеквадратической ошибки

При исследовании свойств поверхности среднеквадратической ошибки MSE-функцию удобно выражать в терминах, смещенных относительно $\mathbf{h}_{N,o}$ векторов ВК $\Delta\mathbf{h}_N = \mathbf{h}_N - \mathbf{h}_{N,o}$, т. е. в системе координат, началом которых является точка $\mathbf{h}_{N,o}$. В этом случае уравнение (12.26) преобразуется к виду:

$$\xi = \xi_{\min} + \Delta\mathbf{h}_N^T \mathbf{R}_N \Delta\mathbf{h}_N. \quad (12.37)$$

Используя (12.22), уравнение (12.37) можно также выразить как

$$\xi = \xi_{\min} + \mathbf{v}_N^T \mathbf{\Lambda}_N \mathbf{v}_N = \xi_{\min} + \sum_{n=1}^N \lambda_n |v_n|^2, \quad (12.38)$$

где

$$\mathbf{v}_N = \mathbf{Q}_N^T \Delta\mathbf{h}_N \quad (12.39)$$

– вектор ВК в повернутой и смещенной системе координат.

При исследовании свойств адаптивных алгоритмов, базирующихся на использовании MSE-функции, полезно использовать графики линий постоянного уровня $\xi = \text{const}$. Эти графики представляют собой гиперэллипсы, так как уравнение (12.38) при $\xi = \text{const}$ – это уравнение гиперэллипса

$$\frac{|v_1|^2}{a_1^2} + \frac{|v_2|^2}{a_2^2} + \dots + \frac{|v_N|^2}{a_N^2} = 1, \quad (12.40)$$

N главных осей которого определяются как $2a_1, 2a_2, \dots, 2a_N$, где $a_1 = \sqrt{(\xi - \xi_{\min})/\lambda_1}$, $a_2 = \sqrt{(\xi - \xi_{\min})/\lambda_2}$, \dots , $a_N = \sqrt{(\xi - \xi_{\min})/\lambda_N}$.

Сечения MSE-функции для разных значений ξ представляют собой одни и те же концентрические гиперэллипсы в пространствах ВК \mathbf{h}_N , $\Delta\mathbf{h}_N$ и \mathbf{v}_N (см. пример на рис. 12.10 для случая $N = 2$).

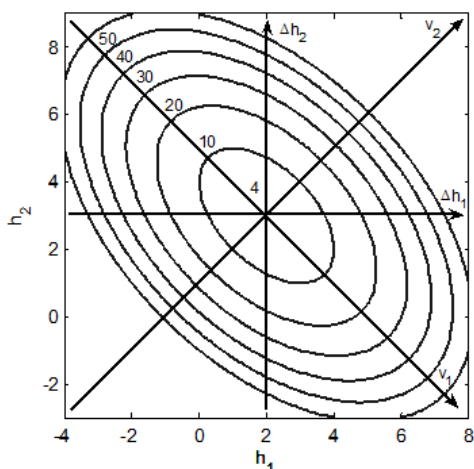


Рис. 12.10. Проекции сечений поверхности среднеквадратической ошибки

На рис. 12.10 показаны сечения MSE-функции, представленной на рис. 12.9. Цифры на поле рис. 12.10 показывают значения MSE-функции ξ на линиях постоянного уровня. В исходной системе координат центры концентрических гиперэллипсов находятся в точке $\mathbf{h}_{N,o}$. В смещенной, а также в повернутой и смещенной системах координат центры гиперэллипсов расположены в начале координат этих систем.

Разброс собственных чисел корреляционной матрицы \mathbf{R}_N , характеризуемый параметром $\chi = \lambda_{\max} / \lambda_{\min}$, определяет форму гиперэллипсов $\xi = const$. Гиперэллипсы вытягиваются вдоль оси повернутой и смещенной системы координат, определяемой меньшим значением собственного числа λ , и сужаются вдоль оси, определяемой большим значением собственного числа λ .

Таким образом, при стационарных входных и требуемом сигналах MSE-функция линейного сумматора является квадратичной, что позволяет найти координаты минимума этой функции, соответствующие винеровскому решению. Далее

рассматриваются два метода поиска винеровского решения в случае, когда параметры MSE-функции \mathbf{R}_N , \mathbf{r}_N и σ_d^2 являются известными. Это алгоритм Ньютона и алгоритм наискорейшего спуска.

12.9. Алгоритм Ньютона

Алгоритм Ньютона для поиска минимума MSE-функции может быть получен следующим образом. Если обе части уравнения градиента (12.28) MSE-функции умножить на обратную корреляционную матрицу

$$\mathbf{R}_N^{-1}(-2\mathbf{r}_N + 2\mathbf{R}_N\mathbf{h}_N) = \mathbf{R}_N^{-1}\nabla_{\mathbf{h}_N}\xi,$$

то, используя уравнение оптимального винеровского решения (12.30), можно получить уравнение $-2\mathbf{h}_{N,o} + 2\mathbf{h}_N = \mathbf{R}_N^{-1}\nabla_{\mathbf{h}_N}\xi$. Из этого уравнения следует, что вектор ВК, соответствующий винеровскому решению, можно вычислить как

$$\mathbf{h}_{N,o} = \mathbf{h}_N - \frac{1}{2}\mathbf{R}_N^{-1}\nabla_{\mathbf{h}_N}\xi, \quad (12.41)$$

где \mathbf{h}_N – вектор произвольных ВК линейного сумматора.

Однако на практике \mathbf{R}_N^{-1} и $\nabla_{\mathbf{h}_N}\xi$ редко являются известными. Поэтому уравнение (12.41) обычно модифицируется как

$$\mathbf{h}_N(k) = \mathbf{h}_N(k-1) - \mu \mathbf{R}_N^{-1}\nabla_{\mathbf{h}_N^*}\xi(k-1), \quad (12.42)$$

где μ – параметр, называемый шагом сходимости. Уравнение (12.42) представляет собой алгоритм Ньютона, с помощью которого итерационно находятся координаты минимума MSE-функции.

Используя уравнение градиента (12.28) $\nabla_{\mathbf{h}_N}\xi = -2\mathbf{r}_N + 2\mathbf{R}_N\mathbf{h}_N$, уравнение (12.42) можно преобразовать к виду

$$\mathbf{h}_N(k) = (1 - 2\mu) \mathbf{h}_N(k-1) + 2\mu \mathbf{h}_{N,o}. \quad (12.43)$$

На рис. 12.11 показан пример поиска действительных ВК с помощью уравнения (12.42) из трех начальных значений действительных ВК: $\mathbf{h}_N(0) = [0, 0]^T$, $\mathbf{h}_N(0) = [-6, 2]^T$ и $\mathbf{h}_N(0) = [1, 8]^T$. На каждой итерации алгоритма Ньютона изменения ВК направлены непосредственно в сторону минимума MSE-функции.

Функция $\xi(k)$, вычисляемая с помощью уравнения (12.26) как $\xi(k) = \sigma_d^2 - \mathbf{h}_N^T(k) \mathbf{r}_N - \mathbf{r}_N^T \mathbf{h}_N(k) + \mathbf{h}_N^T(k) \mathbf{R}_N \mathbf{h}_N(k)$, называется обучающей кривой. Такая кривая для движения алгоритма Ньютона из начальной точки $\mathbf{h}_N(0) = [0, 0]^T$ (рис. 12.11) представлена на рис. 12.12. Видно, что обучающая кривая имеет участок, именуемый переходным процессом, где $\xi(k)$ меняется «заметно» от итерации к итерации, и участок, именуемый установившимся состоянием, где $\xi(k)$ от итерации к итерации остается постоянным или «практически постоянным».

