

УЧЕБНИКИ НГТУ

В. Н. ВАСЮКОВ

**ЦИФРОВАЯ
ОБРАБОТКА СИГНАЛОВ
И СИГНАЛЬНЫЕ
ПРОЦЕССОРЫ
В СИСТЕМАХ ПОДВИЖНОЙ
РАДИОСВЯЗИ**



Министерство образования Российской Федерации
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

В.Н. ВАСЮКОВ

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ
И СИГНАЛЬНЫЕ ПРОЦЕССОРЫ
В СИСТЕМАХ
ПОДВИЖНОЙ РАДИОСВЯЗИ

Учебник



НОВОСИБИРСК
2006

УДК 621.372.083.92(075.8)
В 201

Рецензенты: профессор, д-р техн. наук *А. Г. Вострецов*
профессор, д-р техн. наук *И. С. Грузман*

Подготовлено на кафедре теоретических основ
радиотехники НГТУ

Васюков, В.Н.

В 201 Цифровая обработка сигналов и сигнальные процессоры
в системах подвижной радиосвязи: Учебник / В.Н. Васюков. –
Новосибирск: Изд-во НГТУ, 2006. – 292 с. – («Учебники НГТУ»).

ISBN 5-7782-0314-4

В учебнике излагаются основы теории цифровой обработки сигналов и вопросы ее применения. Рассмотрены методы анализа дискретных цепей при детерминированных и случайных воздействиях, синтеза и реализации цифровых фильтров, особенности архитектуры сигнальных процессоров и применение цифровой обработки сигналов в системах подвижной радиосвязи и других областях. Изложение опирается на современный математический аппарат и сопровождается большим количеством примеров и упражнений.

Учебник предназначен для студентов, обучающихся по специальности 201200 – «Средства связи с подвижными объектами» и может быть использован при изучении цифровой обработки сигналов студентами и магистрантами радиотехнических и близких специальностей. Книга будет полезна инженерам, интересующимся цифровой обработкой сигналов.

УДК 621.372.083.92(075.8)

ISBN 5-7782-0314-4

© Васюков В.Н., 2006
© Новосибирский государственный
технический университет, 2006

ПРЕДИСЛОВИЕ

Сегодня трудно представить себе, что всего 30 лет назад цифровая обработка сигналов представляла собой узкую область исследований, основным назначением которой было моделирование аналоговых радиотехнических цепей и сигналов на универсальных цифровых электронных вычислительных машинах (ЭВМ). Современная цифровая обработка сигналов (ЦОС) – это глубокая теория и широчайшее многообразие практических приложений, среди которых радиолокация, радионавигация, прикладная геодезия, медицина, автоматика, телемеханика, компьютерная томография и, разумеется, мобильная связь. Мобильная станция (сотовый телефон) является наглядным примером практического использования достижений теории и техники цифровой обработки сигналов, без которой сотовая связь осталась бы привилегией очень состоятельных людей.

Бурное развитие ЦОС началось с появлением больших интегральных схем и, в частности, микропроцессоров. Развитие элементной базы стимулировало теоретические исследования, а они, в свою очередь, питали новыми идеями разработчиков аппаратных и программных средств. Несомненно, этот процесс будет продолжаться и в обозримом будущем, поэтому современный инженер должен хорошо знать как теоретические основы, так и практические методы применения цифровой обработки сигналов. Для инженера, специализирующегося в области мобильной связи, это жизненно необходимо.

С момента возникновения ЦОС как самостоятельного раздела теории цепей и сигналов до настоящего времени опубликовано множество теоретических и прикладных работ, в том числе знаменитые книги Б. Голда и Ч. Рейдера [1], А. Оппенгейма и Р. Шафера [2], Л. Рабинера и Б. Гоулда [3], ставшие классическими, а также работы Л.М. Гольденберга и др. [4, 5]. К сожалению, эти книги издавались слишком давно и незначительными тиражами и стали к настоящему времени недоступны для большинства читателей. Последние издания [6 – 8] не являются учебниками и малопригодны

для первоначального изучения цифровой обработки сигналов. Предлагаемый вниманию читателя учебник является попыткой восполнить недостаток учебной литературы в области ЦОС.

Для усвоения материала учебника достаточно знать математику в объеме вузовского курса, включая линейную алгебру, основы математического анализа, элементы теории функций комплексного переменного, основы теории случайных процессов. Учебник построен следующим образом.

В кратком введении (разд. 1) приводятся некоторые сведения о возникновении и развитии ЦОС, преимуществах цифровой обработки сигналов перед аналоговой, а также даются определения дискретных и аналоговых сигналов.

В разд. 2, озаглавленном «Математические основы ЦОС», излагаются вопросы математического описания дискретных сигналов и линейных инвариантных к сдвигу (стационарных) дискретных цепей. Содержание этого раздела имеет принципиальное значение для понимания остальных разделов учебника. Стержнем изложения теории ЦОС в учебнике являются операция дискретной свертки на множестве последовательностей и соответствующая ей операция умножения на множестве непрерывных функций на единичной окружности. С этой точки зрения z -преобразование, преобразование Фурье, дискретное преобразование Фурье представляются совершенно естественными и, пожалуй, единственно адекватными инструментами анализа дискретных сигналов и линейных стационарных дискретных цепей.

Раздел 3 содержит изложение теории дискретных линейных инвариантных к сдвигу цепей (ЛИС-цепей), включая их описание разностными уравнениями и передаточными функциями, структурные схемы, нуль-полюсные диаграммы, строение импульсной характеристики ЛИС-цепи конечного порядка, особенности цепей с линейной фазочастотной характеристикой, минимально-фазовых и всепропускающих цепей.

В разд. 4 кратко излагается корреляционно-спектральная теория стационарных случайных последовательностей и их взаимодействия с дискретными ЛИС-цепями.

Раздел 5 посвящен введению в теорию многомерных дискретных сигналов и цепей.

Раздел 6 содержит изложение методов синтеза дискретных (цифровых) фильтров, а именно: методов оконного взвешивания, частотной выборки и быстрой свертки для фильтров с конечной импульсной характеристикой, а также методов аналого-цифровой трансформации (инвариантности импульсной характеристики и

билинейного преобразования) для фильтров с бесконечной импульсной характеристикой. Рассмотрены также методы частотных преобразований для аналоговых фильтров-прототипов и цифровых фильтров.

Вопросам реализации цифровых фильтров посвящен разд. 7. Рассмотрены различия между цифровыми и дискретными цепями, специфические эффекты, связанные с переходом от аналоговых сигналов к цифровым (квантование и дрожание отсчетов), эффекты квантования параметров дискретных фильтров при их реализации в цифровой форме, а также различные формы реализации.

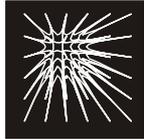
В разд. 8 рассматриваются некоторые основные идеи цифрового спектрального анализа, который находит применение в современной цифровой обработке сигналов и, в частности, в системах мобильной связи.

Краткий обзор архитектурных особенностей сигнальных процессоров и средств разработки приложений, а также некоторые сведения о программируемых логических интегральных схемах содержится в разд. 9.

Завершает учебник разд. 10, в котором рассматриваются основные применения ЦОС в системах подвижной радиосвязи: цифровая фильтрация, перенос спектра, цифровая модуляция и демодуляция, цифровое формирование диаграммы направленности антенны и обработка речевых сигналов.

Автор надеется, что учебник окажется полезным тем читателям, которые пожелают понять теоретические основы ЦОС и познакомиться с ее практическими применениями. Все замечания, пожелания и предложения читателей будут с благодарностью приняты и учтены.

Автор выражает искреннюю благодарность коллегам, взявшим на себя труд чтения учебника, что способствовало его улучшению.



1. ВВЕДЕНИЕ

Обработка сигналов представляет собой совокупность операций, направленных на наиболее эффективную передачу, хранение и извлечение информации. Вплоть до последней четверти XX века в большинстве случаев обработка сигналов производилась в аналоговой форме, так как цифровая техника была недостаточно малогабаритной, быстродействующей и дешевой.

Теоретические основы цифровой обработки сигналов были заложены еще в 20 – 30-х годах, когда в работах Х. Найквиста и В.А. Котельникова была доказана возможность взаимно однозначного представления аналогового сигнала с ограниченной полосой частот последовательностью его значений (отсчетов), взятых через определенные промежутки времени. В 50 – 60-е годы универсальные цифровые вычислительные машины стали применяться для цифрового моделирования аналоговых устройств и систем; были разработаны основы цифровой фильтрации. Революция в методах и средствах обработки сигналов началась с открытия в 1965 году класса эффективных («быстрых») алгоритмов дискретного преобразования Фурье и появления в 1971 году микропроцессоров, что в совокупности сделало возможной цифровую обработку сигналов в реальном времени.

Стремительное развитие элементной базы, повышение быстродействия, снижение массогабаритных характеристик и стоимости больших интегральных схем, в частности микропроцессоров, обеспечило возможность реализации цифровыми средствами многих алгоритмов обработки сигналов, которые было невозможно или крайне трудно реализовать в аналоговой технике. Это, в свою очередь, стимулировало рост интереса к возможностям ЦОС и появление новых алгоритмов и привело к созданию теории дискретных и цифровых сигналов и цепей, которая еще далека от завершения. Тем не менее ясно, что эта теория весьма существенно отличается от теории аналоговых сигналов и цепей, хотя между ними имеется и несомненное сходство. Следует иметь в виду, что прямое перене-

сение взглядов, обычных для аналоговой теории, на цифровые цепи и сигналы чревато серьезными ошибками.

Цифровой обработке сигналов свойственны следующие широко известные преимущества перед аналоговой обработкой:

- принципиальная возможность реализации практически *любых* алгоритмов обработки (в аналоговой технике могут быть реализованы далеко не всякие алгоритмы); развитие элементной базы обеспечивает реализуемость все более широкого класса алгоритмов обработки в *реальном масштабе времени*;

- потенциально сколь угодно высокая *точность* реализации алгоритмов;

- принципиальная возможность *безошибочного* воспроизведения сигналов при передаче и хранении на основе помехоустойчивого кодирования, которое применимо только к дискретным сигналам.

Реализация перечисленных преимуществ на практике возможна лишь на основе глубокого понимания теории дискретных сигналов и цепей, которая базируется на таких разделах математики, как линейная алгебра, теория функций комплексного переменного, теория вероятностей и т.д.

Центральным понятием теории ЦОС является понятие дискретного сигнала. Математической моделью дискретного сигнала служит решетчатая функция, или *последовательность* $x[n]$, где n – аргумент, принимающий значения из дискретного множества, а функция $x[\cdot]$ может принимать значения из непрерывного множества вещественных или комплексных чисел (впрочем, в некоторых приложениях, таких как обработка пространственно-временных сигналов или цветных изображений, функция $x[\cdot]$ принимает векторные значения). Аналоговые сигналы описываются функциями непрерывных аргументов. Как известно, аналоговый сигнал $x_a(t)$ с *финитным* (ограниченным по ширине) спектром, сосредоточенным в полосе частот $(-\omega_b, \omega_b)$, может быть без потери информации заменен дискретной последовательностью своих значений, взятых с шагом $T < \pi/\omega_b$. В таком случае эта последовательность $x[n] = x_a(nT)$ представляет собой дискретный сигнал. С другой стороны, многие дискретные сигналы (последовательности) не имеют аналогового прототипа и изначально являются дискретными. Теория ЦОС не нуждается в предположении о существовании аналогового сигнала как источника возникновения

последовательности. Однако если возникает надобность, любой последовательности можно сопоставить аналоговый сигнал с конечным спектром, из которого она могла бы быть получена путем дискретизации.

В дальнейшем, если не сказано иное, предполагается, что аргумент n последовательности $x[n]$ принимает целые значения от $-\infty$ до $+\infty$ (обозначается $n = \overline{-\infty, +\infty}$). Чтобы отличать дискретный аргумент от непрерывного, принято заключать его в квадратные скобки.

Дискретные сигналы часто изображают графиками так, как показано на рис. 1.1. Следует иметь в виду, что хотя ось абсцисс представлена на графике сплошной линией, дискретный сигнал *не определен* для нецелых значений аргумента. Для краткости значения дискретного сигнала в дальнейшем именуется *отсчетами*.

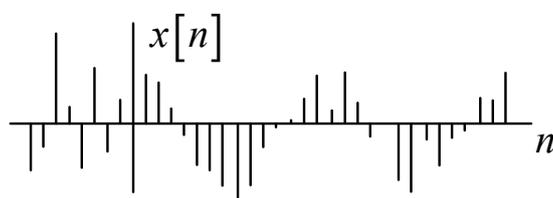
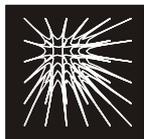


Рис. 1.1. Графическое изображение дискретного сигнала

В теории ЦОС различают понятия дискретного и цифрового сигналов. Под цифровым сигналом понимают дискретный сигнал, *квантованный* по уровню. Другими словами, цифровой сигнал — это последовательность, принимающая значения из дискретного (к тому же, как правило, конечного) множества. Это связано с тем, что цифровые устройства всегда имеют ограниченную разрядность и сигналы, подлежащие *цифровой* обработке, неизбежно округляются (квантуются). Следует отметить, что существуют устройства *дискретной* обработки сигналов (например, известны фильтры на поверхностных акустических волнах — ПАВ), не предполагающие квантования сигналов [9]; в учебнике такие устройства не рассматриваются. Для изучения большинства разделов теории *цифровой* обработки сигналов модель *дискретного* сигнала удобнее, поэтому всюду, где это возможно, используется именно такая модель. Моделью цифрового сигнала пользуются обычно лишь в тех случаях, когда рассматриваются специфические эффекты, связанные с квантованием сигнала, округлением промежуточных результатов, ограничением разрядной сетки цифрового устройства и т.п. Эти вопросы рассмотрены в разд. 7.



2. МАТЕМАТИЧЕСКИЕ ОСНОВЫ ЦОС

2.1. ЭЛЕМЕНТАРНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

Рассмотрим некоторые последовательности, играющие важную роль в теории ЦОС.

1. δ -Последовательность, определяемая выражением

$$\delta[n] = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0, \end{cases}$$

представлена на рис. 2.1. Такая же последовательность, задержанная на k шагов,

$$\delta[n - k] = \begin{cases} 1, & n = k, \\ 0, & n \neq k, \end{cases}$$

изображена на рис. 2.2.

2. Единичная ступенчатая последовательность

$$u[n] = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0, \end{cases}$$

представленная на рис. 2.3, очевидно, связана с $\delta[n]$ соотношениями

$$u[n] = \sum_{k=-\infty}^n \delta[k] = \sum_{k=0}^{\infty} \delta[n - k], \quad \delta[n] = u[n] - u[n - 1],$$

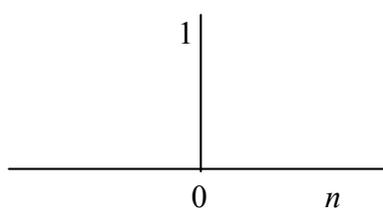


Рис. 2.1. δ -Последовательность

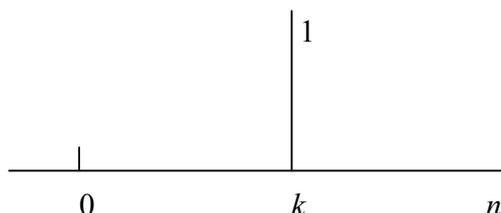


Рис. 2.2. Задержанная δ -последовательность

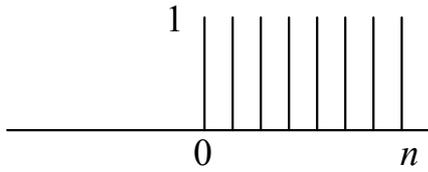


Рис. 2.3. Единичная ступенчатая последовательность

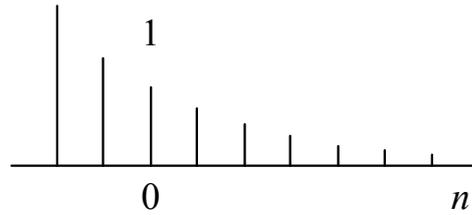


Рис. 2.4. Вещественная экспоненциальная последовательность

представляющими дискретные аналоги известных формул

$$\sigma(t) = \int_{-\infty}^t \delta(\tau) d\tau, \quad \delta(t) = \frac{d}{dt} \sigma(t)$$

для функции единичного скачка Хевисайда $\sigma(t)$ и δ -функции Дирака $\delta(t)$.

3. Экспоненциальная последовательность

$$e[n] = a^n,$$

где $a = re^{j\omega}$ – комплексное число, может быть представлена в виде

$$e[n] = r^n e^{j\omega n} = r^n (\cos \omega n + j \sin \omega n).$$

В частном случае, при $\omega = 0$, получается вещественная экспоненциальная последовательность, представленная для $0 < r < 1$ на рис. 2.4.

При $r = 1$ вещественная и мнимая части $e[n]$ представляют собой гармонические последовательности $c[n] = \text{Re}\{e[n]\} = \cos \omega n$ и $s[n] = \text{Im}\{e[n]\} = \sin \omega n$ (рис. 2.5, 2.6).

4. Последовательность $\tilde{x}[n]$ является *периодической*, если при некотором целом N , называемом периодом, $\tilde{x}[n] = \tilde{x}[n + N]$.

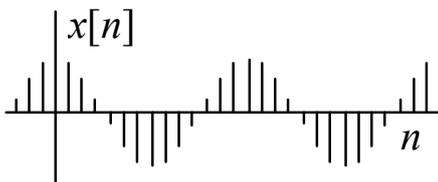


Рис. 2.5. Косинусоидальная последовательность

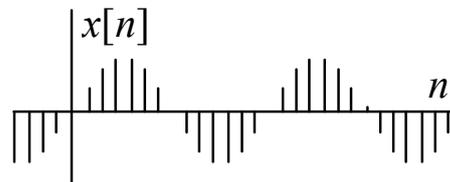


Рис. 2.6. Синусоидальная последовательность

Очевидно, что в отличие от аналоговых прототипов $\cos \omega t$ и $\sin \omega t$ дискретные косинусоида и синусоида могут быть *непериодическими* (в качестве упражнения читателю предлагается определить условия периодичности и непериодичности последовательностей $c[n]$ и $s[n]$).

Ясно, что можно задать бесконечное множество различных последовательностей. Для того чтобы изучать последовательности как элементы этого множества, т. е. с общих позиций, нужно ввести характеристики, определяющие степень сходства и различия между ними. Различие между последовательностями, выражаемое числом, можно уподобить расстоянию между точками (векторами) в некотором пространстве. Удобной для изучения алгебраической моделью множества последовательностей является, таким образом, линейное (векторное) пространство.

2.2. ПРОСТРАНСТВО ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Линейное пространство – это по определению множество M элементов, называемых *векторами*, обладающее следующими свойствами:

А. Для любых векторов x и y из M определена сумма, т.е. существует вектор $(x + y)$, также принадлежащий M (множество *замкнуто* относительно сложения). Кратко это записывается так: $\forall x, y \in M \quad \exists (x + y) \in M$ (\forall читается «для всех», \exists читается «существует»). При этом:

а1) сложение ассоциативно: $x + (y + z) = (x + y) + z$;

а2) в M имеется (единственный) нулевой вектор:

$$\exists \vec{0} \in M : \forall x \in M \quad x + \vec{0} = x ;$$

а3) для любого вектора x в M существует обратный ему вектор, обозначаемый $-x$:

$$\forall x \in M \quad \exists (-x) \in M : \quad x + (-x) = \vec{0} ;$$

а4) сложение коммутативно: $x + y = y + x \quad \forall x, y \in M$.

Условия а1 – а3 известны в алгебре как аксиомы *группы*. Вместе с условием а4 они определяют *коммутативную (абелеву) группу* относительно операции сложения.

Б. Имеется множество (поле*) \mathbb{F} элементов (скаляров), и определена операция умножения вектора на скаляр, так что $\forall x \in M \quad \forall \alpha \in \mathbb{F}: \alpha x \in M$, причем:

б1) умножение вектора на скаляр ассоциативно:

$$\forall \alpha, \beta \in \mathbb{F} \quad \forall x \in M: \alpha(\beta x) = (\alpha\beta)x;$$

б2) в поле \mathbb{F} имеется единичный скаляр:

$$\exists 1 \in \mathbb{F}: \forall x \in M \quad 1x = x;$$

б3) в поле \mathbb{F} имеется нулевой скаляр:

$$\exists 0 \in \mathbb{F}: \forall x \in M \quad 0x = \vec{0};$$

б4) умножение вектора на скаляр дистрибутивно:

$$\forall \alpha, \beta \in \mathbb{F} \quad \forall x, y \in M: \begin{cases} \alpha(x+y) = \alpha x + \alpha y, \\ (\alpha + \beta)x = \alpha x + \beta x. \end{cases}$$

Множество, для которого определены операции, удовлетворяющие аксиомам а1 – а4 и б1 – б4, называется *линейным* (векторным) *пространством над полем* \mathbb{F} .

Нетрудно убедиться, что множество последовательностей, отсчеты которых принимают вещественные значения (вещественных последовательностей), с операциями сложения и умножения на вещественное число (скаляр) удовлетворяет всем перечисленным условиям. Поэтому говорят, что множество вещественных последовательностей образует линейное (векторное) пространство над полем \mathbb{R} вещественных чисел. Аналогично множество комплексных последовательностей образует векторное пространство над полем \mathbb{C} комплексных чисел. В дальнейшем для общности рассматривается именно это пространство.

Поскольку определены операции суммирования векторов (последовательностей) и умножения вектора (последовательности) на скаляр, то определена конечная сумма (*линейная комбинация*)

* Полем в алгебре называется множество, являющееся коммутативной группой относительно сложения и одновременно коммутативной группой относительно умножения, причем для нулевого (по сложению) элемента не существует обратный по умножению элемент (не определено «деление на нуль»). Множество \mathbb{R} всех вещественных чисел и множество \mathbb{C} всех комплексных чисел с обычными сложением и умножением представляют собой наиболее известные примеры поля.

$x = \sum_{k=1}^K \alpha_k x_k$ для любой совокупности K векторов $x_k \in M$, $k = \overline{1, K}$

со скалярными коэффициентами $\alpha_k \in \mathbb{C}$, $k = \overline{1, K}$.

Если зафиксировать векторы x_k , $k = \overline{1, K}$, то получаемые *при всевозможных* наборах $\{\alpha_k \in \mathbb{C}, k = \overline{1, K}\}$ линейные комбинации образуют *линейную оболочку* совокупности векторов $\{x_k, k = \overline{1, K}\}$.

Множество векторов $\{x_k, k = \overline{1, K}\}$ *линейно независимо*, если равенство $\sum_{k=1}^K \alpha_k x_k = 0$ возможно только при $\alpha_k = 0 \quad \forall k$.

Линейная оболочка множества K линейно независимых векторов (последовательностей) $\{x_k, k = \overline{1, K}\}$ образует K -мерное *подпространство* M_K пространства M ; тогда $\{x_k, k = \overline{1, K}\}$ – *базис* подпространства M_K . (Подпространством называется такое подмножество некоторого пространства, которое само по себе является линейным пространством меньшей размерности.)

Базис $\{x_k, k = \overline{1, K}\}$ не единствен для M_K ; любая совокупность из K линейно независимых векторов (последовательностей), которые все принадлежат M_K , является для M_K базисом; любой вектор из M_K можно представить линейной комбинацией базисных векторов.

Увеличивая размерность K подпространства M_K до бесконечности, в пределе можно получить линейную оболочку бесконечной совокупности базисных векторов (последовательностей) – бесконечномерное пространство. Строго говоря, необходимо убедиться в том, что любая линейная комбинация *бесконечного* множества линейно-независимых последовательностей принадлежит пространству M . Соответствующее свойство бесконечномерных линейных пространств называется *полнотой*. Для пространств последовательностей, которые будут использоваться далее, свойство полноты выполняется.

Пример 2.1. Один из возможных базисов для пространства последовательностей состоит из сдвинутых δ -последовательностей

$$\delta[n-k], \quad k = \overline{-\infty, \infty}.$$

Можно убедиться, что набор из любого количества таких последовательностей линейно независим; нетрудно видеть, что произвольную последовательность $x[n], n = \overline{-\infty, \infty}$ можно представить линейной комбинацией

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k]. \quad \blacktriangleleft \quad (2.1)$$

(Символ \blacktriangleleft здесь и далее обозначает конец примера.)

Пример 2.2. Рассмотрим множество M_N всех последовательностей конечной длины N , определенных при $n = \overline{0, N-1}$. Каждая такая последовательность однозначно определяется своими N значениями (отсчетами) $x[n], n = \overline{0, N-1}$, и, следовательно, может быть представлена вектором в N -мерном пространстве. Базис пространства M_N можно задать, например, в виде следующего набора векторов (последовательностей) длины N :

$$\begin{aligned} e_0 &= \delta[n], \quad n = \overline{0, N-1}, \\ e_1 &= \delta[n-1], \quad n = \overline{0, N-1}, \\ &\dots\dots\dots \\ e_{N-1} &= \delta[n-N+1], \quad n = \overline{0, N-1}. \quad \blacktriangleleft \end{aligned}$$

Пример 2.3. Рассмотрим множество всех последовательностей бесконечной длины таких, что они могут принимать ненулевые значения лишь при $n = \overline{0, N-1}$. Каждая такая последовательность однозначно определяется своими N значениями (отсчетами) $x[n], n = \overline{0, N-1}$, и, следовательно, может быть представлена вектором в N -мерном пространстве. Базис пространства M_N можно задать, например, в виде следующего набора векторов:

$$\begin{aligned} e_0 &= \delta[n], \quad n = \overline{-\infty, \infty}, \\ e_1 &= \delta[n-1], \quad n = \overline{-\infty, \infty}, \\ &\dots\dots\dots \\ e_{N-1} &= \delta[n-N+1], \quad n = \overline{-\infty, \infty}. \end{aligned}$$

Линейная оболочка этого базиса, очевидно, является подпространством (N -мерным) бесконечномерного пространства последовательностей M . ◀

Важным шагом в определении структуры пространства последовательностей является задание *скалярного произведения* для любых двух последовательностей $x[n]$ и $y[n]$ выражением

$$(x, y) = \sum_{n=-\infty}^{\infty} x[n]y^*[n],$$

где $(\cdot)^*$ – символ комплексного сопряжения.

В частности, скалярное умножение последовательности на себя дает

$$(x, x) = \sum_{n=-\infty}^{\infty} x[n]x^*[n] = \sum_{n=-\infty}^{\infty} |x[n]|^2 = E_x,$$

где E_x имеет смысл *энергии* сигнала (последовательности) $x[n]$.

Часто выделяют для рассмотрения только такие последовательности, для которых $E < \infty$. Последовательности ограниченной энергии образуют бесконечномерное пространство, обозначаемое l_2 . Каждой последовательности из l_2 соответствует определенное вещественное число

$$\sqrt{E_x} = \sqrt{(x, x)} = \|x\|_2, \quad (2.2)$$

называемое *нормой*. Пространство, в котором введена норма, называется нормированным. Заметим, что норму можно задать и другими способами; от определения нормы зависят свойства пространства (далее, в п. 2.5, будет введена другая норма). Норму, определенную выражением (2.2), будем называть l_2 -нормой. Понятие нормы является обобщением понятия длины (модуля) вектора. Для произвольных последовательностей $x, y \in l_2$ справедливо *неравенство Шварца*

$$|(x, y)|^2 \leq (x, x)(y, y).$$

Две последовательности x и y называются *ортогональными*, если $(x, y) = 0$. Можно ввести понятие угла φ между вещественными

последовательностями (векторами) выражением $\frac{(x, y)}{\|x\|_2 \cdot \|y\|_2} = \cos \varphi$.

Неравенство Шварца, переписанное в форме $|(x, y)| \leq \|x\|_2 \cdot \|y\|_2$, по существу, означает, что косинус угла между векторами не превосходит по модулю единицы.

Пример 2.4. Две последовательности конечной длины N , определенные выражениями $x[n] = \cos\left(\frac{2\pi}{N}n\right)$ и $y[n] = \sin\left(\frac{2\pi}{N}n\right)$ при $n = \overline{0, N-1}$, ортогональны, так как

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi}{N}n\right) \sin\left(\frac{2\pi}{N}n\right) = \sum_{n=0}^{N-1} \frac{1}{2} \sin\left(\frac{4\pi}{N}n\right) = 0.$$

Здесь использовано то обстоятельство, что $\sin\left(\frac{4\pi}{N}n\right) = \text{Im}\left(e^{j\frac{4\pi}{N}n}\right)$,

поэтому $\sum_{n=0}^{N-1} \sin\left(\frac{4\pi}{N}n\right) = \text{Im}\left\{\sum_{n=0}^{N-1} e^{j\frac{4\pi}{N}n}\right\}$, а сумма в фигурных скоб-

ках равна нулю как сумма N векторов единичной длины на комплексной плоскости, каждый из которых отстоит от соседних векторов на угол $4\pi/N$. ◀

В нормированном пространстве l_2 можно для каждой пары последовательностей $x, y \in l_2$ определить расстояние между ними путем задания функции (метрики) выражением $d_2(x, y) = \|x - y\|_2$. Пространство, в котором определена метрика, называется *метрическим*.

Подводя итог, можно сказать, что пространство l_2 последовательностей ограниченной энергии бесконечномерно, снабжено скалярным произведением, порождающим норму и метрику, а также полно (доказательство полноты см., например, в [10]). Пространства со скалярным произведением и порожденной им нормой, обладающие свойством полноты, называются *гильбертовыми*. Таким образом, l_2 – гильбертово пространство.

2.3. БАЗИСЫ ПРОСТРАНСТВА ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Для любого линейного пространства можно задать базис многими способами, однако при этом возникает вопрос, как найти коэффициенты для представления произвольного вектора в заданном базисе. В гильбертовых пространствах благодаря наличию скалярного произведения эта задача решается достаточно просто (путем построения так называемого *взаимного* или *сопряженного* базиса [11]).

Пусть в гильбертовом пространстве M задан произвольный базис $\{v_i, i = \overline{1, \infty}\}$, составленный из линейно-независимых векторов. Сопряженный базис $\{w_i, i = \overline{1, \infty}\}$ должен удовлетворять условию

$$(v_i, w_j) = \delta_{ij} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

где δ_{ij} – символ Кронекера. Найдем скалярное произведение произвольного вектора x , представленного в виде линейной комбинации

$x = \sum_{i=1}^{\infty} \alpha_i v_i$, и вектора w_k из сопряженного базиса:

$$(x, w_k) = \left(\sum_{i=1}^{\infty} \alpha_i v_i, w_k \right) = \sum_{i=1}^{\infty} \alpha_i (v_i, w_k) = \sum_{i=1}^{\infty} \alpha_i \delta_{ik} = \alpha_k.$$

Таким образом, если известен сопряженный базис, то произвольный коэффициент разложения вектора в заданном базисе можно найти как скалярное произведение $\alpha_k = (x, w_k)$ данного вектора и соответствующего вектора из сопряженного базиса.

Особенно удобно, если все векторы базиса попарно ортогональны и имеют норму, равную 1. Такие базисы называются *ортонормальными* (ортонормированными). Существует простая процедура, называемая *процедурой Грама – Шмидта*, позволяющая построить ортонормальный базис гильбертова пространства на основе линейно-независимой совокупности векторов.

Пусть в гильбертовом пространстве M задана совокупность $\{v_i, i = \overline{1, \infty}\}$ линейно-независимых векторов. Чтобы построить совокупность $\{u_i, i = \overline{1, \infty}\}$ ортонормальных векторов, следует дейст-

зовать в соответствии со схемой ($\{w_i, i = \overline{1, \infty}\}$ – вспомогательная совокупность векторов):

$$1) w_1 = v_1, u_1 = \frac{1}{\|w_1\|_2} w_1;$$

$$2) w_2 = v_2 - (v_2, u_1)u_1, u_2 = \frac{1}{\|w_2\|_2} w_2;$$

$$3) w_3 = v_3 - (v_3, u_1)u_1 - (v_3, u_2)u_2, u_3 = \frac{1}{\|w_3\|_2} w_3$$

и т.д. Таким образом, на каждом шаге процедуры из очередного вектора линейно-независимой системы вычитаются его проекции на все уже построенные на предыдущих шагах векторы ортонормального базиса, после чего вектор нормируется (первый и второй шаги процедуры Грама – Шмидта показаны на рис. 2.7.).

В качестве упражнения читателю предлагается непосредственной проверкой убедиться, что все векторы $\{u_i, i = \overline{1, 3}\}$ нормированы и попарно ортогональны.

Очевидно, при различной нумерации исходной системы векторов $\{v_i, i = \overline{1, \infty}\}$ процедура Грама – Шмидта приводит к различным базисам.

Для бесконечномерных пространств базисы должны обладать свойством *полноты*. Если базис полон, то в данном пространстве

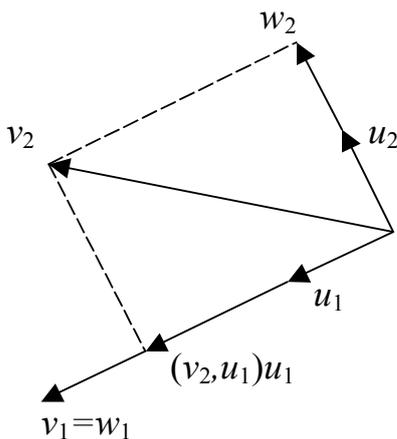


Рис. 2.7. Первый и второй шаги процедуры Грама – Шмидта

не существует векторов, ортогональных *всем* векторам базиса. Полнота базиса означает, что, взяв достаточно большое, но конечное число N базисных векторов, можно аппроксимировать данный вектор x с любой точностью конечной суммой

$$x_N = \sum_{i=1}^N \alpha_i u_i.$$

Для полного ортонормального базиса справедливо *неравенство Бесселя*

$$\|x\|_2^2 \geq \sum_{i=1}^N |\alpha_i|^2.$$

Представление $x = \sum_{i=1}^{\infty} \alpha_i u_i$ произвольной последовательности x относительно ортонормального полного базиса $\{u_i, i = \overline{1, \infty}\}$ называется *обобщенным рядом Фурье*. Совокупность $\{\alpha_i, i = \overline{1, \infty}\}$ называется *спектром* последовательности x относительно базиса $\{u_i, i = \overline{1, \infty}\}$.

Как уже отмечалось, существует бесконечно много базисов, однако предпочтительными являются ортонормальные базисы, которые имеют преимущество, состоящее в удобстве нахождения спектральных коэффициентов. В самом деле, найдем скалярное произведение вектора x с базисным вектором u_k :

$$(x, u_k) = \left(\sum_{i=1}^{\infty} \alpha_i u_i, u_k \right) = \sum_{i=1}^{\infty} \alpha_i (u_i, u_k) = \sum_{i=1}^{\infty} \alpha_i \delta_{ik} = \alpha_k.$$

Таким образом, $\alpha_i = (x, u_i)$. Ортонормальный базис удовлетворяет условию

$$(u_i, u_j) = \delta_{ij} = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases}$$

поэтому его также называют *самосопряженным*.

В дальнейшем, как правило, будем полагать, что базисные последовательности пронумерованы целыми числами от $-\infty$ до ∞ , и записывать обобщенный ряд Фурье в виде $x = \sum_{i=-\infty}^{\infty} \alpha_i u_i$.

Пользуясь выражением для обобщенного ряда Фурье, найдем квадрат нормы последовательности (энергию дискретного сигнала):

$$\begin{aligned} \|x\|_2^2 &= (x, x) = \left(\sum_{i=-\infty}^{\infty} \alpha_i u_i, \sum_{k=-\infty}^{\infty} \alpha_k u_k \right) = \\ &= \sum_{i=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \alpha_i \alpha_k^* (u_i, u_k) = \sum_{i=-\infty}^{\infty} \alpha_i \alpha_i^* = \sum_{i=-\infty}^{\infty} |\alpha_i|^2. \end{aligned}$$

С другой стороны, $\|x\|_2^2 = \sum_{n=-\infty}^{\infty} |x[n]|^2$, поэтому

$$\sum_{n=-\infty}^{\infty} |x[n]|^2 = \sum_{i=-\infty}^{\infty} |\alpha_i|^2.$$

Это выражение представляет собой одну из форм *равенства Парсевала*. Смысл равенства Парсевала состоит в том, что квадрат нормы вектора представляет собой сумму квадратов модулей компонент (координат) этого вектора относительно произвольного *ортонормального полного* базиса. Заметим, что отсчеты последовательности в левой части равенства представляют собой ее координаты (спектральные коэффициенты) относительно ортонормального базиса, определяемого выражением (2.1).

Обобщение равенства Парсевала, называемое также *обобщенной формулой Рэлея*, относится к скалярному произведению двух последовательностей

$$\begin{aligned} (x, y) &= \left(\sum_{i=-\infty}^{\infty} \alpha_i u_i, \sum_{k=-\infty}^{\infty} \beta_k u_k \right) = \\ &= \sum_{i=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \alpha_i \beta_k^* (u_i, u_k) = \sum_{k=-\infty}^{\infty} \alpha_k \beta_k^*. \end{aligned}$$

Равенство Парсевала, в частности, позволяет пользоваться понятиями нормы и скалярного произведения последовательностей, представленных в ортонормальном полном базисе, не конкретизируя вид этого базиса.

2.4. ПОСЛЕДОВАТЕЛЬНОСТИ И ДИСКРЕТНЫЕ ЦЕПИ

Любой алгоритм обработки дискретного сигнала (последовательности) в самом общем виде можно описать как отображение множества (пространства) последовательностей в некоторое множество решений. Так, например, если целью обработки является *обнаружение* в рассматриваемой последовательности полезной составляющей («сигнала»), то множество решений содержит две «точки»: "0" (сигнала нет) и "1" (сигнал есть). Если цель обработки состоит в *фильтрации* последовательности (подавление шума и/или помехи), строение пространства решений совпадает со строением пространства исходных последовательностей. В таких

случаях алгоритм обработки и цепь (устройство), которая его выполняет, математически представляются отображением (оператором) $T\{\cdot\}$, так что

$$y[n] = T\{x[n]\},$$

где $x[n]$ и $y[n]$ называются соответственно *прообразом* и *образом* (при отображении $T\{\cdot\}$). Это можно иллюстрировать структурной схемой, приведенной на рис. 2.8.

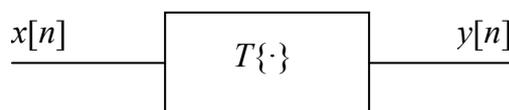


Рис. 2.8. Оператор цепи

Не существует общего метода исследования всех возможных цепей, поэтому для получения полезных для практики результатов приходится накладывать на оператор цепи дополнительные условия. Широкий класс дискретных цепей составляют цепи (операторы), подчиняющиеся так называемому *принципу суперпозиции*, выражаемому формулой

$$T\{\alpha x_1[n] + \beta x_2[n]\} = \alpha T\{x_1[n]\} + \beta T\{x_2[n]\},$$

верной при любых (комплексных) числах α, β и любых последовательностях x_1, x_2 .

Такие цепи (операторы, алгоритмы) называются *линейными*. Цепи (операторы, алгоритмы), не удовлетворяющие принципу суперпозиции, называются *нелинейными*.

Ограничение класса рассматриваемых цепей лишь линейными цепями позволяет применить для их анализа представление входной последовательности в виде взвешенной суммы (линейной комбинации) элементарных (базисных) последовательностей. В самом деле, в силу принципа суперпозиции реакцию (отклик) цепи на такой сигнал можно найти как взвешенную *таким же образом* сумму откликов на элементарные (базисные) последовательности. Приняв, например, в качестве базиса совокупность сдвинутых δ -последовательностей $\{\delta[n-k], n, k = -\infty, \infty\}$, можно произвольную последовательность представить в форме обобщенного ряда Фурье (2.1)

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k],$$

где $x[k]$ – весовой (спектральный) коэффициент при δ -последовательности, задержанной на k шагов.

Руководствуясь принципом суперпозиции, запишем для выходной последовательности

$$y[n] = T\{x[n]\} = T\left\{\sum_{k=-\infty}^{\infty} x[k]\delta[n-k]\right\} = \sum_{k=-\infty}^{\infty} x[k]T\{\delta[n-k]\}.$$

Таким образом, задача *анализа* (нахождения выходной последовательности по известной входной последовательности и известному описанию цепи) сводится для линейных цепей к определению откликов цепи на элементарные (базисные) последовательности. Обозначив отклик цепи $T\{\delta[n-k]\}$ через $h[n,k]$, получим

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n,k],$$

где $h[n,k] = h_k[n]$ имеет смысл отклика цепи на δ -последовательность, задержанную на k шагов.

Важный подкласс класса линейных цепей составляют цепи, для которых вид оператора $T\{\cdot\}$ не зависит от сдвига k , поэтому изменение реакции такой цепи на сдвинутую последовательность при разных сдвигах $k = -\infty, +\infty$ выражается также лишь в сдвиге отклика. Для таких цепей, называемых *линейными инвариантными к сдвигу* (сокращенно *ЛИС-цепями*), или *стационарными*,

$$h[n,k] = h_k[n] = h[n-k].$$

Инвариантность к сдвигу отражает «беспристрастность» цепи к моменту прихода сигнала на ее вход. Многие цепи можно считать (приближенно или точно) ЛИС-цепями. С другой стороны, анализ ЛИС-цепей намного проще, чем анализ цепей, не удовлетворяющих этому свойству. Поэтому далее в этом разделе рассмотрение ограничивается только ЛИС-цепями. Для ЛИС-цепей связь входной и выходной последовательностей задается выражением

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k], \quad (2.3)$$

называемым *дискретной сверткой*, а $h[\cdot]$ представляет собой функцию одного дискретного аргумента (последовательность) и называется *импульсной характеристикой* дискретной цепи. Свертка однозначно определяет преобразование последовательности ЛИС-цепью. Следовательно, импульсная характеристика составляет исчерпывающее описание ЛИС-цепи. Для краткости часто используется символическая форма записи свертки (2.3) в виде $y = x \otimes h$.

Таким образом, любая ЛИС-цепь *полностью* определяется своей импульсной характеристикой, которая представляет собой отклик цепи на δ -последовательность

$$h[n] = \sum_{k=-\infty}^{\infty} \delta[k] h[n-k], \quad n = \overline{-\infty, +\infty}.$$

Часто (особенно в системах ЦОС, работающих в реальном масштабе времени) на импульсную характеристику накладывается условие *каузальности* (причинности)

$$h[n] = 0 \text{ при } n < 0.$$

Свойство каузальности означает, что отклик цепи не может зависеть от «будущих» значений входной последовательности. Требование каузальности импульсной характеристики представляется естественным для цепей, выполняющих обработку сигнала *в темпе поступления его отсчетов*, т.е. «в реальном времени», когда выходные значения по необходимости могут зависеть лишь от «прошлого» и «настоящего» входного сигнала. Если допускается большая временная задержка выходной последовательности относительно входной, то входная последовательность может быть записана в буферную память и устройство обработки имеет доступ как к «прошлым», так и к «будущим» ее значениям. Более того, во многих приложениях ЦОС, например в обработке изображений, понятия прошлого и будущего вообще теряют смысл, так как аргументы последовательности представляют собой пространственные координаты. В таких случаях импульсная характеристика устройства обработки может не быть каузальной. Иногда каузальность называют физической реализуемостью, что не совсем удачно, так как в указанных *физически реализуемых* системах выполнение условия каузальности не является обязательным.

2.5. ДИСКРЕТНАЯ СВЕРТКА И ЕЕ СВОЙСТВА

Произведя в выражении дискретной свертки (2.3) замену переменных $n - k = m$, $k = n - m$, получим

$$y[n] = \sum_{m=-\infty}^{\infty} x[n-m]h[m] = \sum_{k=-\infty}^{\infty} h[k]x[n-k].$$

Таким образом, установлено, что свертка обладает свойством коммутативности, что символически можно выразить равенством $x \otimes h = h \otimes x$.

Очевидно, это свойство свертки как бинарной (двухместной) операции над последовательностями сближает свертку с такой бинарной операцией, как умножение. (Далее будет показано, что операции свертки и умножения действительно очень близки.)

Коммутативность свертки позволяет сделать следующий очень важный вывод. Входная последовательность и импульсная характеристика ЛИС-цепи играют в анализе *равноправные* роли: отклик цепи с импульсной характеристикой $h[n]$ на последовательность $x[n]$ (рис. 2.9) равен отклику цепи с импульсной характеристикой $x[n]$ на последовательность $h[n]$ (рис. 2.10).

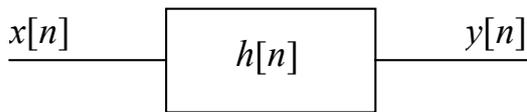


Рис. 2.9. Отклик цепи на последовательность $x[n]$

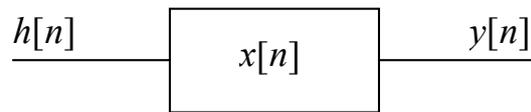


Рис. 2.10. Иллюстрация коммутативности свертки

Этот факт позволяет, в частности, считать *любую* последовательность $x[n]$ импульсной характеристикой некоторой *воображаемой* (гипотетической) цепи, т.е. откликом этой цепи на последовательность $\delta[n]$ (рис. 2.11).

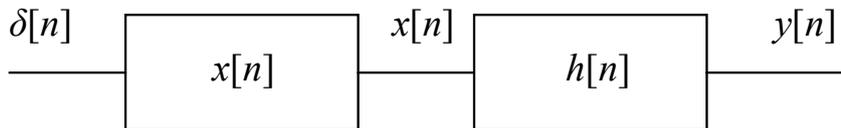


Рис. 2.11. Последовательность $x[n]$ как отклик гипотетической цепи на последовательность $\delta[n]$

Рассмотрим свертку произвольной последовательности $v[n]$ с последовательностью $w[n]$, которая сама равна свертке $w = x \otimes y$:

$$\begin{aligned} w \otimes v &= (x \otimes y) \otimes v = \sum_{k=-\infty}^{\infty} w[k]v[n-k] = \\ &= \sum_{k=-\infty}^{\infty} v[n-k] \left(\sum_{s=-\infty}^{\infty} x[s]y[k-s] \right) = \sum_{s=-\infty}^{\infty} x[s] \left(\sum_{k=-\infty}^{\infty} v[n-k]y[k-s] \right) = \\ &= \sum_{s=-\infty}^{\infty} x[s] \sum_{r=-\infty}^{\infty} v[n-r-s]y[r] = x \otimes (v \otimes y). \end{aligned}$$

С учетом коммутативности можно записать $(x \otimes y) \otimes v = x \otimes (y \otimes v)$.

Таким образом, свертка обладает свойством *ассоциативности*, что углубляет ее сходство с умножением.

Ассоциативность означает, что две цепи, представленные на рис. 2.12, эквивалентны.

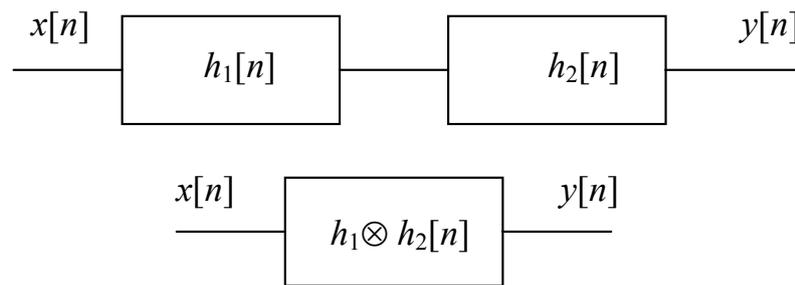


Рис. 2.12. Ассоциативность свертки

Нетрудно убедиться непосредственной подстановкой в том, что свертка удовлетворяет также условиям *дистрибутивности*

$$x \otimes (y + u) = x \otimes y + x \otimes u ,$$

$$\alpha(x \otimes y) = \alpha x \otimes y = x \otimes \alpha y ,$$

где α – комплексный коэффициент.

Благодаря «равноправью» входных последовательностей и импульсных характеристик ЛИС-цепей мы можем рассматривать те и другие как элементы единого пространства последовательностей, которое, во-первых, является *линейным пространством*, а во-вторых – снабжено бинарной операцией свертки, удовлетворяющей аксиомам ассоциативности и дистрибутивности. Такие мате-

матические структуры называются *алгебрами*, а операция (в данном случае свертка) называется обобщенным *умножением*. Если в алгебре есть элемент e , удовлетворяющий условию $x \otimes e = x$ для всех x , то он называется *единицей* алгебры. Легко убедиться в том, что роль единицы для алгебры со сверткой в качестве обобщенного умножения играет последовательность $\delta[n]$.

Итак, мы установили, что линейное пространство последовательностей с операцией свертки в качестве обобщенного умножения представляет собой коммутативную *алгебру с единицей*.

Для завершения построения математической модели ограничим круг рассматриваемых последовательностей *абсолютно суммируемыми* последовательностями, для которых конечна сумма (норма)

$$\|x\|_1 = \sum_{n=-\infty}^{\infty} |x[n]| < \infty.$$

Множество l_1 таких последовательностей является *нормированной алгеброй* относительно свертки. В то же время l_1 – линейное пространство относительно сложения и умножения на комплексный скаляр.

Обычно среди всех ЛИС-цепей выделяют цепи, удовлетворяющие условию устойчивости. Оно означает, что *ограниченная* по амплитуде входная последовательность не должна вызывать *бесконечного* отклика. Заметим, что существуют и другие определения устойчивости. Устойчивость в смысле приведенного определения иногда называют *ОВОВ-устойчивостью* (от слов «ограниченный вход – ограниченный выход»).

Покажем, что если импульсная характеристика имеет конечную l_1 -норму, т.е. является абсолютно суммируемой, то цепь устойчива в указанном смысле. В самом деле, если входная последовательность $x[n]$ ограничена, т.е. $|x[n]| \leq K_0$, где K_0 – некоторое положительное число, то для отклика $y[n]$ цепи с импульсной характеристикой $h[n]$ справедливо выражение

$$|y[n]| = \left| \sum_{k=-\infty}^{\infty} x[n-k]h[k] \right| \leq K_0 \sum_{k=-\infty}^{\infty} |h[k]|,$$

а поскольку $\sum_{k=-\infty}^{\infty} |h[k]| < \infty$, то и $|y[n]| < \infty$. Таким образом, если

$\|h\|_1 < \infty$, то цепь устойчива.

Справедливо и обратное утверждение: если $\|h\|_1 = \infty$, то цепь неустойчива. Доказать это можно, показав, что при $\|h\|_1 = \infty$ ограниченное воздействие (хотя бы и специально подобранное) может вызвать неограниченный отклик. Воздействие вида

$$x[n] = \begin{cases} \frac{h^*[-n]}{|h[-n]|}, & h[-n] \neq 0, \\ 0, & h[-n] = 0, \end{cases}$$

очевидно, ограничено; предполагая, что $\|h\|_1 = \infty$, получаем при $n = 0$ отклик

$$y[0] = \sum_{k=-\infty}^{\infty} x[-k]h[k] = \sum_{k=-\infty}^{\infty} \frac{h[k]h^*[k]}{|h[k]|} = \sum_{k=-\infty}^{\infty} |h[k]| = \infty.$$

Итак, для устойчивости цепи *необходимо и достаточно*, чтобы ее импульсная характеристика была абсолютно суммируемой, т.е. принадлежала l_1 . Можно сказать, что l_1 «состоит из импульсных характеристик устойчивых ЛИС-цепей».

Заметим, что если $\|x\|_1 < \infty$, то $\|x\|_2 < \infty$; обратное в общем случае неверно. Таким образом, пространство l_1 «беднее», чем l_2 . При этом l_1 не является подпространством l_2 , так как в l_1 и l_2 поразному определены нормы. Более того, l_2 – гильбертово пространство, а l_1 – нет.

Пример последовательности, принадлежащей l_2 , но не входящей в l_1 , будет рассмотрен позднее (см. пример 2.9).

2.6. ПРЕДСТАВЛЕНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ПОЛИНОМАМИ И ФОРМАЛЬНЫМИ РЯДАМИ

При всем сходстве операций умножения и свертки последовательностей между ними существует заметное отличие. Умножение является *локальной* операцией, так как произведение двух последовательностей при некотором значении аргумента определяется отсчетами последовательностей-сомножителей при *этом же* значении аргумента. Напротив, результат свертки при любом зна-

чении аргумента получается с учетом значений исходных последовательностей *при всех* значениях аргументов.

Известный в математике прием состоит в том, что математической модели процесса (явления, объекта) ставится в соответствие другая модель, в каком-либо отношении более удобная для исследований. В нашем случае подходящей *моделью* для алгебры l_1 могла бы стать другая алгебра (обозначим ее Z), в которой бинарная алгебраическая операция представляла бы обычное умножение. Соответствие может быть установлено отображением $\Psi: l_1 \rightarrow Z$, которое должно сохранять структурные свойства линейного пространства l_1 и поэтому удовлетворять для всех $x, y \in l_1$, $\alpha \in \mathbb{C}$ условиям

$$\Psi\{x + y\} = \Psi\{x\} + \Psi\{y\},$$

$$\Psi\{\alpha x\} = \alpha \Psi\{x\},$$

$$\Psi\{x \otimes y\} = \Psi\{x\} \Psi\{y\}.$$

Такое отображение называется гомоморфизмом. Если отображение к тому же взаимно однозначно, то оно называется *изоморфизмом*. Тогда существует обратное отображение $\Psi^{-1}: Z \rightarrow l_1$, которое также взаимно однозначно. В таком случае вместо выполнения свертки $x \otimes y$ в алгебре l_1 может оказаться более удобным перейти к образам последовательностей $\Psi\{x\}$ и $\Psi\{y\}$, перемножить их и выполнить обратное отображение, так как

$$x \otimes y = \Psi^{-1}[\Psi\{x\} \cdot \Psi\{y\}].$$

Как будет показано далее, такое отображение на самом деле существует и широко используется как при анализе дискретных цепей и сигналов, так и в практике цифровой обработки сигналов. Рассмотрим вначале случай последовательностей конечной длины: $x[n]$, $n = 0, N$, $y[n]$, $n = 0, M$, где N и M – целые числа.

Поставим в соответствие последовательности $x[n]$ полином

$$X(s) = x[0] + x[1]s + \dots + x[N-1]s^{N-1} + x[N]s^N,$$

а последовательности $y[n]$ – полином

$$Y(s) = y[0] + y[1]s + \dots + y[M-1]s^{M-1} + y[M]s^M$$

относительно некоторой формальной переменной s .

Составим прямоугольную таблицу произведений коэффициентов полиномов (являющихся также значениями исходных последовательностей) так, что произведение $x[k]y[m]$ находится на пересечении k -й строки и m -го столбца (рис. 2.13).

$x[0]y[0]$	$x[0]y[1]$	$x[0]y[2]$...	$x[0]y[M]$
$x[1]y[0]$	$x[1]y[1]$	$x[1]y[2]$...	$x[1]y[M]$
$x[2]y[0]$	$x[2]y[1]$	$x[2]y[2]$...	$x[2]y[M]$
⋮	⋮	⋮		⋮
⋮	⋮	⋮		⋮
$x[N]y[0]$	$x[N]y[1]$	$x[N]y[2]$...	$x[N]y[M]$

Рис. 2.13. Умножение полиномов

Результат умножения $W(s) = X(s)Y(s)$ этих полиномов – снова полином (относительно s) степени $M + N$, причем коэффициент при k -й степени s находится суммированием по k -й диагонали таблицы, так что

$$w[k] = \sum_{m=0}^k x[m]y[k - m].$$

Нетрудно видеть, что коэффициенты полинома-произведения находятся путем вычисления *свертки* последовательностей коэффициентов полиномов-сомножителей.

Поэтому, отображая указанным способом множество последовательностей конечной длины на множество полиномов, мы получаем желаемую модель – алгебру полиномов с *умножением* полиномов вместо *свертки* в исходном множестве последовательностей.

Рассмотрим теперь бесконечные последовательности $x[n]$, $n = -\infty, \infty$, $y[n]$, $n = -\infty, \infty$, и поставим им в соответствие ряды

$$\sum_{n=-\infty}^{\infty} x[n]s^n \tag{2.4}$$

и $\sum_{n=-\infty}^{\infty} y[n]s^n$. Перемножая почленно эти ряды и группируя коэффициенты при одинаковых степенях s , получим ряд

$$\sum_{n=-\infty}^{\infty} w[n]s^n,$$

где
$$w[k] = \sum_{m=-\infty}^{\infty} x[m]y[k-m]. \quad (2.5)$$

Способ суммирования аналогичен изображенному на рис. 2.13, за тем исключением, что теперь таблица имеет неограниченную протяженность по всем направлениям. Очевидно, отображение множества последовательностей бесконечной длины на множество рядов вида (2.4) является гомоморфизмом.

Ряды вида (2.4) называются *формальными*, поскольку не затрагивается вопрос об их сходимости. Сходимость не имеет значения при установлении взаимно однозначного соответствия между отдельно взятой последовательностью и ее представлением в виде формального ряда (2.4). Однако, рассматривая свертку двух последовательностей или произведение двух формальных рядов, видим, что выражение (2.5) содержит сумму бесконечного числа слагаемых, поэтому может оказаться $w[k] = \infty$. Применительно к воздействию сигнала на цепь это означает, что при ограниченном сигнале и ограниченной импульсной характеристике отклик может быть неограниченным (возникает неустойчивость).

Ранее, при обсуждении вопроса устойчивости, было показано, что для ограниченности свертки двух последовательностей достаточно, чтобы хотя бы одна из них принадлежала l_1 , а другая была, по крайней мере, ограниченной. Более полное изучение условий ограниченности свертки (2.5) связано с вопросом о сходимости рядов вида (2.4).

2.7. z-ПРЕОБРАЗОВАНИЕ

Заменив в формальном ряде (2.4) переменную s на z^{-1} , где z – комплексное переменное, получим ряд

$$\sum_{n=-\infty}^{\infty} x[n]z^{-n}.$$

Если этот ряд сходится, то функция комплексного переменного

$$X(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n} \quad (2.6)$$

называется *z-образом* последовательности $x[n]$, $n = -\infty, \infty$, а выражение (2.6) носит название *z-преобразования* (иногда его называют преобразованием Лорана). *z-Преобразование*, если оно существует (ряд в (2.6) сходится), является взаимно однозначным (тогда существует *обратное z-преобразование*), поэтому отображение, осуществляемое выражением (2.6), представляет собой изоморфизм.

Пользуясь изоморфизмом между множеством последовательностей и множеством формальных рядов, можно вместо последовательностей оперировать их *z-образами*, при этом *свертке последовательностей* соответствует более простая в осуществлении операция *умножения z-образов*. В частности, на этом основан широко применяемый на практике метод фильтрации дискретных сигналов, называемый методом быстрой свертки и подробно рассматриваемый в разд. 6. Для того чтобы установить пределы (условия) применимости указанного изоморфизма, необходимо изучить свойства *z-образов* как функций комплексного переменного z . Некоторые из основных понятий теории функций комплексного переменного приведены в *Приложении 1*.

В качестве простейшего примера рассмотрим *z-образ* δ -последовательности. Очевидно,

$$\sum_{n=-\infty}^{\infty} \delta[n] z^{-n} = 1.$$

Этот результат не должен казаться неожиданным. Действительно, в алгебре l_1 именно δ -последовательность играет роль *единицы относительно свертки*, *z-преобразование изоморфно отображает* алгебру l_1 со сверткой на алгебру Z функций (*z-образов*) с умножением, поэтому свертке некоторой последовательности с $\delta[n]$ соответствует умножение *z-образа* этой последовательности на 1.

z-Преобразование сходится не для всех последовательностей и не во всех точках комплексной *z-плоскости*. Совокупность точек *z-плоскости*, в которых сумма (2.6) конечна, называется *областью сходимости z-преобразования* данной последовательности.

При нахождении z -преобразований некоторых последовательностей будем пользоваться известной формулой суммирования геометрической прогрессии

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a} \quad \text{при} \quad |a| < 1.$$

Замечание. В этом выражении знак равенства нельзя понимать буквально, так как левая и правая части суть разные математические объекты, совпадающие не при всех значениях аргумента a . В литературе по цифровой обработке сигналов это обстоятельство обычно не акцентируется, что нередко приводит к определенным затруднениям при изучении теории z -преобразования.

Пример 2.5. Для ступенчатой последовательности

$$u[n] = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0 \end{cases}$$

z -преобразование $\sum_{n=0}^{\infty} z^{-n}$ сходится к функции $U(z) = \frac{1}{1-z^{-1}} = \frac{z}{z-1}$

в области сходимости $|z^{-1}| < 1$, т.е. при $|z| > 1$ (рис. 2.14). ◀

Функция $U(z)$ имеет простой (однократный) полюс (см. Приложение 1) при $z=1$, обозначенный на диаграмме z -плоскости крестиком.

Кроме полюсов для описания поведения функций комплексного переменного z используются такие характерные точки, как нули. Нулем функции $f(z)$ называется такое значение аргумента $z = z_0$, что $f(z_0) = 0$. Функция $U(z)$ имеет, очевидно, один нуль в точке $z = 0$ (на диаграмме нуль обозначается кружком).

Здесь уместно еще раз напомнить о различии между рядом в формуле z -преобразования и z -образом последовательности. В этом примере при $z = 0$ ряд расходится, а функция $U(z)$, представляющая собой z -образ (но лишь в области сходимости!), имеет вполне определенное значение 0.

Пример 2.6. Каузальная экспоненциальная последовательность $x[n] = a^n u[n]$ имеет z -преобразование

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

при $|az^{-1}| < 1$, т.е. при $|z| > |a|$.

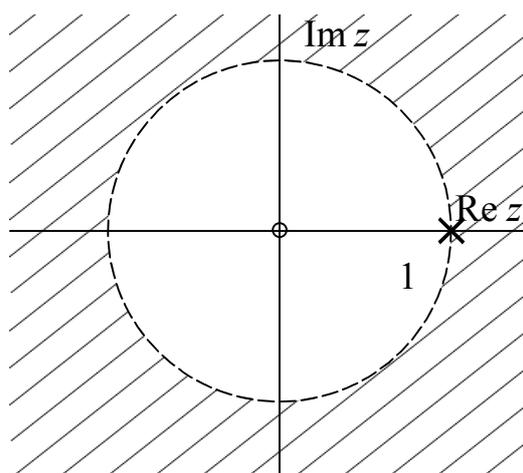


Рис. 2.14. Нуль-полюсная диаграмма функции $U(z)$

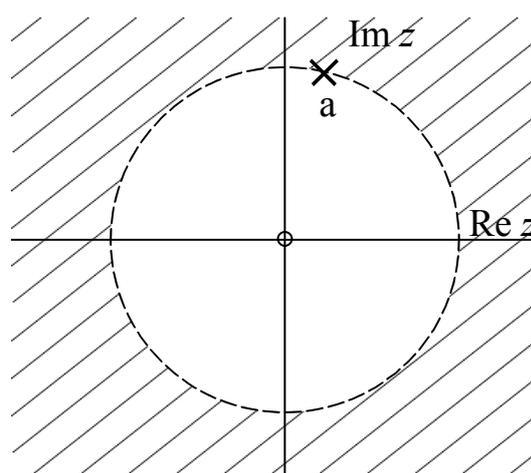


Рис. 2.15. Нуль-полюсная диаграмма z-образа каузальной экспоненциальной последовательности

Область сходимости (рис. 2.15) представляет собой внешность круга радиуса $|a|$. Единственный полюс z-образа $X(z)$ равен a ; единственный нуль находится в точке $z = 0$. ◀

Пример 2.7. Антикаузальная (равная нулю при $n > 0$) экспоненциальная последовательность $y[n] = b^n u[-n]$ имеет z-образ

$$Y(z) = \sum_{n=-\infty}^0 b^n z^{-n} = \sum_{n=0}^{\infty} b^{-n} z^n = \frac{1}{1 - b^{-1}z} = \frac{b}{b - z}$$

с полюсом в точке b и областью сходимости в виде внутренности круга $|z| < |b|$ (рис. 2.16). ◀

Замечание. Форма области сходимости зависит от скорости убывания значений последовательности при $|n| \rightarrow \infty$.

Если последовательность убывает на бесконечности экспоненциально (как в рассмотренных примерах) или быстрее, область сходимости представляет собой открытое (не содержащее своей границы) множество точек z-плоскости. В этом

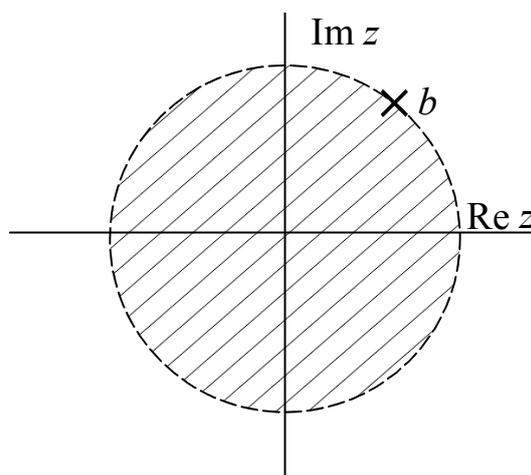


Рис. 2.16. Нуль-полюсная диаграмма z-образа антикаузальной экспоненциальной последовательности

подразделе мы ограничимся рассмотрением только таких «быстро убывающих» последовательностей. Для них справедливы утверждения:

1. Для любой *каузальной* последовательности областью сходимости является *внешность* круга *минимального* радиуса с центром в точке $z = 0$ (не включая его границу!), содержащего *все полюсы* (рис. 2.17);
2. Для *антикаузальной* последовательности область сходимости – *внутренность* *максимального* круга (исключая границу), *не содержащего* полюсов (рис. 2.18).

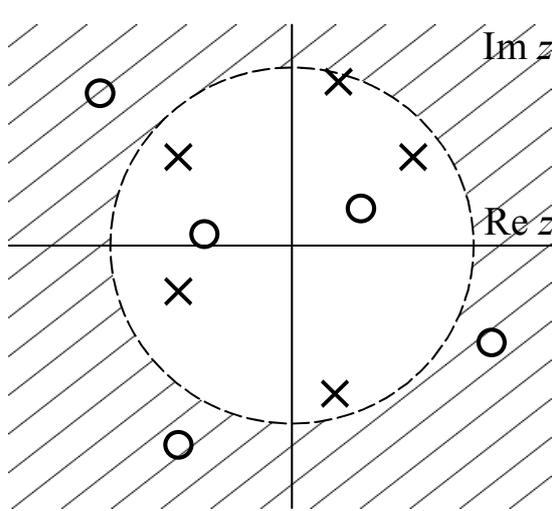


Рис. 2.17. Область сходимости z -преобразования каузальной последовательности

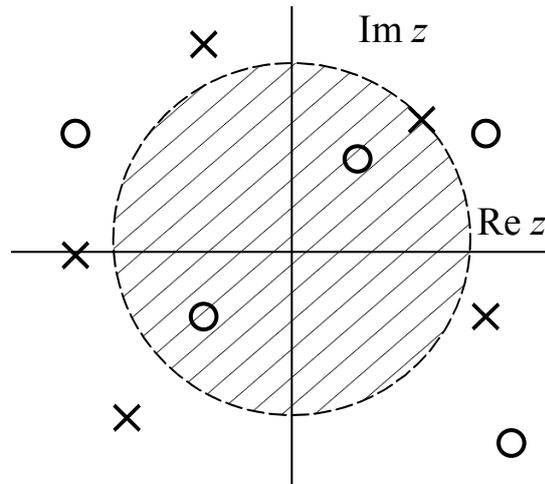


Рис. 2.18. Область сходимости z -преобразования антикаузальной последовательности

Во многих случаях рассматриваемую последовательность можно представить линейной комбинацией (взвешенной суммой) других, более простых последовательностей. При этом следует иметь в виду, что z -преобразование линейно, т.е. удовлетворяет принципу суперпозиции (это легко проверить подстановкой в формулу z -преобразования взвешенной суммы $\alpha x_1[n] + \beta x_2[n]$).

Для того чтобы сходилось z -преобразование суммы, достаточно, чтобы сходились z -преобразования каждого из слагаемых. Отсюда следует, что область сходимости z -преобразования суммы последовательностей содержит *пересечение* областей сходимости z -преобразований всех слагаемых.

Пример 2.8. Некаузальная биэкспоненциальная последовательность, показанная на рис. 2.19, а, описывается выражением

$$w[n] = \begin{cases} a^n, & n \geq 0, \\ b^n, & n < 0. \end{cases}$$

Очевидно, что $w[n] = x[n] + y[n] - \delta[n]$, где $x[n]$ и $y[n]$ – последовательности из примеров 2.6 и 2.7. Пользуясь свойством линейности z-преобразования, запишем z-образ

$$\begin{aligned} W(z) &= \frac{z}{z-a} + \frac{b}{b-z} - 1 = \frac{z}{z-a} + \frac{z}{b-z} = \\ &= \frac{bz - az}{(z-a)(b-z)} = \frac{(a-b)z}{(z-a)(z-b)} \end{aligned}$$

с полюсами a и b , нулем в точке $z = 0$ и областью сходимости в виде кольца $|a| < |z| < |b|$, которое непусто лишь при $|b| > |a|$ (рис. 2.19, б). ◀

В общем случае область сходимости z-преобразования произвольной *быстро убывающей* последовательности имеет вид *открытого* кольца $R_- < |z| < R_+$, где, возможно, $R_- = 0$ или $R_+ = \infty$. Очевидно, существуют последовательности, для которых z-преобразование не сходится нигде ($R_+ \leq R_-$). Такова, например, последо-

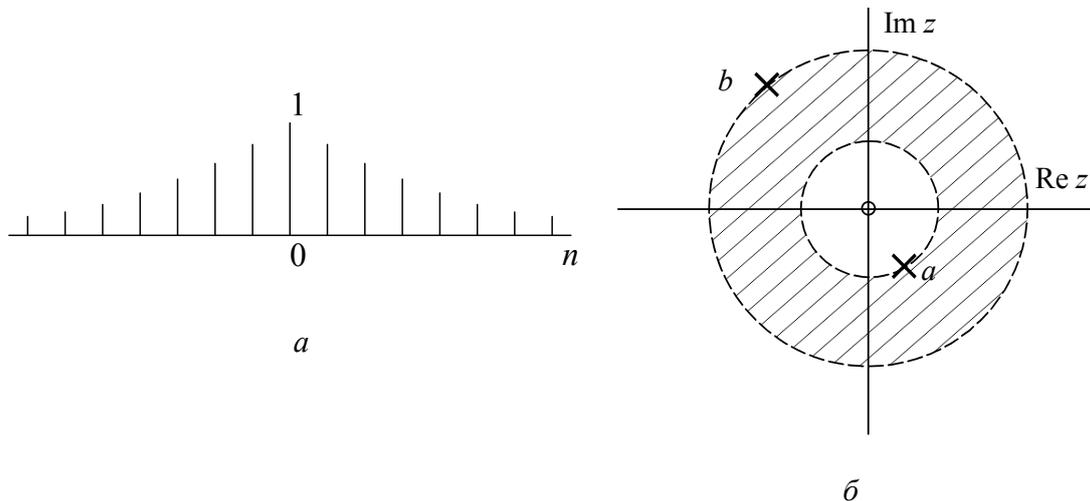


Рис. 2.19. Некаузальная последовательность (а) и область сходимости ее z-преобразования (б)

вательность из примера 2.8 при $|b| < 1 < |a|$, показанная на рис. 2.20 для вещественных a и b .

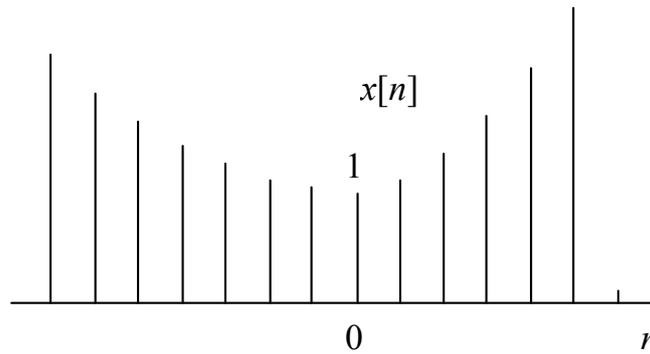


Рис. 2.20. Последовательность с пустой областью сходимости z -преобразования

Если для некоторой последовательности z -преобразование сходится в открытом кольце, то в этом кольце z -образ является аналитической функцией, а формула (2.6) описывает ее представление рядом Лорана. Очевидно, все полюсы z -образа находятся за пределами кольца или на его границах.

Свертке последовательностей соответствует перемножение их z -образов. Чтобы произведение двух функций $X(z)$ и $Y(z)$ было ограниченным, *достаточно*, чтобы каждая из функций-сомножителей была ограничена; поэтому область сходимости произведения *содержит* пересечение областей сходимости сомножителей $X(z)$ и $Y(z)$. Однако условие ограниченности сомножителей не является *необходимым*: если в некоторой точке z -плоскости имеет место полюс одного сомножителя и нуль другого, произведение в этой точке может не иметь особенности. Таким образом, область сходимости произведения может быть *шире* пересечения областей сходимости сомножителей за счет взаимной компенсации нулей и полюсов.

Основные свойства z -преобразования приведены в *Приложении 2*.

2.8. ЕДИНИЧНАЯ ОКРУЖНОСТЬ

Особую роль в анализе дискретных сигналов и цепей играет поведение z -образов последовательностей на единичной окружности (1-окружности) – окружности радиуса 1 с центром в точке $z = 0$.

Очевидно, на единичной окружности z -преобразование некоторой последовательности

$$X(z) \Big|_{|z|=1} = \sum_{n=-\infty}^{\infty} x[n] z^{-n} \Big|_{|z|=1}$$

сходится, если сходится ряд

$$\sum_{n=-\infty}^{\infty} |x[n] z^{-n}| \Big|_{|z|=1} = \sum_{n=-\infty}^{\infty} |x[n]|.$$

Итак, для любой *абсолютно суммируемой* последовательности z -преобразование *сходится на 1-окружности*, и наоборот, если область сходимости z -преобразования последовательности включает 1-окружность, то последовательность *абсолютно суммируема*. Следовательно, устойчивость ЛИС-цепи прямо связана с поведением z -образа ее импульсной характеристики на единичной окружности.

Таким образом, l_1 – это пространство (а также алгебра со сверткой в качестве умножения) всех *абсолютно суммируемых* последовательностей, т. е. последовательностей, z -преобразования которых сходятся на 1-окружности. Множество Z функций комплексного переменного z , представимых сходящимися на 1-окружности z -преобразованиями, есть алгебра с обычным умножением, *изоморфная* l_1 , причем изоморфизм устанавливается формулами прямого и обратного z -преобразований. Поэтому алгебра функций Z может служить моделью алгебры l_1 . Это означает, что любому истинному утверждению относительно функций из множества Z соответствует истинное утверждение относительно соответствующих последовательностей из l_1 , при этом только слово «умножение» нужно заменить словом «свертка».

Пользуясь этим обстоятельством, мы, например, сразу можем утверждать, что результат свертки абсолютно суммируемых последовательностей снова абсолютно суммируем (так как пересечение областей сходимости, содержащих 1-окружность, очевидно, также ее содержит). Тем самым подтверждается тот факт, что пространство l_1 *замкнуто* относительно операции свертки.

Ранее мы рассматривали последовательности, z -преобразования которых сходятся в кольце $R_- < |z| < R_+$. Чтобы такие последовательности принадлежали l_1 , очевидно, должно выполняться усло-

вие $R_- < 1 < R_+$, т.е. кольцо сходимости должно включать 1-окружность. Следует отметить, что в l_1 имеется целый класс последовательностей, z -преобразования которых сходятся *только* на 1-окружности. Эти последовательности убывают на бесконечности (при $|n| \rightarrow \infty$) медленнее экспоненциальных, но являются абсолютно суммируемыми.

Пример 2.9. Последовательность $x[n] = \frac{\sin^2(\omega_0 n)}{(\omega_0 n)^2}$, где ω_0 – вещественное число, имеет скорость убывания $1/n^2$, поэтому является абсолютно суммируемой. Ряд $X(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$ равномерно сходится при $|z|=1$ и расходится при $|z| \neq 1$. Пределом суммы ряда при $|z|=1$ является непрерывная функция, определенная только на

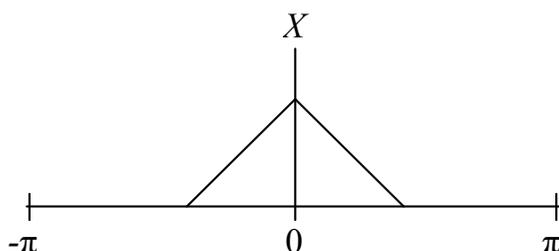


Рис. 2.21. Непрерывная функция – z -образ последовательности, описанной в примере 2.9

единичной окружности и показанная на рис. 2.21 графиком, по оси абсцисс которого откладывается угол ω , указывающий положение точки z на 1-окружности. ◀

Вообще, для того чтобы последовательность была абсолютно суммируемой, она должна убывать при $|n| \rightarrow \infty$ со скоростью не менее, чем $1/n^v$, $v > 1$ (см., например, [12]).

2.9. ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Поскольку во многих случаях интерес представляет поведение z -преобразования не на всей z -плоскости, а только на единичной окружности, то можно в самом выражении z -преобразования задать 1-окружность в качестве области изменения комплексного

переменного z . Рассмотрим z -преобразование абсолютно суммируемой последовательности $x[n]$ на 1-окружности. Переменная z , принадлежащая 1-окружности, в показательной форме имеет вид $z = \exp(j\omega)$, где ω – угол между вещественной осью z -плоскости и вектором, изображающим комплексное число z . Подставив $z = \exp(j\omega)$ в выражение (2.6), получим

$$X(z) \Big|_{z=\exp(j\omega)} = X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}. \quad (2.7)$$

Ряд в правой части выражения (2.7) сходится для *абсолютно суммируемой* последовательности *равномерно* к *непрерывной* функции аргумента ω .