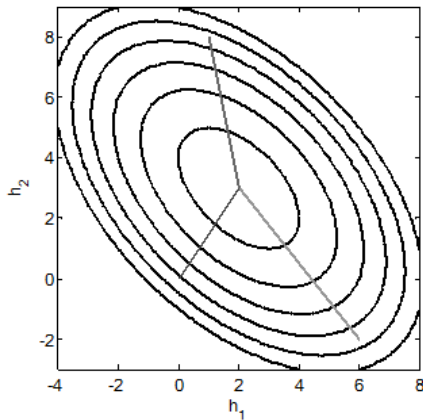


Рис. 12.11. Нахождение винеровского решения с помощью алгоритма Ньютона

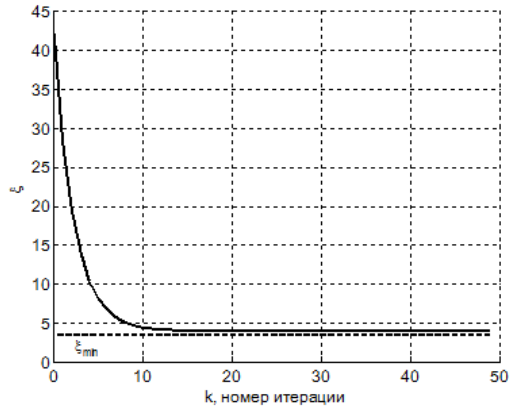


Рис. 12.12. Обучающая кривая алгоритма Ньютона: $\mathbf{h}_N(0) = [0, 0]^T$

Переходные процессы часто отображают в терминах ВК. Графики таких переходных процессов для рассмотренного случая приведены на рис.12.13. Видно, что движение вектора ВК осуществляется из начальной точки $\mathbf{h}_N(0) = [0, 0]^T$ в направлении вектора винеровского решения $\mathbf{h}_{N,o} = [2, 3]^T$.

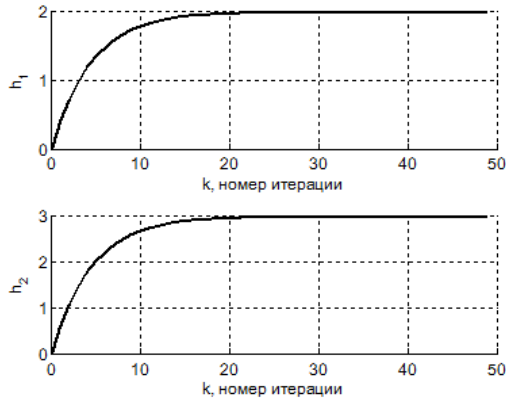


Рис. 12.13. Переходные процессы вычисления ВК в алгоритме Ньютона

Можно показать, что уравнение (12.43) на k -й итерации имеет вид

$$\mathbf{h}_N(k) = \mathbf{h}_{N,o} + (1-2\mu)^k [\mathbf{h}_N(0) - \mathbf{h}_{N,o}]. \quad (12.44)$$

Чтобы при $k \rightarrow \infty$ значение $\mathbf{h}_N(k) \rightarrow \mathbf{h}_{N,o}$, необходимо, чтобы $(1-2\mu)^k \rightarrow 0$, для чего должно выполняться условие $|1-2\mu| < 1$, откуда допустимые границы параметра μ определяются как

$$0 < \mu < 1. \quad (12.45)$$

12.10. Постоянные времени алгоритма Ньютона

Скорость движения алгоритма Ньютона в направлении $\mathbf{h}_N(k) \rightarrow \mathbf{h}_{N,o}$ можно оценить с помощью постоянной времени переходного процесса вычисления элементов вектора $\mathbf{h}_N(k)$. Из уравнения (12.44) следует, что этот переходной процесс определяется множителем $(1-2\mu)^k$ во втором слагаемом. Так как данный процесс носит затухающий характер, то его можно трактовать как экспоненциально затухающий, т. е.

$(1-2\mu)^k [\mathbf{h}_N(0) - \mathbf{h}_{N,o}] = C e^{-\frac{k}{k_{\tau_h}}}$, где k_{τ_h} – постоянная времени переходного процесса в терминах числа отсчетов обрабатываемых сигналов, а $C = [\mathbf{h}_N(0) - \mathbf{h}_{N,o}] = \text{const}$.

Постоянная времени уравнения (12.44) k_{τ_h} определяется как

$$k_{\tau_h} = 1/(2\mu). \quad (12.46)$$

Постоянную времени переходного процесса обучающей кривой $\xi(k)$ алгоритма Ньютона можно оценить следующим образом. Из (12.44) определим

$$\Delta \mathbf{h}_N(k) = \mathbf{h}_N(k) - \mathbf{h}_{N,o} = (1 - 2\mu)^k [\mathbf{h}_N(0) - \mathbf{h}_{N,o}] = (1 - 2\mu)^k \Delta \mathbf{h}_N(0) \quad . \quad (12.47)$$

и затем используем в уравнении MSE-функции (12.37). Тогда получим, что

$$\xi(k) = \xi_{\min} + \Delta \mathbf{h}_N^T(k) \mathbf{R}_N \Delta \mathbf{h}_N(k) = \xi_{\min} + (1 - 2\mu)^{2k} \Delta \mathbf{h}_N^T(0) \mathbf{R}_N \Delta \mathbf{h}_N(0). \quad (12.48)$$

Из (12.48) видно, что переходный процесс обучающей кривой $\xi(k)$ определяется множителем $(1 - 2\mu)^{2k}$. Поэтому постоянная времени обучающей кривой k_{τ_ξ} определяется как

$$k_{\tau_\xi} = 1/(4\mu). \quad (12.49)$$

Таким образом, из (12.45) следует, что в алгоритме Ньютона значение шага сходимости, гарантирующее устойчивость этого алгоритма, не зависит от собственных чисел корреляционной матрицы \mathbf{R}_N обрабатываемых сигналов. Постоянные времени (12.46) и (12.49) также не зависят от собственных чисел корреляционной матрицы.

12.11. Алгоритм наискорейшего спуска

Другим способом поиска вектора ВК адаптивного линейного сумматора, соответствующего минимуму MSE-функции, является алгоритм наискорейшего спуска, который также представляет собой итерационную процедуру. Данный алгоритм на каждой итерации изменяет вектор ВК в направлении, противоположном градиенту этой функции в точке многомерного пространства, определяемой текущими значениями ВК. Другими словами, движение к минимуму MSE-функции осуществляется в направлении, обратном максимальной скорости изменения целевой функции, что и определило его название, алгоритм наискорейшего спуска.

Вычисление ВК в алгоритме наискорейшего спуска осуществляется в соответствии с уравнением

$$\mathbf{h}_N(k) = \mathbf{h}_N(k-1) - \mu \nabla_{\mathbf{h}_N} \xi(k-1). \quad (12.50)$$

Используя уравнение градиента MSE-функции (12.28), (12.50) преобразуется как

$$\mathbf{h}_N(k) = \mathbf{h}_N(k-1) + 2\mu \mathbf{r}_N - 2\mu \mathbf{R}_N \mathbf{h}_N(k-1). \quad (12.51)$$

Для анализа условий стабильности алгоритма наискорейшего спуска, уравнение (12.51) невозможно привести к виду, подобному уравнению (12.44), так как матрица \mathbf{R}_N не является диагональной. В то же время этот анализ можно выполнить, если уравнение (12.51) выразить в терминах ВК (12.39) $\mathbf{v}_N(k) = \mathbf{Q}_N^T \Delta \mathbf{h}_N(k)$ в повернутой и смещенной системе координат, т. е. как

$$\mathbf{v}_N(k) = \left[\mathbf{I}_N - 2\mu \mathbf{\Lambda}_N \right] \mathbf{v}_N(k-1). \quad (12.52)$$

Уравнение (12.51) также можно привести к виду

$$\begin{aligned} \mathbf{v}_N(k) &= (\mathbf{I}_N - 2\mu \mathbf{\Lambda}_N)^k \mathbf{v}_N(0) = \\ &= \begin{bmatrix} (1-2\mu\lambda_1)^k & 0 & \cdots & 0 \\ 0 & (1-2\mu\lambda_2)^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (1-2\mu\lambda_N)^k \end{bmatrix} \mathbf{v}_N(0). \end{aligned} \quad (12.53)$$