Полученное выражение представляет собой *преобразование Фурье* последовательности $x[n]$ – *сужение* z -преобразования на 1-окружность. С другой стороны, выражение (2.7) описывает разложение непрерывной функции $X(e^{j\omega})$, определенной на 1-окружности, т.е. при $-\pi < \omega \leq \pi$, в ряд Фурье по комплексным экспонентам $\{\exp(-j\omega n), n = -\infty, \infty\}$, рассматриваемым при каждом n , как функции переменной ω . Тогда значения $x[n]$ – не что иное, как *коэффициенты* этого разложения. Поэтому последовательность $x[n]$ определяется выражением для нахождения коэффициентов ряда Фурье

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega. \quad (2.8)$$

Это выражение называется *обратным преобразованием Фурье* функции $X(e^{j\omega})$. Переменная ω как параметр комплексной экспоненциальной последовательности $e^{j\omega n}$ имеет смысл *круговой частоты*; функцию $X(e^{j\omega})$, входящую в (2.8) и определяемую выражением (2.7), называют *спектральной плотностью* последовательности $x[n]$.

Обратное преобразование Фурье (2.8) задает представление последовательности $x[n]$ в виде *суперпозиции* (интегральной суммы) «сплошного» множества комплексных экспоненциальных последо-

вательностей $\{(e^{j\omega n}, n = -\infty, \infty, -\pi < \omega \leq \pi)\}$ с весами («амплитудами»), определяемыми спектральной плотностью $X(e^{j\omega})$.

Предположим, что комплексная экспоненциальная последовательность $x[n] = e^{j\omega_0 n}$ с частотой ω_0 поступает на вход ЛИС-цепи с импульсной характеристикой $h[n]$. Тогда выходная последовательность

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k] x[n-k] = \sum_{k=-\infty}^{\infty} h[k] e^{j\omega_0(n-k)} = \\ &= \sum_{k=-\infty}^{\infty} h[k] e^{j\omega_0 n} e^{-j\omega_0 k} = e^{j\omega_0 n} \sum_{k=-\infty}^{\infty} h[k] e^{-j\omega_0 k} = e^{j\omega_0 n} H(e^{j\omega_0}). \end{aligned}$$

Очевидно, прохождение последовательности $e^{j\omega_0 n}$ через ЛИС-цепь сопровождается лишь ее масштабным изменением – умножением на некоторое комплексное число $H(e^{j\omega_0})$, определяемое значением ω_0 .

Функция

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h[n] e^{-j\omega n},$$

определенная через преобразование Фурье импульсной характеристики $h[n]$ цепи, называется ее *комплексной частотной характеристикой* (КЧХ).

Функция $H(z) = \sum_{n=-\infty}^{\infty} h[n] z^{-n}$ называется *передаточной функцией* цепи. Комплексная частотная характеристика может рассматриваться как *сужение* передаточной функции на 1-окружность.

Вектор, который под воздействием данного оператора лишь умножается на комплексное число, называется *собственным вектором* оператора, а упомянутое число – *собственным значением*, соответствующим этому собственному вектору. Таким образом, комплексная экспоненциальная последовательность $e^{j\omega n}$ является *собственной* последовательностью для оператора *любой* ЛИС-цепи с собственным значением $H(e^{j\omega n})$, вообще говоря комплексным. Следует подчеркнуть, что это справедливо при любом значении

частоты ω , поэтому совокупность $\{(e^{j\omega n}, n = \overline{-\infty, \infty}), -\pi < \omega \leq \pi\}$ всех последовательностей $e^{j\omega n}$ при всевозможных значениях частоты представляет собой базис пространства l_2 , образованный собственными векторами ЛИС-оператора.

Итак, формула обратного преобразования Фурье (2.8) определяет разложение произвольной последовательности $x[n]$ в *собственном* (для оператора любой ЛИС-цепи) *базисе*, образованном всеми последовательностями $\{(e^{j\omega n}, n = \overline{-\infty, \infty}), -\pi < \omega \leq \pi\}$; тогда выходная последовательность складывается из тех же слагаемых, но только домноженных на $H(e^{j\omega})$:

$$y[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) H(e^{j\omega}) e^{j\omega n} d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} Y(e^{j\omega}) e^{j\omega n} d\omega,$$

где $Y(e^{j\omega}) = X(e^{j\omega})H(e^{j\omega})$. Таким образом, спектральная плотность $Y(e^{j\omega})$ выходной последовательности находится умножением спектральной плотности входной последовательности на комплексную частотную характеристику. В этом состоит *спектральный метод анализа* дискретных ЛИС-цепей.

Комплексную частотную характеристику цепи и спектральную плотность последовательности можно рассматривать как сужения соответствующих z -образов на 1-окружность. С другой стороны, зная КЧХ цепи или спектральную плотность последовательности, если они зависят от частоты только через $e^{j\omega}$ и являются *аналитическими* функциями, можно получить соответственно передаточную функцию цепи и z -образ последовательности заменой $e^{j\omega} = z$ (эта операция называется *аналитическим продолжением*).

Замечание. Обратное преобразование Фурье (2.8) можно применить и к разрывной функции переменной ω , если соответствующие интегралы конечны. Однако получаемая при этом последовательность коэффициентов ряда не будет *абсолютно суммируемой*, и ряд (2.7) поэтому не будет *равномерно* сходиться. Это объясняется тем, что ряд (2.7) можно представить как предел

при $N \rightarrow \infty$ конечной суммы $S_N = \sum_{n=-N}^N x[n] e^{-j\omega n}$; конечная сумма при любом N представляет собой непрерывную функцию, поэто-

му сходимость к разрывной функции не может быть равномерной [10]. В некоторых случаях все же пользуются рядом (2.7), понимая его как ряд, сходящийся в *среднеквадратическом* смысле. При этом следует отдавать себе отчет в том, что функция частоты $X(e^{j\omega})$ не аналитическая и, следовательно, не может быть аналитически продолжена на всю z -плоскость заменой $e^{j\omega} = z$.

Пример 2.10. Идеальный *фильтр нижних частот* (ФНЧ) с П-образной КЧХ ($\omega_{\text{ср}}$ – частота среза)

$$H_{\Pi}(e^{j\omega}) = \begin{cases} 1, & |\omega| \leq \omega_{\text{ср}}, \\ 0, & \omega_{\text{ср}} < |\omega| < \pi, \end{cases}$$

показанной на рис. 2.22, а, имеет импульсную характеристику

$$h_{\Pi}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_{\Pi}(e^{j\omega}) e^{j\omega n} d\omega = \frac{1}{\pi} \int_0^{\omega_{\text{ср}}} \cos \omega n d\omega = \frac{\omega_{\text{ср}}}{\pi} \frac{\sin \omega_{\text{ср}} n}{\omega_{\text{ср}} n},$$

$$n = -\infty, \infty,$$

изображенную на рис. 2.22, б. ◀

Эта импульсная характеристика убывает на бесконечности со скоростью $1/|n|$, поэтому она не является абсолютно суммируемой. Следовательно, такой фильтр, будь он реализован, был бы *неустойчивым*. Кроме того, реализация такого фильтра потребовала бы бесконечного объема памяти для учета бесконечного числа как прошлых, так и будущих значений входной последовательности.

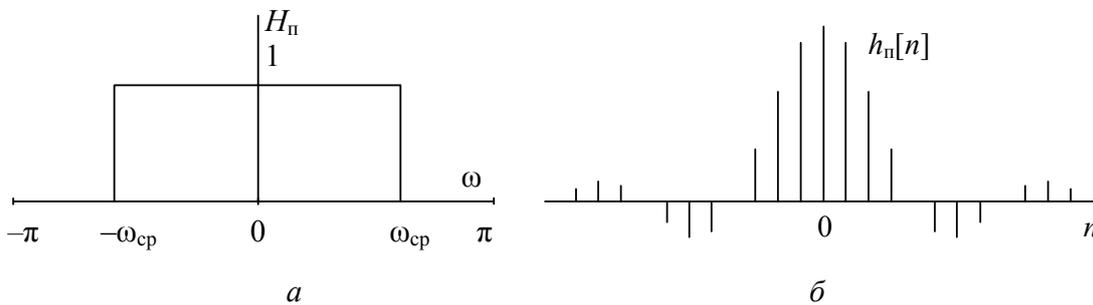


Рис. 2.22. Амплитудно-частотная (а) и импульсная (б) характеристики идеального ФНЧ

Таким образом, идеальный фильтр с П-образной амплитудно-частотной характеристикой не может быть реализован *точно*. Приближенная реализация возможна, если отбросить «хвосты» импульсной характеристики, оставив конечное число отсчетов. Такая *усеченная* импульсная характеристика, очевидно, абсолютно суммируема.

Комплексная частотная характеристика фильтра с усеченной импульсной характеристикой

$$H_N(e^{j\omega}) = \sum_{n=-N}^N h[n] e^{-j\omega n}$$

как конечная сумма ряда Фурье является аналитической функцией и при $N \rightarrow \infty$ сходится к идеальной (П-образной) КЧХ в среднеквадратическом смысле, при этом КЧХ «усеченного» фильтра содержит выбросы (осцилляции) определенной амплитуды, которая не уменьшается с ростом N (рис. 2.23).

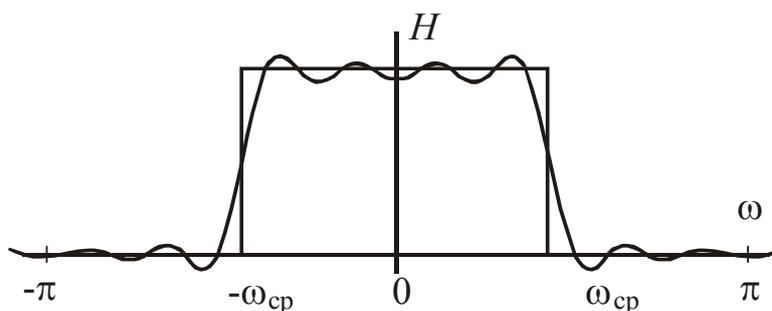


Рис. 2.23. Явление Гиббса

Такое поведение конечной суммы ряда Фурье при увеличении числа слагаемых известно как *явление Гиббса*. Оно обусловлено, по существу, слишком медленным (со скоростью $1/|n|$) убыванием отсчетов импульсной характеристики при $|n| \rightarrow \infty$. Импульсная характеристика П-образного ФНЧ является не абсолютно, а лишь *квадратично* суммируемой, т. е. обладает конечной l_2 -нормой:

$\sum_{n=-\infty}^{\infty} |h[n]|^2 < \infty$. Соответствующий ряд Фурье *не сходится равномерно* ни к какому пределу, хотя он сходится к П-образной КЧХ в *среднеквадратическом* смысле. Такая сходимость гарантирует стремление к нулю среднеквадратического отклонения ε^2 комплексной

частотной характеристики $H_N(e^{j\omega})$ фильтра с усеченной импульсной характеристикой от идеальной П-образной КЧХ $H(e^{j\omega})$:

$$\varepsilon^2 = \int_{-\pi}^{\pi} [H_N(e^{j\omega}) - H(e^{j\omega})]^2 d\omega \rightarrow 0.$$

Стремление к нулю уровня осцилляций имело бы место при *равномерной* сходимости ряда (2.7), для чего требуется более быстрое убывание отсчетов импульсной характеристики с ростом $|n|$. Поэтому для уменьшения гиббсовских осцилляций комплексной частотной характеристики при реализации фильтра усеченную импульсную характеристику умножают на дополнительную весовую последовательность («окно») конечной длины, более или менее быстро убывающую с ростом $|n|$. В таком случае вид комплексной частотной характеристики фильтра зависит от вида окна, подбором которого можно весьма существенно ослабить осцилляции КЧХ (но только ценой уменьшения крутизны спада характеристики в окрестности частот среза $\pm\omega_{cp}$). Подробнее этот вопрос рассмотрен в разд. 6.

2.10. ДИСКРЕТНОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Если последовательности $x[n]$ и $h[n]$ заданы в виде замкнутых выражений («целиком») и существуют их z -преобразования (сходятся соответствующие ряды), то свертка $x \otimes h$ может быть найдена как результат обратного z -преобразования, примененного к произведению z -образов $X(z)H(z)$. Можно для той же цели воспользоваться прямым и обратным преобразованиями Фурье, если они существуют.

Однако если последовательность $x[n]$ задана своими значениями, как это бывает при обработке реальных сигналов, замкнутое выражение для $x[n]$ неизвестно и найти z -образ $X(z)$ как сумму ряда (2.6) и, следовательно, как функцию от z невозможно. В таком случае встает вопрос о *вычислении* z -преобразования для произвольной последовательности. Очевидно, что *вычислить* z -преобразование можно только для последовательности *конечной* длины и лишь в *конечном* множестве точек z -плоскости (в против-

ном случае потребовалось бы выполнить *бесконечное* количество арифметических операций).

Рассмотрим последовательность $x[n]$ конечной длины N такую, что $x[n]=0$ при $n < 0$ и $n \geq N$. Для такой последовательности z -образ

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

представляет собой полином от z^{-1} степени $N-1$.

Для того чтобы однозначно задать полином степени $N-1$, например, с помощью интерполяционной формулы Лагранжа, достаточно указать значения

этого полинома в N различных точках z -плоскости. Отсюда ясно, что для точного задания z -образа последовательности, содержащей N отсчетов, достаточно вычислить z -преобразование в N различных точках z -плоскости, причем расположение этих точек можно выбрать произвольно. Главным мотивом в выборе точек является удобство вычисления прямого и обратного z -преобразований. С этой точки зрения наиболее выгодно разместить N точек равномерно на 1-окружности (рис. 2.24, здесь $N=6$). Таким образом, z -преобразование вычисляется в точках

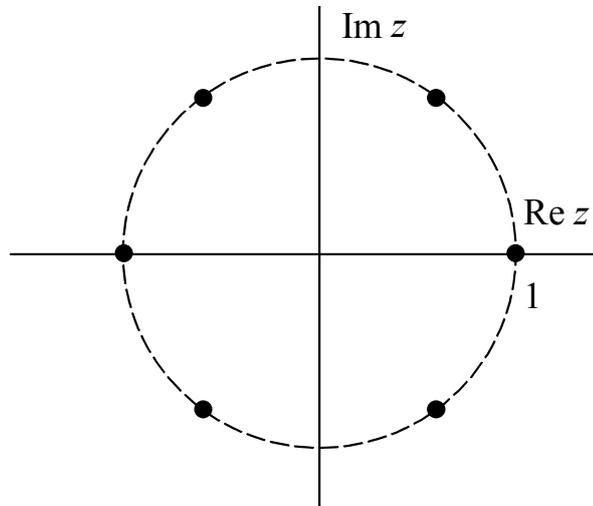


Рис. 2.24. Точки вычисления дискретного преобразования Фурье для $N=6$

$$z_k = e^{j\frac{2\pi}{N}k}, \quad k = \overline{0, N-1}.$$

Значения z -преобразования в точках z_k

$$X(z_k) = X\left(e^{j\frac{2\pi}{N}k}\right) = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}$$

при $k = \overline{0, N-1}$ образуют последовательность конечной длины, поэтому обозначим $X(z_k) = X[k]$, тогда

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}, \quad k = \overline{0, N-1}. \quad (2.9)$$

Это выражение носит название дискретного преобразования Фурье (ДПФ). Обратное ДПФ (ОДПФ) определяется выражением (см. Приложение 3)

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi}{N}kn}, \quad n = \overline{0, N-1}. \quad (2.10)$$

Нетрудно видеть, что ОДПФ представляет собой разложение последовательности $x[n]$ длины N в базисе, образованном N орто-

гональными последовательностями $\left\{ \varphi_k[n] = \frac{1}{N} e^{j\frac{2\pi}{N}kn}, k = \overline{0, N-1} \right\}$

такой же длины. Значения $X[k]$ являются коэффициентами этого разложения.

Скалярное произведение пары базисных последовательностей

$$\begin{aligned} (\varphi_k, \varphi_m) &= \sum_{n=0}^{N-1} \frac{1}{N} e^{j\frac{2\pi}{N}kn} \frac{1}{N} e^{-j\frac{2\pi}{N}mn} = \\ &= \frac{1}{N^2} \sum_{n=0}^{N-1} e^{j\frac{2\pi}{N}(k-m)n} = \frac{1}{N} \delta_{km}, \end{aligned}$$

где δ_{km} – символ Кронекера.

Видно, что базис ортогонален, но не ортонормален: норма базисной последовательности равна $|\varphi_k| = \frac{1}{\sqrt{N}} \quad \forall k$. Выполнив нормировку

$$\psi_k[n] = \sqrt{N} \varphi_k[n], \quad k = \overline{0, N-1},$$

получаем ортонормальный базис

$$\left\{ \left(\psi_k[n] = \frac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}kn}, n = \overline{0, N-1} \right), k = \overline{0, N-1} \right\}.$$

Очевидно, $(\psi_k, \psi_m) = \delta_{km}$, что и означает ортонормальность базиса $\{\psi_k, k = \overline{0, N-1}\}$.

По отношению к этому базису прямое и обратное ДПФ принимают симметричную форму

$$X_\psi[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}, \quad k = \overline{0, N-1},$$

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_\psi[k] e^{j\frac{2\pi}{N}kn}, \quad n = \overline{0, N-1}.$$

Поскольку эта пара ДПФ–ОДПФ соответствует ортонормальному базису, можно записать равенство Парсеваля

$$\sum_{n=0}^{N-1} |x[n]|^2 = \sum_{k=0}^{N-1} |X_\psi[k]|^2. \quad (2.11)$$

Для неортонормального базиса $\{\phi_k[n], k = \overline{0, N-1}\}$ равенство Парсеваля имеет вид $\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2$. Это следует из выражения (2.11), если учесть, что $X_\psi[k] = \frac{1}{\sqrt{N}} X[k]$. Тот же результат можно получить прямым вычислением энергии последовательности:

$$\begin{aligned} \sum_{n=0}^{N-1} |x[n]|^2 &= \sum_{n=0}^{N-1} x[n] x^*[n] = \\ &= \sum_{n=0}^{N-1} \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi}{N}kn} \frac{1}{N} \sum_{m=0}^{N-1} X^*[m] e^{-j\frac{2\pi}{N}mn} = \\ &= \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} X[k] X^*[m] \sum_{n=0}^{N-1} e^{j\frac{2\pi}{N}n(k-m)} = \\ &= \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} X[k] X^*[m] N \delta_{km} = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2. \end{aligned}$$

Несмотря на преимущества ортонормального базиса перед ненормированным ортогональным, в настоящее время в литературе чаще встречается пара ДПФ–ОДПФ в форме (2.9), (2.10). Иногда используют выражения, аналогичные (2.9), (2.10), за тем исключением, что нормирующий множитель $1/N$ входит в выражение прямого ДПФ, а не обратного (при этом коэффициент $X[0]$ имеет физический смысл *постоянной составляющей* исходной последовательности $x[n]$).

Итак, прямое и обратное ДПФ устанавливают взаимно однозначное соответствие между последовательностью конечной длины N и ее z -образом, заданным значениями в N точках z -плоскости. Более того, полученные выражения не просто устанавливают соответствие, но дают *способ вычисления* и, следовательно, оказываются полезными при *реализации* алгоритмов цифровой обработки сигналов. Что особенно важно, существуют весьма эффективные (так называемые *быстрые*) алгоритмы реализации ДПФ – ОДПФ, широко применяемые в практике цифровой обработки сигналов.

Обратное ДПФ (2.10), как было показано, определяет разложение последовательности конечной длины N в базисе, образованном комплексными экспоненциальными последовательностями

$$\left\{ \left(\phi_k[n] = \frac{1}{N} e^{j \frac{2\pi}{N} kn}, n = \overline{0, N-1} \right), k = \overline{0, N-1} \right\}.$$

Базисные последовательности имеют частоты $\omega_k = \frac{2\pi}{N} k$, $k = \overline{0, N-1}$, находящиеся в рациональном соотношении с величиной 2π , поэтому все они являются периодическими. Нетрудно заметить, что выражения для базисных последовательностей при расширении области определения (при $n < 0$ и $n > N-1$) описывают периодические последовательности бесконечной протяженности, при этом величина N равна или кратна их периодам. Поэтому и выражение ОДПФ, рассматриваемое на всей временной оси, можно понимать как представление периодической последовательности $\tilde{x}[n]$ в базисе периодических последовательностей бесконечной длины

$$\left\{ \left(\tilde{\phi}_k[n] = \frac{1}{N} e^{j \frac{2\pi}{N} kn}, n = \overline{-\infty, \infty} \right), k = \overline{0, N-1} \right\}.$$

Последовательность $\tilde{x}[n]$ совпадает при $n = \overline{0, N-1}$ с последовательностью $x[n]$ так же, как все N базисных последовательностей $\tilde{\varphi}_k$ совпадают при $n = \overline{0, N-1}$ с соответствующими последовательностями конечной длительности φ_k . Ясно, что $\tilde{x}[n]$ может принимать всего N различных значений и поэтому полностью определяется N значениями $X[k]$, $k = \overline{0, N-1}$. Величина N либо равна, либо кратна периоду последовательности $\tilde{x}[n]$.

С другой стороны, значения $X[k]$ определяются z -преобразованием последовательности $x[n]$ (или $\tilde{x}[n]$) в точках, *равномерно* расположенных на 1-окружности. Поэтому можно считать, что $X[k]$, $k = \overline{0, N-1}$ представляет собой лишь фрагмент периодической (с периодом N) последовательности $\tilde{X}[k]$, $k = \overline{-\infty, +\infty}$, получаемой периодическим обходом тех же N точек 1-окружности.

Используя ДПФ–ОДПФ, всегда необходимо иметь в виду указанную периодичность последовательностей – как «временных», так и частотных. Как при дискретизации аналогового сигнала, финитность (ограниченность полосы частот) спектра является условием отсутствия наложения спектральных «хвостов» при периодическом повторении спектральной плотности (см. разд. 6), так и при использовании ДПФ – ОДПФ нужно учитывать возможность перекрытия последовательностей во временной и частотной областях при их периодическом продолжении. Особенно важно учитывать возможность такого наложения при реализации алгоритмов цифровой фильтрации в частотной области методом быстрой свертки, которые рассмотрены в разд. 6.

Переход от спектральной плотности к последовательности отсчетов является обратимым. Пусть $x[n]$, $n = \overline{0, N-1}$, – последовательность длины N ; ей соответствует z -образ

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

Обозначая $a = e^{j\frac{2\pi}{N}k} z^{-1}$ и принимая во внимание, что

$$\begin{aligned} \sum_{n=0}^{N-1} a^n &= \sum_{n=0}^{\infty} a^n - \sum_{n=N}^{\infty} a^n = \sum_{n=0}^{\infty} a^n - \sum_{k=0}^{\infty} a^k a^N = \\ &= \frac{1}{1-a} - \frac{a^N}{1-a} = \frac{1-a^N}{1-a}, \end{aligned}$$

получим

$$X(z) = \sum_{k=0}^{N-1} \frac{X[k]}{N} \frac{1-z^{-N}}{1-e^{j\frac{2\pi}{N}k} z^{-1}}.$$

Это выражение определяет z -образ последовательности длины N через заданные его значения в N точках z -плоскости (на 1-окружности). Подставив $z = e^{j\omega}$, получим

$$\begin{aligned} X(e^{j\omega}) &= \sum_{k=0}^{N-1} \frac{X[k]}{N} \frac{1-e^{-j\omega N}}{1-e^{-j\left(\omega-k\frac{2\pi}{N}\right)}} = \\ &= \sum_{k=0}^{N-1} \frac{X[k]}{N} \frac{e^{-j\frac{\omega N}{2}} \sin\left(\frac{\omega N}{2}\right)}{e^{-j\left(\frac{\omega}{2}-k\frac{\pi}{N}\right)} \sin\left(\frac{\omega}{2}-k\frac{\pi}{N}\right)}. \end{aligned}$$

Это выражение позволяет восстановить (интерполировать) спектральную плотность как функцию непрерывной частоты ω по N известным ее отсчетам $X[k]$, $k=0, N-1$. Таким образом, полученное выражение является аналогом известного ряда Котельникова (см., например, [13]).

В некоторых случаях полезной оказывается матричная (операторная) интерпретация дискретного преобразования Фурье. С этой точки зрения дискретное преобразование Фурье представляет собой линейное преобразование вектора \mathbf{x} , составленного из отсчетов исходной последовательности $(x[0], x[1], \dots, x[N-1])^T$, где $(\cdot)^T$ – символ транспонирования, в вектор \mathbf{X} спектральных отсчетов

$(X[0], X[1], \dots, X[N-1])^T$. Выражение (2.9) описывает при каждом $k = \overline{0, N-1}$ вычисление одной k -й компоненты результирующего вектора спектральных отсчетов $(X[0], X[1], \dots, X[N-1])^T$. Таким образом, преобразование исходного вектора в результирующий может быть описано матричным выражением $\mathbf{X} = \mathbf{W} \cdot \mathbf{x}$, которое в развернутом виде имеет форму

$$\begin{pmatrix} X[0] \\ X[1] \\ X[2] \\ \dots \\ X[N-1] \end{pmatrix} = \begin{pmatrix} W_N^0 & W_N^0 & \dots & \dots & W_N^0 \\ W_N^0 & W_N^1 & \dots & \dots & W_N^{N-1} \\ W_N^0 & W_N^2 & \dots & \dots & W_N^{2(N-1)} \\ \dots & \dots & \dots & \dots & \dots \\ W_N^0 & W_N^{N-1} & & & W_N^{(N-1)^2} \end{pmatrix} \begin{pmatrix} x[0] \\ x[1] \\ x[2] \\ \dots \\ x[N-1] \end{pmatrix}.$$

В этом выражении использовано общепринятое обозначение

$$W_N = e^{-j\frac{2\pi}{N}}.$$

Итак, прямое дискретное преобразование Фурье определяется квадратной матрицей \mathbf{W} , составленной из различных степеней единственного комплексного числа W_N . Заметим, что k -й отсчет ДПФ-спектра $X[k]$ находится как скалярное произведение k -й строки матрицы \mathbf{W} и исходного вектора \mathbf{x} . Анализ строк матрицы \mathbf{W} показывает, что каждая строка представляет собой отрезок длины N комплексной экспоненциальной последовательности с частотой, определяемой номером строки. Так, первая строка при $k=0$ составлена из единиц, строка, соответствующая $k=1$, представляет собой один период комплексной экспоненциальной по-

следовательности $e^{j\frac{2\pi}{N}n}$, $n = \overline{0, N-1}$, и т. д. Заметим, что все строки матрицы \mathbf{W} взаимно ортогональны. То же справедливо и в отношении столбцов, так как матрица \mathbf{W} симметрична (инвариантна к транспонированию). Как известно, линейный оператор, действующий на вектор в конечномерном пространстве, определяет координаты вектора в новом базисе, составленном из столбцов матрицы, описывающей этот оператор. Следовательно, дискретное преобразование Фурье – это преобразование вектора к новому ортогональному базису. Обратное преобразование описывается об-

ратной матрицей \mathbf{W}^{-1} , которая также состоит из ортогональных столбцов. Элементы обратной матрицы являются комплексно-сопряженными по отношению к элементам матрицы \mathbf{W} , кроме того, учитывается нормирующий коэффициент $1/N$:

$$\mathbf{W}^{-1} = \frac{1}{N} \begin{pmatrix} W_N^0 & W_N^0 & \dots & \dots & W_N^0 \\ W_N^0 & W_N^{-1} & \dots & \dots & W_N^{-(N-1)} \\ W_N^0 & W_N^{-2} & \dots & \dots & W_N^{-2(N-1)} \\ \dots & \dots & \dots & \dots & \dots \\ W_N^0 & W_N^{-(N-1)} & \dots & \dots & W_N^{-(N-1)^2} \end{pmatrix}.$$

Новый базис можно сделать ортонормальным, умножив матрицу \mathbf{W} на нормирующий коэффициент $1/\sqrt{N}$, тогда такой же коэффициент должен быть применен и к матрице обратного преобразования. В этом последнем случае матрицы прямого и обратного преобразований удовлетворяют условию *эрмитовой сопряженности* $\mathbf{W}^{-1} = (\mathbf{W}^*)^T$. Столбцы каждой из матриц взаимно ортогональны и нормированы, поэтому матрицы описывают преобразования между двумя ортонормальными базисами. Такие преобразования называются *унитарными*. Для унитарных преобразований выполняются равенство Парсеваля и обобщенная формула Рэлея.

Упражнения

1. Убедитесь непосредственной проверкой в том, что множество всех последовательностей с вещественными значениями замкнуто относительно сложения и умножения на вещественный скаляр, но не замкнуто относительно умножения на комплексное число.

2. Проверьте линейную независимость любого набора последовательностей из множества $\{\delta[n-k], k = \overline{-\infty, \infty}\}$.

3. Убедитесь в том, что при различной нумерации исходной системы векторов $\{v_i, i = \overline{1, \infty}\}$ процедура Грама – Шмидта приводит к различным базисам.

4. Докажите, что роль единицы для алгебры l_1 со сверткой в качестве обобщенного умножения играет δ -последовательность $\delta[n]$.

5. Докажите справедливость выражения (2.1).
6. Докажите, что последовательность, принадлежащая l_1 , принадлежит также l_2 .
7. Докажите, что l_2 -норма в трехмерном пространстве совпадает с обычным геометрическим определением длины (модуля) вектора.
8. Докажите, что отображение множества последовательностей бесконечной длины на множество рядов вида (2.4) является гомоморфизмом.
9. Докажите коммутативность свертки.
10. Докажите дистрибутивность свертки.
11. Докажите линейность z -преобразования.
12. Докажите, что область сходимости суммы последовательностей содержит пересечение областей сходимости всех слагаемых.
13. Запишите выражения для последовательностей, соответствующих трем вариантам выбора области сходимости z -преобразования, показанным на рис. П 2.2.
14. Докажите, что z -преобразование последовательности из примера 2.9 расходится при любом z , не лежащем на 1-окружности.
15. Определите параметры функции, показанной на рис. 2.21 (максимальное значение функции и значение аргумента, при котором функция становится равной нулю). Воспользуйтесь формулой обратного преобразования Фурье.

16. Докажите ортонормальность базиса

$$\left\{ \psi_k[n] = \frac{1}{\sqrt{N}} e^{j \frac{2\pi}{N} kn}, k = \overline{0, N-1} \right\} \text{ при условии } n = \overline{0, N-1}.$$

17. Для прямого ДПФ, имеющего форму

$$X[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} kn}, k = \overline{0, N-1}, \text{ найдите выражение обратного ДПФ.}$$

Вопросы для самоконтроля

1. Что такое скалярное произведение?
2. Что такое норма? В чем ее физический смысл?
3. Что такое метрика?
4. Сколько различных базисов можно задать в пространстве последовательностей?

5. Какие пространства называются гильбертовыми?
6. Что такое ортонормальный базис?
7. В чем состоят преимущества ортонормального базиса?
8. Сколько различных ортонормальных базисов можно задать в пространстве последовательностей?
9. Что такое обобщенный ряд Фурье?
10. В чем состоит смысл процедуры Грама – Шмидта?
11. Что такое равенство Парсеваля, в чем его смысл?
12. В чем состоит значение принципа суперпозиции для анализа цепей?
13. Что такое линейная цепь? Что такое ЛИС-цепь?
14. Что такое импульсная характеристика цепи?
15. Почему термин «каузальность» предпочтительнее термина «физическая реализуемость»?
16. Что такое «обработка в реальном времени»?
17. В чем состоит условие устойчивости цепи?
18. Какое из утверждений верно: $l_1 \subset l_2$ или $l_2 \subset l_1$?
19. Какую роль играет последовательность $\delta[n]$ в алгебре l_1 ?
20. Можно ли утверждать, что l_1 – гильбертово пространство?
21. Справедливо ли утверждение: если последовательность абсолютно суммируема, то ее z -преобразование сходится?
22. Существуют ли последовательности, для которых z -преобразование сходится во всех точках z -плоскости?
23. Что такое z -образ?
24. Если точка a – единственный полюс z -образа последовательности $x[n]$, можно ли утверждать, что z -преобразование $x[n]$ сходится всюду, за исключением точки a ?
25. Какой вид имеет область сходимости z -преобразования каузальной быстро убывающей последовательности?
26. Какой вид имеет область сходимости z -преобразования антикаузальной быстро убывающей последовательности?
27. Какой вид имеет область сходимости z -преобразования некаузальной быстро убывающей последовательности?
28. Какой вид имеет последовательность, если область сходимости ее z -преобразования – открытое кольцо, содержащее 1-окружность?
29. Какова область сходимости z -преобразования последовательности, убывающей на бесконечности медленнее экспоненты, но быстрее $1/|n|$?

30. Всегда ли z -образ абсолютно суммируемой последовательности представляет собой аналитическую функцию?
31. С чем связана особая роль 1-окружности в теории z -преобразования?
32. С какой скоростью должна убывать на бесконечности абсолютно суммируемая последовательность?
33. Как связано преобразование Фурье последовательности с ее z -преобразованием?
34. Что можно сказать о типах сходимости преобразования Фурье последовательностей из l_1 и l_2 ?
35. Охарактеризуйте базис, в котором представляется последовательность выражением обратного преобразования Фурье.
36. С чем связана особая роль преобразования Фурье в анализе сигналов и цепей?
37. Всегда ли можно аналитически продолжить функцию с 1-окружности на ее окрестность?
38. Почему нельзя реализовать идеальный (Π -образный) фильтр нижних частот?
39. Что такое явление Гиббса? В чем его причина?
40. Что такое дискретное преобразование Фурье?
41. Как связаны преобразование Фурье последовательности конечной длины и ДПФ этой же последовательности? Является ли эта связь взаимно однозначной?
42. Какими свойствами обладают преобразование Фурье и ДПФ вещественной последовательности?
43. Какой геометрический смысл имеет матрица ДПФ, ее столбцы?



3. СТРУКТУРНЫЕ СХЕМЫ И РАЗНОСТНЫЕ УРАВНЕНИЯ ЛИС-ЦЕПЕЙ

Дискретные цепи, рассмотренные ранее как абстрактные математические модели, находят физическое воплощение в устройствах цифровой обработки сигналов и в алгоритмах работы специализированных сигнальных процессоров. В первом случае описанию дискретной цепи в виде импульсной характеристики необходимо сопоставить способ соединения элементарных устройств (элементов), определяемый структурной схемой. Во втором случае необходимо указать последовательность элементарных действий, которые должны выполняться программируемым вычислителем, чтобы преобразовать входную последовательность в выходную. И тот и другой способы представления цепи, очевидно, должны быть эквивалентны, если они соответствуют одной импульсной характеристике. Рассмотрим представление дискретных цепей структурными схемами, которые имеют особенно простой вид для линейных цепей, инвариантных к сдвигу.

3.1. СТРУКТУРА ЦЕПИ С КОНЕЧНОЙ ИМПУЛЬСНОЙ ХАРАКТЕРИСТИКОЙ

Как было установлено в п. 2.4, ЛИС-цепь однозначно определяется своей импульсной характеристикой. Предположим для начала, что импульсная характеристика $h[n]$ имеет конечную длину N , т.е. $h[n] \neq 0$ при $n = \overline{0, N-1}$. Примем также для простоты, что на вход цепи поступает последовательность $x[n]$ длины M , т.е. $x[n] \neq 0$ при $n = \overline{0, M-1}$. Выходная последовательность $y[n]$ в этом случае находится как свертка

$$y[n] = \sum_{k=0}^{N-1} h[k]x[n-k]. \quad (3.1)$$

Это выражение определяет способ взвешенного суммирования значений последовательности $x[n]$ с коэффициентами $h[\cdot]$, как это показано на рис. 3.1, где треугольниками изображены масштабирующие звенья (усилители), а прямоугольниками – звенья задержки на 1 шаг.

Очевидно, что до момента $n = 0$ все отсчеты выходной последовательности равны нулю, $y[n] = 0$, так как равны нулю значения $x[n]$, $x[n-1]$, $x[n-2]$ и т.д.

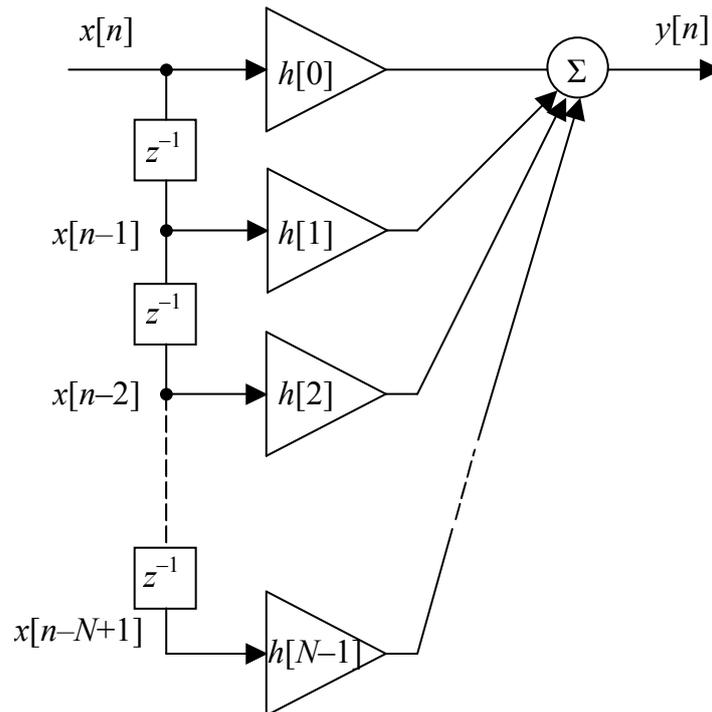


Рис. 3.1. Структура цепи с конечной импульсной характеристикой

При $n = 0$ значение выходной последовательности $y[0] = x[0]h[0]$, при $n = 1$ $y[1] = x[1]h[0] + x[0]h[1]$ и т.д. Последним ненулевым значением выходной последовательности $y[n]$ при $n = N + M - 2$ будет значение $y[N + M - 2] = x[M - 1]h[N - 1]$; таким образом, последовательность $y[n]$ имеет длину $N + M - 1$, т.е. $y[n] \neq 0$ при $n = \overline{0, N + M - 2}$. (Здесь и далее в подобных случаях под длиной последовательности понимается количество ее ненулевых отсчетов, точнее, количество точек «интервала» дискретной временной оси, вне которого отсчеты последовательности заведомо равны 0; при этом некоторые внутренние отсчеты могут равняться нулю.)

Видно, что схема на рис. 3.1 действительно представляет ЛИС-цепь с импульсной характеристикой $h[n]$, $n = \overline{0, N-1}$, причем параметры цепи очевидным образом связаны со значениями импульсной характеристики: коэффициент передачи каждого масштабного звена равен соответствующему отсчету импульсной характеристики.

Нетрудно видеть, что ограничение длины входной последовательности на самом деле не является принципиальным и схема, приведенная на рис. 3.1, реализует свертку импульсной характеристики $h[n]$ с последовательностью произвольной длины (в том числе и бесконечной!); при этом значения $y[n]$ вычисляются в темпе поступления значений входной последовательности (в реальном времени). Задержка выходного сигнала относительно входного определяется при аппаратной (схемной) реализации скоростью передачи сигнала через усилитель и сумматор.

Конечная длина импульсной характеристики, напротив, имеет в данном случае принципиальное значение. Именно длиной импульсной характеристики определяется количество масштабирующих и задерживающих звеньев, т.е. сложность цепи.

Дискретные цепи такой структуры получили название цепей с конечной импульсной характеристикой (*КИХ-цепей*).

3.2. РАЗНОСТНОЕ УРАВНЕНИЕ КИХ-ЦЕПИ

Все операции, выполняемые элементами, показанными на схеме, очевидно, могут быть выполнены цифровым вычислителем (процессором) в соответствии с определенным алгоритмом. Напомним, что пока рассматриваются *дискретные* цепи, точность вычислений предполагается неограниченной.

Поскольку значения импульсной характеристики цепи неизменны (цепь инвариантна к сдвигу, т.е. стационарна), можно для краткости ввести обозначения $h[0] = b_0$, $h[1] = b_1$, ..., $h[N-1] = b_{N-1}$. Тогда, очевидно, входная и выходная последовательности КИХ-цепи связаны уравнением

$$y[n] = b_0 x[n] + b_1 x[n-1] + \dots + b_{N-1} x[n-N+1], \quad (3.2)$$

которое называется разностным уравнением с постоянными коэффициентами. По существу, разностное уравнение (3.2) – это просто другая форма записи свертки (3.1). Как было установлено в разд. 2, z -преобразование заменяет свертку (последовательностей) умно-

жением (z -образов), поэтому разностному уравнению, связывающему последовательности, должно соответствовать более простое алгебраическое уравнение связи z -образов. Это уравнение может быть найдено применением z -преобразования к обеим частям разностного уравнения (3.2). Поскольку правая часть разностного уравнения (3.2) представляет собой линейную комбинацию сдвинутых копий одной последовательности и в силу линейности z -преобразования правая часть алгебраического уравнения будет линейной комбинацией их z -образов. Выясним влияние задержки последовательности на ее z -образ.

Предположим, что $X(z)$ есть z -образ последовательности $x[n]$, а $x_1[n] = x[n-1]$. Тогда z -образ последовательности $x_1[n]$ равен

$$X_1(z) = \sum_{n=-\infty}^{\infty} x[n-1] z^{-n} = z^{-1} \sum_{k=-\infty}^{\infty} x[k] z^{-k} = z^{-1} X(z).$$

Итак, задержка последовательности на 1 шаг эквивалентна умножению ее z -образа на z^{-1} . Поэтому на структурных схемах дискретных цепей звено задержки обозначается символом z^{-1} .

С учетом этого факта, а также свойства линейности z -преобразования запишем на основании разностного уравнения (3.2) уравнение, связывающее z -образы последовательностей

$$Y(z) = b_0 X(z) + b_1 z^{-1} X(z) + \dots + b_{N-1} z^{-(N-1)} X(z), \quad (3.3)$$

или $Y(z) = X(z)H(z)$, где $H(z) = \sum_{k=0}^{N-1} b_k z^{-k} = \sum_{k=0}^{N-1} h[k] z^{-k}$ – передаточная функция КИХ-цепи, имеющая вид полинома относительно $s = z^{-1}$. Как и следовало ожидать, КИХ-цепь выполняет умножение z -образа входной последовательности на полином степени $N-1$.

Если входная последовательность принадлежит l_1 , т.е. является абсолютно суммируемой (напомним, что при этом ее z -преобразование сходится на 1-окружности), то и выходная последовательность тоже является абсолютно суммируемой (это следует из того, что алгебра l_1 замкнута относительно свертки, т.е. свертка двух любых последовательностей из l_1 дает последовательность снова из l_1 , а импульсная характеристика *любой* КИХ-цепи как последо-

вательность конечной длины принадлежит l_1). В алгебре Z функций комплексного переменного, представимых сходящимися на 1-окружности z -преобразованиями, это означает, что произведение функции $X(z)$ из Z и полинома $H(z)$ снова принадлежит Z .

Разностному уравнению (3.2), связывающему последовательности, соответствует уравнение (3.3) относительно их z -образов. (Заметим, что при анализе аналоговых цепей преобразование Лапласа играет похожую роль, поскольку заменяет дифференциальное уравнение алгебраическим [13].) Следует отметить, что алгебраическое уравнение (3.3) применяется главным образом для анализа цепей, в то время как разностное уравнение (3.2) представляет собой алгоритм работы процессора, реализующего цепь с заданной импульсной характеристикой программным способом.

3.3. РАЗНОСТНОЕ УРАВНЕНИЕ И СТРУКТУРНАЯ СХЕМА РЕКУРСИВНОЙ ЦЕПИ

Очевидно, КИХ-цепи составляют лишь ограниченный класс ЛИС-цепей. В самом деле, импульсная характеристика устойчивой ЛИС-цепи должна быть абсолютно суммируемой, но не обязательно конечной последовательностью. Произвольная функция из алгебры Z представляется на 1-окружности равномерно сходящимся степенным рядом с бесконечным в общем случае числом членов. Следует заметить, что в силу равномерной сходимости ряда эту функцию можно аппроксимировать конечной суммой ряда с *наперед заданной точностью*, однако аппроксимация произвольной устойчивой ЛИС-цепи КИХ-цепью может привести к неприемлемо большому количеству членов ряда, определяющему количество ветвей (*порядок*) структурной схемы, т.е. к чрезмерным аппаратурным или временным затратам.

Существуют ЛИС-цепи *конечного* порядка (содержащие конечное число элементов) с импульсными характеристиками бесконечной длины. Такие цепи не умножают, а *делят* z -образ входной последовательности на полином и описываются уравнением

$$Y(z) = \frac{X(z)}{A(z)},$$

где $A(z) = \sum_{k=0}^{M-1} \alpha_k z^{-k}$.

Умножая обе части уравнения на $A(z)$, получим алгебраическое уравнение, связывающее z -образы,

$$\alpha_0 Y(z) + \alpha_1 z^{-1} Y(z) + \dots + \alpha_{M-1} z^{-(M-1)} Y(z) = X(z),$$

откуда обратным z -преобразованием получаем разностное уравнение

$$\begin{aligned} \alpha_0 y[n] &= x[n] - \alpha_1 y[n-1] - \dots - \alpha_{M-1} y[n-M+1] = \\ &= x[n] - \sum_{k=1}^{M-1} \alpha_k y[n-k]. \end{aligned}$$

Поделив обе части уравнения на α_0 и введя новые обозначения для коэффициентов, получим уравнение

$$y[n] = \beta x[n] + \sum_{k=1}^{M-1} a_k y[n-k], \quad (3.4)$$

соответствующее структурной схеме, представленной на рис. 3.2.

Особенность схемы состоит в том, что при вычислении значений выходной последовательности $y[n]$ используются ранее вычисленные значения той же последовательности, поэтому часто

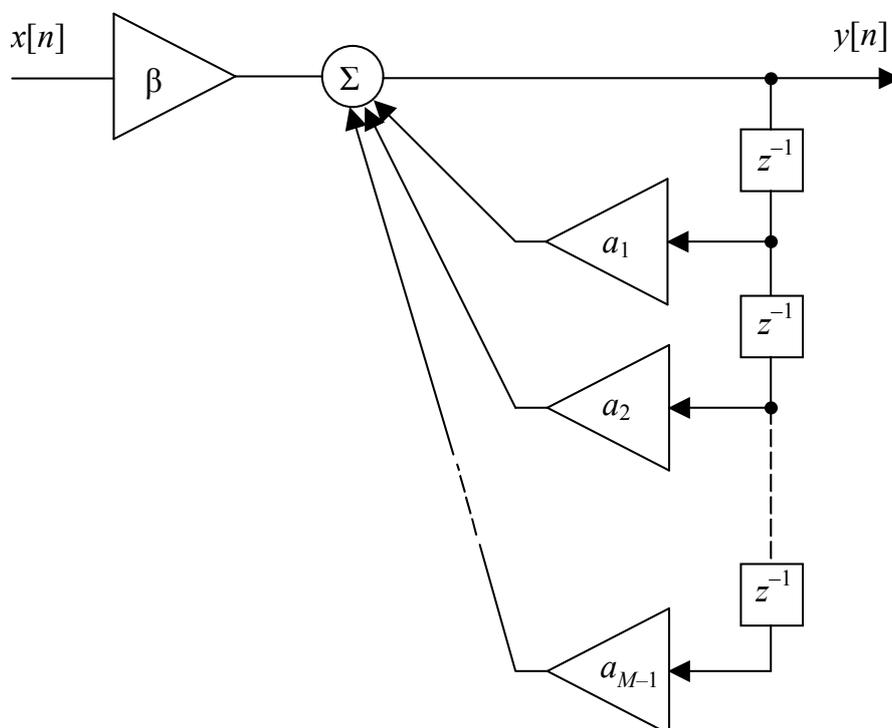


Рис. 3.2. Структура рекурсивной цепи

такие цепи называют *рекурсивными*^{*)}, а КИХ-цепи – соответственно *нерекурсивными*, или *трансверсальными*. Характерным для структуры рекурсивной цепи является наличие ветвей обратной связи.

Здесь и далее предполагается, что для любого разностного уравнения заданы соответствующие *начальные условия*. Для уравнения (3.4), например, при $n = 0, 1, 2, \dots$ начальными условиями являются значения выходной последовательности $y[-M + 1]$, $y[-M + 2]$, ..., $y[-1]$.

В том, что рекурсивные цепи обладают импульсными характеристиками бесконечной длины, можно убедиться на следующем примере.

Пример 3.1. Простейшая рекурсивная цепь описывается уравнением $y[n] = x[n] + ay[n - 1]$ и имеет структурную схему, представленную на рис. 3.3.

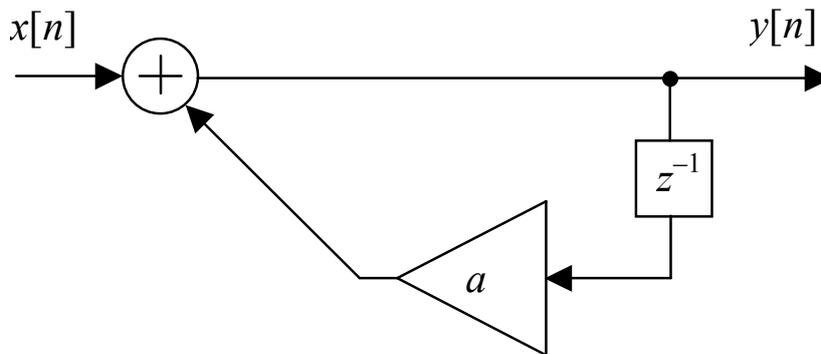


Рис. 3.3. Структура простейшей рекурсивной цепи

Нетрудно видеть, что передаточная функция этой цепи $Y(z) = \frac{1}{1 - az^{-1}}$. Импульсную характеристику $h[n]$ простейшей рекурсивной цепи можно вычислить рекуррентно, подставляя $\delta[n]$ в качестве $x[n]$ при нулевых начальных условиях и определяя $y[n]$ для $n = 0, 1, 2, 3, \dots$ и т.д.

Очевидно, $y[n] = 0$ при $n < 0$ (подразумеваются нулевые начальные условия). При $n = 0$ имеет место воздействие $\delta[n] = 1$ и поэтому $y[0] = 1$. При $n = 1$ сигнал на входе $\delta[n] = 0$, но для задержанной выходной последовательности $y[n - 1] = y[0] = 1$, поэтому

^{*)} Recursio (лат.) – возвращение.

сигнал на выходе $y[n] = y[1] = a$. Рассуждая аналогично, получим $y[n] = h[n] = a^n$ для всех $n \geq 0$. Таким образом, импульсная характеристика простейшей рекурсивной цепи представляет собой комплексную экспоненциальную последовательность

$$h[n] = a^n = r^n e^{j\omega n}, \quad n \geq 0,$$

где $re^{j\omega} = a$. ◀

Импульсную характеристику рекурсивной цепи можно вычислить и другим способом. Следует вспомнить, что z -образ последовательности конечной длины есть полином, а z -образ бесконечной последовательности – степенной ряд, причем степень переменной $s = z^{-1}$ фиксирует порядковый номер соответствующего отсчета последовательности. Перемножение полиномов эквивалентно свертке последовательностей конечной длины. Полиномы можно не только умножать, но и делить, пользуясь обычным алгоритмом деления «углом». Поэтому деление единицы (z -образа δ -последовательности) на полином $1 - az^{-1}$ (знаменатель передаточной функции простейшей рекурсивной цепи) дает ряд по степеням $s = z^{-1}$, последовательность коэффициентов которого представляет собой импульсную характеристику:

$$\begin{array}{r} 1 \qquad \qquad \qquad | \underline{1 - az^{-1}} \\ 1 - az^{-1} \qquad \qquad | \quad 1 + az^{-1} + a^2 z^{-2} + a^3 z^{-3} + \dots \\ az^{-1} \\ \underline{az^{-1} - a^2 z^{-2}} \\ \qquad \qquad \qquad a^2 z^{-2} \\ \qquad \qquad \qquad \underline{a^2 z^{-2} - a^3 z^{-3}} \\ \qquad \qquad \qquad \qquad \qquad a^3 z^{-3} \\ \qquad \qquad \qquad \qquad \qquad \underline{a^3 z^{-3} - a^4 z^{-4}} \quad \text{и т.д.} \end{array}$$

Из начала полученного ряда видно, что последовательность значений импульсной характеристики $1, a, a^2, a^3, \dots$ соответствует формуле $h[n] = a^n$ при $n \geq 0$.

Наконец, импульсную характеристику произвольной ЛИС-цепи можно найти через обратное z -преобразование передаточной функции (см. Приложение 2).

Итак, рекурсивные цепи имеют импульсные характеристики бесконечной длины и относятся к классу цепей с бесконечными импульсными характеристиками, называемых *БИХ-цепями*. Бесконечная импульсная характеристика в отличие от конечной уже не является автоматически абсолютно суммируемой.

Поэтому необходимо установить, при каких условиях импульсная характеристика рекурсивной цепи является абсолютно суммируемой, т.е. принадлежит l_1 . Другими словами, речь идет об установлении *условий устойчивости* рекурсивных цепей.

Поскольку рекурсивная цепь делит z -образ входной последовательности на полином, то вопрос об устойчивости цепи математически эквивалентен вопросу о возможности деления в алгебре Z функций, представимых равномерно сходящимися на 1-окружности степенными рядами (и, соответственно, *обобщенного деления* как операции, обратной свертке, в алгебре l_1). Действительно, если при делении z -образа $X(z)$ входной абсолютно суммируемой последовательности $x[n]$ на $A(z)$ получается функция $Y(z)$ из Z , то выходная последовательность $y[n]$ абсолютно суммируема.

К сожалению, частное от деления двух непрерывных функций не всегда является непрерывной функцией, т.е. если $A(z)$ непрерывна, то $1/A(z)$ может и не быть таковой. Легко видеть, в частности, что для любой функции $A_0(z)$, непрерывной на 1-окружности и обращающейся в нуль в одной ее точке, функция $1/A_0(z)$ терпит в этой точке разрыв. Поэтому рекурсивная цепь с передаточной функцией $H(z) = 1/A_0(z)$ будет неустойчивой. Рекурсивная цепь может быть неустойчивой и при отсутствии нулей знаменателя передаточной функции на 1-окружности.

В терминах алгебры l_1 указанному факту соответствует утверждение, что не для всякой абсолютно суммируемой последовательности x найдется такая абсолютно суммируемая последовательность \bar{x} , что $x \otimes \bar{x} = \delta$ (это означает, что не всякому устойчивому фильтру соответствует устойчивый *обратный* фильтр, см., например, разд. 10).

3.4. РАЗНОСТНОЕ УРАВНЕНИЕ И СТРУКТУРА ЛИС-ЦЕПИ КОНЕЧНОГО ПОРЯДКА

Разностное уравнение конечного порядка общего вида получается объединением уравнений для рекурсивной и нерекурсивной цепей, тогда текущее значение выходной последовательности $y[n]$ зависит как от «входа» $x[n]$, так и от значений «выхода» $y[\cdot]$ при других значениях аргумента («времени»):

$$\begin{aligned} y[n] - a_1 y[n-1] - \dots - a_{M-1} y[n-M+1] = \\ = b_0 x[n] + b_1 x[n-1] + \dots + b_{N-1} x[n-N+1]. \end{aligned} \quad (3.5)$$

Коэффициент a_0 при $y[n]$ принят равным 1; если это не так, то всегда можно разделить все коэффициенты разностного уравнения на a_0 . (Заметим, что текущее значение $y[n]$ зависит только от текущего значения $x[n]$ входа и *предшествующих* значений входной и выходной последовательностей, т.е. это уравнение описывает *каузальную* или *неупреждающую* цепь.) Очевидно, вводя зависимость от «будущих» значений ($x[n+1]$, $y[n+1]$, $x[n+2]$, $y[n+2]$, ... и т.д.), можно записать разностное уравнение и для некаузальной цепи.

Из уравнения (3.5), переписанного в форме

$$y[n] = \sum_{k=0}^{N-1} b_k x[n-k] + \sum_{r=1}^{M-1} a_r y[n-r],$$

следует структурная схема ЛИС-цепи конечного порядка (рис. 3.4).

Переходя от последовательностей к их z -образам, получим алгебраическое уравнение

$$Y(z) \left[1 - \sum_{r=1}^{M-1} a_r z^{-r} \right] = X(z) \sum_{k=0}^{N-1} b_k z^{-k},$$

из которого найдем передаточную функцию

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{N-1} b_k z^{-k}}{1 - \sum_{r=1}^{M-1} a_r z^{-r}} = B(z) \frac{1}{A(z)}. \quad (3.6)$$

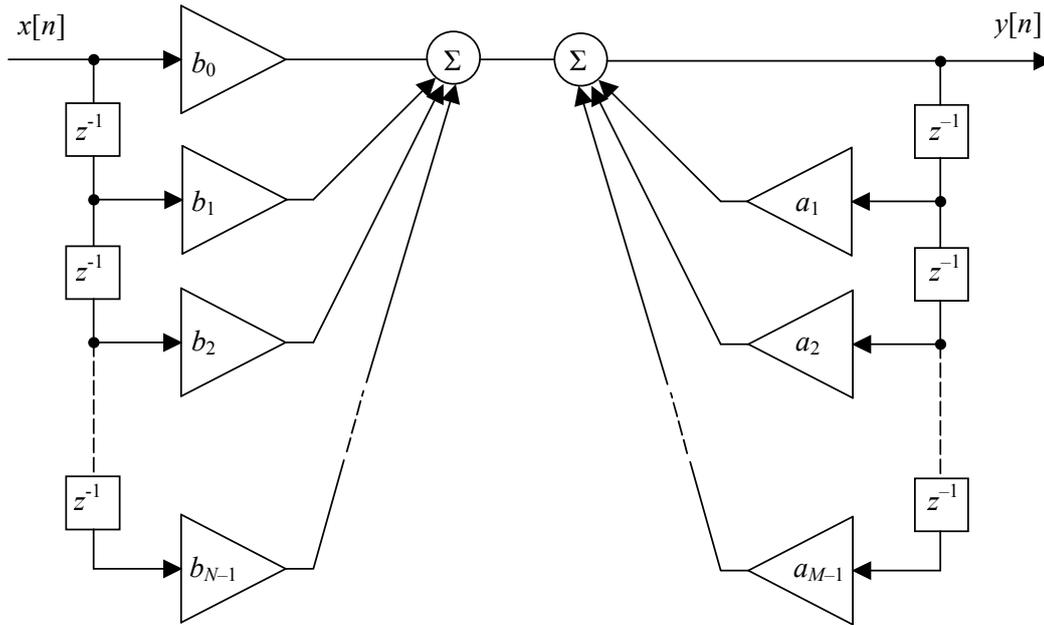


Рис. 3.4. Структура ЛИС-цепи конечного порядка

Из полученного уравнения видно, что умножение z -образа входной последовательности на $H(z)$ можно заменить умножением на $B(z)$ с последующим делением на $A(z)$. Поэтому ЛИС-цепь конечного порядка (т.е. ЛИС-цепь, состоящая из конечного числа элементов) в общем случае может быть представлена последовательным (каскадным) соединением КИХ- и БИХ-цепей (нерекурсивной и рекурсивной цепей), уже рассмотренных ранее. Следовательно, импульсная характеристика этой цепи равна свертке импульсных характеристик ее нерекурсивной и рекурсивной частей и имеет в общем случае *бесконечную* длину. Исключением является случай, когда полином $B(z)$ делится на $A(z)$ без остатка, т.е. $B(z) = A(z)C(z)$, где $C(z)$ – некоторый полином относительно z^{-1} . В этом случае цепь фактически имеет передаточную функцию в виде полинома $C(z)$ и, следовательно, импульсную характеристику конечной длины, хотя по структуре является рекурсивной цепью.

Пример 3.2. Передаточная функция цепи имеет вид

$$H(z) = \frac{1 - z^{-5}}{1 - z^{-1}}.$$

Выполним деление углом полиномов числителя и знаменателя

$$\begin{array}{r}
 1 - z^{-5} \quad \left| \begin{array}{l} 1 - z^{-1} \\ \hline 1 + z^{-1} + z^{-2} + z^{-3} + z^{-4} \\ \hline z^{-1} - z^{-5} \\ \hline z^{-1} - z^{-2} \\ \hline z^{-2} - z^{-5} \\ \hline z^{-2} - z^{-3} \\ \hline z^{-3} - z^{-5} \\ \hline z^{-3} - z^{-4} \\ \hline z^{-4} - z^{-5} \\ \hline z^{-4} - z^{-5} \\ \hline 0 \end{array} \right.
 \end{array}$$

Таким образом, передаточная функция может быть записана в виде

$$H(z) = 1 + z^{-1} + z^{-2} + z^{-3} + z^{-4} = \sum_{n=0}^4 z^{-n},$$

и цепь имеет конечную импульсную характеристику

$$h[n] = \begin{cases} 1, & 0 \leq n \leq 4; \\ 0 & \text{в противном случае.} \quad \blacktriangleleft \end{cases}$$

3.5. СТРОЕНИЕ ИМПУЛЬСНОЙ ХАРАКТЕРИСТИКИ ЛИС-ЦЕПИ КОНЕЧНОГО ПОРЯДКА

Передаточная функция, равная частному двух полиномов, является *дробно-рациональной* функцией. Ясно, что класс дробно-рациональных функций шире класса полиномов, так как полином – частный случай дробно-рациональной функции при единичном знаменателе. Соответственно множество всех импульсных харак-

теристик цепей с дробно-рациональными передаточными функциями содержит *подмножество* всех импульсных характеристик конечной длины. Все остальные импульсные характеристики имеют, очевидно, бесконечную длину.

Импульсные характеристики конечной длины могут иметь произвольные отсчеты, и при этом они остаются конечными импульсными характеристиками, каждой из которых соответствует некоторая трансверсальная цепь конечного порядка с передаточной функцией полиномиального вида, *устойчивая по определению*. Импульсные характеристики бесконечной длины, напротив, не всегда могут быть реализованы цепями конечного порядка. Вопрос о том, какой вид *может* иметь бесконечная импульсная характеристика ЛИС-цепи конечного порядка, представляет практический интерес, поскольку непосредственно связан с реализуемостью алгоритмов ЦОС.

Пусть передаточная функция ЛИС-цепи имеет вид

$$H(z) = \frac{\sum_{k=0}^{N-1} b_k z^{-k}}{1 - \sum_{r=1}^{M-1} a_r z^{-r}} = \frac{\sum_{k=0}^{N-1} b_k z^{-k}}{\sum_{r=0}^{M-1} a_r z^{-r}} = \frac{B(z)}{A(z)}, \quad (3.7)$$

где $a_0 = 1$.

Известно, что полином с вещественными коэффициентами порядка Q имеет Q корней, которые могут либо быть вещественными, либо составлять комплексно-сопряженные пары. Корни могут также совпадать (быть кратными). Следовательно, знаменатель передаточной функции (3.7) имеет $(M - 1)$ корней.

Рассмотрим три частных случая.

А. Простые вещественные корни

Допустим, что знаменатель $A(z)$ передаточной функции (3.7) имеет только вещественные корни, причем все они различны. Обозначим корни d_1, d_2, \dots, d_{M-1} .

Предположим вначале, что порядок полинома-числителя $(N - 1)$ меньше порядка $(M - 1)$ полинома-знаменателя, т. е. передаточная функция представляет собой правильную дробь.

Образую сумму элементарных дробей вида

$$\frac{C_1}{1 - d_1 z^{-1}} + \frac{C_2}{1 - d_2 z^{-1}} + \dots + \frac{C_{M-1}}{1 - d_{M-1} z^{-1}}. \quad (3.8)$$

Приводя слагаемые к общему знаменателю и складывая, получим в знаменателе согласно основной теореме алгебры полином $A(z)$, а в числителе – полином степени не выше $(M - 2)$. Приравнивая полученную дробь передаточной функции, можно найти неопределенные коэффициенты C_1, C_2, \dots, C_{M-1} . Тем самым доказано, что дробно-рациональная передаточная функция, у которой степень числителя *меньше* степени знаменателя, может быть представлена в виде суммы элементарных дробей

$$H(z) = \frac{C_1}{1 - d_1 z^{-1}} + \frac{C_2}{1 - d_2 z^{-1}} + \dots + \frac{C_{M-1}}{1 - d_{M-1} z^{-1}}, \quad (3.9)$$

а сама ЛИС-цепь – в виде параллельного соединения простейших рекурсивных цепей (рис. 3.5). Импульсная характеристика этой цепи равна сумме экспоненциальных импульсных характеристик простейших рекурсивных цепей, соединенных параллельно. Заметим, что основаниями экспонент являются *полюсы* передаточной функции цепи.

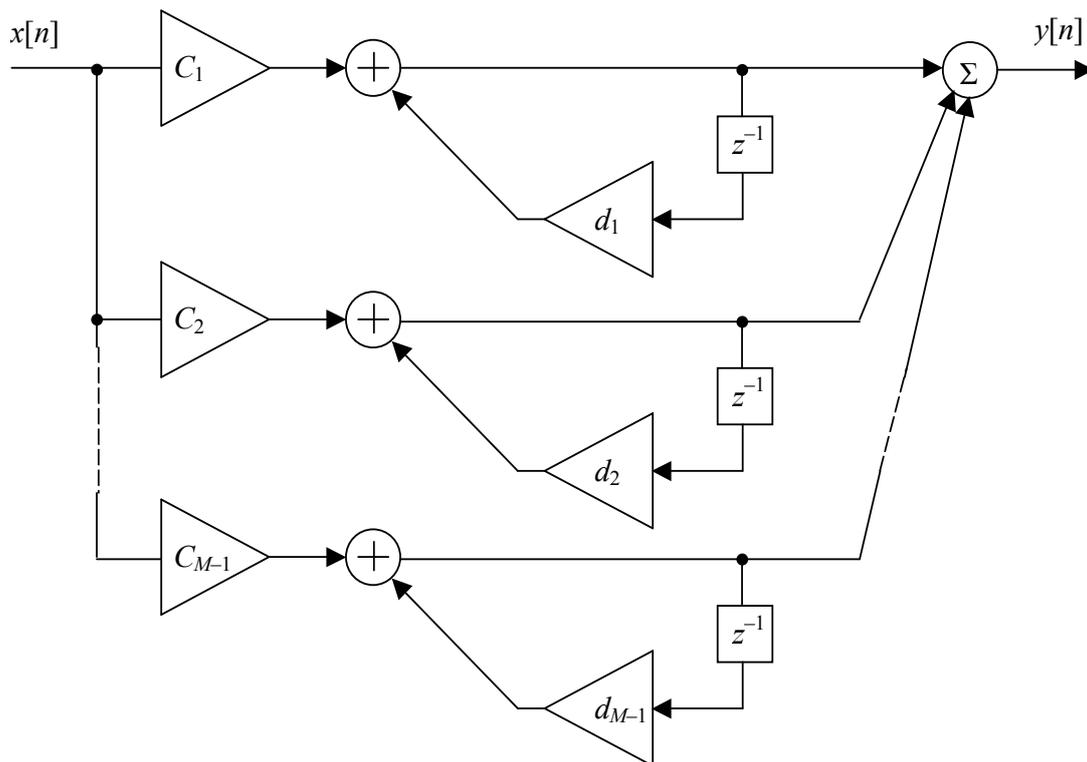


Рис. 3.5. Структура ЛИС-цепи в виде параллельного соединения простейших рекурсивных цепей

Предположим теперь, что степень числителя $(N - 1)$ не меньше степени знаменателя $(M - 1)$. Тогда числитель $B(z)$ можно поделить на знаменатель $A(z)$ и записать передаточную функцию в виде суммы

$$H(z) = D(z) + \frac{B'(z)}{A(z)},$$

где $D(z)$ – целая часть передаточной функции – полином степени $(N - M)$; $B'(z)$ – остаток от деления. Второе слагаемое представляет собой правильную дробь, у которой степень числителя строго меньше степени знаменателя, и соответствует ранее рассмотренному случаю. Суммирование передаточных функций соответствует параллельному соединению цепей. Такой случай иллюстрируется следующим примером.

Пример 3.3. Пусть передаточная функция цепи имеет вид

$$H(z) = \frac{8z^{-3} + 4z^{-2} + z^{-1} - 16}{4z^{-2} + z^{-1} + 8}.$$

Выполняя деление «углом» полинома-числителя на полином-знаменатель

$$\begin{array}{r} 8z^{-3} + 4z^{-2} + z^{-1} - 16 \quad \Big| \quad 4z^{-2} + z^{-1} + 8 \\ \underline{8z^{-3} + 2z^{-2} + 16z^{-1}} \qquad \quad 2z^{-1} + 0,5 \\ 2z^{-2} - 15z^{-1} - 16 \\ \underline{2z^{-2} + 0,5z^{-1} + 4} \\ -15,5z^{-1} - 20, \end{array}$$

получим результат деления с остатком

$$2z^{-1} + 0,5 - \frac{15,5z^{-1} + 20}{4z^{-2} + z^{-1} + 8},$$

поэтому передаточная функция может быть записана в виде суммы полинома и правильной дроби, у которой степень числителя строго меньше степени знаменателя:

$$H(z) = 2z^{-1} + 0,5 - \frac{15,5z^{-1} + 20}{4z^{-2} + z^{-1} + 8}. \quad \blacktriangleleft$$

Итак, легко видеть, что если степень $(N - 1)$ числителя $B(z)$ передаточной функции цепи не меньше степени $(M - 1)$ знаменателя $A(z)$, то схема, реализующая передаточную функцию, может быть получена параллельным соединением КИХ-цепи с передаточной функцией, равной целой части исходной передаточной функции, и $(M - 1)$ параллельно соединенных простейших рекурсивных цепей, имеющих в совокупности передаточную функцию, равную правильной дроби. Поэтому импульсная характеристика ЛИС-цепи с дробно-рациональной передаточной функцией может быть представлена суммой конечной последовательности длины $(N - M + 1)$, соответствующей целой части передаточной функции, и бесконечно протяженных (экспоненциальных) последовательностей – импульсных характеристик простейших рекурсивных цепей в количестве, равном степени $(M - 1)$ знаменателя передаточной функции.

Б. Кратные вещественные корни

Если корень d_k знаменателя $A(z)$ имеет кратность m , то в сумме (3.9) присутствуют члены

$$\frac{C_{kr}}{(1 - d_k z^{-1})^r}$$

для всех $r = \overline{1, m}$, каждый из которых вносит свой вклад в импульсную характеристику цепи. Рассмотрим особенности этого случая на простых примерах.

Пример 3.4. Передаточная функция

$$H(z) = \frac{1}{(1 - az^{-1})^2}$$

имеет знаменатель с корнем a кратности 2. Разделим числитель передаточной функции на знаменатель:

$$\begin{array}{r}
 1 \qquad \qquad \qquad \left| \frac{1 - 2az^{-1} + a^2z^{-2}}{1 - 2az^{-1} + a^2z^{-2}} \right. \\
 \frac{1 - 2az^{-1} + a^2z^{-2}}{1 - 2az^{-1} + a^2z^{-2}} \quad 1 + 2az^{-1} + 3a^2z^{-2} + 4a^3z^{-3} + \dots \\
 \qquad \qquad \qquad \frac{2az^{-1} - a^2z^{-2}}{2az^{-1} - a^2z^{-2}} \\
 \qquad \qquad \qquad \frac{2az^{-1} - 4a^2z^{-2} + 2a^3z^{-3}}{3a^2z^{-2} - 2a^3z^{-3}} \\
 \qquad \qquad \qquad \frac{3a^2z^{-2} - 6a^3z^{-3} + 3a^4z^{-4}}{4a^3z^{-3} - 3a^4z^{-4}} \quad \text{и т.д.}
 \end{array}$$

Импульсная характеристика

$$h[n] = (n+1)a^n, \quad n \geq 0,$$

представляет собой экспоненту с основанием a , умноженную на линейно растущую последовательность. ◀

Пример 3.5. Передаточная функция

$$H(z) = \frac{1}{(1 - az^{-1})^3};$$

знаменатель имеет корень a кратности 3. Полином в знаменателе передаточной функции имеет вид $A(z) = 1 - 3az^{-1} + 3a^2z^{-2} - a^3z^{-3}$.

Выполняя деление $1/A(z)$, получаем степенной ряд

$$1 + 3az^{-1} + 6a^2z^{-2} + 10a^3z^{-3} + 15a^4z^{-4} + 21a^5z^{-5} + 28a^6z^{-6} + \dots$$

Импульсная характеристика

$$h[n] = \frac{(n+1)(n+2)}{2} a^n, \quad n \geq 0. \quad \blacktriangleleft$$

В общем случае слагаемое вида $\frac{1}{(1-az^{-1})^r}$ в разложении передаточной функции соответствует слагаемому

$$\frac{(n+1)(n+2)\dots(n+r-1)}{(r-1)!} a^n, \quad n \geq 0 \quad (3.10)$$

в импульсной характеристике. Поэтому при наличии полюса кратности m импульсная характеристика цепи включает слагаемые вида (3.10) при всех $r = \overline{1, m}$.

В. Комплексно-сопряженные пары корней

Допустим, что среди корней знаменателя передаточной функции имеются комплексные; для обеспечения вещественности импульсной характеристики цепи необходимо, чтобы комплексные корни были попарно комплексно-сопряженными. Рассмотрим два комплексно-сопряженных корня d_k и d_k^* . В разложении передаточной функции на элементарные дроби этим корням соответствуют слагаемые

$$\frac{C_k}{1-d_k z^{-1}} \quad \text{и} \quad \frac{C_k^*}{1-d_k^* z^{-1}},$$

которые описывают простейшие рекурсивные цепи с импульсными характеристиками $C_k (d_k)^n$ и $C_k^* (d_k^*)^n$ при $n \geq 0$.

Сумма импульсных характеристик указанных цепей

$$\begin{aligned} C_k (d_k)^n + C_k^* (d_k^*)^n &= |C_k| e^{j\varphi_k} \left\{ |d_k| e^{j\omega_k} \right\}^n + |C_k| e^{-j\varphi_k} \left\{ |d_k| e^{-j\omega_k} \right\}^n = \\ &= |C_k| \cdot |d_k|^n \cdot 2 \cos(\omega_k n + \varphi_k). \end{aligned}$$

Таким образом, каждая пара комплексно-сопряженных корней знаменателя передаточной функции дает вклад в импульсную характеристику цепи в виде косинусоидальной последовательности с экспоненциально изменяющейся огибающей.

Комбинируя результаты пп. *Б* и *В*, можно получить случай кратных комплексно-сопряженных корней.

Суммируя рассмотренные случаи, можно утверждать, что импульсная характеристика произвольной цепи с дробно-рациональной передаточной функцией складывается из конечной последователь-

ности, определяемой целой частью передаточной функции, и экспонент (домноженных на степенные последовательности в случае кратных корней знаменателя и/или на косинусоидальные последовательности для комплексно-сопряженных пар корней) в количестве, равном числу корней знаменателя.

Конечная последовательность абсолютно суммируема всегда, а каузальная экспонента – при условии, что модуль основания меньше единицы (экспонента затухает). Поэтому для устойчивости каузальной ЛИС-цепи конечного порядка необходимо и достаточно, чтобы все полюсы передаточной функции (они же – корни знаменателя) находились внутри единичной окружности.

Все сказанное в этом разделе очевидным образом переносится на *антикаузальные* ЛИС-цепи конечного порядка. Для этих цепей текущий отсчет выходной последовательности не может зависеть от предшествующих отсчетов (цепь действует «в обратном времени»). Разностное уравнение антикаузальной цепи имеет вид

$$y[n] = \sum_{k=0}^{N-1} b_k x[n+k] + \sum_{r=1}^{M-1} a_r y[n+r].$$

Очевидно, для устойчивости антикаузальной ЛИС-цепи конечного порядка необходимо и достаточно, чтобы все полюсы ее передаточной функции находились *вне* 1-окружности.

Случай *некаузальной* ЛИС-цепи, для которой текущий отсчет выходной последовательности зависит как от прошлых, так и от будущих значений, более сложен и будет рассмотрен в разд. 7.

3.6. КОМПЛЕКСНАЯ ЧАСТОТНАЯ ХАРАКТЕРИСТИКА ЛИС-ЦЕПИ КОНЕЧНОГО ПОРЯДКА

Комплексная частотная характеристика связана с импульсной характеристикой парой преобразований Фурье, поэтому каждая из этих характеристик в отдельности описывает ЛИС-цепь исчерпывающим образом. Выбор конкретной характеристики определяется исключительно удобством ее применения в рамках решаемой задачи. Поэтому наряду со строением импульсной характеристики интерес представляет вид комплексной частотной характеристики ЛИС-цепи конечного порядка.

Рассмотрим ЛИС-цепь конечного порядка с передаточной функцией (3.6). В соответствии с основной теоремой алгебры по-

линомы в числителе и знаменателе передаточной функции могут быть представлены в виде

$$\sum_{k=0}^{N-1} b_k z^{-k} = b_0 (1 - c_1 z^{-1})(1 - c_2 z^{-1}) \dots (1 - c_{N-1} z^{-1}),$$

$$1 - \sum_{r=1}^{M-1} a_r z^{-r} = (1 - d_1 z^{-1})(1 - d_2 z^{-1}) \dots (1 - d_{M-1} z^{-1}),$$

поэтому передаточная функция ЛИС-цепи конечного порядка однозначно с точностью до коэффициента b_0 определяется совокупностью нулей c_1, c_2, \dots, c_{N-1} и полюсов d_1, d_2, \dots, d_{M-1} :

$$H(z) = b_0 \frac{\prod_{k=1}^{N-1} (1 - c_k z^{-1})}{\prod_{r=1}^{M-1} (1 - d_r z^{-1})}. \quad (3.11)$$

Можно сказать, что ЛИС-цепь однозначно (с точностью до константы) определяется своей нуль-полюсной диаграммой.