Из (12.53) следует, что для того чтобы при $k \rightarrow \infty$ значение $\mathbf{v}_N(k) \rightarrow \mathbf{0}_N$ при $\mathbf{v}_N(0) \neq \mathbf{0}_N$ или, что тоже самое, $\mathbf{h}_N(k) \rightarrow \mathbf{h}_{N,o}$, необходимо, чтобы $(1-2\mu\lambda_n)^k \rightarrow 0$ для всех $n=1,2,\dots,N$, для чего должно выполняться условие $|1-2\mu\lambda_n| < 1$. Общее условие для всех значений λ_n можно сформулировать как

$$0 < \mu < 1/\lambda_{\max}. \quad (12.54)$$

12.12. Постоянные времени алгоритма наискорейшего спуска

Скорость движения алгоритма наискорейшего спуска в направлении $\mathbf{h}_N(k) \rightarrow \mathbf{h}_{N,o}$ можно оценить с помощью постоянных времени переходных процессов вычисления элементов вектора $\mathbf{v}_N(k)$ в уравнении (12.52). Эта постоянная времени для ВК v_n определяется как

$$k_{\tau_{n,v}} = 1/(2\mu\lambda) \quad (12.55)$$

при $\mu \ll 1$.

Переход к исходной системе координат можно осуществить следующим образом. Покажем, что

$$\mathbf{h}_N(k) = \mathbf{h}_{N,o} + \sum_{m=1}^N \mathbf{q}_m v_m(k). \quad (12.56)$$

Так как, согласно (12.52), $v_m(k) = (1 - 2\mu\lambda_m)^k v_m(0)$, то

$$h_n(k) = h_{o,n} + \sum_{m=1}^N q_{m,n} (1 - 2\mu\lambda_m)^k v_m(0), \quad (12.57)$$

где $q_{m,n}$ – это n -й элемент m -го собственного вектора \mathbf{q}_m корреляционной матрицы \mathbf{R}_N . Из уравнения (12.57) следует, что каждый ВК $h_n(k)$ в алгоритме наискорейшего спуска сходится к соответствующему ВК винеровского решения $h_{o,n}(k)$

пропорционально $\sum_{m=1}^N (1 - 2\mu\lambda_m)^k$. Определить из-за этой суммы

постоянную времени $k_{\tau_{n,h}}$ для ВК $h_n(k)$ не представляется возможным. Тем не менее, для корреляционной матрицы \mathbf{R}_N с разными собственными числами и $\mu = \text{const}$ сходимость уравнения (12.57) будет определяться значением λ_{\min} , так как слагаемое $(1 - 2\mu\lambda_{\min})^k \rightarrow 0$ медленнее всех других аналогичных

слагаемых, содержащих $\lambda_n \neq \lambda_{\min}$ в этом уравнении. Другими словами, согласно (12.55) и (12.57), значение каждой из постоянных времени $k_{\tau_{n,h}}$ находится в пределах

$$\frac{1}{2\mu\lambda_{\max}} \leq k_{\tau_{n,h}} \leq \frac{1}{2\mu\lambda_{\min}}. \quad (12.58)$$

Таким образом, если корреляционная матрица \mathbf{R}_N имеет разные, сильно различающиеся собственные числа, т. е. $\chi = \lambda_{\max}/\lambda_{\min} \gg 1$, то длительность переходного процесса вычисления ВК адаптивного линейного сумматора, использующего алгоритм наискорейшего спуска, будет определяться наименьшим собственным числом.

Переходный процесс в терминах обучающей кривой $\xi(k)$ можно оценить следующим образом. Если в уравнении (12.38) использовать уравнение (12.52), то

$$\xi(k) = \xi_{\min} + \sum_{n=1}^N (1 - 2\mu\lambda_n)^{2k} \lambda_n v_n^2(0). \quad (12.59)$$

Поэтому аналогично уравнению (12.58) можно определить, что значение

$$\frac{1}{4\mu\lambda_{\max}} \leq k_{\tau_{n,\xi}} \leq \frac{1}{4\mu\lambda_{\min}}. \quad (12.60)$$

Таким образом, согласно (12.58) и (12.60), сходимость ВК и MSE-функции к оптимальным значениям определяется минимальным собственным числом корреляционной матрицы входных сигналов \mathbf{R}_N адаптивного линейного сумматора, а также зависит от шага сходимости μ . Чем меньше μ , тем больше постоянные времени переходных процессов и тем медленнее происходит движение вектора ВК в сторону оптимального винеровского решения.

Результаты моделирования алгоритма наискорейшего спуска в условиях, аналогичных алгоритму Ньютона (см. подраздел 12.9), представлены на рис. 12.14 – 12.16.

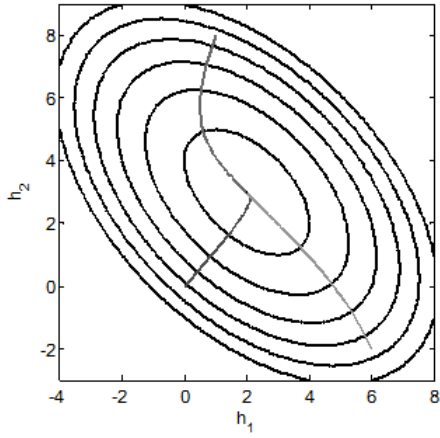


Рис. 12.14. Нахождение винеровского решения с помощью алгоритма наискорейшего спуска

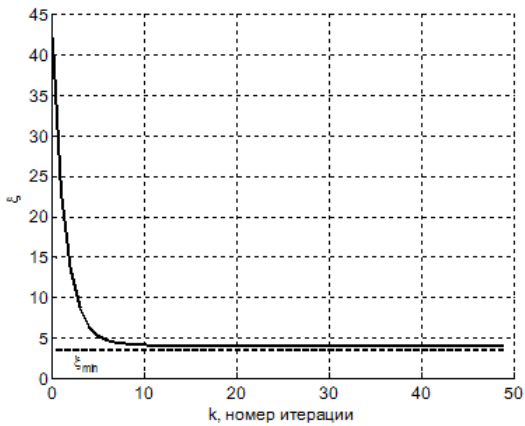


Рис. 12.15. Обучающая кривая алгоритма наискорейшего спуска:

$$\mathbf{h}_N(0) = [0, 0]^T$$

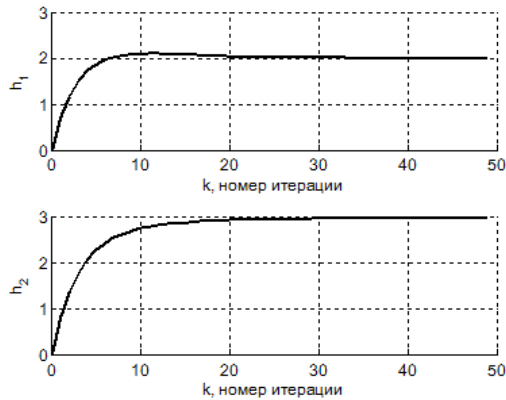


Рис. 12.16. Переходные процессы вычисления ВК в алгоритме наискорейшего спуска: $\mathbf{h}_N(0) = [0, 0]^T$

12.13. LMS-алгоритм

Алгоритмы Ньютона и наискорейшего спуска являются прообразами алгоритмов адаптивной фильтрации. На практике обычно используются упрощенные версии этих алгоритмов, например LMS-алгоритм, принадлежащий к классу алгоритмов стохастического градиентного поиска. Термин «стохастический» используется для того, чтобы отличать LMS-алгоритм с упрощенным градиентом, от алгоритма наискорейшего спуска, использующего точное значение градиента целевой MSE-функции. Упрощение градиента является следствием упрощения целевой функции. Поскольку упрощенный градиент является не точным, он подвержен случайным флуктуациям. Поэтому движение вектора ВК к оптимальному решению носит в некотором смысле случайный характер, от чего и используется термин «стохастический». При обработке стационарных сигналов и соответствующем выборе шага LMS-алгоритм сходится к винеровскому решению.

Если на k -м временном отсчете имеются оценки матрицы \mathbf{R}_N и вектора \mathbf{r}_N , полученные на окне всего в один отсчет

обрабатываемых сигналов, т. е. $\hat{\mathbf{R}}_N(k) = \mathbf{x}_N(k)\mathbf{x}_N^T(k)$ и $\hat{\mathbf{r}}_N = \mathbf{x}_N(k)d(k)$, то градиент $\hat{\nabla}_{\mathbf{h}_N} \xi(k-1)$ (12.28) можно выразить как

$$\hat{\nabla}_{\mathbf{h}_N} \xi(k-1) = -2\mathbf{x}_N(k)\alpha(k), \quad (12.61)$$

где $\alpha(k)$ – априорная ошибка моделирования (оценивания) сигнала $d(k)$ (см. уравнение (12.2)).

Используя уравнения (12.61) в уравнении (12.50), получим, что вектор ВК можно определять как

$$\mathbf{h}_N(k) = \mathbf{h}_N(k-1) + 2\mu \mathbf{x}_N(k)\alpha(k). \quad (12.62)$$

Уравнение (12.62) – это LMS-алгоритм поиска ВК адаптивного фильтра. Название Least Mean Square происходит от того, что алгоритм минимизирует средний квадрат мгновенной ошибки, т. е. ошибки, определяемой по одному отсчету обрабатываемых сигналов.

Поскольку в уравнении (12.62) используются не точные оценки градиента, то в адаптивном процессе возникает шум, в результате чего траектория движения вектора $\mathbf{h}_N(k)$ к вектору $\mathbf{h}_{N,o}$ не совпадает с траекторией движения в алгоритме наискорейшего спуска (12.50). Этот шум уменьшается по мере движения к стационарной точке $\mathbf{h}_{N,o}$, так как приращения вектора

ВК, равные $\mu \left[-\hat{\nabla}_{\mathbf{h}_N} \xi(k-1) \right]$, вблизи точки $\mathbf{h}_{N,o}$ уменьшаются по сравнению с приращениями в начальной точке $\mathbf{h}_N(0) \neq \mathbf{h}_{N,o}$. Это происходит из-за того, что значения градиента вблизи стационарной точки $\mathbf{h}_{N,o}$ стремятся к нулю, т. е. $\nabla_{\mathbf{h}_N} \xi(k-1) \rightarrow 0$, а значит, и значения оценок этого градиента также стремятся к нулю, т. е. $\hat{\nabla}_{\mathbf{h}_N} \xi(k-1) \rightarrow 0$.

LMS-алгоритм является самым простым среди алгоритмов адаптивной фильтрации с точки зрения вычислительной и

алгоритмической сложности. Оценка вычислительной сложности этого алгоритма равна $2N$ умножений и $2N$ сложений на одну итерацию.

12.14. Переходные процессы в LMS-алгоритме

Основной характеристикой всех алгоритмов адаптивной фильтрации является сходимость вектора ВК к оптимальному значению $\mathbf{h}_{N,o}$, при котором MSE-функция достигает своего минимума. ВК LMS-алгоритма, несмотря на его простоту, также в среднем сходятся к $\mathbf{h}_{N,o}$.

В LMS-алгоритме оценка градиента $\hat{\nabla}_{\mathbf{h}_N} \xi(k-1)$ при постоянном значении вектора ВК является несмещенной, т. е.

$$E \left\{ \hat{\nabla}_{\mathbf{h}_N} \xi(k-1) \right\} = \nabla_{\mathbf{h}_N} \xi . \quad (12.63)$$

Согласно (12.51), уравнение вычисления вектора ВК с помощью алгоритма наискорейшего спуска представимо в виде

$$\mathbf{h}_N(k) = [\mathbf{I}_N - 2\mu \mathbf{R}_N] \mathbf{h}_N(k-1) + 2\mu \mathbf{r}_N . \quad (12.64)$$

Можно показать, что в среднем уравнение (12.62) сходится к уравнению (12.64), т. е.

$$E \left\{ \mathbf{h}_N(k) \right\} = [\mathbf{I}_N - 2\mu \mathbf{R}_N] E \left\{ \mathbf{h}_N(k-1) \right\} + 2\mu \mathbf{r}_N . \quad (12.65)$$

Таким образом, среднее значение вектора ВК (12.65) в LMS-алгоритме совпадает со значением вектора ВК (12.64) в алгоритме наискорейшего спуска при $0 < \mu < 1/\lambda_{\max}$ (см. (12.54)). Поэтому в среднем ВК LMS-алгоритма сходится к винеровскому решению, если выполняется условие

$$0 < \mu < 1/\lambda_{\max} . \quad (12.66)$$

На практике при использовании АФ определять собственные числа корреляционной матрицы обычно не представляется

возможным. Однако из свойства шести собственных чисел корреляционной матрицы (12.24) $\text{tr}\{\mathbf{R}_N\} = \sum_{n=1}^N R_{nn} = \sum_{n=1}^N \lambda_n$

следует, что $\lambda_{\max} \leq \text{tr}\{\mathbf{R}_N\} = \sum_{n=1}^N R_{nn}$. С учетом этого соотношения, (12.66) можно представить в виде несколько более жесткого условия устойчивости LMS-алгоритма:

$$0 < \mu < \frac{1}{\sum_{n=1}^N R_{nn}}. \quad (12.67)$$

Постоянная времени $k_{\tau_{i,h}}$ среднего значения действительных ВК LMS-алгоритма находится в пределах

$$\frac{1}{2\mu\lambda_{\max}} \leq k_{\tau_{n,h}} \leq \frac{1}{2\mu\lambda_{\min}}, \quad (12.68)$$

т. е. совпадет с аналогичными значениями для алгоритма наискорейшего спуска.

Следовательно, постоянные времени переходных процессов среднего значения ВК в LMS-алгоритме совпадают с соответствующими постоянными времени алгоритма наискорейшего спуска.

На рис. 12.17 показаны результаты моделирования алгоритма наискорейшего спуска и LMS-алгоритма при решении задачи линейного предсказания авторегрессионного процесса при $\chi = 5$. Темной линией показан переходной процесс в терминах ВК алгоритма наискорейшего спуска, а светлой – LMS-алгоритма при двух значениях параметра μ , удовлетворяющих условиям (12.54) и (12.66).

Постоянная времени обучающей кривой LMS-алгоритма $\xi(k) = E\{|\alpha(k)|^2\}$ определяется как

$$\frac{1}{4\mu\lambda_{\max}} \leq k_{\tau_{n,\xi}} \leq \frac{1}{4\mu\lambda_{\min}}. \quad (12.69)$$

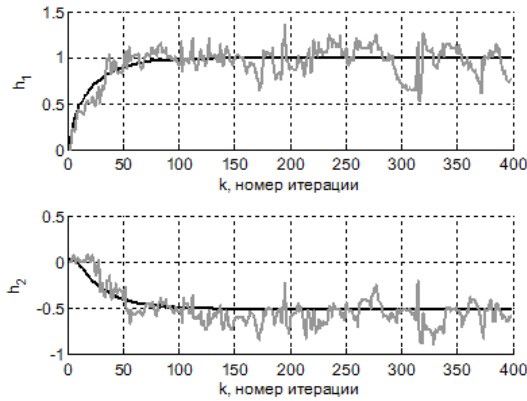


Рис. 12.17. Переходные процессы вычисления ВК при $\mu = 0.025$
(одна реализация)

Таким образом, в среднем ВК LMS-алгоритма сходятся к винеровскому решению $\mathbf{h}_{N,o}$. Однако, даже находясь вблизи $\mathbf{h}_{N,o}$, ВК на каждой итерации отличаются от этого решения как

$$\Delta \mathbf{h}_N(k) = \mathbf{h}_N(k) - \mathbf{h}_{N,o}, \quad (12.70)$$

что обусловлено грубой оценкой вектора градиента (12.61), содержащей шумовую составляющую. Это, в свою очередь, обуславливает наличие дополнительной среднеквадратической ошибки ξ_{exc} , т. е.