Подставляя в (3.11) выражение $z = e^{j\omega}$, получаем сужение передаточной функции на 1-окружность – комплексную частотную характеристику:

$$H(e^{j\omega}) = b_0 \frac{\prod_{k=1}^{N-1} (1 - c_k e^{-j\omega})}{\prod_{r=1}^{M-1} (1 - d_r e^{-j\omega})} = b_0 e^{j\omega(M-N)} \frac{\prod_{k=1}^{N-1} (e^{j\omega} - c_k)}{\prod_{r=1}^{M-1} (e^{j\omega} - d_r)}.$$

Модуль КЧХ представляет собой амплитудно-частотную характеристику (АЧХ), а аргумент – фазочастотную характеристику (ФЧХ). Таким образом, амплитудно-частотная характеристика

$$|H(e^{j\omega})| = |b_0| \frac{\prod_{k=1}^{N-1} |e^{j\omega} - c_k|}{\prod_{m=1}^{M-1} |e^{j\omega} - d_m|} \quad (3.12)$$

и фазочастотная характеристика

$$\begin{aligned} \phi(\omega) &= \arg H(e^{j\omega}) = \\ &= \arg b_0 + \omega(M - N) + \sum_{k=1}^{N-1} \arg(e^{j\omega} - c_k) - \sum_{m=1}^{M-1} \arg(e^{j\omega} - d_m) \quad (3.13) \end{aligned}$$

как функции частоты ω определяются взаимным положением нулей и полюсов передаточной функции и вектора на комплексной плоскости, конец которого пробегает единичную окружность при ω , изменяющемся от $-\pi$ до π . Из выражения (3.12) видно, что числитель АЧХ есть произведение сомножителей, каждый из которых представляет собой модуль вектора, проведенного из нуля передаточной функции в точку $e^{j\omega}$ на 1-окружности, определяемому значением частоты. То же справедливо и для знаменателя.

Таким образом, изменение частоты ω приводит к изменению длин указанных векторов и через них влияет на значение АЧХ. Из выражения (3.13) ясно, что на фазочастотную характеристику оказывают влияние углы между указанными векторами, соединяющими нули (полюсы) передаточной функции с точкой $e^{j\omega}$, и направлением вещественной оси комплексной z -плоскости. Кроме того, в ФЧХ входит линейное слагаемое, определяемое соотношением порядков числителя и знаменателя, а также аргумент коэффициента b_0 (как правило, рассматриваются цепи с вещественными импульсными характеристиками, которым соответствуют разностные уравнения с вещественными коэффициентами, тогда $\arg b_0 = 0$).

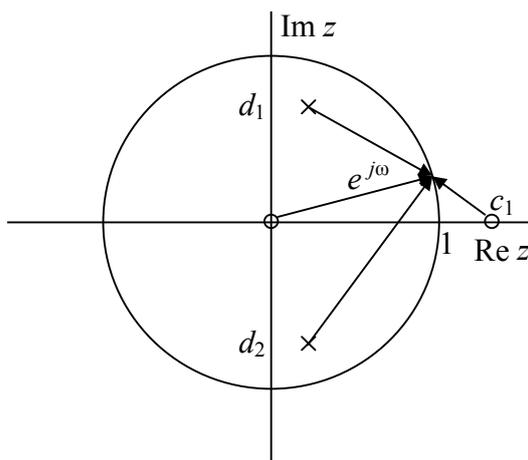


Рис. 3.6. Нуль-полюсная диаграмма ЛИС-цепи 2-го порядка

Пример 3.6. ЛИС-цепь второго порядка с КЧХ

$$\begin{aligned} H(e^{j\omega}) &= \frac{1 + b_1 e^{-j\omega}}{1 - a_1 e^{-j\omega} - a_2 e^{-j2\omega}} = \\ &= e^{j\omega} \frac{e^{j\omega} - c_1}{(e^{j\omega} - d_1)(e^{j\omega} - d_2)} \end{aligned}$$

имеет нуль-полюсную диаграмму, изображенную на рис. 3.6 и содержащую два полюса d_1 и d_2 , один нуль $c_1 = -b_1$ и еще один нуль в точке $z = 0$. На ри-

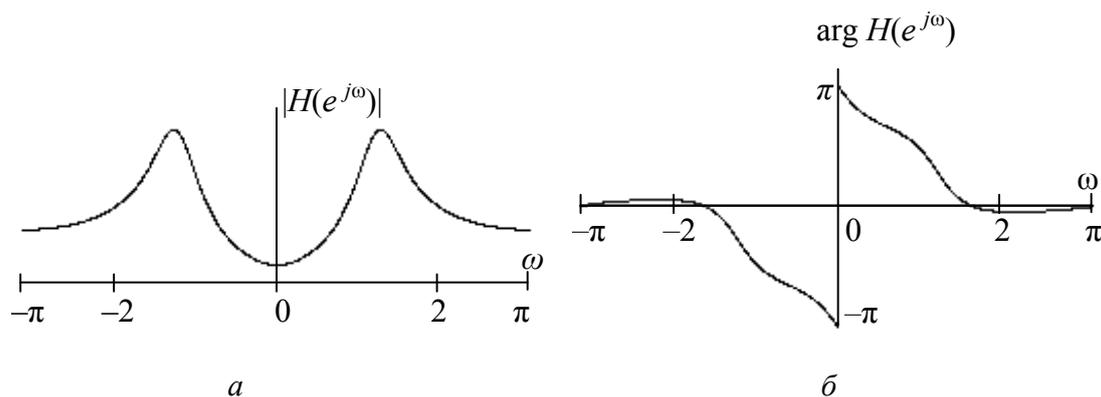


Рис. 3.7. Амплитудно-частотная (а) и фазочастотная (б) характеристики ЛИС-цепи 2-го порядка

сунке показаны разностные векторы, соответствующие числителю и знаменателю КЧХ. Амплитудно-частотная и фазочастотная характеристики представлены на рис. 3.7 ◀

3.7. ВСЕПРОПУСКАЮЩИЕ И МИНИМАЛЬНО-ФАЗОВЫЕ ЦЕПИ

Рассмотрим каузальную устойчивую цепь, у которой передаточная функция имеет один нуль и один полюс, причем нуль находится в точке $c = re^{j\varphi}$, а полюс – в точке $d = \frac{1}{r}e^{j\varphi}$ (рис. 3.8).

Комплексная частотная характеристика такой цепи

$$H(e^{j\omega}) = \frac{e^{j\omega} - re^{j\varphi}}{e^{j\omega} - \frac{1}{r}e^{j\varphi}} = r \frac{\cos \omega - r \cos \varphi + j(\sin \omega - r \sin \varphi)}{r \cos \omega - \cos \varphi + j(r \sin \omega - \sin \varphi)}$$

Путем несложных преобразований можно показать, что АЧХ такой цепи постоянна и равна r , поэтому цепь является *все пропускающей*, т. е. не обладает частотной избирательностью. Фазочастотная характеристика данной цепи $\phi(\omega) = \arg H(e^{j\omega})$ определяется конкретными значениями r и φ .

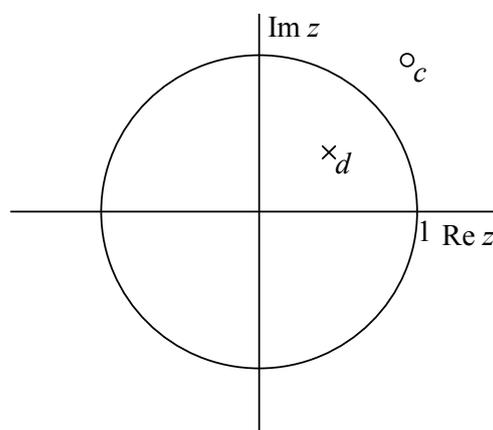


Рис. 3.8. Нуль-полюсная диаграмма все пропускающей цепи 1-го порядка

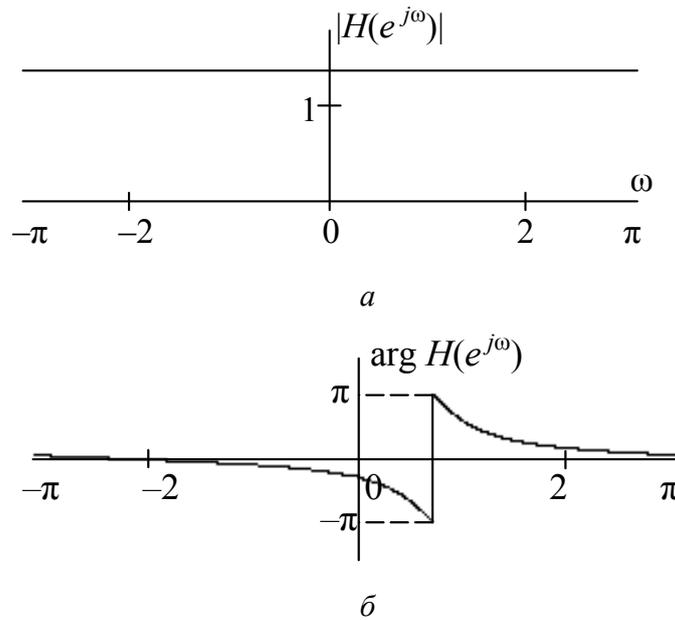


Рис. 3.9. Амплитудно-частотная (а) и фазочастотная (б) характеристики всепропускающей цепи 1-го порядка

Пример АЧХ и ФЧХ всепропускающей цепи при $r = \sqrt{2}$, $\varphi = \pi/4$ приведен на рис. 3.9.

Таким образом, всепропускающая цепь воздействует только на фазовый спектр сигнала, поэтому ее называют также *фазовым контуром* или *фазовращателем*. Отметим, что такая цепь в силу каузальности может только *увеличить* задержку (и, следовательно, фазовый сдвиг) проходящего через нее сигнала.

Рассмотрим каузальную устойчивую цепь, у которой передаточная функция имеет полюсы d_1, d_2, \dots, d_{M-1} , а также расположенные внутри единичной окружности нули c_1, c_2, \dots, c_{N-1} . Представим себе, что каскадно с этой цепью включается всепропускающая цепь с полюсом, в точности совпадающим с одним из нулей первой цепи, например c_1 . Ноль этой всепропускающей *устойчивой* цепи будет располагаться вне единичной окружности в точке $1/c_1^*$. В результате такого соединения двух цепей образуется устойчивая каузальная цепь, обладающая такой же АЧХ, как исходная цепь (с точностью до константы), но обеспечивающая бóльшую задержку сигнала (бóльший «набег фазы») по сравнению с исходной цепью вследствие неупреждающего (каузального) характера фазового контура. Действуя подобным образом, можно «переместить» все нули изнутри единичной окружности наружу, при

этом будет изменяться только фазочастотная характеристика. Следовательно, цепь, имеющая нули только *внутри* единичной окружности, допускает изменения фазочастотной характеристики без изменения формы АЧХ только в сторону *увеличения* фазового набега. Поэтому каузальные устойчивые цепи, не имеющие нулей вне единичной окружности, называются *минимально-фазовыми* цепями. Аналогично антикаузальные цепи, не имеющие нулей внутри 1-окружности, называют *максимально-фазовыми* цепями.

3.8. ЛИС-ЦЕПИ С ЛИНЕЙНОЙ ФАЗОЧАСТОТНОЙ ХАРАКТЕРИСТИКОЙ

Иногда при синтезе ЛИС-цепей (фильтров) в качестве непременного условия выдвигается требование обеспечения линейной фазочастотной характеристики. Линейность ФЧХ обеспечивает сохранение *формы* сигнала, если его спектр полностью укладывается в полосу пропускания фильтра, где коэффициент передачи постоянен. Ниже будет показано, что для линейности фазочастотной характеристики ЛИС-цепи достаточно, чтобы импульсная характеристика была четно-симметрична относительно своей середины.

Предположим, что импульсная характеристика цепи $h[n]$, $n = \overline{0, N-1}$ содержит *четное* число N вещественных отсчетов, причем выполняется указанное условие симметрии: $h[n] = h[N-1-n]$ для $n = \overline{0, N/2-1}$. Запишем комплексную частотную характеристику цепи в виде двух слагаемых

$$\begin{aligned} H(e^{j\omega}) &= \sum_{n=0}^{\frac{N-1}{2}} h[n] e^{-j\omega n} + \sum_{n=\frac{N}{2}}^{N-1} h[n] e^{-j\omega n} = \\ &= \sum_{n=0}^{\frac{N-1}{2}} h[n] e^{-j\omega n} + \sum_{k=0}^{\frac{N-1}{2}} h\left[k + \frac{N}{2}\right] e^{-j\omega\left(k + \frac{N}{2}\right)} = \end{aligned}$$

(рассматривая слагаемые второй суммы в обратном порядке)

$$= \sum_{n=0}^{\frac{N-1}{2}} h[n] e^{-j\omega n} + \sum_{n=0}^{\frac{N-1}{2}} h[N-1-n] e^{-j\omega(N-1-n)} .$$

С учетом четной симметрии импульсной характеристики имеем $h[0] = h[N-1]$, $h[1] = h[N-2]$ и т. д., поэтому

$$H(e^{j\omega}) = \sum_{n=0}^{\frac{N-1}{2}} h[n] \left(e^{-j\omega n} + e^{-j\omega(N-1-n)} \right).$$

Вынесем за знак суммы множитель $e^{-j\omega \frac{N-1}{2}}$, тогда

$$\begin{aligned} H(e^{j\omega}) &= e^{-j\omega \frac{N-1}{2}} \sum_{n=0}^{\frac{N-1}{2}} h[n] \left(e^{-j\omega(n-\frac{N-1}{2})} + e^{j\omega(n-\frac{N-1}{2})} \right) = \\ &= 2e^{-j\omega \frac{N-1}{2}} \sum_{n=0}^{\frac{N-1}{2}} h[n] \cos \left[\omega \left(n - \frac{N-1}{2} \right) \right]. \end{aligned}$$

Из полученного выражения видно, что комплексная частотная характеристика представляет собой произведение вещественной функции частоты (АЧХ) и фазового множителя, аргумент которого (ФЧХ) *линейно* зависит от частоты.

Заметим, что задержка, которую приобретает сигнал при прохождении такой цепи, равна нецелому числу $(N-1)/2$. Если требуется обеспечить целую задержку, то число отсчетов N должно быть нечетным. Тогда

$$\begin{aligned} H(e^{j\omega}) &= \sum_{n=0}^{\frac{N-1}{2}-1} h[n] \left(e^{-j\omega n} + e^{j\omega(N-1-n)} \right) + h \left[\frac{N-1}{2} \right] e^{-j\omega \frac{N-1}{2}} = \\ &= e^{-j\omega \frac{N-1}{2}} \left\{ h \left[\frac{N-1}{2} \right] + 2 \sum_{n=0}^{\frac{N-1}{2}-1} h[n] \cos \left[\omega \left(n - \frac{N-1}{2} \right) \right] \right\}. \end{aligned}$$

Итак, если импульсная характеристика КИХ-цепи четно-симметрична относительно своей середины, то ФЧХ цепи линейна, причем угол наклона ФЧХ определяется положением (номером) среднего отсчета. В частности, если импульсная характеристика некаузальна, так что $h[n] = h[-n]$ при всех n , то ФЧХ тождественно равна 0.

Из сказанного следует, в частности, что каузальные БИХ-цепи с линейными ФЧХ существовать не могут. Поэтому каузальный

цифровой фильтр со строго линейной ФЧХ может быть только КИХ-фильтром. (Среди БИХ-фильтров известны фильтры Бесселя с *приближенно* линейной ФЧХ.)

Заслуживает внимания вопрос о виде нуль-полюсной диаграммы фильтра с линейной ФЧХ. Рассмотрим некаузальный КИХ-фильтр, импульсная характеристика которого удовлетворяет условию

$$h[n] = h[-n] \text{ при } n = \overline{-M, M}.$$

Это означает, что если в выражении $h[n]$ n заменить на $-n$, то импульсная характеристика не изменится, а поскольку она составляет исчерпывающее описание ЛИС-цепи, то не изменятся и другие характеристики цепи, в том числе нуль-полюсная диаграмма. Передаточная функция должна в таком случае быть инвариантной (безразличной) к такой замене, поэтому

$$\begin{aligned} H(z) &= \sum_{n=-M}^M h[n]z^{-n} = \sum_{n=-M}^M h[-n]z^{-n} = \\ &= \sum_{k=-M}^M h[k]z^k = \sum_{k=-M}^M h[k](z^{-1})^{-k} = H(z^{-1}) = H\left(\frac{1}{z}\right). \end{aligned}$$

Итак, передаточная функция (а значит, и нуль-полюсная диаграмма) рассматриваемой цепи не изменится, если z заменить на z^{-1} , поэтому каждому нулю d_k передаточной функции соответствует нуль $1/d_k$ (убедитесь, что кроме нулей передаточная функция такой цепи содержит M -кратный полюс в точке $z = 0$). Таким образом, для некаузальной цепи с четной импульсной характеристикой нуль-полюсная диаграмма удовлетворяет условию симметрии относительно замены z на z^{-1} . На рис. 3.10 показан пример нуль-полюсной диаграммы цепи с линейной ФЧХ [2]. Эта цепь имеет вещественную импульсную характеристику, поэтому нуль-полюсная диаграм-

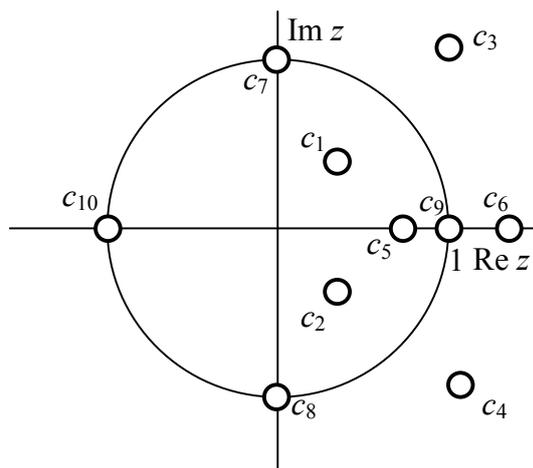


Рис. 3.10. Нуль-полюсная диаграмма цепи с линейной ФЧХ

ма симметрична относительно вещественной оси z -плоскости. Благодаря этому с учетом условия линейности фазочастотной характеристики множество всех нулей распадается на четыре группы. Первая группа образована нулями общего вида, не лежащими на единичной окружности или осях z -плоскости. Такими нулями являются в данном случае c_1 , $c_2 = c_1^*$, $c_3 = 1/c_2$ и $c_4 = 1/c_1 = c_3^*$. Объединение этих нулей в группу обеспечивает ее замкнутость относительно операций комплексного сопряжения и замены каждого нуля на обратный, поэтому такая группа может быть реализована цепью с вещественными коэффициентами и линейной ФЧХ. Нули на мнимой оси являются взаимно обратными и одновременно комплексно-сопряженными $c_7 = 1/c_8 = c_8^*$. Каждый из нулей c_9 и c_{10} является комплексно-сопряженным и обратным самому себе. Наконец, нули c_5 и c_6 являются вещественными взаимно обратными. Каждая из перечисленных групп нулей сама по себе соответствует цепи с вещественными коэффициентами и линейной ФЧХ. Каскадное соединение этих цепей образует цепь с нуль-поллюсной диаграммой, представленной на рис. 3.10.

Известно, что сдвиг последовательности соответствует умножению ее z -образа на соответствующую степень переменной z^{-1} , поэтому при сдвиге импульсной характеристики ее нуль-поллюсная диаграмма (НПД) не изменяется нигде, кроме точек $z = 0$ и $z = \infty$ (что происходит в этих точках?).

Суммируя сказанное, можно утверждать, что:

- 1) БИХ-цепь с линейной ФЧХ может быть только некаузальной;
- 2) КИХ-цепь с линейной ФЧХ имеет импульсную характеристику, четно-симметричную относительно своей середины;
- 3) нуль-поллюсная диаграмма цепи с линейной ФЧХ симметрична относительно преобразования z -плоскости вида $z \rightarrow z^{-1}$ (кроме, может быть, точек $z = 0$ и $z = \infty$).

Еще одна интересная для практики разновидность линейной ФЧХ соответствует импульсной характеристике, удовлетворяющей условию

$$h[n] = -h[N - 1 - n] \quad \text{для } n = 0, \frac{N}{2} - 1,$$

т. е. нечетно-симметричной относительно середины. Запишем комплексную частотную характеристику такой цепи при четном N :

$$\begin{aligned} H(e^{j\omega}) &= \sum_{n=0}^{\frac{N-1}{2}} h[n]e^{-j\omega n} + \sum_{n=\frac{N}{2}}^{N-1} h[n]e^{-j\omega n} = \\ &= \sum_{n=0}^{\frac{N-1}{2}} h[n]e^{-j\omega n} + \sum_{k=0}^{\frac{N-1}{2}} h[k + \frac{N}{2}]e^{-j\omega(k + \frac{N}{2})} = \end{aligned}$$

(рассматривая слагаемые второй суммы в обратном порядке)

$$= \sum_{n=0}^{\frac{N-1}{2}} h[n]e^{-j\omega n} + \sum_{n=0}^{\frac{N-1}{2}} h[N-1-n]e^{-j\omega(N-1-n)}.$$

С учетом нечетной симметрии импульсной характеристики имеем

$$H(e^{j\omega}) = \sum_{n=0}^{\frac{N-1}{2}} h[n](e^{-j\omega n} - e^{-j\omega(N-1-n)}).$$

Вынесем за знак суммы множитель $e^{-j\omega \frac{N-1}{2}}$, тогда

$$\begin{aligned} H(e^{j\omega}) &= e^{-j\omega \frac{N-1}{2}} \sum_{n=0}^{\frac{N-1}{2}} h[n] \left(e^{-j\omega \left(n - \frac{N-1}{2} \right)} - e^{j\omega \left(n - \frac{N-1}{2} \right)} \right) = \\ &= -2e^{-j\omega \left(\frac{N-1}{2} \right) + j\frac{\pi}{2}} \sum_{n=0}^{\frac{N-1}{2}} h[n] \sin \left[\omega \left(n - \frac{N-1}{2} \right) \right]. \end{aligned}$$

При нечетном N нечетная симметрия импульсной характеристики требует, чтобы средний отсчет был равен нулю, поэтому комплексная частотная характеристика

$$\begin{aligned} H(e^{j\omega}) &= \sum_{n=0}^{\frac{N-1}{2}-1} h[n](e^{-j\omega n} - e^{j\omega(N-1-n)}) = \\ &= e^{-j\omega \frac{N-1}{2} - j\frac{\pi}{2}} \left\{ 2 \sum_{n=0}^{\frac{N-1}{2}-1} h[n] \sin \left[\omega \left(n - \frac{N-1}{2} \right) \right] \right\}. \end{aligned}$$

Таким образом, график полученной фазочастотной характеристики представляет собой прямую линию, проходящую через точку $(0; -\pi/2)$. Производная ФЧХ, равная групповой задержке [2], определяется положением середины импульсной характеристики. В частности, если импульсная характеристика некаузальна и справедливо условие $h[n] = -h[-n]$ при всех n , то ФЧХ равна

$$\varphi(\omega) = \begin{cases} -\pi/2, & \omega \in (0, \pi), \\ \pi/2, & \omega \in (-\pi, 0). \end{cases}$$

При отрицательных частотах учитывается дополнительный фазовый сдвиг, равный π и возникающий вследствие перемены знака вещественного сомножителя КЧХ. Фазовую характеристику такого вида имеет, в частности, преобразователь Гильберта, применяемый в системах связи (см. разд. 10).

3.9. РЕКУРСИВНАЯ ВЫЧИСЛИМОСТЬ

Разностное уравнение, как можно было убедиться, представляет собой лишь *одну из форм описания* дискретной цепи, наряду с импульсной характеристикой, передаточной функцией, комплексной частотной характеристикой, нуль-полюсной диаграммой и структурной схемой. Однако в случае программной реализации дискретной цепи разностное уравнение выступает на первый план, так как оно представляет собой *алгоритм*, согласно которому программируемое цифровое устройство последовательного действия (например, сигнальный процессор) вычисляет отсчеты выходной последовательности с использованием как входной последовательности, так и в общем случае уже вычисленных выходных отсчетов.

Одно из неперемных условий, которым должен удовлетворять любой алгоритм – условие *рекурсивной вычислимости*, именуемой далее просто вычислимостью. Это условие состоит в том, что все значения переменных и констант, используемые на каждом шаге исполнения алгоритма, должны быть к этому моменту известны – либо заданы в виде начальных условий, либо вычислены на предыдущих шагах. Вычислимость представляет собой совершенно естественное требование, однако не всякая ЛИС-цепь удовлетворяет ему. Проверить вычислимость можно, например, по виду разностного уравнения.

Рассмотрим в качестве примера разностное уравнение

$$y[n] = \sum_{i=0}^N b_i x[n-i] + \sum_{j=1}^M a_j y[n-j], \quad n = \overline{0, \infty}.$$

Это уравнение соответствует ЛИС-цепи общего вида конечного порядка, и по его виду можно заключить, что оно вычислимо, если заданы начальные условия – значения $x[-i]$, $i = \overline{1, N}$ и $y[-j]$, $j = \overline{1, M}$. Действительно, при расчете значения выходной последовательности в момент n используются ранее рассчитанные значения выходной последовательности, а также текущее и прошлые значения входной последовательности. Иными словами, цепь является неупреждающей (каузальной) и потому вычислимой. Однако каузальность не есть синоним вычислимости. Вычислимой будет также антикаузальная цепь, описываемая в общем случае разностным уравнением

$$y[n] = \sum_{i=0}^N b_i x[n+i] + \sum_{j=1}^M a_j y[n+j], \quad n = \overline{0, -\infty},$$

правда, вычисления должны осуществляться «в обратном времени», и в качестве начальных условий должны быть заданы значения $x[i]$, $i = \overline{1, N}$ и $y[j]$, $j = \overline{1, M}$. Разумеется, речь идет не о реальном обращении физического времени, а лишь о том обстоятельстве, что входная последовательность может быть записана в запоминающее устройство (буфер) и затем считана в обратном порядке. Если же цепь *некаузальна*, т. е. не является ни каузальной, ни антикаузальной, то для ее реализации необходимы специальные меры. Для произвольной устойчивой ЛИС-цепи этот вопрос решается путем разложения на две цепи, вычисляемые по отдельности, следующим образом.

Рассмотрим некаузальную ЛИС-цепь с абсолютно суммируемой импульсной характеристикой. Область сходимости z -преобразования импульсной характеристики включает единичную окружность и представляет собой открытое кольцо. Передаточная функция цепи

$$H(z) = \frac{B(z)}{A(z)},$$

где $A(z)$ и $B(z)$ – полиномы.

Пользуясь основной теоремой алгебры, можно разложить оба полинома на простейшие множители (биномы, или двучлены):

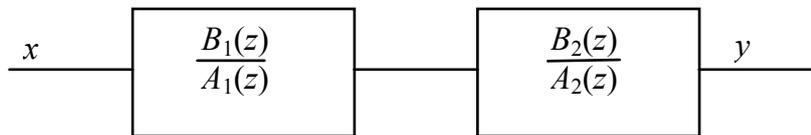
$$H(z) = b_0 \frac{(1 - c_1 z^{-1})(1 - c_2 z^{-1}) \dots (1 - c_N z^{-1})}{(1 - d_1 z^{-1})(1 - d_2 z^{-1}) \dots (1 - d_M z^{-1})}, \quad (3.14)$$

где N и M – степени числителя и знаменателя. Сгруппируем все нули, лежащие внутри 1-окружности, а также нули, лежащие вне 1-окружности, и так же поступим с полюсами. Тогда можно записать передаточную функцию в виде

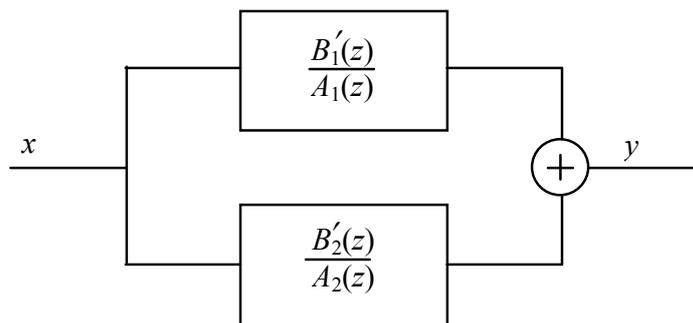
$$H(z) = \frac{B_1(z)B_2(z)}{A_1(z)A_2(z)} = \frac{B_1(z)}{A_1(z)} \frac{B_2(z)}{A_2(z)},$$

где первая дробь имеет нули и полюсы внутри единичного круга, т. е. является передаточной функцией *каузальной минимально-фазовой* цепи, а вторая дробь представляет *антикаузальную максимально-фазовую* цепь. Первая из цепей вычислима в прямом времени, а вторая – в обратном (при соответствующих начальных условиях). С учетом этого исходную некаузальную ЛИС-цепь можно реализовать в виде каскадного соединения, представленного на рис.3.11, а.

На практике реализация обработки входной последовательности такой цепью выполняется в два этапа. Первый этап состоит в



а



б

Рис. 3.11. Каскадная (а) и параллельная (б) реализации вычислимой некаузальной цепи

решении разностного уравнения прямого времени, соответствующего первой дроби. Результирующая последовательность используется в качестве входной при решении в обратном времени разностного уравнения, соответствующего второй дроби.

Альтернативный способ решения проблемы вычислимости для некаузальной цепи состоит в разложении

$$H(z) = \frac{B_1'(z)}{A_1(z)} + \frac{B_2'(z)}{A_2(z)},$$

где, очевидно, $A_2(z)B_1'(z) + A_1(z)B_2'(z) = B(z)$. Таким образом, исходная некаузальная ЛИС-цепь может быть реализована в виде параллельного соединения двух ЛИС-цепей, одна из которых каузальная (но не обязательно минимально-фазовая), а другая – антикаузальная (не обязательно максимально-фазовая) (рис. 3.11, б). Обе цепи вычислимы, но снова одна в прямом, а другая – в обратном времени.

3.10. ОПИСАНИЕ ЦЕПЕЙ НАПРАВЛЕННЫМИ ГРАФАМИ И МАТРИЧНЫМИ РАЗНОСТНЫМИ УРАВНЕНИЯМИ

Традиционный способ наглядного представления дискретных цепей в виде структурных схем, включающих звенья задержки, масштабные звенья и сумматоры, не является ни единственным, ни самым удобным. С точки зрения удобства анализа целесообразно дальнейшее упрощение: замена структурной схемы сигнальным *графом* – топологической моделью цепи. При этом цепь представляется множеством вершин (узлов), соединенных ветвями (ребрами), причем каждая ветвь имеет направление и весовой коэффициент, равный масштабному коэффициенту или величине z^{-1} , в зависимости от того, соответствуют этой ветви в схеме масштабное звено или элемент задержки. Ветви графа помечаются стрелками, показывающими направление передачи сигналов между узлами. Если в узле графа сходятся несколько ветвей, считается, что сигналы в этом узле суммируются. Если из узла выходят несколько ветвей, то считается, что по ним распространяются копии сигнала. Таким образом, по структурной схеме можно построить граф, и, наоборот, по графу можно составить структурную схему.

Пример структурной схемы и соответствующего сигнального графа приведен на рис. 3.12. На первый взгляд переход от схемы к

сигнальному графу не дает ничего нового, однако такая замена позволяет применять к анализу цепей математический аппарат теории графов и получать результаты, не очевидные в рамках привычных представлений.

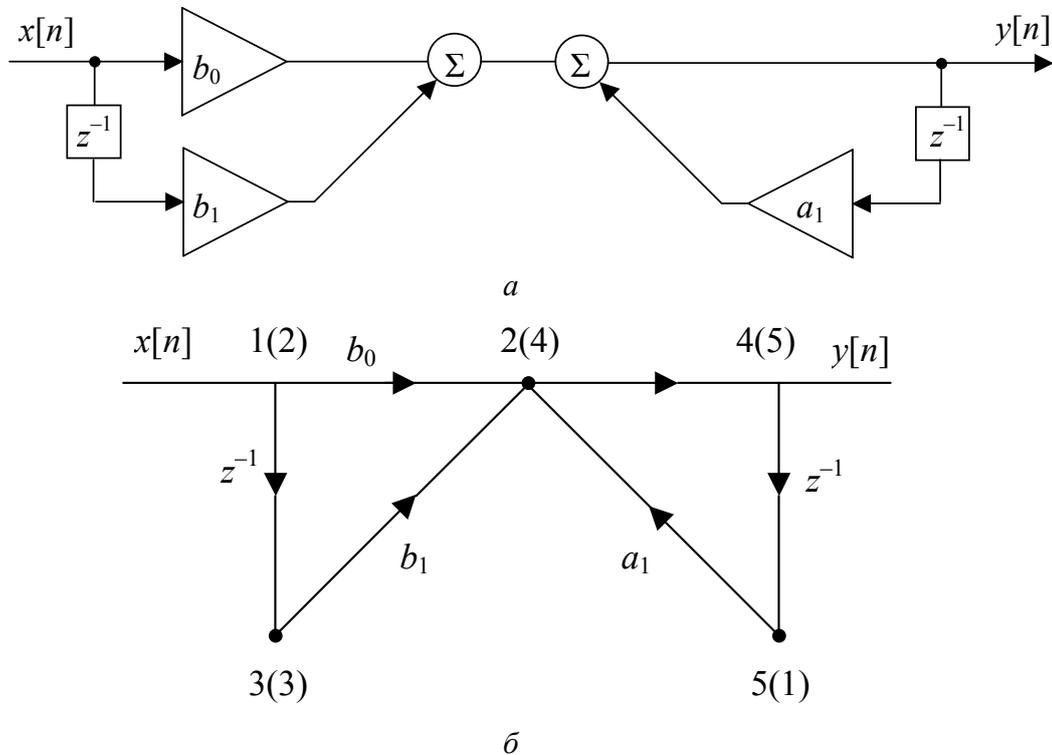


Рис. 3.12. Структурная схема цепи (а) и ее сигнальный граф (б)

При анализе сигнальных графов принято выделять *истоковые* узлы (такие, из которых ветви только исходят), а также *стоковые* узлы (такие, в которые ветви только входят). Все остальные узлы графа являются *внутренними*.

Припишем всем внутренним узлам номера от 1 до N , сигналы в этих узлах обозначим $w_i, i = \overline{1, N}$. Аналогично сигналы во всех истоковых узлах обозначим $x_j, j = \overline{1, M}$. Наконец, выходные сигналы в соответствии с номерами стоковых узлов получают обозначения $y_k, k = \overline{1, P}$.

Удобно считать, что все узлы графа соединены ветвями со всеми узлами (т. е. граф является *полным*), просто некоторые ветви (иногда почти все) имеют нулевую передачу. Обозначим b_{ji} передачу от j -го истокового узла к i -му внутреннему узлу, f_{il} — переда-

чу от i -го внутреннего узла к l -му узлу, c_{ik} – передачу от i -го внутреннего узла к k -му стоковому.

Тогда, очевидно, сигнал в i -м узле

$$w_i = \sum_{j=1}^M b_{ji} x_j + \sum_{l=1}^N f_{li} w_l, \quad i = \overline{1, N}.$$

Для стоковых узлов

$$y_k = \sum_{i=1}^N c_{ik} w_i, \quad k = \overline{1, P}.$$

Для ЛИС-цепи все коэффициенты не зависят от времени n , поэтому уравнения для z -образов сигналов можно записать в виде

$$W_i(z) = \sum_{j=1}^M b_{ji} X_j(z) + \sum_{l=1}^N f_{li} W_l(z), \quad i = \overline{1, N}, \quad (3.15)$$

$$Y_k(z) = \sum_{i=1}^N c_{ik} W_i(z), \quad k = \overline{1, P}. \quad (3.16)$$

Для схемы, приведенной на рис. 3.12, ненулевыми являются коэффициенты $b_{11} = 1$, $c_{41} = 1$, $f_{12} = b_0$, $f_{13} = z^{-1}$, $f_{32} = b_1$, $f_{24} = 1$, $f_{45} = z^{-1}$, $f_{52} = a_1$ (узлы графа пронумерованы цифрами, стоящими вне скобок).

Выражения (3.15), (3.16), связывающие сигналы в узлах графа, можно переписать в матричной форме:

$$\mathbf{W}(z) = \mathbf{F}^T(z) \mathbf{W}(z) + \mathbf{B}^T \mathbf{X}(z), \quad (3.17)$$

$$\mathbf{Y}(z) = \mathbf{C}^T \mathbf{W}(z), \quad (3.18)$$

где $(\cdot)^T$ – символ транспонирования, $(N \times N)$ -матрица \mathbf{F} составлена из коэффициентов f_{ij} , $(M \times N)$ -матрица \mathbf{B} составлена из коэффициентов b_{ij} , $(N \times P)$ -матрица \mathbf{C} – из коэффициентов c_{ij} (во всех случаях i – номер строки, а j – номер столбца).

Очевидно, что такое описание цепи является исчерпывающим, так как оно однозначно следует из графа (а значит, из структурной

схемы), и, наоборот, по матричным уравнениям можно восстановить граф и структурную схему.

Элементами матриц служат коэффициенты передачи ветвей графа, поэтому, как правило, большинство элементов матрицы \mathbf{F} равны 0; элементы матриц \mathbf{B} и \mathbf{C} равны 1 для ветвей, соединяющих внутренние узлы с истоковыми и стоковыми узлами, и 0 в противном случае.

Для схемы, приведенной на рис. 3.12,

$$\mathbf{F} = \begin{pmatrix} 0 & b_0 & z^{-1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & b_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & z^{-1} \\ 0 & a_1 & 0 & 0 & 0 \end{pmatrix}; \quad \mathbf{C} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix};$$

$\mathbf{B} = (1 \ 0 \ 0 \ 0 \ 0)$; тогда матричные уравнения имеют вид

$$\begin{pmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \\ W_5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ b_0 & 0 & b_1 & 0 & a_1 \\ z^{-1} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & z^{-1} & 0 \end{pmatrix} \cdot \begin{pmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \\ W_5 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \cdot X; \quad (3.19)$$

$$\mathbf{Y} = (0 \ 0 \ 0 \ 1 \ 0) \cdot \begin{pmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \\ W_5 \end{pmatrix}. \quad (3.20)$$

Для простоты опущена зависимость от z для всех z -образов узловых переменных (сигналов в узлах).

Из уравнений (3.17), (3.18), очевидно, следует

$$\mathbf{W}(z) - \mathbf{F}^T(z)\mathbf{W}(z) = \mathbf{B}^T \mathbf{X}(z),$$

или

$$\left[\mathbf{I} - \mathbf{F}^T(z) \right] \mathbf{W}(z) = \mathbf{B}^T \mathbf{X}(z),$$

откуда

$$\mathbf{W}(z) = \left[\mathbf{I} - \mathbf{F}^T(z) \right]^{-1} \mathbf{B}^T \mathbf{X}(z).$$

Здесь \mathbf{I} – единичная $(N \times N)$ -матрица, $(\cdot)^{-1}$ – символ обращения квадратной невырожденной матрицы. С учетом (3.18) можно записать

$$\mathbf{Y}(z) = \mathbf{C}^T \left[\mathbf{I} - \mathbf{F}^T(z) \right]^{-1} \mathbf{B}^T \mathbf{X}(z) = \mathbf{T}(z) \mathbf{X}(z),$$

где матрица $\mathbf{T}(z)$ описывает передачу от истоков к стокам графа и является *матричной передаточной функцией*.