$$\xi = \xi_{\min} + \xi_{exc}. \quad (12.71)$$

12.15. Качество адаптивной фильтрации с помощью LMS-алгоритма

Одним из параметров, используемых для оценки качества адаптивной фильтрации с помощью LMS-алгоритма, является относительное значение дополнительной среднеквадратической ошибки (misadjustment):

$$M = \frac{\xi_{exc}}{\xi_{\min}} = \frac{\xi - \xi_{\min}}{\xi_{\min}}. \quad (12.72)$$

На практике шаг сходимости μ обычно выбирается так, чтобы значение (12.72) не превышало 10%, т. е. $M \leq 0.1$. Это значение достигается, если $\mu \ll 1$. В этом случае уравнение (12.72) определяется как

$$M \approx \frac{\mu \sum_{n=1}^N \lambda_n}{1 - \mu \sum_{n=1}^N \lambda_n} = \frac{\mu \operatorname{tr}\{\mathbf{R}_N\}}{1 - \mu \operatorname{tr}\{\mathbf{R}_N\}}. \quad (12.73)$$

Ранее уже отмечалось, что условие сходимости в среднем вектора ВК адаптивного фильтра $\mathbf{h}_N(k)$ на базе LMS-алгоритма к винеровскому решению $\mathbf{h}_{N,o}$ определяется уравнением (12.66). Однако в силу того что оценка вектора градиента (12.61) $\hat{\nabla}_{\mathbf{h}_N} \xi(k-1) = -2\mathbf{x}_N(k)\alpha(k)$ подвержена случайным флуктуациям, верхняя граница шага сходимости μ , при которой гарантируется устойчивое поведение LMS-алгоритма, должна быть несколько меньше. Эта граница определяется как

$$\mu_{\max} = \frac{1}{3 \sum_{n=1}^N \lambda_n} = \frac{1}{3 \operatorname{tr}\{\mathbf{R}_N\}}. \quad (12.74)$$

12.16. NLMS-алгоритм

LMS-алгоритм используется в приложениях, где можно априори оценить предельно допустимое значение шага сходимости μ , гарантирующее стабильность алгоритма. Как следует из (12.74), такая оценка возможна, если известны статистические характеристики входного сигнала $x(k)$. В приложениях, где эти характеристики неизвестны или меняются со временем (например, при обработке нестационарных сигналов), часто используется нормализованный LMS-алгоритм (Normalized LMS, NLMS), являющийся одной из разновидностей LMS-алгоритма.

Существует несколько подходов к получению NLMS-алгоритма, один из которых заключается в том, что NLMS-алгоритм можно рассматривать и как упрощение алгоритма Ньютона. В результате этого упрощения может быть получено уравнение вычисления ВК адаптивного фильтра:

$$\mathbf{h}_N(k) = \mathbf{h}_N(k-1) + \frac{2\hat{\mu}}{\mathbf{x}_N^T(k)\mathbf{x}_N(k) + \delta^2} \mathbf{x}_N(k)\alpha(k), \quad (12.75)$$

где, согласно (12.45), $0 < \hat{\mu} < 1$ – масштабирующий множитель динамического шага сходимости $[\mathbf{x}_N^T(k)\mathbf{x}_N(k) + \delta^2]^{-1}$.

Таким образом, из (12.75) следует, что градиент в приращении вектора ВК в NLMS-алгоритме нормализуется к значению $\mathbf{x}_N^T(k)\mathbf{x}_N(k)$, что послужило источником слова «нормализованный» в названии этого алгоритма. Оценка вычислительной сложности алгоритма (12.75), включая операцию вычисления $\alpha(k)$ (см. (12.2)), равна $3N$ умножениям, $3N$ сложениям и одному делению на одну итерацию.

В случае обработки нестационарных сигналов вектор $\mathbf{x}_N(k)$ может сильно варьироваться от итерации к итерации, т. е. может меняться норма этого вектора:

$$\|\mathbf{x}_N(k)\|_2 = \sqrt{\sum_{n=1}^N x^2(k-n+1)},$$

а значит, и градиент $-2\mathbf{x}_N(k)\alpha(k)$ и приращение вектора ВК:

$$\Delta\mathbf{h}_N(k) = -2\mu \mathbf{x}_N(k)\alpha(k) \quad (12.76)$$

в LMS-алгоритме (12.62).

В NLMS-алгоритме это приращение определяется как

$$\Delta\mathbf{h}_N(k) = \frac{2\hat{\mu}}{\mathbf{x}_N^T(k)\mathbf{x}_N(k) + \delta^2} \mathbf{x}_N(k)\alpha(k) = \frac{2\hat{\mu}}{\|\mathbf{x}_N(k)\|_2^2 + \delta^2} \mathbf{x}_N(k)\alpha(k), \quad (12.77)$$

т. е. оно обратно пропорционально $\|\mathbf{x}_N(k)\|_2^2$. Нормировка градиента «сглаживает» его резкие перепады из-за изменения $\mathbf{x}_N(k)$. Поэтому при фиксированных значениях параметров μ и $\hat{\mu}$ поведение NLMS-алгоритма при обработке нестационарных сигналов является более устойчивым, чем поведение LMS-алгоритма.

К более функционально эффективным алгоритмам адаптивной фильтрации относится также RLS-алгоритм.

12.17. Задача наименьших квадратов и ее решение

Адаптивная RLS-фильтрация сигналов базируется на минимизации квадратичной целевой функции. Для получения решения этой задачи используется обобщенный метод наименьших квадратов – LS. Фильтрация по методу наименьших квадратов является детерминированным аналогом винеровской фильтрации. Последняя базируется на усреднении по ансамблю реализаций стационарных в широком смысле стохастических процессов, в то время как LS-фильтрация использует усреднение по времени, в результате чего находятся ВК оптимального фильтра, которые различаются для каждой реализации стохастического процесса.

Ранее было показано, что адаптивные алгоритмы могут быть получены путем реализации стратегии градиентного поиска минимума целевой функции. Достоинством таких алгоритмов является их простота, недостатком – медленная и неравномерная сходимость вектора $\mathbf{h}_N(k)$ к оптимальному вектору $\mathbf{h}_{N,o}$, а также то, что даже по мере приближения к оптимальному решению вектор $\mathbf{h}_N(k)$ обычно хаотично движется в окрестности этого решения, а не сходится к нему.

Рассматриваемый здесь способ поиска вектора $\mathbf{h}_{N,o}$ базируется на другом подходе. Он заключается в том, что решение обеспечивается не только для текущего отсчета k , но и является оптимальным и для всех предыдущих отсчетов.

Оценка вектора $\mathbf{h}_{N,o}$, т. е. вектора \mathbf{h}_N , находится путем решения системы линейных уравнений

$$\mathbf{h}_N^T \mathbf{x}_N(i) = d(i), \quad (12.78)$$

где $i = N, N+1, \dots, k$. При $k \geq 2N$ выражение (12.78) представляет собой переопределенную систему линейных уравнений, так как число уравнений $k - N + 1$ больше числа неизвестных N . Решение системы уравнений (12.78) может быть получено в среднеквадратическом приближении оценок

$$y(i) = \mathbf{h}_N^T \mathbf{x}(i) \quad (12.79)$$

к значениям сигнала $d(i)$, т. е. путем решения задачи минимизации среднеквадратического значения ошибок:

$$e(i) = d(i) - y(i). \quad (12.80)$$

Переменные (12.79) и (12.80) можно интерпретировать как сигналы КИХ-фильтра (рис. 12.18). Поэтому задачу нахождения вектора $\mathbf{h}_{N,o}$ допустимо сформулировать как задачу нахождения вектора ВК \mathbf{h}_N фильтра (рис. 12.18), при которых его выходной сигнал $y(i)$ наилучшим образом воспроизводит, т. е. моделирует, сигнал $d(i)$. В этой постановке предполагается, что число ВК N модели линейной системы известно, и оно совпадает с числом ВК фильтра.

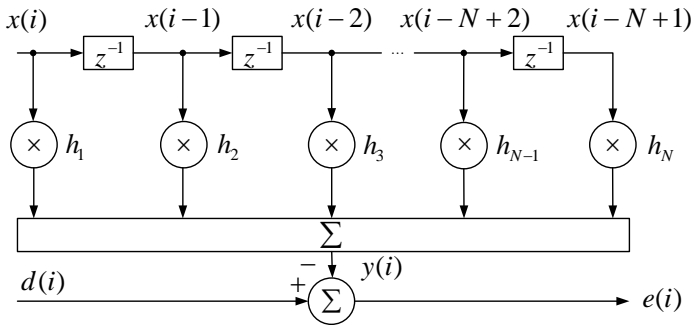


Рис. 12.18. Решение LS-задачи с помощью линейного КИХ-фильтра

12.18. Рекурсивная задача наименьших квадратов

В п. 12.2 отмечалось, что адаптивные алгоритмы, как правило, носят рекурсивный характер. Рекурсивное решение задачи наименьших квадратов, т. е. поиска вектора ВК $\mathbf{h}_N(k)$, можно найти путем минимизации целевой функции:

$$E_N(k) = \sum_{i=1}^k e_N^2(i). \quad (12.81)$$

Обрабатываемые сигналы в целевой функции (12.81) часто взвешиваются с помощью параметра λ (см. рис. 12.19), т. е. как

$$E_N(k) = \sum_{i=1}^k \lambda^{k-i} e_N^2(i). \quad (12.82)$$

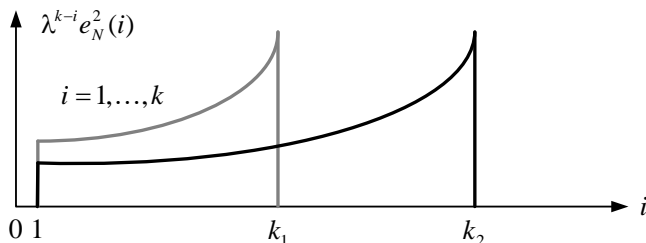


Рис. 12.19. Экспоненциальное взвешивание обрабатываемых сигналов

Здесь и далее параметр λ не следует путать с собственными числами корреляционной матрицы λ_n , где $n = 1, 2, \dots, N$.

Взвешивание сигналов с помощью параметра λ позволяет учитывать небольшие изменения в статистике сигналов, если эти сигналы нестационарные, т. е. их статистические параметры меняются во времени. Параметр λ называется параметром экспоненциального взвешивания, или параметром «забывания» («forgetting factor»). Его значение обычно ограничено как $(1 - 0.4/N) \leq \lambda \leq 1$.

12.19. Решение рекурсивной задачи наименьших квадратов

Целевую функцию (12.82) можно представить как

$$\begin{aligned} E_N(k) &= \sum_{i=1}^k \lambda^{k-i} e_N^2(i) = \\ &= E_d(k) - \mathbf{r}_N^T(k) \mathbf{h}_N(k) - \mathbf{h}_N^T(k) \mathbf{r}_N + \mathbf{h}_N^T(k) \mathbf{R}_N(k) \mathbf{h}_N(k), \end{aligned} \quad (12.83)$$

где

$$E_d(k) = \sum_{i=1}^k \lambda^{k-1} d^2(i) = \lambda E_d(k-1) + d^2(k), \quad (12.84)$$

$$\mathbf{r}_N(k) = \sum_{i=1}^k \lambda^{k-1} \mathbf{x}_N(i) d(i) = \lambda \mathbf{r}_N(k-1) + \mathbf{x}_N(k) d(k), \quad (12.85)$$

$$\mathbf{R}_N(k) = \sum_{i=1}^k \lambda^{k-1} \mathbf{x}_N(i) \mathbf{x}_N^T(i) = \lambda \mathbf{R}_N(k-1) + \mathbf{x}_N(k) \mathbf{x}_N^T(k). \quad (12.86)$$

Выражения (12.85) и (12.86) будем именовать «вектором взаимной корреляции» и «корреляционной матрицей» соответственно.

Значение функции (12.83) минимально при

$$\nabla_{\mathbf{h}_N} E_N(k) = \frac{\partial E_N(k)}{\partial \mathbf{h}_N(k)} = -2\mathbf{r}_N(k) + 2\mathbf{R}_N(k) \mathbf{h}_N(k) = \mathbf{0}_N, \quad (12.87)$$

откуда

$$\mathbf{h}_N(k) = \mathbf{R}_N^{-1}(k) \mathbf{r}_N(k). \quad (12.88)$$

Рекурсивное решение LS-задачи минимизации функций (12.83), т. е. Recursive Least Squares (RLS) осуществляется следующим образом. Используя (12.85) и (12.86), можно для

каждого k решать LS-задачу, т. е. вычислять $\mathbf{h}_N(k) = \mathbf{R}_N^{-1}(k)\mathbf{r}_N(k)$, не вычисляя $\mathbf{R}_N^{-1}(k)$ каждый раз заново.

Для этого, используя уравнение (12.86), матрица $\mathbf{R}_N^{-1}(k)$ представима в виде:

$$\mathbf{R}_N^{-1}(k) = \lambda^{-1} \left[\mathbf{R}_N^{-1}(k-1) - \mathbf{g}_N(k)\mathbf{x}_N^H(k)\mathbf{R}_N^{-1}(k-1) \right], \quad (12.89)$$

где

$$\mathbf{g}_N(k) = \mathbf{R}_N^{-1}(k)\mathbf{x}_N(k) = \frac{\mathbf{R}_N^{-1}(k-1)\mathbf{x}_N(k)}{\lambda + \mathbf{x}_N^H(k)\mathbf{R}_N^{-1}(k-1)\mathbf{x}_N(k)} \quad (12.90)$$

– вектор коэффициентов Калмана.

Используя (12.85), (12.89) и (12.90), вместо уравнения (12.88) можно получить уравнение рекурсивного вычисления вектора $\mathbf{h}_N(k)$. Это уравнение имеет вид

$$\mathbf{h}_N(k) = \mathbf{h}_N(k-1) + \mathbf{g}_N(k)\alpha_N(k), \quad (12.91)$$

где $\alpha_N(k)$ – априорная ошибка оценивания процесса (см. (12.2)).

Таким образом, RLS-алгоритм вычисления вектора ВК адаптивного фильтра, минимизирующего функцию (12.83), определяется как последовательность вычислений в соответствии с уравнениями (12.90), (12.89), (12.2), (12.91), вычисляемыми на каждой итерации k .

Инициализация вектора ВК в RLS-алгоритме осуществляется как $\mathbf{h}_N(0) = \mathbf{0}_N$, так как при $k=0$ переменные $\mathbf{x}_N(0) = \mathbf{0}_N$ и $d(0) = 0$ вектор $\mathbf{r}_N(0) = \mathbf{x}_N(0)d(0) = \mathbf{0}_N$, а значит, и $\mathbf{h}_N(0) = \mathbf{R}_N^{-1}(0)\mathbf{r}_N(0) = \mathbf{0}_N$.

Начальное значение (инициализация) матрицы $\mathbf{R}_N(k)$ может быть задано, например, как $\mathbf{R}_N(0) = \delta^2\mathbf{I}_N$, следовательно, и начальное значение обратной матрицы – как

$$\mathbf{R}_N^{-1}(0) = \delta^{-2}\mathbf{I}_N. \quad (12.92)$$

Переход в RLS-алгоритме от комплексной арифметики к арифметике действительных чисел является тривиальным и сводится к исключению операций комплексного сопряжения переменных в уравнениях алгоритма.

Вычислительная сложность RLS-алгоритма пропорциональна $O(N^2)$ арифметическим операциям. Вычислительная сложность RLS-алгоритма равна $2N^2 + 4N$ операциям умножения, $1.5N^2 + 3.5N$ операциям сложения и одной операции деления.