Элементы матрицы \mathbf{F} равны либо числовым константам, либо z^{-1} , поэтому можно записать $\mathbf{F}(z) = \mathbf{F}_1 + z^{-1} \mathbf{F}_2$, где \mathbf{F}_1 и \mathbf{F}_2 – матрицы, содержащие в качестве элементов только константы (включая нули). Тогда

$$\mathbf{W}(z) = \mathbf{F}_1^T \mathbf{W}(z) + z^{-1} \mathbf{F}_2^T \mathbf{W}(z) + \mathbf{B}^T \mathbf{X}(z)$$

и, переходя от z -образов к последовательностям, получаем

$$\mathbf{w}[n] = \mathbf{F}_1^T \mathbf{w}[n] + \mathbf{F}_2^T \mathbf{w}[n-1] + \mathbf{B}^T \mathbf{x}[n], \quad (3.21)$$

$$\mathbf{y}[n] = \mathbf{C}^T \mathbf{w}[n]. \quad (3.22)$$

Анализ матричного уравнения (3.19) показывает, что для вычисления величины W_2 переменная W_3 уже должна быть известна, что, очевидно, делает невозможным последовательное вычисление компонент вектора \mathbf{W} . Указанное затруднение, однако, легко преодолеть в данном случае перенумерацией узлов графа (или, что эквивалентно, одновременной перестановкой строк и столбцов матрицы \mathbf{F}^T). Вообще, для того чтобы можно было вычислить все компоненты вектора \mathbf{W} последовательно, необходимо, чтобы ненулевые элементы матрицы \mathbf{F}_1^T находились ниже главной диагонали. Если такая перестановка возможна, цепь является вычислимой.

На рис. 3.12, *б* новые номера указаны в скобках. Соответствующие матрицы имеют вид

$$\mathbf{F}_1^T = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ a_1 & b_0 & b_1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{F}_2^T = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Замечания.

1. Вычислимость определяется только матрицей \mathbf{F}_1^T , но не \mathbf{F}_2^T , поскольку матрица \mathbf{F}_2^T описывает влияние «прошлого» на «настоящее», а прошлое по предположению либо вычислено на предыдущем шаге, либо задано в виде начальных условий. Однако перестановка строк и столбцов должна выполняться для матрицы \mathbf{F}^T в целом, так как означает перенумерацию узлов графа.

2. Разложение матрицы \mathbf{F}^T на две матрицы – \mathbf{F}_1^T и \mathbf{F}_2^T – соответствует разложению (декомпозиции) графа на два подграфа, в одном из которых оставлены только ветви, не содержащие задержек, а в другом – только ветви с задержками. Отсутствие в первом подграфе циклов (замкнутых путей) является необходимым и достаточным условием вычислимости графа [2]. Подграфы для цепи, приведенной на рис. 3.12, изображены на рис. 3.13.

Систему (3.21), (3.22) можно переписать в виде

$$\left[\mathbf{I} - \mathbf{F}_1^T \right] \mathbf{w}[n] = \mathbf{F}_2^T \mathbf{w}[n-1] + \mathbf{B}^T \mathbf{x}[n],$$

$$\mathbf{y}[n] = \mathbf{C}^T \mathbf{w}[n],$$

или

$$\mathbf{w}[n] = \left[\mathbf{I} - \mathbf{F}_1^T \right]^{-1} \mathbf{F}_2^T \mathbf{w}[n-1] + \left[\mathbf{I} - \mathbf{F}_1^T \right]^{-1} \mathbf{B}^T \mathbf{x}[n],$$

$$\mathbf{y}[n] = \mathbf{C}^T \mathbf{w}[n].$$

Введя обозначения

$$\mathbf{A} = \left[\mathbf{I} - \mathbf{F}_1^T \right]^{-1} \mathbf{F}_2^T, \quad \mathbf{D} = \left[\mathbf{I} - \mathbf{F}_1^T \right]^{-1} \mathbf{B}^T,$$

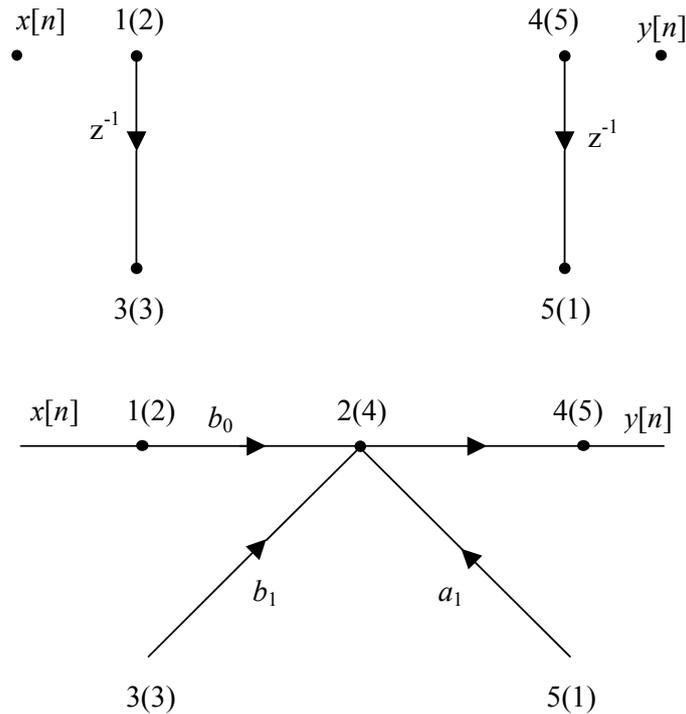


Рис. 3.13. Подграфы для цепи, приведенной на рис. 3.12

получим уравнения, которые в литературе принято называть уравнениями состояния:

$$\mathbf{w}[n] = \mathbf{A}\mathbf{w}[n - 1] + \mathbf{D}\mathbf{x}[n],$$

$$\mathbf{y}[n] = \mathbf{C}^T \mathbf{w}[n].$$

Уравнения состояния можно записать только в том случае, если матрица $[\mathbf{I} - \mathbf{F}_1^T]^{-1}$ существует, что эквивалентно вычислимости графа. Таким образом, каждая вычислимая цепь может быть описана уравнениями состояния.

3.11. ФОРМЫ РЕАЛИЗАЦИИ ЛИС-ЦЕПЕЙ

Выше было показано, что любая дискретная ЛИС-цепь имеет передаточную функцию дробно-рационального вида. В то же время одна и та же дробно-рациональная функция может быть представлена в различных формах, что соответствует разным формам реализации ЛИС-цепи.

А. Прямая форма

Прямая форма реализации ЛИС-цепи соответствует представлению передаточной функции в виде произведения (3.7) полинома $B(z)$ и дроби $1/A(z)$. Поскольку произведение передаточных функций соответствует каскадному соединению цепей, то прямая форма реализации ЛИС-цепи конечного порядка представляет собой каскадное соединение трансверсальной части с передаточной функцией $B(z)$ и рекурсивной части с передаточной функцией $1/A(z)$. Структурная схема и граф ЛИС-цепи в прямой форме показаны на рис. 3.14.

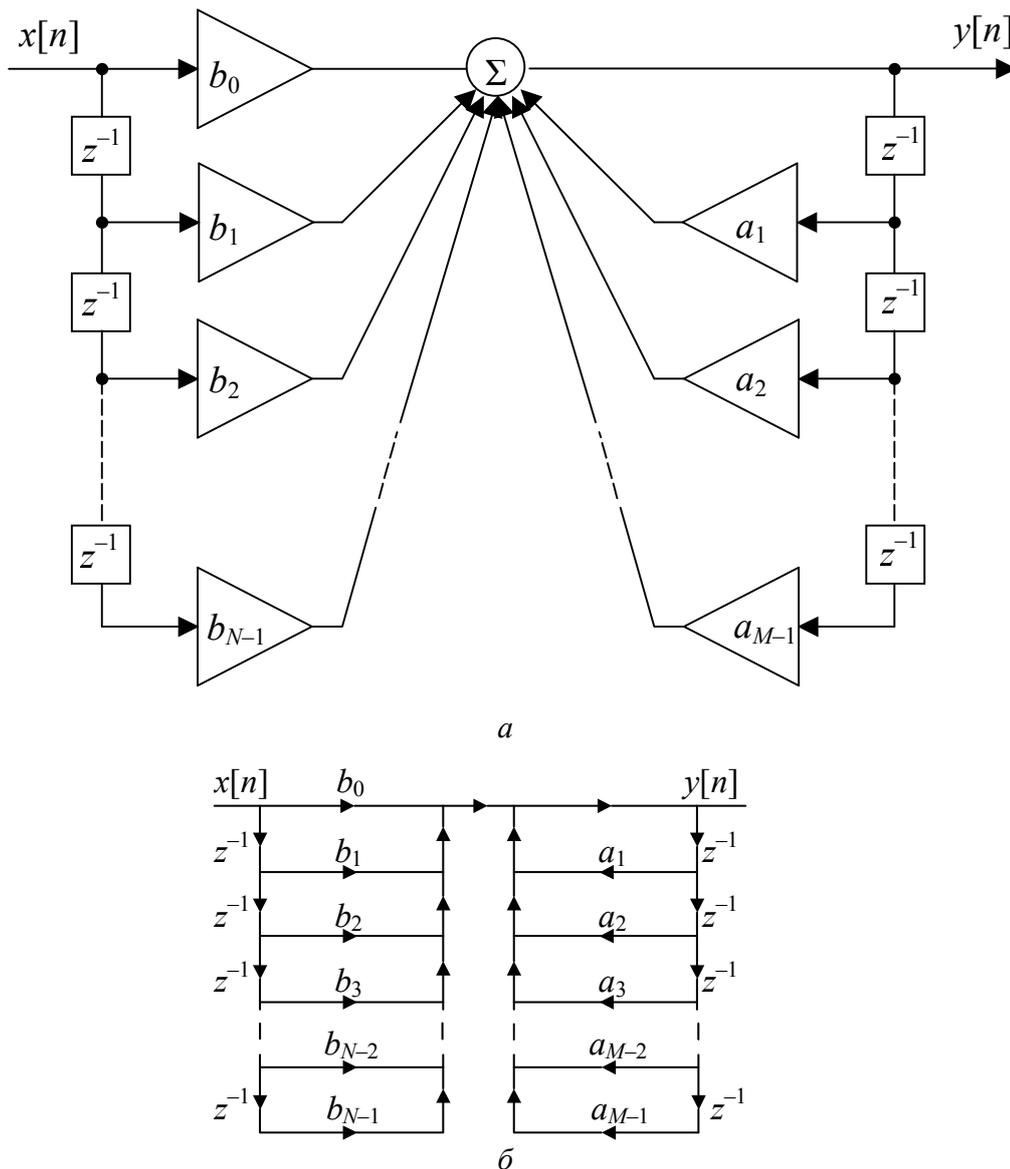


Рис. 3.14. Структурная схема ЛИС-цепи в прямой форме (а) и соответствующий ей сигнальный граф (б)

Б. Каноническая форма

Поменяв местами сомножители передаточной функции и, соответственно, трансверсальную и рекурсивную части структурной схемы, получаем каноническую форму реализации цепи. Каноническая форма имеет преимущество перед прямой формой, состоящее в том, что она позволяет отождествить ветви рекурсивной и трансверсальной частей, содержащие элементы задержки, так как по этим ветвям распространяется один и тот же сигнал. Структурная схема и граф ЛИС-цепи в канонической форме показаны на рис. 3.15. Здесь принято $N = M$; если это не так, схема и граф изменяются очевидным образом.

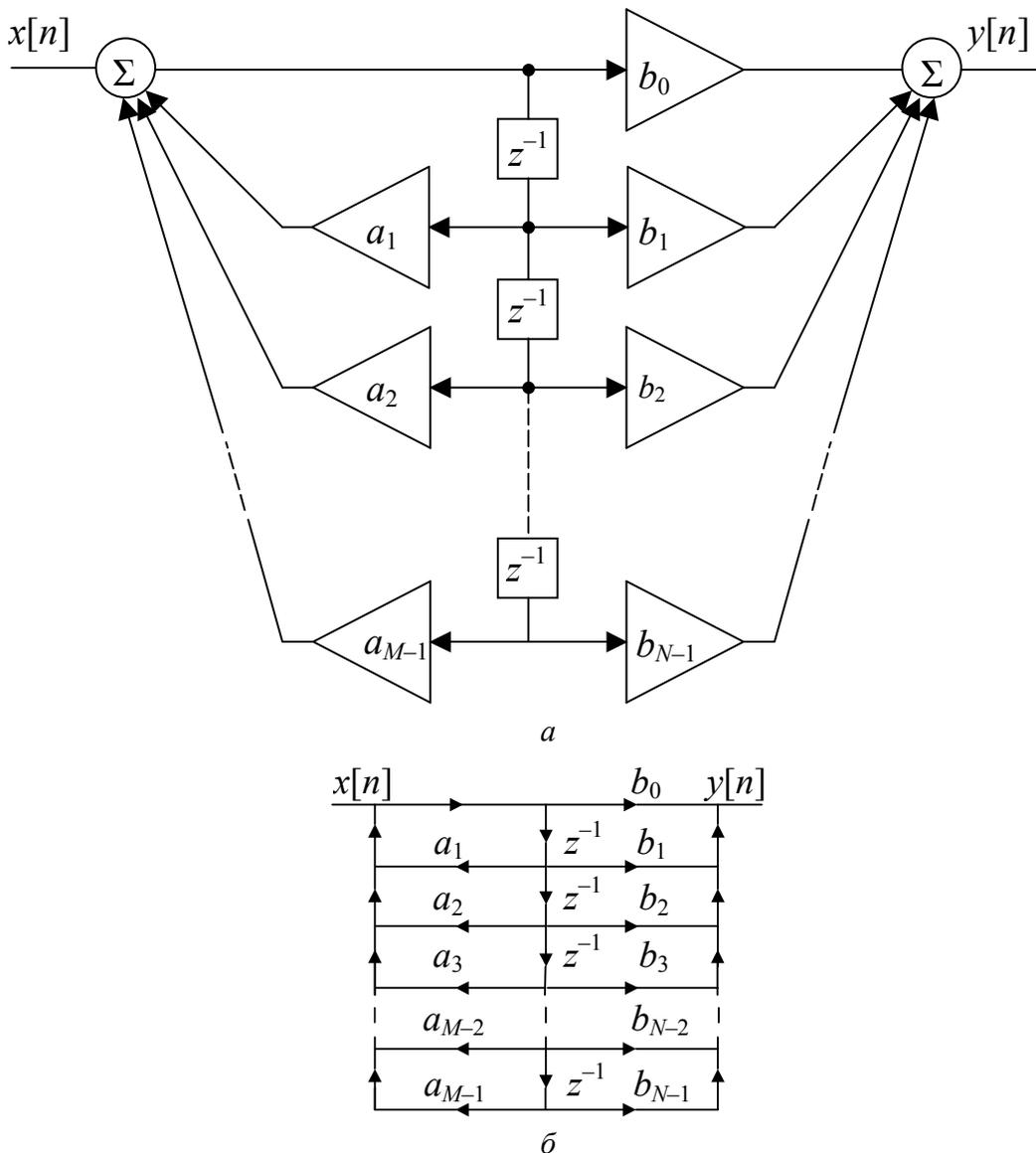


Рис. 3.15. Структурная схема ЛИС-цепи в канонической форме (а) и соответствующий ей сигнальный граф (б)

В. Параллельная форма

Параллельная форма реализации ЛИС-цепи соответствует разложению передаточной функции в сумму элементарных дробей, уже рассмотренному в п. 3.5. Структурная схема цепи показана на рис. 3.5, а соответствующий ей граф приведен на рис. 3.16.

Если степень числителя передаточной функции цепи не меньше степени знаменателя, то структурная схема и сигнальный граф дополняются трансверсальной цепью соответствующего порядка, подключаемой параллельно цепи, показанной на рис. 3.5 и 3.16.

Г. Каскадная форма

Разложению (3.14) числителя и знаменателя передаточной функции в произведения двучленов соответствует каскадная форма реализации ЛИС-цепи. Структурная схема (и сигнальный граф) состоит из звеньев двух типов: трансверсальных звеньев первого порядка и простейших рекурсивных звеньев. Сигнальный граф каскадной формы ЛИС-цепи показан на рис. 3.17 (структурная схема не приводится из-за ее громоздкости).

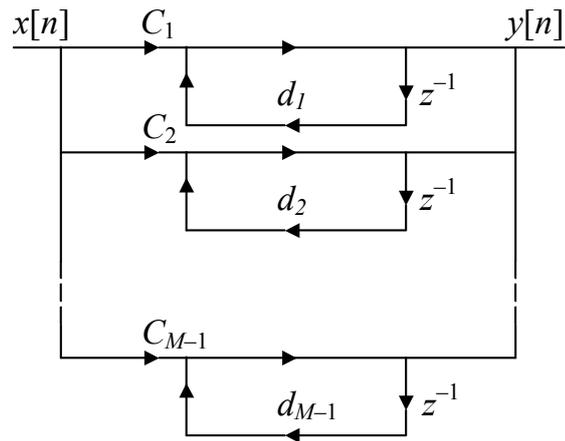


Рис. 3.16. Сигнальный граф ЛИС-цепи в параллельной форме

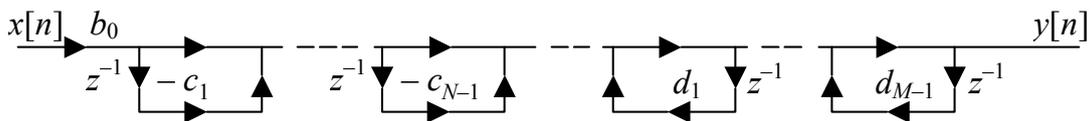


Рис. 3.17. Сигнальный граф ЛИС-цепи в каскадной форме

Д. Обращенные формы

Представление цепей сигнальными графами позволяет получать новые формы реализации цепей путем применения к известным формам определенных преобразований. Одно из таких преобразований состоит в замене всех направлений ветвей на противоположные и называется *транспозицией* графа. При этом вход и выход цепи также меняются местами. Для цепи с одним входом и одним выходом передаточная функция цепи, полученной транспозицией графа, совпадает с передаточной функцией исходной цепи [2].

Следует иметь в виду, что в каскадной и параллельной формах реализации ЛИС-цепей, получаемых соединением цепей первого порядка, коэффициенты масштабных звеньев могут быть комплексными. Это неудобно с точки зрения практической реализации, поэтому обычно простейшие звенья с комплексно-сопряженными коэффициентами группируют; при этом получаются звенья второго порядка с вещественными коэффициентами. Группирование проводится путем сложения передаточных функций двух параллельно соединенных простейших звеньев или умножения передаточных функций двух каскадно соединенных простейших звеньев с последующим приведением к общему знаменателю или раскрытием скобок.

Рассмотренные формы реализации ЛИС-цепей не исчерпывают всех возможных вариантов. Существуют и другие формы, среди которых наиболее известными являются так называемые волновые и лестничные структуры (подробности см. в [14]).

Следует иметь в виду, что разные формы реализации одной передаточной функции эквивалентны, если цепь является *дискретной*. Различные формы реализации *цифровых* цепей оказываются неэквивалентными, так как имеют разную чувствительность к округлению коэффициентов масштабных звеньев, а также к квантованию промежуточных результатов. Подробнее этот вопрос будет рассмотрен в разд. 7.

Упражнения

1. Докажите, что любая КИХ-цепь устойчива.
2. Докажите, что ЛИС-цепь конечного порядка имеет в общем случае бесконечно длинную импульсную характеристику.
3. Начертите структурную схему цепи из примера 12 в виде параллельного соединения трансверсальной цепи и простейших рекурсивных цепей и запишите ее импульсную характеристику.
4. Исследуйте поведение передаточной функции (3.6) при $z = 0$ и при $z = \infty$. Выведите формулу для определения кратности нулей и полюсов в этих точках.
5. Докажите, что каузальные БИХ-цепи с линейными фазочастотными характеристиками не существуют.
6. Докажите, что цепь с нуль-полюсной диаграммой, представленной на рис. 3.8, действительно является всепропускающей.
7. Запишите уравнения состояния для цепи, показанной на рис. 3.12.

8. Найдите импульсную характеристику цепи $h[n]$, если передаточная функция $H(z) = \frac{1}{1 - z^{-N}}$, $|z| < 1$. Постройте нуль-полюсную диаграмму.

9. Найдите передаточную функцию цепи с импульсной характеристикой

$$h[n] = \begin{cases} 1, & |n| \leq 2, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Постройте нуль-полюсную диаграмму.

10. Найдите z -преобразование последовательности

$$x[n] = \begin{cases} 1, & n = 0, N, 2N, \dots, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Постройте нуль-полюсную диаграмму.

11. Передаточная функция цепи

$$H(z) = \frac{1 - z^{-4}}{1 - z^{-1}}.$$

Найдите импульсную характеристику. Постройте нуль-полюсную диаграмму.

12. Найдите свертку $x[n]$ и $h[n]$ с использованием z -преобразования, если

$$x[n] = \begin{cases} 1, & 0 \leq n \leq 3, \\ 0 & \text{в противном случае,} \end{cases}$$

$$h[n] = \begin{cases} 1, & 2 \leq n \leq 4, \\ 0 & \text{в противном случае.} \end{cases}$$

13. Найдите $x[n] \otimes y[n]$, если

$$x[n] = \begin{cases} 1, & n = 0, \\ -a, & n = 1, \\ 0 & \text{в остальных случаях,} \end{cases}$$

$$y[n] = a^n, \quad n \geq 0.$$

14. Преобразуйте цепь, показанную на рис. 3.18, к параллельному виду.

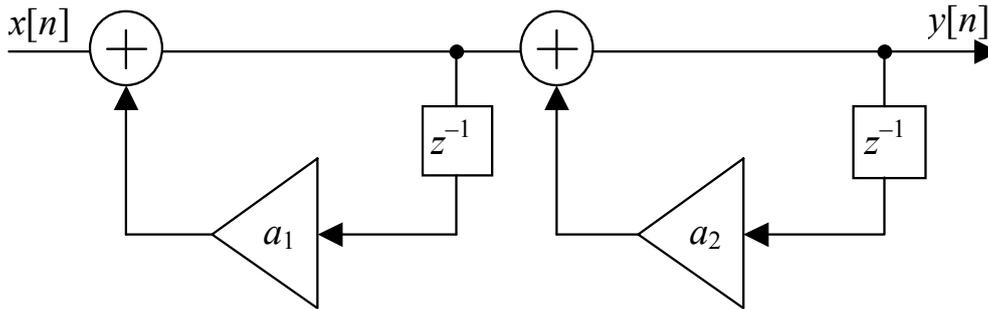


Рис. 3.18. К упражнению 14

15. Найдите импульсную характеристику цепи, нуль-полюсная диаграмма которой показана на рис. 3.19.

16. Найдите импульсную характеристику цепи, нуль-полюсная диаграмма которой показана на рис. 3.20.

17. Для дискретной цепи, описываемой разностным уравнением вида

$$y[n] = x[n] - x[n - 5] + y[n - 1],$$

постройте структурную схему. Найдите эквивалентную нерекурсивную цепь, постройте структурную схему. Найдите отклик этой

цепи на последовательность $x[n] = \begin{cases} 1, & 0 \leq n \leq 5, \\ 0, & n < 0 \text{ или } n > 5. \end{cases}$

18. Выполните транспозицию сигнальных графов, соответствующих прямой, канонической, параллельной и каскадной формам реализации ЛИС-цепей. Сравните полученные формы с известными.

19. Покажите, что при группировании двух простейших звеньев с комплексно-сопряженными коэффициентами получается звено второго порядка с вещественными коэффициентами.

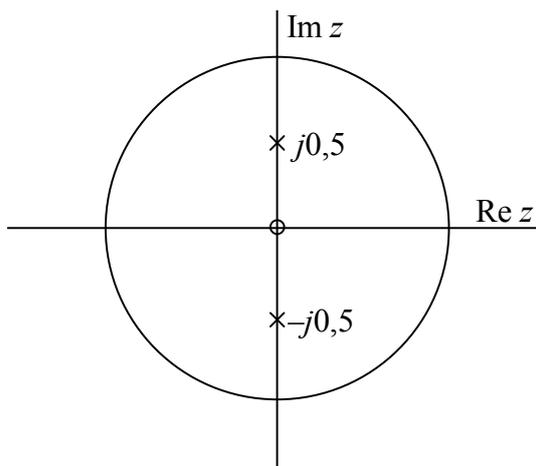


Рис. 3.19. К упражнению 15

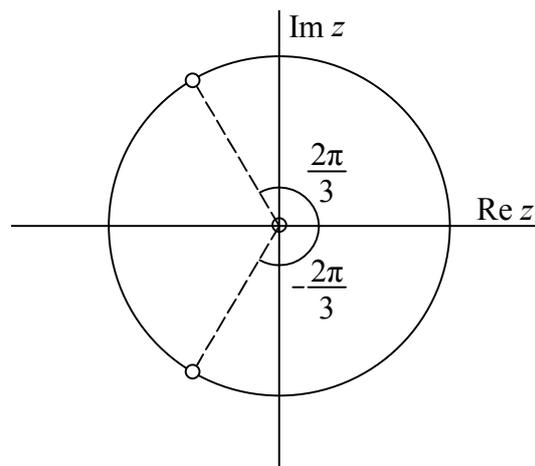
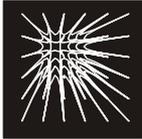


Рис. 3.20. К упражнению 16

Вопросы для самоконтроля

1. Какие цепи выполняют умножение и деление z -образа входной последовательности на полином?
2. Как преобразовать разностное уравнение в передаточную функцию цепи?
3. Полностью ли описывается динамика дискретной цепи разностным уравнением?
4. Какие особенности имеет разностное уравнение каузальной цепи?
5. Может ли дробно-рациональная передаточная функция соответствовать конечной импульсной характеристике?
6. Может ли рекурсивная цепь конечного порядка иметь конечную импульсную характеристику? Объясните ответ с позиций деления полиномов.
7. Как влияет на вид импульсной характеристики ЛИС-цепи наличие: а) кратных полюсов? б) комплексно-сопряженных пар полюсов?
8. Как влияет на вид импульсной характеристики ЛИС-цепи превышение степени числителя передаточной функции над степенью знаменателя?
9. Однозначно ли связаны передаточная функция и нуль-полюсная диаграмма цепи?
10. Как называется цепь, все нули которой находятся внутри единичной окружности? В чем заключается особенность таких цепей?
11. Можно ли построить ЛИС-цепь с импульсной характеристикой вида $h[n] = 1/(n + 1)$, $n \geq 0$. Если да, то как? Если нет, то почему?
12. Всегда ли передаточная функция ЛИС-цепи дробно-рациональна?
13. Какие ограничения существуют на расположение нулей и полюсов?
14. Как можно истолковать требование о том, что все полюсы устойчивой каузальной цепи должны находиться внутри 1-окружности?
15. Как должны располагаться полюсы устойчивой антикаузальной цепи?
16. Каков смысл комплексной частотной характеристики ЛИС-цепи?
17. Можно ли реализовать цепь со строго постоянной АЧХ? Если можно, то как?
18. Что такое вычислимость? Что необходимо для вычислимости?
19. Как можно обеспечить вычислимость некаузальной цепи?
20. Что такое сигнальный граф?
21. Перечислите возможные формы реализации цифровых цепей.
22. Что такое транспозиция и для чего она применяется?



4. СЛУЧАЙНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ И ЛИС-ЦЕПИ

До сих пор предполагалось, что все последовательности являются детерминированными (неслучайными). Были рассмотрены различные описания последовательностей (временное, частотное и представление в виде z -образа), а также методы анализа воздействия детерминированных последовательностей на ЛИС-цепи. На практике *все* сигналы являются случайными, так как если сигнал детерминирован (предопределен, известен точно), то нет нужды его принимать, а следовательно, и передавать. Сигнал как переносчик информации всегда случаен; помеха также всегда случайна (в противном случае ее можно было бы скомпенсировать и тем самым избавиться от ее вредного воздействия). Вместе с тем анализ детерминированных сигналов проще, и, кроме того, для решения многих задач модель детерминированного сигнала вполне приемлема. Однако многие практические задачи не могут быть решены в рамках детерминированного подхода. Таковы, например, задачи, связанные с анализом помехоустойчивости систем связи, с анализом шумов квантования и эффектов округления промежуточных результатов в системах цифровой обработки сигналов и т.п. Поэтому важно научиться анализировать дискретные случайные сигналы (случайные последовательности) и их воздействие на дискретные цепи.

4.1. СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

Случайная последовательность представляется совокупностью (ансамблем) детерминированных последовательностей, каждая из которых может реализоваться в эксперименте (наблюдении) в соответствии с некоторым случайным механизмом. Ансамбль реализаций можно описать как функцию двух переменных, одна из

которых – n – дискретна, а другая – λ – может быть дискретной или непрерывной и принимать значения из соответствующего множества Λ ; переменную λ удобно использовать в виде индекса («номера» реализации в ансамбле):

$$\{x_\lambda[n], \quad n = \overline{-\infty, \infty}; \quad \lambda \in \Lambda\}.$$

Случайный механизм порождения реализаций описывается распределением P_λ вероятности, заданным на множестве Λ .

Иногда полезно рассматривать случайную последовательность через «сечения»: при фиксированном λ сечение $x_\lambda[n]$, $n = \overline{-\infty, \infty}$, является обычной детерминированной последовательностью, а при фиксированном n совокупность значений $x_\lambda[n]$ совместно с распределением вероятности на Λ представляет собой случайную величину.

Как всякая случайная величина, случайная последовательность в сечении n исчерпывающе описывается *функцией распределения* вероятности

$$F(x, n) = P\{x_\lambda[n] \leq x\}.$$

Далее, как правило, будем полагать, что Λ – множество вещественных чисел (λ – непрерывная переменная). В таком случае эквивалентный способ описания случайной последовательности состоит в задании плотности распределения вероятности

$$w(x, n) = \frac{dF(x, n)}{dx}.$$

Плотность распределения вероятности можно формально ввести и для дискретных случайных величин, если допустить, что плотность может быть *сингулярной* (содержать дельта-функции).

Часто не требуется полное описание сечения случайной последовательности и можно ограничиться числовыми характеристиками. Простейшая числовая характеристика – математическое ожидание (среднее значение, или просто *среднее*) случайной последовательности

$$\bar{x}[n] = \mathbf{E}\{x[n]\} = \int_{-\infty}^{\infty} x[n]w(x, n)dx$$

определяется как 1-й начальный момент плотности распределения вероятности $w(x, n)$. Символ $\mathbf{E}\{\cdot\}$ используется для обозначения

оператора усреднения по ансамблю. Дисперсия (средняя мощность флюктуационной составляющей) случайной последовательности определяется как второй центральный момент

$$D_x[n] = \sigma_x^2[n] = \mathbf{E} \left\{ |x[n] - \bar{x}[n]|^2 \right\} = \int_{-\infty}^{\infty} |x[n] - \bar{x}[n]|^2 w(x, n) dx .$$

Очевидно, среднее и дисперсия представляют собой *детерминированные* последовательности. Полное описание случайной последовательности в некотором сечении эквивалентно заданию *всех* начальных и центральных моментов, определяемых соответственно выражениями

$$m_k[n] = \mathbf{E} \left\{ x^k[n] \right\} = \int_{-\infty}^{\infty} x^k[n] w(x, n) dx ,$$

$$M_k[n] = \mathbf{E} \left\{ (x[n] - \bar{x}[n])^k \right\} = \int_{-\infty}^{\infty} (x[n] - \bar{x}[n])^k w(x, n) dx .$$

Характеристики случайной последовательности, относящиеся к единственному сечению при фиксированном n , ничего не говорят о ее *динамических* (частотных) свойствах. Для описания динамики случайной последовательности следует рассмотреть степень статистической связи (зависимости) различных сечений случайной последовательности. Исчерпывающее описание случайной последовательности дается N -мерной (совместной) плотностью распределения вероятности для N значений последовательности, относящихся к N различным значениям аргумента n

$$w_N(x[n_1], x[n_2], \dots, x[n_N])$$

при произвольном N . Во многих случаях полное описание не требуется и достаточно ограничиться двумерной плотностью распределения вероятности $w_2(x[n_1], x[n_2])$ или даже только моментами первого и второго порядков, включая второй центральный смешанный (корреляционный) момент, этой плотности

$$\begin{aligned} r_x[n_1, n_2] &= \mathbf{E} \left\{ (x[n_1] - \bar{x}[n_1]) (x^*[n_2] - \bar{x}^*[n_2]) \right\} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x[n_1] - \bar{x}[n_1]) (x^*[n_2] - \bar{x}^*[n_2]) w_2(x_1, x_2, n_1, n_2) dx_1 dx_2 . \end{aligned}$$

Широкий и удобный для многих практических целей класс случайных последовательностей образуют *стационарные* случайные последовательности. Различают два вида стационарности.

Стационарной в узком смысле называют случайную последовательность, у которой N -мерная совместная плотность распределения вероятности не меняется при одновременном сдвиге всех N сечений влево или вправо по оси «времени». Стационарность в широком смысле означает инвариантность к сдвигу лишь *моментов не выше второго порядка* двумерной плотности распределения случайной последовательности. Таким образом, случайная последовательность стационарна в широком смысле, если ее среднее $\bar{x}[n] = \bar{x}$ и дисперсия $D_x[n] = D_x = \sigma_x^2$ не зависят от n , а корреляционный момент для двух сечений при $n = n_1$ и $n = n_2$ зависит только от разности $(n_2 - n_1)$, т.е. представляет собой функцию одной (дискретной) переменной, называемую автокорреляционной последовательностью (АКП):

$$r_x[m] = \mathbf{E}\{(x[n] - \bar{x})(x^*[n+m] - \bar{x}^*)\}.$$

Заметим, что $r_x[0] = D_x$.

Нетрудно видеть, что случайная последовательность, стационарная в узком смысле, стационарна и в широком смысле; обратное в *общем случае* неверно.

Описание случайной последовательности в терминах среднего и АКП (в терминах корреляционной теории) в общем случае не является полным, так как игнорирует моменты высших порядков (только совокупность *всех* моментов, в том числе смешанных, содержит полную информацию о многомерной совместной плотности распределения вероятности [15]). Тем не менее простота корреляционного описания случайных последовательностей и возможность решать на его основе содержательные задачи (например, задачу анализа воздействия стационарных случайных последовательностей на ЛИС-цепи) обеспечили широкое применение корреляционной теории случайных последовательностей. Этому способствует также то обстоятельство, что для *гауссовской* модели, которая удовлетворительно описывает многие реальные случайные последовательности, корреляционное описание является исчерпывающим (моменты порядков выше второго однозначно выражаются через моменты 1-го и 2-го порядков), а понятия стационарности в узком и широком смысле *совпадают*.

Помимо характеристик, относящихся к одной случайной последовательности, можно рассматривать взаимно корреляционную последовательность (ВКП)

$$r_{xy}[n_1, n_2] = \mathbf{E}\{(x[n_1] - \bar{x}[n_1])(y^*[n_2] - \bar{y}^*[n_2])\}$$

как простейшую меру статистической зависимости двух различных последовательностей $x[n]$ и $y[n]$ в сечениях n_1 и n_2 . Если ВКП зависит только от разности $(n_2 - n_1)$, то последовательности называются *стационарно связанными*.

4.2. ВОЗДЕЙСТВИЕ СТАЦИОНАРНЫХ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ НА ЛИС-ЦЕПИ

При рассмотрении воздействия на ЛИС-цепь каждой реализации случайной последовательности в отдельности никаких новых задач не возникает, так как отдельная реализация случайной последовательности есть детерминированная последовательность. Особенность *статистического* анализа состоит в том, что входная и выходная случайные последовательности рассматриваются как *ансамбли* с соответствующими многомерными плотностями распределения вероятности. В общем случае получение полного вероятностного описания выходной последовательности (N -мерной совместной плотности распределения вероятности) представляет собой непростую задачу; достаточно просто можно получить решение задачи анализа для стационарных (в широком смысле) случайных последовательностей, если рассматривать лишь моменты 1-го и 2-го порядков (напомним, что для гауссовских случайных последовательностей они составляют исчерпывающее описание).

В рамках корреляционной теории стационарная в широком смысле случайная последовательность $x[n]$, воздействующая на вход ЛИС-цепи, описывается средним \bar{x} и АКП $r_x[m]$. Очевидно, можно ставить вопросы об определении среднего \bar{y} и автокорреляционной последовательности $r_y[m]$ выходной случайной последовательности $y[n]$, а также взаимно корреляционной последовательности $r_{xy}[m]$. ЛИС-цепь задается своей импульсной характеристикой $h[n]$, которая предполагается абсолютно суммируемой.

Учитывая линейность цепи, можно рассматривать воздействие на цепь среднего $\bar{x}[n]$ входной последовательности $x[n]$ и ее

флюктуационной составляющей $x[n] - \bar{x}[n]$ по отдельности. Поскольку $\bar{x}[n] = \bar{x}$ – детерминированная последовательность, постоянная в силу стационарности $x[n]$, можно записать математическое ожидание $\bar{y}[n]$ выходной последовательности как свертку

$$\bar{y}[n] = \sum_{k=-\infty}^{\infty} h[k] \bar{x}[n-k] = \bar{x} \sum_{k=-\infty}^{\infty} h[k] = \bar{x} H(e^{j\omega}) \Big|_{\omega=0}.$$

Как и следовало ожидать, коэффициент усиления постоянной составляющей равен значению КЧХ на нулевой частоте (конечному вследствие абсолютной суммируемости $h[n]$). Далее для простоты анализа среднее входной последовательности полагается равным нулю.

Анализ воздействия стационарной случайной последовательности с нулевым средним на ЛИС-цепь может быть проведен для каждой реализации в отдельности с последующим усреднением по ансамблю реализаций выходной случайной последовательности. Трудность состоит в том, что (в силу предположения о стационарности) реализации случайной последовательности $x[n]$ с вероятностью 1 не являются ни абсолютно, ни даже квадратично суммируемыми, поэтому для них не существуют ни z -преобразование, ни преобразование Фурье. Однако поскольку нас будут интересовать усредненные характеристики последовательностей (АКП и ВКП), указанная трудность достаточно просто преодолевается: следует только рассматривать выражения для реализаций как формальные, а физическую интерпретацию производить лишь после усреднения по ансамблю.

Запишем выражение для реализации случайной последовательности $y[n]$ в виде свертки

$$y[n] = \sum_{k=-\infty}^{\infty} x[k] h[n-k].$$

Автокорреляционная последовательность для $y[n]$ находится усреднением

$$\begin{aligned} r_y[m] &= \mathbf{E}\{y[n]y^*[n+m]\} = \\ &= \mathbf{E}\left\{ \sum_{k=-\infty}^{\infty} h[k]x[n-k] \sum_{l=-\infty}^{\infty} h^*[l]x^*[n+m-l] \right\} = \end{aligned}$$

$$\begin{aligned}
 &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h[k]h^*[l]\mathbf{E}\{x[n-k]x^*[n+m-l]\} = \\
 &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h[k]h^*[l]r_x[m+k-l] = \\
 &= \sum_{\mu=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} h[k]h^*[k+\mu]r_x[m-\mu] = \sum_{\mu=-\infty}^{\infty} r_h[\mu]r_x[m-\mu].
 \end{aligned}$$

Здесь использовано свойство линейности оператора усреднения $\mathbf{E}\{\cdot\}$. Окончательное выражение

$$r_y[m] = \sum_{\mu=-\infty}^{\infty} r_h[\mu]r_x[m-\mu]$$

имеет форму свертки АКП входной последовательности с последовательностью

$$r_h[\mu] = \sum_{k=-\infty}^{\infty} h[k]h^*[k+\mu],$$

которую можно трактовать как автокорреляционную последовательность импульсной характеристики, полученную усреднением по «времени», в роли которого выступает индекс суммирования k .