12.20. Качество адаптивной фильтрации с помощью RLS-алгоритма

RLS-алгоритм характеризуется разным качеством при $\lambda = 1$ и при $\lambda < 1$. Можно показать, что при $\lambda = 1$ и больших k , т. е. в установившемся состоянии, среднеквадратическая ошибка $\xi(k)$ определяется как

$$\xi(k) = \xi_{\min} + \frac{\xi_{\min}}{k} N. \quad (12.93)$$

Поэтому при $k \rightarrow \infty$ значение

$$\xi(\infty) = \xi_{\min}. \quad (12.94)$$

Это означает, что

$$M = \frac{\xi_{exc}}{\xi_{\min}} = \frac{\xi(\infty) - \xi_{\min}}{\xi_{\min}} = \frac{\xi_{\min} - \xi_{\min}}{\xi_{\min}} = 0, \quad (12.95)$$

т. е. при $\lambda = 1$, относительное значение дополнительной СКО $\xi_{exc} = \xi(\infty) - \xi_{\min}$ RLS-алгоритма в установившемся состоянии равно нулю.

При $\lambda < 1$ и больших k

$$\xi(\infty) = \xi_{\min} \left[1 + N \frac{1 - \lambda}{1 + \lambda} \right]. \quad (12.96)$$

Это означает, что

$$M = \frac{\xi_{exc}}{\xi_{min}} = \frac{\xi(\infty) - \xi_{min}}{\xi_{min}} = N \frac{1 - \lambda}{1 + \lambda}. \quad (12.97)$$

Таким образом, в RLS-алгоритме значение $\xi_{exc} = \xi(\infty) - \xi_{min}$ в установившемся режиме и значение M не зависят от разброса собственных чисел корреляционной матрицы входных сигналов АФ. Это отличает RLS-алгоритм от LMS-алгоритма, в котором такая зависимость наблюдается (см. (12.73)).

Параметр экспоненциального взвешивания λ в RLS-алгоритме был введен с целью обеспечения возможности слежения за «медленно меняющейся статистикой» обрабатываемых адаптивным фильтром сигналов. Эффективность использования параметра λ для обработки сигналов с меняющейся статистикой уменьшается по мере роста N , т. к. при $\lambda \rightarrow 1$ увеличивается «память алгоритма». Поэтому, если $\lambda \rightarrow 1$ то, согласно (12.95) и (12.97), $M \rightarrow 0$, что хорошо с точки зрения «нулевой» дополнительной среднеквадратической ошибки ξ_{exc} в установившемся состоянии. Однако таким образом не обеспечивается возможность слежения АФ за изменяющимися условиями функционирования.

Поэтому на практике значения λ выбирают из компромисса, исходя из вышеизложенных свойств RLS-алгоритма, а также его устойчивости, которая гарантируется, если параметр λ удовлетворяет соотношению $(1 - 0.4/N) \leq \lambda \leq 1$, зависящему от числа ВК адаптивного фильтра N . При значениях λ , меньших указанной границы, RLS-алгоритм и его разновидности, как правило, являются неустойчивыми.

12.21. Заключительные замечания по адаптивной обработке сигналов

В настоящее время для вычисления весовых коэффициентов адаптивных фильтров в основном используются LMS-алгоритм, его нормализованная версия – NLMS-алгоритм, а также RLS-алгоритм.

Градиентные LMS-, NLMS-алгоритмы и различные их модификации являются очень популярными в силу их низкой

вычислительной сложности (наименьшей среди всех алгоритмов адаптивной фильтрации) и простоты математических выражений, описывающих эти алгоритмы.

Кроме рассмотренного RLS-алгоритма, существуют разновидности RLS-алгоритмов адаптивной фильтрации на базе QR-разложения матрицы входных сигналов АФ. Эти алгоритмы характеризуются высокой устойчивостью. Кроме того, QR RLS-алгоритмы также обладают квадратичной вычислительной сложностью.

Вычислительная сложность уменьшается до линейной в так называемых быстрых (вычислительно эффективных) RLS-алгоритмах, к которым относятся и так называемые лестничные (часто встречается и название «решетчатые») адаптивные алгоритмы, получившие свое название из-за структуры вычислений переменных, напоминающей лестницу (решетку).

Промежуточным классом между LMS- и RLS-алгоритмами являются алгоритмы аффинных проекций, представляющие собой разновидность блочного NLMS-алгоритма. При обработке нестационарных сигналов алгоритм аффинных проекций демонстрирует эффективность (в терминах длительности переходного процесса), близкую к RLS-алгоритмам. Быстрая версия алгоритма аффинных проекций, в случае использования в АФ с большим числом ВК, характеризуется линейной вычислительной сложностью, близкой к сложности NLMS-алгоритма.

Известно, что если вместо КИХ-фильтров с постоянными ВК использовать БИХ-фильтры, то число ВК можно существенно уменьшить, что заметно снижает требования к вычислительным ресурсам при реализации БИХ-фильтров. Это в равной мере относится и к адаптивным БИХ-фильтрам, которые сегодня не нашли ни широкого практического использования, ни большого теоретического интереса, что обусловлено не столько их потенциальной неустойчивостью, сколько многоэкстремальным характером целевой функции в таких фильтрах, не гарантирующим нахождения глобального минимума этой функции.

Таким образом, адаптивная обработка сигналов как современное научно-техническое направление и дисциплина в ЦОС, радиотехнике и связи гораздо шире, чем круг вопросов, рассмотренных выше.

12.22. Контрольные вопросы

1. Когда требуется использовать адаптивные фильтры?
2. В чем состоят два принципиальных отличия между адаптивным и неадаптивным фильтрами?
3. Чему равна длительность итерации вычисления весовых коэффициентов в АФ?
4. Какие основные требования предъявляются к АФ?
5. Какими по структуре бывают АФ?
6. Чему равен выходной сигнал АФ?
7. В чем состоит различие между априорными и апостериорными ошибками моделирования требуемого сигнала АФ?
8. Что такое адаптивный алгоритм?
9. Что такое целевая функция АФ?
10. Какие целевые функции находятся в основе работы АФ?
11. Что такое адаптивная идентификация неизвестной системы?
12. Каким образом формируется требуемый сигнал в задаче идентификации неизвестной системы?
13. Что такое межсимвольная интерференция?
14. Как решается задача адаптивного выравнивания характеристик электрических каналов связи?
15. Чем объясняется способность адаптивной антенной решетки подавлять сигналы источников внешних помех?
16. Какой сигнал образуется на выходе АФ в задаче шумоочистки сигналов?
17. Какой сигнал образуется на выходе сигнала ошибки в задаче шумоочистки сигналов?
18. На каком свойстве основана возможность адаптивной фильтрации узкополосных сигналов?

19. Какой сигнал образуется на выходе АФ в задаче адаптивной фильтрации узкополосных сигналов?
20. Какой сигнал образуется на выходе сигнала ошибки в задаче адаптивной фильтрации узкополосных сигналов?
21. Что такое линейный сумматор?
22. Что такое корреляционная матрица?
23. Что такое положительная определенность или полуопределенность корреляционной матрицы?
24. Что такое характеристическое уравнение?
25. Что такое собственные числа и собственные векторы корреляционной матрицы?
26. Сколько собственных чисел имеет корреляционная матрица?
27. Чему равны собственные числа корреляционной матрицы, возведенной в степень положительного целого числа?
28. Что такое линейная независимость собственных векторов корреляционной матрицы?
29. Что означает ортогональность собственных векторов корреляционной матрицы?
30. Чему равна диагональная матрица собственных чисел корреляционной матрицы?
31. Чему равно разложение корреляционной матрицы по собственным векторам и собственным числам?
32. Чему равно разложение обратной корреляционной матрицы по собственным векторам и собственным числам?
33. Чему равен след корреляционной матрицы?
34. Чему равен определитель корреляционной матрицы?
35. Через какие переменные определяется функция среднеквадратической ошибки между требуемым и выходным сигналами фильтра линейного сумматора?
36. Чему равен градиент функции среднеквадратической ошибки?
37. Как называется и чему равен вектор весовых коэффициентов линейного сумматора в точке многомерного пространства, соответствующей минимуму функции среднеквадратической ошибки?
38. Что такое фильтр Винера?
39. Чему равно минимальное значение функции среднеквадратической ошибки?