Отсюда видно, что если АКП входной последовательности и АКП импульсной характеристики абсолютно суммируемы, такова же будет и АКП выходной последовательности. В этом случае сходятся на 1-окружности соответствующие z -преобразования и можно записать

$$R_y(z) = R_h(z)R_x(z), \quad (4.1)$$

где

$$\begin{aligned}
 R_y(z) &= \sum_{m=-\infty}^{\infty} r_y[m]z^{-m}, & R_x(z) &= \sum_{m=-\infty}^{\infty} r_x[m]z^{-m}, \\
 R_h(z) &= \sum_{m=-\infty}^{\infty} r_h[m]z^{-m}.
 \end{aligned}$$

Характеристикой цепи в выражении (4.1) является функция $R_h(z)$, но, как известно, цепь исчерпывающе описывается своей

передаточной функцией. Поэтому желательно выяснить связь функции $R_h(z)$ с передаточной функцией цепи $H(z)$. Рассмотрим $r_h[m]$ как свертку

$$r_h[m] = \sum_{k=-\infty}^{\infty} h[k]h_1[m-k],$$

где введено обозначение $h_1[m-k] = h^*[k+m]$. Очевидно, $h_1[-k] = h^*[k]$ или $h_1[k] = h^*[-k]$.

Найдем для $h_1[k]$ z -образ

$$\begin{aligned} H_1(z) &= \sum_{k=-\infty}^{\infty} h_1[k]z^{-k} = \sum_{k=-\infty}^{\infty} h^*[-k]z^{-k} = \sum_{n=-\infty}^{\infty} h^*[n]z^n = \\ &= \left\{ \sum_{n=-\infty}^{\infty} h[n](z^*)^n \right\}^* = \left\{ \sum_{n=-\infty}^{\infty} h[n](1/z^*)^{-n} \right\}^* = H^*(1/z^*). \end{aligned}$$

Свертке последовательностей соответствует произведение z -образов, поэтому

$$R_h(z) = H(z)H_1(z) = H(z)H^*(1/z^*),$$

следовательно, выражение (4.1) можно переписать в виде

$$R_y(z) = H(z)H^*(1/z^*)R_x(z).$$

В силу абсолютной суммируемости автокорреляционных последовательностей области сходимости соответствующих z -преобразований содержат 1-окружность и можно рассматривать частотные описания случайных последовательностей как сужения на 1-окружность соответствующих z -образов

$$\begin{aligned} R_y(e^{j\omega}) &= H(e^{j\omega})H^*(1/e^{-j\omega})R_x(e^{j\omega}) = \\ &= |H(e^{j\omega})|^2 R_x(e^{j\omega}). \end{aligned}$$

Для частотной функции $R_x(e^{j\omega})$ можно записать

$$R_x(e^{j\omega}) = \sum_{m=-\infty}^{\infty} r_x[m]e^{-j\omega m}, \quad (4.2)$$

и обратно

$$r_x[m] = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_x(e^{j\omega}) e^{j\omega m} d\omega. \quad (4.3)$$

Аналогично связаны между собой $r_y[m]$ и $R_y(e^{j\omega})$.

Пара выражений (4.2) и (4.3) представляет собой формулировку *теоремы Винера – Хинчина* для случайных последовательностей.

Для уяснения физического смысла функции $R_x(e^{j\omega})$ положим в (4.3) $m = 0$. Тогда

$$r_x[0] = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_x(e^{j\omega}) d\omega = D_x$$

есть мощность случайной последовательности $x[n]$. Поэтому функция $R_x(e^{j\omega})$ описывает распределение мощности случайной последовательности по оси частот (точнее, по 1-окружности) и является, следовательно, *спектральной плотностью мощности* (СПМ) случайной последовательности.

Рассмотрим взаимную корреляцию входной и выходной последовательностей

$$\begin{aligned} r_{xy}[m] &= \mathbf{E}\{x[n]y^*[n+m]\} = \mathbf{E}\left\{x[n] \sum_{k=-\infty}^{\infty} h^*[k]x^*[n+m-k]\right\} = \\ &= \sum_{k=-\infty}^{\infty} h^*[k] \mathbf{E}\{x[n]x^*[n+m-k]\} = \sum_{k=-\infty}^{\infty} h^*[k]r_x[m-k]. \end{aligned}$$

Переходя к z -образам, получим

$$R_{xy}(z) = H^*(z^*)R_x(z).$$

Здесь использовано свойство z -преобразования для комплексно-сопряженной последовательности (см. Приложение 2)

$$\sum_{n=-\infty}^{\infty} h^*[n]z^{-n} = \left(\sum_{n=-\infty}^{\infty} h[n](z^*)^{-n} \right)^* = H^*(z^*).$$

Аналогично найдем

$$r_{yx}[m] = \mathbf{E}\{y[n]x^*[n+m]\} = \mathbf{E}\left\{ \sum_{k=-\infty}^{\infty} h[k]x[n-k]x^*[n+m] \right\} =$$

$$= \sum_{k=-\infty}^{\infty} h[k]r_x[m+k] = \sum_{k=-\infty}^{\infty} h[k]r_{x1}[m-k],$$

где $r_{x1}[k] = r_x[-k]$.

Поскольку для последовательности, обращенной во времени, z -образ

$$\sum_{n=-\infty}^{\infty} r_x[-n]z^{-n} = \sum_{n=-\infty}^{\infty} r_x[n]\left(\frac{1}{z}\right)^{-n} = R_x\left(\frac{1}{z}\right),$$

то справедливо выражение

$$R_{yx}(z) = H(z)R_x\left(\frac{1}{z}\right).$$

Выражения для взаимных спектральных плотностей мощности как функций частоты получаются сужением соответствующих z -образов на 1-окружность (подстановкой $z = e^{j\omega}$).

Упражнения

1. Запишите выражения для взаимных спектральных плотностей мощности входной и выходной последовательностей через СПМ входной последовательности и КЧХ цепи.

2. Докажите, что случайная последовательность, стационарная в узком смысле, стационарна и в широком смысле.

3. Докажите, что гауссовская случайная последовательность, стационарная в широком смысле, стационарна и в узком смысле.

Вопросы для самоконтроля

1. Что такое усреднение по ансамблю?
2. Что такое усреднение по времени?
3. Что такое моменты и для чего они применяются?
4. Какие последовательности называются стационарными в широком смысле? В узком смысле? Когда эти понятия эквивалентны?



5. МНОГОМЕРНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ И ЛИС-ЦЕПИ

5.1. МНОГОМЕРНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

До сих пор рассматривались одномерные последовательности – функции одной дискретной (целой) переменной. Обычно аргументу одномерной последовательности приписывается физический смысл дискретного времени. Особенность цифровой обработки сигналов состоит в том, что она успешно применяется к последовательностям с аргументами произвольной физической природы (временем, длиной, волновым числом, углом и т.п.), а также произвольной размерности. Здесь будут рассмотрены способы описания *многомерных* последовательностей и анализ их воздействия на многомерные ЛИС-цепи на примере двумерных цепей и последовательностей. Распространение полученных выводов на общий случай достаточно элементарно. Дело в том, что теория двумерных последовательностей и цепей *принципиально* отличается в некоторых своих чертах от одномерной теории; дальнейшее же увеличение размерности приводит лишь к количественному усложнению выражений. Поэтому в рассуждениях о многомерных сигналах и цепях в целях наглядности будут фигурировать, как правило, двумерные последовательности (чаще всего дискретизированные изображения) и двумерные дискретные цепи.

Двумерная последовательность $x[n_1, n_2]$ может описывать, например, яркость (или оптическую плотность) полутонного черно-белого изображения как функцию целых координат n_1, n_2 на плоскости, принимающую вещественные значения. Таким образом, двумерная последовательность описывается функцией, заданной на двумерной решетке, образованной точками с целыми координатами. Введя для аргументов последовательности векторное обозначение

$$\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} = (n_1, n_2)^T,$$

можно записать обозначение двумерной последовательности в виде $x[n_1, n_2] = x[\mathbf{n}]$. Такое обозначение удобно тем, что оно не содержит явного указания на размерность аргумента и получаемые при рассмотрении двумерных последовательностей выражения могут быть прямо обобщены на случай произвольной размерности – для этого достаточно \mathbf{n} интерпретировать как вектор-столбец с требуемым количеством координат

$$\mathbf{n} = (n_1, n_2, \dots, n_N)^T.$$

Рассмотрим некоторые элементарные двумерные последовательности.

1. Двумерная δ -последовательность

$$\delta[n_1, n_2] = \begin{cases} 1, & \text{если } n_1 = n_2 = 0, \\ 0 & \text{в противном случае.} \end{cases}$$

Очевидно, двумерная δ -последовательность может быть выражена через одномерную δ -последовательность

$$\delta[n_1, n_2] = \delta[n_1] \delta[n_2].$$

Здесь и далее мы используем одинаковое обозначение $\delta[\cdot]$ для одно- и двумерных δ -последовательностей, надеясь, что это не приведет к недоразумениям, так как из размерности аргумента всегда ясно, какая функция имеется в виду.

Вообще, если двумерная последовательность $x[n_1, n_2]$ может быть представлена в виде произведения одномерных последовательностей

$$x[n_1, n_2] = x_1[n_1] x_2[n_2],$$

то она называется *разделимой*.

Очевидно, двумерная δ -последовательность разделима; то же справедливо и для δ -последовательности произвольной размерности N

$$\delta[\mathbf{n}] = \prod_{i=1}^N \delta[n_i].$$

Множество точек двумерной решетки, на котором некоторая последовательность отлична от нуля, называется опорной областью или *носителем* этой последовательности. Очевидно, носителем δ -последовательности является единственная точка $\mathbf{n} = \mathbf{0}$.

2. Двумерная ступенчатая последовательность

$$u[n_1, n_2] = \begin{cases} 1, & n_1 \geq 0, n_2 \geq 0, \\ 0, & n_1 < 0, n_2 < 0 \end{cases}$$

также разделима:

$$u[n_1, n_2] = u[n_1]u[n_2].$$

Носителем $u[n_1, n_2]$ является множество точек, включающее начало координат, положительные полуоси и все точки решетки, лежащие в первом квадранте плоскости (рис. 5.1).

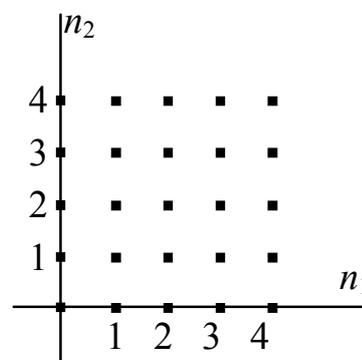


Рис. 5.1. Носитель двумерной ступенчатой последовательности

3. Двумерная экспоненциальная последовательность

$$e[n_1, n_2] = a^{n_1} b^{n_2}, \quad n_1, n_2 = \overline{-\infty, \infty},$$

очевидно, разделима (a и b – произвольные комплексные числа).

На самом деле разделимость последовательностей – скорее исключение, чем правило. Так, двумерная экспонента утрачивает разделимость, если изменить ее опорную область, например, таким образом:

$$e[n_1, n_2] = \begin{cases} a^{n_1} b^{n_2}, & n_1 = \overline{-\infty, \infty}, n_2 \geq n_1, \\ 0 & \text{в противном случае.} \end{cases}$$

4. Периодическая двумерная последовательность.

Понятие двумерной периодичности служит первым примером качественного отличия двумерного случая от одномерного [16].

Двумерная последовательность $\tilde{x}[n_1, n_2]$ периодична, если $\tilde{x}[n_1, n_2] = \tilde{x}[n_1 + N_{11}, n_2 + N_{21}]$ и $\tilde{x}[n_1, n_2] = \tilde{x}[n_1 + N_{12}, n_2 + N_{22}]$. В данном случае удобнее записать аргументы в виде столбца, тогда

$$\tilde{x} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \tilde{x} \begin{bmatrix} n_1 + N_{11} \\ n_2 + N_{21} \end{bmatrix} = \tilde{x} \begin{bmatrix} n_1 + N_{12} \\ n_2 + N_{22} \end{bmatrix}$$

или

$$\tilde{x}[\mathbf{n}] = \tilde{x}[\mathbf{n} + \mathbf{N}_1] = \tilde{x}[\mathbf{n} + \mathbf{N}_2].$$

Видно, что последовательность периодична, если она совпадает с собой при сдвиге по решетке в направлении вектора $\mathbf{N}_1 = \begin{pmatrix} N_{11} \\ N_{21} \end{pmatrix}$

или в направлении вектора $\mathbf{N}_2 = \begin{pmatrix} N_{12} \\ N_{22} \end{pmatrix}$. Очевидно, компоненты

N_{11} , N_{12} , N_{21} , N_{22} обязаны быть целыми. Также очевидно, что $\tilde{x}[\mathbf{n}]$ будет совпадать с собой, если сдвигать ее независимо вдоль \mathbf{N}_1 и \mathbf{N}_2 любое целое число раз, поэтому можно записать

$$\tilde{x}[\mathbf{n}] = \tilde{x}[\mathbf{n} + \mathbf{N}\mathbf{r}],$$

где

$$\mathbf{N} = (\mathbf{N}_1 | \mathbf{N}_2) = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix}$$

– матрица периодичности; $\mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$ – вектор с произвольными целыми компонентами.

На рис. 5.2 представлен пример периодической последовательности $\tilde{x}[\mathbf{n}]$ с матрицей периодичности

$$\mathbf{N} = \begin{pmatrix} 3 & 3 \\ 2 & -3 \end{pmatrix};$$

для наглядности принято, что $\tilde{x}[\mathbf{n}]$ образована периодическим повторением фрагмента с носителем, содержащим 6 точек, выделенных на рисунке. Этот фрагмент, всевозможными сдвигами которого вдоль векторов \mathbf{N}_1 и \mathbf{N}_2 порождается вся последовательность $\tilde{x}[\mathbf{n}]$, называется фундаментальным периодом [16].

Очевидно, что матрица периодичности не определяется однозначно, в частности, последовательности, представленной на

рис. 5.2, соответствуют кроме $\mathbf{N} = \begin{pmatrix} 3 & 3 \\ 2 & -3 \end{pmatrix}$ также матрицы

$$\mathbf{N}^{(1)} = \begin{pmatrix} 3 & -3 \\ 2 & 3 \end{pmatrix}, \mathbf{N}^{(2)} = \begin{pmatrix} 3 & 0 \\ 2 & 5 \end{pmatrix}, \mathbf{N}^{(3)} = \begin{pmatrix} -3 & 0 \\ -2 & 5 \end{pmatrix} \text{ и т.д.}$$

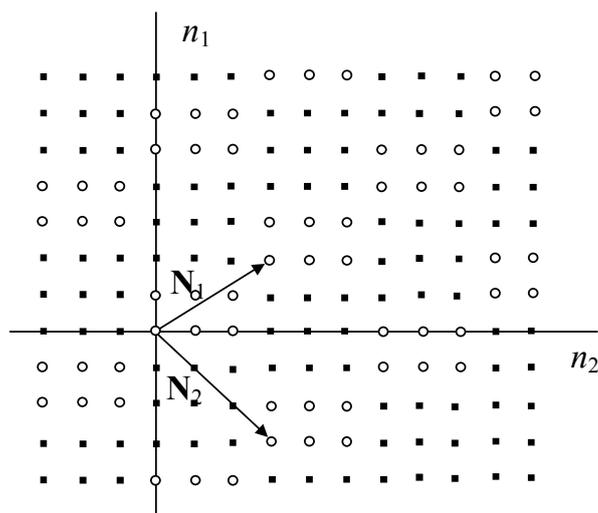


Рис. 5.2. Пример двумерно-периодической последовательности

Нетрудно заметить, что у всех этих матриц есть общая характеристика – модуль определителя

$$|\det \mathbf{N}| = |\det \mathbf{N}^{(1)}| = |\det \mathbf{N}^{(2)}| = |\det \mathbf{N}^{(3)}| = 15,$$

который представляет собой количество точек решетки, составляющих опорную область фундаментального периода последовательности $\tilde{x}[\mathbf{n}]$. Каким бы способом ни задавались матрица периодичности и фундаментальный период для данной периодической последовательности, эта последовательность на всей решетке принимает всего $|\det \mathbf{N}|$ различных значений. (Случай, когда $|\det \mathbf{N}| = 0$, предлагается читателю для самостоятельного рассмотрения.)

Приведенные примеры, разумеется, не исчерпывают всех последовательностей. Очевидно, можно построить бесконечное множество последовательностей, определенных на двумерной решетке точек с целыми координатами. Если определить на этом множестве операции сложения последовательностей и умножения последовательности на скаляр, получится линейное пространство, удовлетворяющее соответствующим аксиомам (см. п. 2.2). В этом пространстве можно ввести скалярное произведение произвольных последовательностей $x[\mathbf{n}]$ и $y[\mathbf{n}]$ выражением

$$(x, y) = \sum_{\mathbf{n}} x[\mathbf{n}]y^*[\mathbf{n}],$$

где при суммировании \mathbf{n} пробегает все точки бесконечной двумерной решетки S_2 . Можно определить также норму последовательности $x[\mathbf{n}]$ как $\|x\|_2 = \sqrt{(x, x)}$ и метрику (расстояние) выражением $d(x, y) = \|x - y\|_2$. Полученное пространство двумерных последовательностей ограниченной нормы есть гильбертово пространство (аналогичное рассмотренному ранее пространству одномерных последовательностей l_2). Таким способом можно построить пространство последовательностей произвольной размерности.

5.2. МНОГОМЕРНЫЕ ЛИС-ЦЕПИ

Многомерная цепь в общем случае выполняет отображение пространства входных последовательностей в пространство выходов. Ограничимся случаем входного и выходного пространств с совпадающей структурой, что соответствует задаче фильтрации. Тогда цепь описывается оператором $T\{\cdot\}$, так что

$$y[\mathbf{n}] = T\{x[\mathbf{n}]\},$$

где $x[\mathbf{n}]$ и $y[\mathbf{n}]$ – входная и выходная многомерные последовательности. Столь общая модель, конечно, слишком широка для получения конкретных результатов, поэтому далее рассматриваем только линейные цепи (операторы), удовлетворяющие принципу суперпозиции

$$T\{\alpha x_1[\mathbf{n}] + \beta x_2[\mathbf{n}]\} = \alpha T\{x_1[\mathbf{n}]\} + \beta T\{x_2[\mathbf{n}]\}.$$

Это означает, что отклик цепи на взвешенную сумму последовательностей равен взвешенной таким же образом сумме откликов. Выбрав в качестве базиса в пространстве последовательностей совокупность всех δ -последовательностей

$$\{\delta[\mathbf{n} - \mathbf{m}], \quad \mathbf{n}, \mathbf{m} \in S_N\},$$

где S_N – N -мерная решетка точек с целыми координатами, можно произвольную последовательность $x[\mathbf{n}]$ записать в виде обобщенного ряда Фурье относительно этого базиса:

$$x[\mathbf{n}] = \sum_{\mathbf{m}} x[\mathbf{m}] \delta[\mathbf{n} - \mathbf{m}].$$

На основании принципа суперпозиции выходная последовательность

$$y[\mathbf{n}] = \sum_{\mathbf{m}} x[\mathbf{m}]h[\mathbf{n}, \mathbf{m}],$$

где $h[\mathbf{n}, \mathbf{m}]$ – весовая последовательность (размерности $2N$), имеющая смысл отклика в точке \mathbf{n} на δ -последовательность с носителем в точке \mathbf{m} .

Предполагая, что $h[\mathbf{n}, \mathbf{m}]$ зависит лишь от разности $\mathbf{n} - \mathbf{m}$, т. е. оператор $T\{\cdot\}$ обладает свойством инвариантности к сдвигу, запишем отклик многомерной ЛИС-цепи

$$y[\mathbf{n}] = \sum_{\mathbf{m}} x[\mathbf{m}]h[\mathbf{n} - \mathbf{m}]. \quad (5.1)$$

Это выражение представляет собой многомерную дискретную свертку.

Нетрудно видеть, что многомерная свертка, как и одномерная, обладает свойством коммутативности, т.е. последовательности $x[\mathbf{n}]$ и $h[\mathbf{n}]$, входящие в выражение (5.1), равноправны. Это позволяет рассматривать их как элементы единого пространства последовательностей. В частности, абсолютно суммируемые последовательности с ограниченной нормой

$$\|x\|_1 = \sum_{\mathbf{n}} |x[\mathbf{n}]| < \infty$$

образуют нормированную алгебру (аналогичную l_1) с операцией многомерной свертки в качестве обобщенного умножения. Как и в одномерном случае, условие

$$\|h\|_1 = \sum_{\mathbf{n}} |h[\mathbf{n}]| < \infty$$

означает устойчивость многомерной цепи с импульсной характеристикой $h[\mathbf{n}]$.

Образует N -мерную гармоническую последовательность

$$x[\mathbf{n}] = \prod_{i=1}^N \exp(j\omega_i n_i) = \exp\left(j \sum_{i=1}^N \omega_i n_i\right) = \exp\{j(\boldsymbol{\omega}, \mathbf{n})\},$$

где $(\boldsymbol{\omega}, \mathbf{n}) = \boldsymbol{\omega}^T \mathbf{n}$ – скалярное произведение частотного $\boldsymbol{\omega} = (\omega_1, \dots, \omega_N)^T$ и координатного $\mathbf{n} \in S_N$ векторов.

Отклик ЛИС-цепи с импульсной характеристикой $h[\mathbf{n}]$ на эту последовательность

$$\begin{aligned} y[\mathbf{n}] &= \sum_{\mathbf{m}} h[\mathbf{m}]x[\mathbf{n} - \mathbf{m}] = \sum_{\vec{m}} h[\mathbf{m}] \exp\{j\boldsymbol{\omega}^T (\mathbf{n} - \mathbf{m})\} = \\ &= \exp(j\boldsymbol{\omega}^T \mathbf{n}) \sum_{\mathbf{m}} h[\mathbf{m}] \exp(-j\boldsymbol{\omega}^T \mathbf{m}) = H(e^{j\boldsymbol{\omega}}) \exp(j\boldsymbol{\omega}^T \mathbf{n}), \end{aligned}$$

где

$$H(e^{j\boldsymbol{\omega}}) = \sum_{\mathbf{m}} h[\mathbf{m}] \exp(-j\boldsymbol{\omega}^T \mathbf{m}) \quad (5.2)$$

– комплексная частотная характеристика многомерной ЛИС-цепи. Выражение (5.2) представляет собой многомерное преобразование Фурье абсолютно суммируемой последовательности $h[\mathbf{n}]$.

Итак, отклик ЛИС-цепи на многомерную гармоническую последовательность равен этой последовательности, умноженной на комплексный скаляр $H(e^{j\boldsymbol{\omega}})$.

Преобразование Фурье входной последовательности $x[\mathbf{n}]$ представляет собой ее спектральную плотность

$$X(e^{j\boldsymbol{\omega}}) = \sum_{\mathbf{n}} x[\mathbf{n}] \exp(-j\boldsymbol{\omega}^T \mathbf{n}). \quad (5.3)$$

Нетрудно убедиться, что функции $H(e^{j\boldsymbol{\omega}})$ и $X(e^{j\boldsymbol{\omega}})$ периодичны по каждой из компонент частотного вектора $\boldsymbol{\omega}$ с периодом 2π (для этого достаточно учесть равенство $\exp(-j\boldsymbol{\omega}^T \mathbf{n}) = \prod_{i=1}^N \exp(-j\omega_i n_i)$).

Разложение периодической функции $X(e^{j\boldsymbol{\omega}})$ в N -мерный ряд Фурье

$$x[\mathbf{n}] = \frac{1}{(2\pi)^N} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} X(e^{j\boldsymbol{\omega}}) \exp(j\boldsymbol{\omega}^T \mathbf{n}) d\boldsymbol{\omega} \quad (5.4)$$

представляет собой обратное преобразование Фурье.

С другой стороны, выражение (5.4) представляет последовательность $x[\mathbf{n}]$ в виде интегральной («сплошной») суммы многомерных гармоник $\exp(j\boldsymbol{\omega}^T \mathbf{n})$ с комплексными амплитудами

$X(e^{j\omega})$. С учетом принципа суперпозиции отклик ЛИС-цепи на такую «смесь» равен

$$\begin{aligned} y[\mathbf{n}] &= \frac{1}{(2\pi)^N} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} H(e^{j\omega}) X(e^{j\omega}) \exp(j\omega^T \mathbf{n}) d\omega = \\ &= \frac{1}{(2\pi)^N} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} Y(e^{j\omega}) \exp(j\omega^T \mathbf{n}) d\omega, \end{aligned}$$

откуда следует

$$Y(e^{j\omega}) = H(e^{j\omega}) X(e^{j\omega}). \quad (5.5)$$

Таким образом, получено *частотное описание* многомерных последовательностей и ЛИС-цепей.

Как многомерная последовательность, импульсная характеристика ЛИС-цепи может быть разделимой. Легко видеть, что в этом случае разделимой оказывается и комплексная частотная характеристика цепи. В таком случае сама цепь называется *разделимой*.

Пример 5.1. Идеальный двумерный фильтр нижних частот имеет комплексную частотную характеристику вида

$$H(e^{j\omega_1}, e^{j\omega_2}) = \begin{cases} 1, & \left\{ \begin{array}{l} |\omega_1| \leq a, \quad |\omega_2| \leq b, \\ 0 \end{array} \right. \text{ в противном случае,} \end{cases}$$

где a, b – параметры, в общем случае различные, определяющие избирательные свойства фильтра.

Импульсная характеристика такого фильтра

$$h[n_1, n_2] = \frac{1}{4\pi^2} \int_{-b}^b \int_{-a}^a e^{j(\omega_1 n_1 + \omega_2 n_2)} d\omega_1 d\omega_2 = \frac{ab}{\pi^2} \frac{\sin an_1}{an_1} \frac{\sin bn_2}{bn_2}$$

является разделимой. ◀

Пример 5.2. Идеальный *изотропный* двумерный фильтр нижних частот имеет комплексную частотную характеристику вида

$$H(e^{j\omega_1}, e^{j\omega_2}) = \begin{cases} 1, & \omega_1^2 + \omega_2^2 \leq R^2, \\ 0 & \text{ в противном случае,} \end{cases}$$

где параметр R определяет двумерную полосу пропускания фильтра, имеющую форму круга A на частотной плоскости.

Импульсная характеристика такого фильтра определяется выражением

$$h[n_1, n_2] = \frac{1}{4\pi^2} \iint_A e^{j(\omega_1 n_1 + \omega_2 n_2)} d\omega_1 d\omega_2.$$

Переходя в частотной плоскости к полярным координатам $\Omega = \sqrt{\omega_1^2 + \omega_2^2}$, $\varphi = \arctg \frac{\omega_2}{\omega_1}$ и вводя вспомогательную переменную

$\theta = \arctg \frac{n_2}{n_1}$, получаем импульсную характеристику

$$\begin{aligned} h[n_1, n_2] &= \frac{1}{4\pi^2} \int_0^R \int_0^{2\pi} \Omega e^{j\Omega \sqrt{n_1^2 + n_2^2} \cos(\theta - \varphi)} d\varphi d\Omega = \\ &= \frac{1}{2\pi} \int_0^R \Omega J_0 \left(\Omega \sqrt{n_1^2 + n_2^2} \right) d\Omega = \frac{R}{2\pi} \frac{J_1 \left(R \sqrt{n_1^2 + n_2^2} \right)}{\sqrt{n_1^2 + n_2^2}}, \end{aligned}$$

где $J_0(\cdot), J_1(\cdot)$ – функции Бесселя 1-го рода нулевого и первого порядков; она разделимой, очевидно, не является. ◀

Изотропность многомерных цепей представляет собой полезное свойство, гарантирующее «беспристрастность» цепи (или, что то же самое, алгоритма обработки) к повороту последовательности, который зачастую является случайным и несущественным фактором. Примерами таких последовательностей являются аэрофото- снимки земной поверхности, фотографии бактерий под микроскопом и т.п. В таких случаях желательно, чтобы алгоритм обработки (например, двумерный фильтр) был безразличен к случайному повороту снимка, т. е. обладал изотропностью.

Разделимость цепей также является полезным свойством, так как разделимая многомерная цепь требует меньших вычислительных затрат, чем неразделимая. В самом деле, если импульсная характеристика цепи разделима, т. е. $h[n_1, n_2] = h_1[n_1]h_2[n_2]$, то двумерная свертка, определяющая отклик такой цепи на произвольную последовательность $x[n_1, n_2]$, равна

$$y[n_1, n_2] = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} h_1[k_1]h_2[k_2]x[n_1 - k_1, n_2 - k_2] =$$

$$\begin{aligned}
&= \sum_{k_1=-\infty}^{\infty} h_1[k_1] \left(\sum_{k_2=-\infty}^{\infty} h_2[k_2] x[n_1 - k_1, n_2 - k_2] \right) = \\
&= \sum_{k_1=-\infty}^{\infty} h_1[k_1] v[n_1 - k_1, n_2],
\end{aligned}$$

где $v[n_1, n_2]$ – промежуточный результат фильтрации последовательности $x[n_1, n_2]$ по координате n_2 . Таким образом, двумерная свертка $x[n_1, n_2]$ с разделимой импульсной характеристикой сводится к выполнению двух одномерных сверток.

Оценим вычислительные затраты на реализацию двумерной свертки в случае разделимой импульсной характеристики. Пусть импульсная характеристика имеет носитель размерами $M \times M$ и используется для фильтрации двумерной последовательности размерами $N \times N$. Вычисление одного значения выходной последовательности требует выполнения $M \times M$ умножений (сложениями как значительно менее трудоемкими и времяемкими операциями в таких случаях пренебрегают). Следовательно, в общем случае вычисление выходной последовательности требует около $M^2 \times N^2$ умножений (для простоты пренебрегаем некоторой экономией вычислений на краях носителя последовательности). В случае разделимой импульсной характеристики выполнение одномерной свертки для вычисления $v[n_1, n_2]$ требует выполнения $M \times N^2$ умножений; столько же требуется для выполнения второй одномерной свертки. Таким образом, всего должно быть выполнено $2M \times N^2$ умножений.

Поскольку изотропность и разделимость цепей являются полезными свойствами, интересен вопрос, возможно ли совместить эти свойства в одной ЛИС-цепи, и если да, то насколько широк класс таких цепей?

Пусть комплексная частотная характеристика двумерной цепи разделима и имеет вид

$$H(e^{j\omega_1}, e^{j\omega_2}) = H_1(e^{j\omega_1})H_2(e^{j\omega_2}). \quad (5.6)$$

Изотропность, в свою очередь, означает, что множество решений уравнения $H(e^{j\omega_1}, e^{j\omega_2}) = C_1$ представляет собой окружность при любом C_1 . Следовательно, такие решения удовлетворяют

условию $\omega_1^2 + \omega_2^2 = R^2$, где R – радиус окружности. Логарифмируя (5.6), получаем

$$\ln \{H_1(e^{j\omega_1})\} + \ln \{H_2(e^{j\omega_2})\} = \ln C_1.$$

Сравнение полученного выражения с уравнением окружности приводит к выводу, что одновременному выполнению свойств изотропности и разделимости удовлетворяют только цепи с КЧХ гауссовского вида $H(e^{j\omega_1}, e^{j\omega_2}) = Ke^{-\omega_1^2/\gamma} e^{-\omega_2^2/\gamma}$, где K, γ – произвольные константы. Однако следует отметить, что в дискретном виде такие цепи *точно* реализовать нельзя.

При условии *абсолютной суммируемости* последовательности $x[\mathbf{n}]$ ряд в формуле (5.3) преобразования Фурье сходится *равномерно* к функции $X(e^{j\omega})$, *непрерывной* по всем компонентам векторной частоты $\omega = (\omega_1, \dots, \omega_N)^T$. Формальная замена $e^{j\omega_i} \rightarrow z_i$ в преобразовании Фурье приводит к многомерному z -преобразованию

$$X(\mathbf{z}) = \sum_{\mathbf{n}} \left(x[\mathbf{n}] \prod_{i=1}^N z_i^{-n_i} \right),$$

сходящемуся на множестве точек N -мерного комплексного пространства (N -мерного аналога комплексной плоскости), удовлетворяющих условию $|z_1| = 1, \dots, |z_N| = 1$.

Это множество аналогично единичной окружности, но уже при $N = 2$ оно представляет собой произведение двух 1-окружностей в 4-мерном пространстве и поэтому немного дает в смысле наглядности. Это одна из причин, ограничивающих применение многомерного z -преобразования. Другая, более серьезная причина состоит в следующем.

Одномерное z -преобразование является мощным инструментом анализа устойчивости рекурсивных цепей. Напомним, что для анализа устойчивости цепи нужно разложить знаменатель передаточной функции на простые множители (или, что равнозначно, найти корни полинома-знаменателя). Расположение полюсов передаточной функции относительно 1-окружности позволяет однозначно решить вопрос об устойчивости цепи (см. разд. 2).

К сожалению, теорема о разложении полинома на простые множители («основная теорема алгебры») *не справедлива* для $N > 1$. Полюсы передаточной функции многомерной рекурсивной

цепи не являются изолированными точками, а образуют множества (гиперповерхности) в пространстве соответствующей размерности. Например, передаточная функция

$$H(z_1, z_2) = \frac{1}{1 - bz_1^{-1}z_2^{-1}}$$

имеет «полюс» в виде гиперповерхности, описываемой уравнением $z_1 z_2 = b$.

По указанным причинам анализ устойчивости многомерных рекурсивных цепей представляет гораздо более сложную задачу, чем такой анализ для одномерных цепей. Некоторые способы решения этой задачи описаны в [16].

5.3. МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ И ЛИС-ЦЕПИ

Многомерная случайная последовательность представляется ансамблем реализаций

$$\{x_\lambda[\mathbf{n}], \mathbf{n} \in S_N, \lambda \in \Lambda\}$$

в совокупности с распределением вероятностей P_λ , заданным на множестве Λ .

При фиксированном λ получается сечение – детерминированная многомерная последовательность, а при фиксированном \mathbf{n} – случайная величина, описываемая функцией распределения

$$F(x, \mathbf{n}) = P\{x_\lambda[\mathbf{n}] \leq x\}.$$

Если λ – непрерывная переменная, то можно определить плотность распределения вероятностей

$$w(x, \mathbf{n}) = \frac{dF(x, \mathbf{n})}{dx}.$$

Как и для одномерных случайных последовательностей, очевидным образом определяются математическое ожидание

$$\bar{x}[\mathbf{n}] = \mathbf{E}\{x[\mathbf{n}]\} = \int_{-\infty}^{\infty} x[\mathbf{n}]w(x, \mathbf{n})dx,$$

дисперсия

$$D_x[\mathbf{n}] = \mathbf{E}\{|x[\mathbf{n}] - \bar{x}[\mathbf{n}]|^2\} = \int_{-\infty}^{\infty} |x[\mathbf{n}] - \bar{x}[\mathbf{n}]|^2 w(x, \mathbf{n}) dx$$

и корреляционный момент

$$\begin{aligned} r_x[\mathbf{n}_1, \mathbf{n}_2] &= \mathbf{E}\{(x[\mathbf{n}_1] - \bar{x}[\mathbf{n}_1])(x^*[\mathbf{n}_2] - \bar{x}^*[\mathbf{n}_2])\} = \\ &= \int_{-\infty}^{\infty} (x[\mathbf{n}_1] - \bar{x}[\mathbf{n}_1])(x^*[\mathbf{n}_2] - \bar{x}^*[\mathbf{n}_2]) w_2(x_1, x_2, \mathbf{n}_1, \mathbf{n}_2) dx_1 dx_2. \end{aligned}$$

Многомерная случайная последовательность *стационарна* (в широком смысле), если $\bar{x}[\mathbf{n}] = \bar{x}$; $D_x[\mathbf{n}] = D_x = \sigma_x^2$, а корреляционный момент $r_x[\mathbf{n}_1, \mathbf{n}_2]$ зависит лишь от разности $\mathbf{n}_1 - \mathbf{n}_2$; тогда он называется автокорреляционной последовательностью:

$$r_x[\mathbf{m}] = \mathbf{E}\{(x[\mathbf{n}] - \bar{x})(x^*[\mathbf{n} + \mathbf{m}] - \bar{x})\},$$

при этом $r_x[\mathbf{0}] = D_x$. Стационарность случайной последовательности в широком смысле означает *однородность* ее статистических характеристик первого и второго порядков *при сдвиге* вдоль любого направления.

В отличие от одномерного случая для многомерных случайных последовательностей выделяют еще одно, более сильное свойство стационарности. Случайная последовательность называется *изотропной*, если

$$\bar{x}[\mathbf{n}] = \bar{x}; D_x[\mathbf{n}] = D_x = \sigma_x^2,$$

а ее автокорреляционная последовательность зависит только от *модуля* разности векторов $|\mathbf{n}_1 - \mathbf{n}_2|$.

Если многомерная случайная последовательность изотропна, то она и стационарна; обратное в общем случае неверно.

Условие изотропности случайной последовательности эквивалентно требованию круговой (сферической) симметрии АКП, что, очевидно, может выполняться лишь приближенно (вследствие не-изотропности решетки S_N), однако модель изотропной случайной последовательности обладает «беспристрастностью» к повороту, естественной для многих задач ЦОС и поэтому достаточно широко

используется. В частности, такая беспристрастность является желательной в моделях случайных последовательностей, используемых в задачах фильтрации аэрокосмических фотоснимков земной поверхности, так как ориентация такого снимка определяется случайным положением летательного аппарата и несущественна с точки зрения содержания снимка. Отметим, что в таких случаях желательна также изотропность алгоритмов обработки.

Аналогично автокорреляционной последовательности определяется взаимно корреляционная последовательность.

Анализ прохождения многомерных случайных последовательностей через ЛИС-цепи аналогичен одномерному случаю. Спектральная плотность мощности случайной последовательности связана с АКП парой преобразований Фурье (теорема Винера – Хинчина)

$$R_x(e^{j\omega}) = \sum_{\mathbf{m}} r_x[\mathbf{m}] e^{-j\omega^T \mathbf{m}},$$

$$r_x[\mathbf{m}] = \frac{1}{(2\pi)^N} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} R_x(e^{j\omega}) e^{j\omega^T \mathbf{m}} d\omega.$$

Спектральная плотность мощности отклика ЛИС-цепи с КЧХ $H(e^{j\omega})$ на многомерную случайную последовательность со спектральной плотностью мощности $R_x(e^{j\omega})$ находится как

$$R_y(e^{j\omega}) = H(e^{j\omega}) H^* \left(\frac{1}{e^{-j\omega}} \right) R_x(e^{j\omega}) = |H(e^{j\omega})|^2 R_x(e^{j\omega}).$$

Остальные соотношения, связывающие авто- и взаимно корреляционные последовательности входа и выхода и импульсную характеристику цепи, могут быть записаны по аналогии с одномерным случаем.

Упражнения

1. Запишите три варианта матрицы периодичности для последовательности, представленной на рис. 5.2 (кроме приведенных в тексте).

2. Запишите выражение для скалярного произведения в пространстве последовательностей размерности 3. Покажите, как на основе этого выражения определить норму и метрику.

3. Проверьте выполнение свойств ассоциативности, дистрибутивности и коммутативности для многомерной свертки.

4. Докажите, что делимость импульсной характеристики влечет делимость КЧХ.