40. Что такое принцип ортогональности сигналов в фильтре Винера?
41. Что собой представляет поверхность функции СКО?
42. Что такое система смещенных координат?
43. Что такое система повернутых и смещенных координат?
44. Что такое линии уровня поверхности функции СКО?
45. Какая форма у линий уровня поверхности функции СКО?
46. От чего и как зависит форма линий уровня поверхности функции СКО?
47. Что такое алгоритм Ньютона?
48. Как выглядит траектория движения весовых коэффициентов АФ, вычисляемых с помощью алгоритма Ньютона, на графике линий уровня функции поверхности СКО?
49. Что такое обучающая кривая?
50. Что такое переходный процесс обучающей кривой?
51. Что такое установившееся значение обучающей кривой?
52. Чему равно предельное значение шага сходимости алгоритма Ньютона?
53. Чему равна постоянная времени алгоритма Ньютона в терминах ВК?
54. Чему равна постоянная времени алгоритма Ньютона в терминах СКО?
55. Что такое алгоритм наискорейшего спуска?
56. Как выглядит траектория движения весовых коэффициентов АФ, вычисляемых с помощью алгоритма наискорейшего спуска, на графике линий уровня функции поверхности СКО?
57. Чему равны предельные значения шага сходимости алгоритма наискорейшего спуска?
58. Как выглядит траектория движения ВК алгоритма наискорейшего спуска на графике линий уровня функции поверхности среднеквадратической ошибки?
59. Чему равна постоянная времени алгоритма наискорейшего спуска в терминах ВК?
60. Чему равна постоянная времени алгоритма наискорейшего спуска в терминах СКО?
61. Что такое LMS-алгоритм?

62. В чем состоит различие между LMS-алгоритмом и алгоритмом наискорейшего спуска?
63. Как определяется и чему равна вычислительная сложность LMS-алгоритма?
64. Как соотносятся в среднем градиенты LMS-алгоритма и алгоритма наискорейшего спуска?
65. Как соотносятся в среднем обучающие кривые LMS-алгоритма и алгоритма наискорейшего спуска?
66. Чему равно предельное значение шага сходимости LMS-алгоритма?
67. Чему равна постоянная времени LMS-алгоритма в терминах ВК?
68. Чему равна постоянная времени LMS-алгоритма в терминах СКО?
69. Как выглядят на одном рисунке обучающие кривые LMS-алгоритма и алгоритма наискорейшего спуска в терминах ВК?
70. Что такое избыточная среднеквадратическая ошибка в LMS-алгоритме?
71. Что такое относительная избыточная СКО в LMS-алгоритме?
72. Как определяется относительная избыточная СКО в LMS-алгоритме?
73. Чему равно уточненное с учетом избыточной СКО предельно допустимое значение шага сходимости LMS-алгоритма?
74. Что такое NLMS-алгоритм?
75. Как определяется шаг сходимости в NLMS-алгоритме?
76. Как формулируется задача наименьших квадратов в теории адаптивной фильтрации?
77. Как формулируется рекурсивная задача наименьших квадратов в теории адаптивной фильтрации?
78. Что такое параметр экспоненциального взвешивания сигналов? Зачем он используется при решении рекурсивной задачи наименьших квадратов?
79. Каковы предельные значения параметра экспоненциального взвешивания сигналов?
80. Как определяется целевая функция в рекурсивной задаче наименьших квадратов?

81. Чему равны координаты минимума целевой функции в рекурсивной задаче наименьших квадратов?
82. Что такое вектор коэффициентов Калмана?
83. Зачем и как задается начальное значение корреляционной матрицы при рекурсивном решении задачи наименьших квадратов?
84. Чему равна оценка вычислительной сложности RLS-алгоритма адаптивной фильтрации?
85. Что такое относительная избыточная СКО в RLS-алгоритме?
86. Как определяется относительная избыточная СКО в RLS-алгоритме?
87. Существует ли различие в определении относительной избыточной СКО для АФ с действительными и комплексными весовыми коэффициентами?
88. Из каких характерных участков состоит обучающая кривая RLS-алгоритма и чем определяется форма этих участков?
89. Чему равна постоянная времени обучающей кривой RLS-алгоритма и на каком участке ее можно корректно определять?
90. Какие бывают разновидности RLS-алгоритмов адаптивной фильтрации.

12.23. Литература

1. Уидроу, Б. Адаптивная обработка сигналов: пер. с англ. под ред. В. В. Шахгильдяна / Б. Уидроу, С. Д. Стирнз. – М.: Радио и связь, 1989. – 440 с.
2. Haykin, S. Adaptive filter theory. – 4-th edition / S. Haykin. – Prentice Hall, 2001. – 936 p.
3. Diniz, P. S. R. Adaptive filtering algorithms and practical implementation. – Third edition / P. S. R. Diniz. – New York: Springer Science + Business Media, 2008. – 627 p.
4. Джиган, В. И. Адаптивная обработка сигналов в радиотехнических системах: учеб. пособие / В. И. Джиган. – М.: МИЭТ, 2012. – 148 с.

5. Джиган, В.И. Адаптивные фильтры и их приложения в радиотехнике и связи. Часть 1 / В. И. Джиган // Современная электроника. – 2009. – № 9. – С. 56–63.
6. Джиган, В.И. Адаптивные фильтры и их приложения в радиотехнике и связи. Часть 2 / В. И. Джиган // Современная электроника. – 2010. – № 1. – С. 72–77.
7. Джиган, В.И. Адаптивные фильтры и их приложения в радиотехнике и связи. Часть 3 / В. И. Джиган // Современная электроника. – 2010. – № 2. – С. 70–77.

ЗАКЛЮЧЕНИЕ

В книге рассмотрены базовые методы цифровой обработки сигналов, включая многоскоростную и адаптивную обработку, которые в последнее время широко используются в радиотехнических и инфокоммуникационных системах различного назначения.

Дано общее представление об адаптивных фильтрах, определены требования к ним, а также рассмотрены примеры применения адаптивных фильтров в радиотехнических системах и системах связи. Их использование придает этим системам ранее недостижимые функциональные и потребительские качества.

Рассмотренные материалы могут быть распространены на случай адаптивных фильтров с комплексными ВК, которые находят применение в ААР, эквалайзерах модемов цифровых систем связи, где комплексная форма обрабатываемых сигналов требует применения адаптивных фильтров с комплексными ВК.

Материал пособия может также являться основой для изучения вопросов цифровой обработки речевых сигналов, изображений, видеопоследовательностей. Кроме того, цифровые методы часто используются для предобработки и постобработки в системах распознавания объектов различной природы. К ним можно отнести системы распознавания речи, идентификации диктора и современные системы видеоаналитики.

Традиционно цифровые методы широко применяются в системах технического зрения, в том числе зрения роботов, которые в последнее время используются в огромном количестве приложений, включая космические и медицинские системы.

Учебное издание

Брюханов Юрий Александрович
Приоров Андрей Леонидович
Джиган Виктор Иванович
Хрящев Владимир Вячеславович

**ОСНОВЫ
ЦИФРОВОЙ ОБРАБОТКИ
СИГНАЛОВ**

Учебное пособие

Редактор, корректор
Л. Н. Селиванова, М. Э. Левакова

Подписано в печать 29.04.13. Формат 60x84 1/16.
Усл. печ. л. 25,80. Уч.-изд. л. 16,43.
Тираж 50 экз. Заказ

Оригинал-макет подготовлен
в редакционно-издательском отделе ЯрГУ.
150000, Ярославль, ул. Советская, 14.

Отпечатано в типографии ООО «Филигрань».
г. Ярославль, ул. Свободы, д. 91.
Тел. (4852) 982705,
pechataet@bk.ru