5. Убедитесь, что функции $H(e^{j\omega})$ и $X(e^{j\omega})$ периодичны по каждой из компонент частотного вектора ω с периодом 2π .

6. Запишите соотношения, связывающие авто- и взаимно корреляционные последовательности входа и выхода и импульсную характеристику многомерной ЛИС-цепи.

Вопросы для самоконтроля

1. Чему равен определитель матрицы периодичности для последовательности, периодичной вдоль оси n_1 и постоянной вдоль оси n_2 ?

2. Что можно сказать о последовательности, матрица периодичности которой диагональна?

3. Какие последовательности называют делимыми?

4. В чем состоят преимущества делимых цепей?

5. Связаны ли свойства делимости и изотропности цепей?

6. Почему не существуют изотропные делимые дискретные цепи?

7. Чем отличается стационарная последовательность от изотропной?

8. Какое из свойств – стационарность или изотропность – является более сильным?

9. Почему изотропные случайные последовательности, строго говоря, не существуют?

10. С чем связана трудность анализа устойчивости многомерных цепей?



6. СИНТЕЗ ЦИФРОВЫХ ФИЛЬТРОВ

Одной из наиболее простых и в то же время широко применяемых на практике операций цифровой обработки сигналов является цифровая фильтрация. Фильтрация аналоговых сигналов цифровыми устройствами долгое время отождествлялась с цифровой обработкой сигналов.

Фильтрация означает подавление нежелательных составляющих некоторого колебания (шумов, помех) при минимальных искажениях полезного сигнала. Ниже под фильтрами подразумеваются линейные стационарные цепи, которые разрабатываются специально для избирательного воздействия на различные гармонические составляющие входного сигнала. Хотя методы цифровой фильтрации применимы к дискретным сигналам вне зависимости от их происхождения, в радиотехнике и связи цифровой фильтрации подвергаются аналоговые сигналы, представленные в дискретной форме.

Прежде чем рассматривать задачи построения цифровых фильтров, необходимо выяснить, как происходит преобразование аналоговых сигналов в дискретные.

6.1. ВЗАИМОСВЯЗЬ АНАЛОГОВЫХ, ИМПУЛЬСНЫХ И ДИСКРЕТНЫХ СИГНАЛОВ (СПЕКТРАЛЬНЫЕ И ЭНЕРГЕТИЧЕСКИЕ СООТНОШЕНИЯ)

Основанием для применения цифровой обработки сигналов в радиотехнике служит, как известно, теорема отсчетов (теорема Котельникова), связывающая аналоговый (континуальный) сигнал с совокупностью его выборочных значений (отсчетов), взятых с достаточно малым шагом. В этом разделе рассматривается взаимосвязь спектральных и энергетических характеристик аналогового исходного сигнала, импульсного сигнала, получаемого в процессе

его дискретизации, а также собственно дискретного сигнала (последовательности).

Будем обозначать аналоговый сигнал $x(t)$, его спектральную плотность $X(\Omega)$, прямое и обратное преобразования Фурье для аналоговых сигналов $F(\cdot)$ и $F^{-1}(\cdot)$, частоту в спектральном описании аналоговых сигналов Ω . Для последовательности отсчетов данного сигнала примем обозначение $x[n]$, спектральную плотность последовательности обозначим $X(e^{j\omega})$. Частота, соответствующая спектральному представлению дискретного сигнала (последовательности), обозначается ω , а пара преобразований Фурье для последовательностей – $F(\cdot)$ и $F^{-1}(\cdot)$.

Для простоты анализа удобно считать, что дискретизация аналогового сигнала $x(t)$ осуществляется путем его стробирования периодической последовательностью коротких импульсов $s(t)$, или, что эквивалентно, при помощи амплитудной модуляции колебания $s(t)$ аналоговым сигналом $x(t)$ (амплитудно-импульсной модуляции, АИМ). Получаемый при этом импульсный сигнал равен $x_{\text{и}}(t) = x(t)s(t)$.

Предположим вначале, что

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

где $\delta(\cdot)$ – дельта-функция Дирака, а T_d – шаг дискретизации. В этом случае

$$x_{\text{и}}(t) = x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_d) = \sum_{n=-\infty}^{\infty} x(nT_d) \delta(t - nT_d). \quad (6.1)$$

Спектральная плотность импульсного сигнала

$$\begin{aligned} X_{\text{и}}(\Omega) &= F\{x_{\text{и}}(t)\} = \sum_{n=-\infty}^{\infty} x(nT_d) F\{\delta(t - nT_d)\} = \\ &= \sum_{n=-\infty}^{\infty} x(nT_d) \int_{-\infty}^{\infty} \delta(t - nT_d) e^{-j\Omega t} dt = \sum_{n=-\infty}^{\infty} x(nT_d) e^{-j\Omega T_d n}. \end{aligned}$$

Итак, $X_{\text{и}}(\Omega) = \sum_{n=-\infty}^{\infty} x(nT_d) e^{-j\Omega T_d n}$. С другой стороны, спектральная плотность последовательности $x[n]$ равна $X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}$. Очевидно, если

$$x[n] = x(nT_d) \text{ и } \omega = \Omega T_d, \quad (6.2)$$

то спектральная плотность $X_{\text{и}}(\Omega)$ импульсного сигнала $x_{\text{и}}(t)$ и спектральная плотность $X(e^{j\omega})$ последовательности $x[n]$ чисел, равных весам («амплитудам») дельта-функций в составе этого импульсного сигнала, совпадают. Отметим, что спектральные плотности $X_{\text{и}}(\Omega)$ и $X(e^{j\omega})$ суть функции *различных аргументов* и, несмотря на одинаковые названия, представляют собой разложения разных функций $x_{\text{и}}(t)$ и $x[n]$ в *различных базисах*, поэтому указанное совпадение является формальным. При этом, очевидно, поскольку $-\pi \leq \omega \leq \pi$, то и $-\pi \leq \Omega T_d \leq \pi$, т. е. указанное совпадение спектральных плотностей имеет место на частотных интервалах, определяемых условиями $-\pi \leq \omega \leq \pi$ и $-\pi/T_d \leq \Omega \leq \pi/T_d$.

Теперь выясним, как связаны спектральные плотности АИМ-сигнала $x_{\text{и}}(t)$ и исходного аналогового сигнала $x(t)$. Положим, что $s(t) = \sum_{n=-\infty}^{\infty} d(t - nT_d)$, где $d(t)$ – короткий импульс известной формы (необязательно дельта-функция!).

Периодическое колебание $s(t)$ можно записать в виде ряда Фурье

$$s(t) = \sum_{k=-\infty}^{\infty} C_k e^{jk\Omega_d t}, \quad (6.3)$$

где $\Omega_d = \frac{2\pi}{T_d}$, C_k – коэффициенты ряда, определяемые формой импульса $d(t)$. Спектральная плотность $S(\Omega) = F\{s(t)\}$, очевидно, равна сумме спектральных плотностей гармонических составляю-

щих (комплексных экспонент) с теми же весовыми коэффициентами, т. е.

$$S(\Omega) = F \left\{ \sum_{k=-\infty}^{\infty} C_k e^{jk\Omega_d t} \right\} = 2\pi \sum_{k=-\infty}^{\infty} C_k \delta(\Omega - k\Omega_d).$$

Поскольку АИМ-сигнал получается умножением $x_{\text{и}}(t) = x(t)s(t)$, его спектральная плотность равна свертке спектральных плотностей сомножителей

$$\begin{aligned} X_{\text{и}}(\Omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\sigma) S(\Omega - \sigma) d\sigma = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\sigma) 2\pi \sum_{k=-\infty}^{\infty} C_k \delta(\Omega - \sigma - k\Omega_d) d\sigma = \sum_{k=-\infty}^{\infty} C_k X(\Omega - k\Omega_d) \end{aligned}$$

и представляет собой совокупность *копий* спектральной плотности исходного аналогового сигнала $X(\cdot)$, сдвинутых по частотной оси на величины $k\Omega_d = k \frac{2\pi}{T_d}$, $k = \overline{-\infty, \infty}$, и умноженных на весовые коэффициенты, определяемые *формой* импульса $d(t)$. (Читателю предлагается проверить, что при $d(t) = \delta(t)$ все коэффициенты равны 1.)

Очевидно, если каждая копия $X(\cdot)$ занимает на частотной оси интервал, ширина которого меньше $\Omega_d = \frac{2\pi}{T_d}$, то копии *не перекрываются* и можно выделить путем фильтрации нижних частот единственную копию $X(\Omega - 0 \cdot \Omega_d) = X(\Omega)$, тем самым восстановив исходный аналоговый сигнал из АИМ-сигнала. Таким образом, получено требование к частоте дискретизации сигнала с *финитной* (ограниченной по ширине) спектральной плотностью: частота дискретизации Ω_d должна быть больше, чем удвоенная верхняя частота спектра сигнала (частота Найквиста). При этом в интервале $-\pi/T_d \leq \Omega \leq \pi/T_d$ спектральная плотность

$$X_{\text{и}}(\Omega) = C_0 X(\Omega). \quad (6.4)$$

Заметим, что для точного восстановления аналогового сигнала из АИМ-сигнала должен быть скомпенсирован весовой коэффициент C_0 . Рассмотрим этот вопрос подробнее.

Предположим, что $d(t)$ – прямоугольный импульс с амплитудой U и длительностью τ , симметрично расположенный относительно момента $t = 0$. Тогда

$$C_k = \frac{U}{T_d} \int_{-\tau/2}^{\tau/2} e^{-jk\Omega_d t} dt = \\ = \frac{U}{T_d} \int_{-\tau/2}^{\tau/2} \cos(k\Omega_d t) \frac{d(k\Omega_d t)}{k\Omega_d} = \frac{U\tau}{T_d} \frac{\sin(k\Omega_d \tau/2)}{k\Omega_d \tau/2}.$$

Коэффициент $C_0 = \frac{U\tau}{T_d}$, поэтому восстанавливающий (интерполирующий) ФНЧ должен иметь прямоугольную АЧХ вида

$$K(\Omega) = \begin{cases} \frac{T_d}{U\tau}, & -\frac{\pi}{T_d} < \Omega < \frac{\pi}{T_d}, \\ 0 & \text{в противном случае} \end{cases}$$

и импульсную характеристику

$$h(t) = F^{-1}\{K(\Omega)\} = \frac{\sin\left(\frac{\pi}{T_d} t\right)}{\frac{\pi}{T_d} t}.$$

Заметим, что полученная функция совпадает с нулевой базисной функцией ряда Котельникова (все остальные базисные функции образуются ее сдвигами на величины, кратные шагу дискретизации):

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT_d) \frac{\sin\left[\frac{\pi}{T_d}(t - nT_d)\right]}{\frac{\pi}{T_d}(t - nT_d)}.$$

Это означает, что для *точного* восстановления (интерполяции) аналогового сигнала требуется в моменты nT_d , $n = -\infty, \infty$ возбуждать восстанавливающий фильтр дельта-функциями, умноженными на значения отсчетов сигнала $x(nT_d)$.

Установим связь между спектральными характеристиками аналогового, импульсного и дискретного сигналов. Если выполнены равенства (6.2), то при $-\frac{\pi}{T_d} \leq \Omega \leq \frac{\pi}{T_d}$

$$X_{\text{и}}(\Omega) = X(e^{j\omega}) \Big|_{\omega=\Omega T_d}, \quad (6.4)$$

с другой стороны, согласно (6.3), в том же диапазоне частот

$$X_{\text{и}}(\Omega) = \frac{U\tau}{T_d} X(\Omega). \quad (6.5)$$

Объединяя (6.4) и (6.5), запишем

$$X(e^{j\omega}) = X_{\text{и}}(\Omega) = \frac{U\tau}{T_d} X(\Omega), \quad -\pi \leq \omega = \Omega T_d \leq \pi.$$

Полагая площадь импульса $U\tau$ равной 1, получим равенство

$$X(e^{j\omega}) = X_{\text{и}}(\Omega) = \frac{1}{T_d} X(\Omega), \quad -\pi \leq \omega = \Omega T_d \leq \pi,$$

справедливое и для идеальной дискретизации с помощью дельта-функций (6.1)

Замечание 1. С точки зрения возможности восстановления исходного сигнала из АИМ-сигнала форма импульса $d(t)$ не играет роли, так как влияет только на весовые коэффициенты ряда (6.3); из них только коэффициент C_0 должен быть учтен при построении восстанавливающего фильтра.

Замечание 2. Переход от АИМ-сигнала к дискретному сигналу (последовательности) означает замену функции непрерывного времени (хотя и отличной от нуля лишь на малых временных интервалах) совокупностью чисел, т. е. требует измерения амплитуд импульсов. Процесс измерения значения функции в определенный момент времени описывается при помощи стробирующего (фильтрующего) свойства дельта-функции:

$$x[n] = x(nT_d) = \int_{-\infty}^{\infty} x(t)\delta(t - nT_d)dt.$$

На практике этот процесс реализуется в устройстве выборки-хранения (УВХ) путем замены δ -функции коротким импульсом, что эквивалентно операции свертки

$$x[n] = x(nT_d) = \int_{-\infty}^{\infty} x(t)v(t - nT_d)dt, \quad (6.6)$$

которой в частотной области соответствует *умножение* спектральной плотности $X(\Omega)$ исходного сигнала на спектральную плотность импульса $v(t)$; ясно, что *искажение* спектра сигнала будет тем сильнее, чем сильнее отличается импульс $v(t)$ от δ -функции.

Рассмотрим энергетические соотношения между аналоговым и дискретным сигналами. Энергия аналогового сигнала

$$E_a = \int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\Omega)|^2 d\Omega = \frac{1}{2\pi} \int_{-\Omega_d/2}^{\Omega_d/2} |X(\Omega)|^2 d\Omega. \quad (6.7)$$

Здесь предполагается, что сигнальная функция квадратично интегрируема, а спектральная плотность финитна.

Энергия последовательности

$$\begin{aligned} E_n &= \sum_{n=-\infty}^{\infty} |x[n]|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(e^{j\omega})|^2 d\omega = \\ &= \frac{T_d}{2\pi} \int_{-\pi/T_d}^{\pi/T_d} |X_n(\Omega)|^2 d\Omega = C_0^2 \frac{T_d}{2\pi} \int_{-\pi/T_d}^{\pi/T_d} |X(\Omega)|^2 d\Omega. \end{aligned}$$

Поскольку $\Omega_d/2 = \pi/T_d$, с учетом (6.7) получим $E_n = C_0^2 T_d E_a$; в частности, если дискретизирующая последовательность образована δ -функциями, энергии связаны простым соотношением

$$E_n = \frac{1}{T_d} E_a.$$

6.2. ЗАДАЧА СИНТЕЗА ЦИФРОВОГО ФИЛЬТРА КАК ЗАДАЧА АППРОКСИМАЦИИ

Обычно требования к фильтру выражаются в форме желаемой амплитудно-частотной характеристики и реже – фазочастотной характеристики, т. е. в общем случае синтез фильтра состоит в получении комплексной частотной характеристики требуемого вида. Однако известно, что передаточная функция дискретной цепи с конечной импульсной характеристикой (КИХ-цепи) конечного порядка представляет собой полином, а передаточная функция БИХ-цепи конечного порядка – дробно-рациональную функцию. Таким образом, не всякая функция может быть комплексной частотной характеристикой *реального* дискретного фильтра. Поэтому синтез дискретного фильтра состоит в нахождении такой реализуемой структуры дискретной цепи, характеристики которой были бы в каком-то смысле наиболее близки к желаемым. Следовательно, с математической точки зрения синтез фильтра – это аппроксимация желаемой характеристики при помощи функции из заданного класса (т. е. при помощи полинома или частного двух полиномов).

Дискретная цепь может быть реализована в виде цифровой цепи (цифрового фильтра) различными способами. Основой *аппаратной* реализации является структурная схема цепи; при этом операции умножения отсчетов сигналов на коэффициенты и сложения отсчетов реализуются специализированными цифровыми устройствами конечной разрядности. *Программная* реализация цифрового фильтра основывается на рекуррентном решении разностного уравнения, при этом операции также производятся с ограниченной точностью, определяемой разрядностью арифметико-логического устройства процессора. Таким образом, при любом способе реализации переход от дискретной цепи к цифровой неизбежно сопровождается погрешностями, которые необходимо учитывать и стремиться свести к минимуму.

Проектирование цифрового фильтра обычно подразделяют на три этапа: 1) определение требований к фильтру; 2) аппроксимацию желаемых характеристик при помощи характеристик реализуемой дискретной КИХ- или БИХ-цепи; 3) реализацию полученной структуры с учетом эффектов квантования входных данных и округления параметров (коэффициентов) фильтра.

Рассмотрим эти этапы более подробно на примере фильтра нижних частот с прямоугольной АЧХ.

Отметим, что желаемая П-образная АЧХ соответствует нереализуемому фильтру (см. п. 2.9), поэтому требования к фильтру

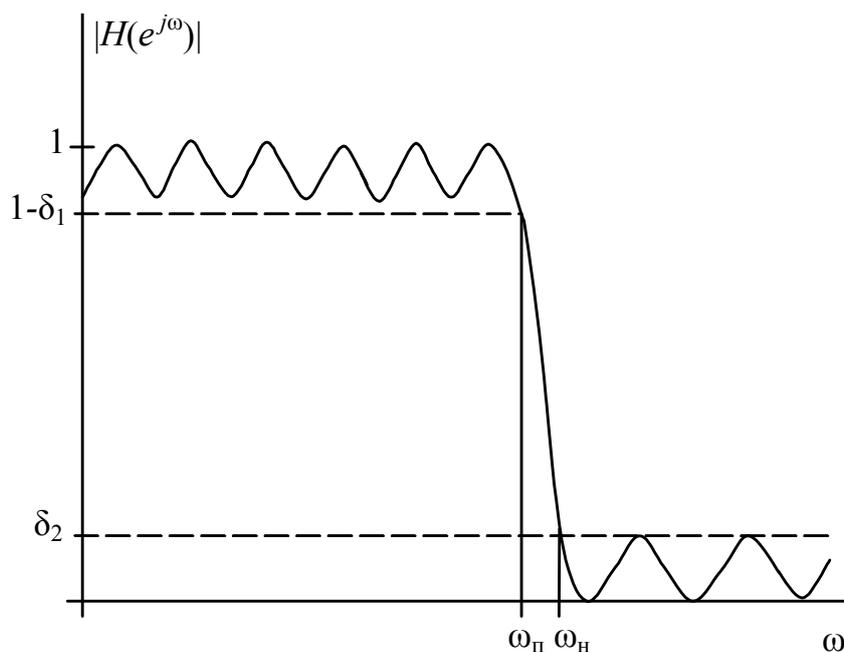


Рис. 6.1. Допуски на амплитудно-частотную характеристику фильтра нижних частот

формулируются обычно в виде полей допуска на отклонение реальной АЧХ от желаемой (рис. 6.1). В полосе *пропускания* (называемой также полосой прозрачности), ограниченной частотами 0 и ω_{Π} , допускаются изменения АЧХ $A(e^{j\omega}) = |H(e^{j\omega})|$ в пределах от $1 - \delta_1$ до 1 (иногда пределы допустимых изменений определяют от $1 - \delta_1$ до $1 + \delta_1$). В полосе *непропускания* (полосе непрозрачности, полосе задерживания), т.е. для частот от $\omega_{\text{Н}}$ до π АЧХ должна быть заключена в пределах от 0 до δ_2 . Параметры δ_1 и δ_2 характеризуют *неравномерность* АЧХ. Переходная полоса от ω_{Π} до $\omega_{\text{Н}}$ необходима, если $1 - \delta_1 > \delta_2$ (а это практически всегда так), потому что иначе АЧХ содержит скачок, а импульсная характеристика не является абсолютно суммируемой, что соответствует неустойчивому фильтру. Таким образом, требования к фильтру нижних частот сводятся к заданию четырех параметров — δ_1 , δ_2 , ω_{Π} и $\omega_{\text{Н}}$.

Второй этап, по существу, составляет *синтез* дискретного фильтра; именно этому вопросу посвящена большая часть раздела. Следует отметить, что методы аппроксимации желаемых характеристик полиномами и дробно-рациональными функциями принципиально различны, поэтому синтез КИХ- и БИХ-фильтров выполняется существенно различными способами. Рассмотренный

ниже *метод взвешивания* состоит в аппроксимации желаемой КЧХ конечной суммой ряда Фурье, «взвешенного» специально подобранной последовательностью (окном). Метод *частотной выборки* базируется на интерполяции КЧХ по заданным ее значениям при помощи интерполяционной формулы Лагранжа. Метод *быстрой свертки* (КИХ-фильтрации в частотной области) предполагает использование алгоритмов быстрого преобразования Фурье. Синтез БИХ-фильтров основан на преобразовании аналоговых фильтров-прототипов в цифровые фильтры (методы *аналого-цифровой трансформации*). Выбор того или иного метода синтеза фильтра должен определяться условиями конкретной практической задачи.

6.3. СИНТЕЗ КИХ-ФИЛЬТРОВ

Фильтры с конечной импульсной характеристикой имеют перед БИХ-фильтрами ряд преимуществ. Во-первых, КИХ-фильтры всегда устойчивы. Во-вторых, только КИХ-фильтр может иметь строго линейную фазочастотную характеристику (см. п. 3.8). Наконец, для КИХ-фильтров наиболее просто решается задача аппроксимации КЧХ желаемого вида реализуемой функцией (тригонометрическим полиномом). Однако КИХ-фильтры имеют существенный недостаток по сравнению с БИХ-фильтрами. Именно, чтобы обеспечить желаемую крутизну АЧХ в переходной полосе частот, требуется КИХ-фильтр в десятки раз более высокого порядка, чем БИХ-фильтр. На практике в зависимости от конкретных обстоятельств применяются фильтры обоих типов. В этом подразделе рассматриваются наиболее распространенные методы синтеза КИХ-фильтров.

6.3.1. МЕТОД ВЗВЕШИВАНИЯ (МЕТОД ФУНКЦИЙ ОКНА)

Комплексная частотная характеристика нерекурсивного (трансервального) дискретного фильтра представляет собой тригонометрический полином, т.е. функцию вида

$$H(e^{j\omega}) = \sum_{n=-M}^M b_n e^{-jn\omega} . \quad (6.8)$$

Здесь не предполагается каузальность фильтра; если каузальность необходима, ее легко можно обеспечить умножением (6.8) на соответствующий фазовый множитель. Если желаемая КЧХ имеет вид $H_{\text{ж}}(e^{j\omega})$, то синтез КИХ-фильтра состоит в нахождении три-

гонометрического полинома, близкого в каком-то смысле к $H_{\text{ж}}(e^{j\omega})$. Обычно в качестве критерия близости выбирается среднеквадратическая ошибка аппроксимации

$$\varepsilon = \int_{-\pi}^{\pi} |H(e^{j\omega}) - H_{\text{ж}}(e^{j\omega})|^2 d\omega,$$

тогда наилучшая аппроксимация обеспечивается, если коэффициентами полинома (6.8) являются коэффициенты разложения желаемой КЧХ в ряд Фурье

$$b_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_{\text{ж}}(e^{j\omega}) e^{j\omega n} d\omega, \quad n = \overline{-M, M}. \quad (6.9)$$

Эти коэффициенты представляют собой отсчеты импульсной характеристики КИХ-фильтра, в общем случае некаузального. После соответствующей задержки получается импульсная характеристика каузального фильтра $h[n], n = \overline{0, N-1}$, где $N = 2M + 1$. Поскольку всякая ЛИС-цепь однозначно определяется своей импульсной характеристикой, на этом синтез КИХ-фильтра можно было бы считать законченным; однако на практике такой фильтр обычно неприемлем из-за гиббсовских осцилляций КЧХ (см. п. 2.9). Поэтому кроме *усечения* ряда Фурье, т. е. ограничения числа отсчетов импульсной характеристики, применяют дополнительное умножение ее на весовую последовательность («окно») подходящей формы. Поскольку причина явления Гиббса заключается в слишком медленном убывании коэффициентов Фурье-разложения разрывной функции, то все применяемые окна убывают от середины к краям. Наиболее широко известны следующие окна, показанные на рис. 6.2 для $N = 15$ (прямоугольное «окно», соответствующее простому усечению импульсной характеристики, приводится для сравнения).

1. Прямоугольное «окно»,

$$w_{\text{п}}[n] = 1, \quad n = \overline{0, N-1}. \quad (6.10)$$

2. Окно Бартлетта (треугольное)

$$w_{\text{б}}[n] = \begin{cases} \frac{2n}{N-1}, & n = \overline{0, \frac{N-1}{2}}, \\ 2 - \frac{2n}{N-1}, & n = \overline{\frac{N-1}{2}, N-1}. \end{cases}$$

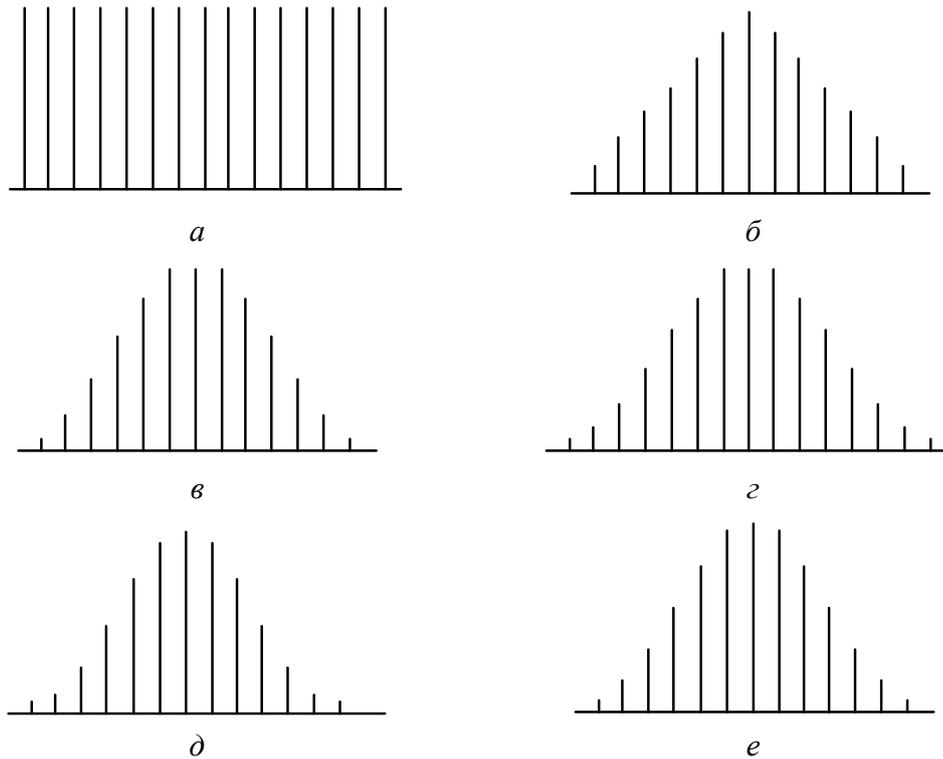


Рис. 6.2. Прямоугольное окно (а) и окна Бартлетта (б), Хэнна (в), Хэмминга (г), Блэкмана (д) и Кайзера (е)

3. Окно Хэнна (фон Ганна)

$$w_x[n] = \frac{1}{2} \left(1 - \cos \frac{2\pi n}{N-1} \right), \quad n = \overline{0, N-1}.$$

4. Окно Хэмминга

$$w_{\text{ХМ}}[n] = 0,54 - 0,46 \cos \frac{2\pi n}{N-1}, \quad n = \overline{0, N-1}.$$

5. Окно Блэкмана

$$w_{\text{БЛ}}[n] = 0,42 - 0,5 \cos \frac{2\pi n}{N-1} + 0,08 \cos \frac{4\pi n}{N-1}, \quad n = \overline{0, N-1}.$$

6. Окно Кайзера

$$w_k[n] = \frac{I_0 \left[\omega_a \sqrt{\left(\frac{N-1}{2} \right)^2 - \left(n - \frac{N-1}{2} \right)^2} \right]}{I_0 \left[\omega_a \frac{N-1}{2} \right]}, \quad n = \overline{0, N-1},$$

где $I_0(\cdot)$ – модифицированная функция Бесселя 0-го порядка, ω_a – параметр, управляющий формой окна (его значение рекомендуется выбирать из условия $4 < \omega_a \frac{N-1}{2} < 9$ [3]). Окно Кайзера, приведенное на рис. 6.2, *e*, соответствует значению $\omega_a = 1$.

Получение простой усеченной импульсной характеристики эквивалентно *умножению* бесконечной импульсной характеристики на окно прямоугольной формы (6.10), поэтому реальная КЧХ равна *свертке* желаемой КЧХ с частотной функцией $W_{\Pi}(e^{j\omega})$, представляющей собой преобразование Фурье окна $w_{\Pi}[n]$. На рис. 6.3, *a* показана частотная функция прямоугольного окна при $N = 15$, а на рис. 6.3, *б* приведен результат ее свертки с П-образной желаемой КЧХ.

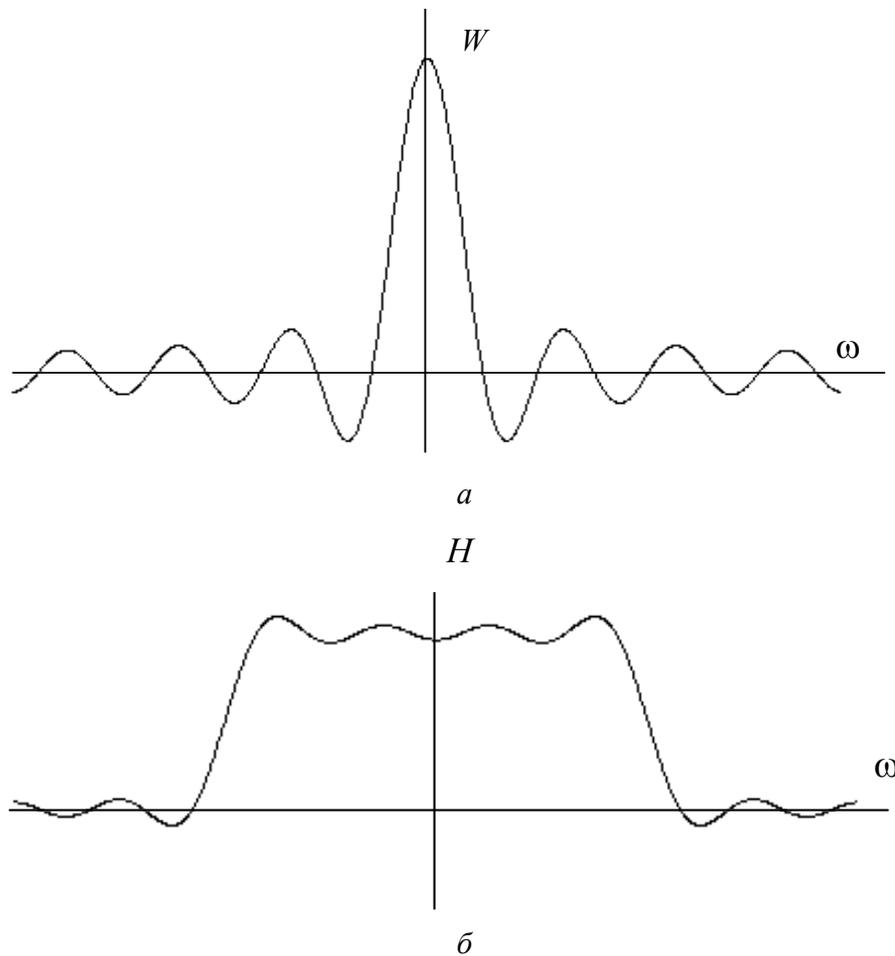


Рис. 6.3. Частотная функция прямоугольного окна (*a*) и результат ее свертки с П-образной АЧХ (*б*)

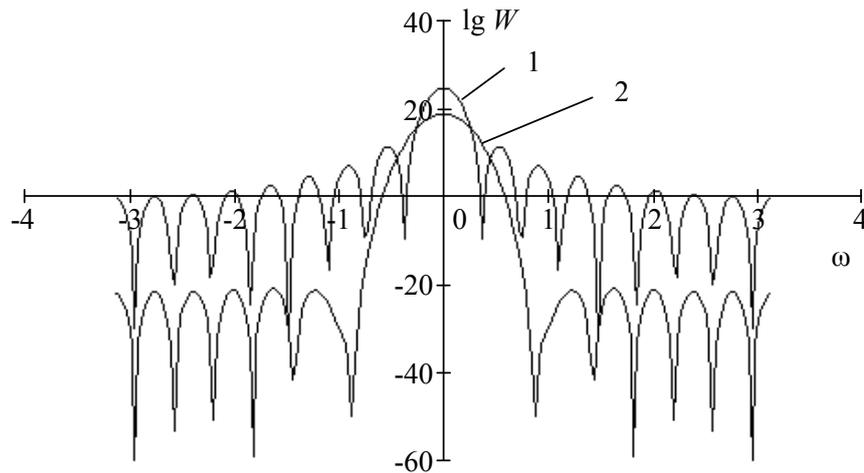


Рис. 6.4. Логарифмы нормированных частотных функций прямоугольного окна и окна Хэмминга

Очевидно, для уменьшения гиббсовских осцилляций КЧХ желательно, чтобы частотная функция окна имела минимальные уровни боковых лепестков. С другой стороны, чтобы КЧХ имела большую крутизну в окрестности частоты среза, частотная функция окна должна иметь как можно более узкий главный лепесток. Эти требования противоречивы, поэтому выбор окна – это выбор компромисса между шириной главного лепестка частотной функции окна, от которой зависит крутизна КЧХ в окрестности частот среза, и уровнем максимального бокового лепестка, определяющего размах гиббсовских осцилляций КЧХ. В таблице приведены данные, позволяющие при синтезе фильтра обоснованно выбрать окно, наиболее подходящее для каждого конкретного случая.

Окно	Ширина главного лепестка (относительно $w_n[n]$)	Относительный уровень первого бокового лепестка АЧХ, дБ
Прямоугольное	1	-21
Бартлетта	2	-25
Хэнна	2	-44
Хэмминга	2	-53
Блэкмана	3	-74
Кайзера	3	около -74

На рис. 6.4 для сравнения приведены логарифмы нормированных частотных функций прямоугольного окна (кривая 1) и окна Хэмминга (кривая 2). На графиках хорошо видны расширение

главного лепестка и уменьшение уровней боковых лепестков при замене прямоугольного окна окном Хэмминга ($N = 17$).

Кроме метода взвешивания иногда применяют другой способ борьбы с гребенчатыми осцилляциями. На этапе формулирования требований к фильтру вводят переходную полосу, в которой задают закон непрерывного изменения АЧХ (например, линейный) [5]. Тогда ряд Фурье для полученной желаемой АЧХ сходится *равномерно* и явление Гиббса отсутствует. Это не означает, что исчезает неравномерность АЧХ, просто осцилляции теперь убывают по амплитуде с увеличением порядка фильтра.

Следует также упомянуть машинные методы синтеза КИХ-фильтров на основе *численной оптимизации*. При этом подбором коэффициентов КИХ-фильтра минимизируются взвешенная среднеквадратическая ошибка

$$\varepsilon_q = \int_{-\pi}^{\pi} q(\omega) \left| H(e^{j\omega}) - H_{\text{ж}}(e^{j\omega}) \right|^2 d\omega,$$

где $q(\omega)$ – весовая функция, позволяющая управлять относительной значимостью ошибок на разных участках частотной оси, или максимальная взвешенная погрешность

$$\varepsilon' = \max_{\omega} \{ q(\omega) | H(e^{j\omega}) - H_{\text{ж}}(e^{j\omega}) | \}.$$

Эти методы позволяют получить меньшие погрешности аппроксимации по сравнению с описанным выше методом оконного взвешивания, но их анализ значительно сложнее [5].

6.3.2. МЕТОД ЧАСТОТНОЙ ВЫБОРКИ

Метод синтеза фильтров с конечной импульсной характеристикой, получивший название метода частотной выборки, основан на аппроксимации желаемой передаточной функции КИХ-фильтра интерполяционным полиномом Лагранжа (*Приложение 4*). При этом предполагается, что точки z -плоскости, в которых задаются значения желаемой передаточной функции, расположены равномерно на 1-окружности и соответствуют точкам частотной оси (отсюда название метода). Очевидно, что при таком выборе узлов интерполяции значения передаточной функции в узлах равны соответствующим значениям КЧХ.

В выражении интерполяционного полинома Лагранжа (П. 4.1)

$$H(z) = \prod_{n=0}^{N-1} (1 - z^{-1} z_n) \sum_{m=0}^{N-1} \frac{A_m}{(1 - z^{-1} z_m)}, \quad (6.11)$$

положим $z_n = e^{j\frac{2\pi}{N}n}$, $n = \overline{0, N-1}$, т. е. узлы интерполяции z_n , $n = \overline{0, N-1}$ разместим равномерно на 1-окружности. Очевидно, все эти точки представляют собой корни степени N из единицы, так как $(z_n)^N = 1, \forall n = \overline{0, N-1}$, следовательно, они являются корнями полинома $1 - z^{-N}$. Тогда произведение в (6.11) можно в соответствии с основной теоремой алгебры заменить на $1 - z^{-N}$. Коэффициенты A_m в формуле (6.11) находятся для этого случая как

$$A_m = \frac{H_m}{\prod_{\substack{k=0 \\ k \neq m}}^{N-1} (1 - z_m^{-1} z_k)} = H_m \lim_{z \rightarrow z_m} \frac{(1 - z^{-1} z_k)}{\prod_{k=0}^{N-1} (1 - z_m^{-1} z_k)}.$$

Раскрывая неопределенность по правилу Лопиталья, получаем

$$\begin{aligned} A_m &= H_m \lim_{z \rightarrow z_m} \frac{(1 - z^{-1} z_k)}{(1 - z^{-N})} = H_m \lim_{z \rightarrow z_m} \frac{(z^{-2} z_k)}{(N z^{-N-1})} = \\ &= H_m \frac{z_m^N}{N} = \frac{H_m}{N} = \frac{1}{N} H(e^{j\frac{2\pi}{N}m}). \end{aligned}$$

На основании всего изложенного можно записать

$$H(z) = \frac{1 - z^{-N}}{N} \sum_{m=0}^{N-1} \frac{H(e^{j\frac{2\pi}{N}m})}{1 - z^{-1} e^{j\frac{2\pi}{N}m}} = (1 - z^{-N}) \sum_{m=0}^{N-1} \frac{A_m}{1 - z^{-1} W_N^{-m}},$$

где $W_N = e^{-j\frac{2\pi}{N}}$. Структура полученного фильтра с частотной выборкой представлена на рис. 6.5. Очевидно, особенно простая структура получится в том случае, когда большинство из заданных

значений $H\left(e^{j\frac{2\pi}{N}m}\right)$, $m = \overline{0, N-1}$, равно 0, что выполняется для фильтров с узкой полосой пропускания. Это обстоятельство и определяет область применения данного метода. Кроме того, при программной реализации на основе сигнальных процессоров (см. разд. 9) фильтры на основе частотной выборки могут давать очень большой выигрыш в быстродействии по сравнению с другими формами реализации КИХ-фильтров. Этот выигрыш связан с рекурсивностью разностного уравнения, описывающего динамику такого фильтра. Преимущества фильтра на основе частотной выборки наглядно иллюстрируются следующим примером [20].

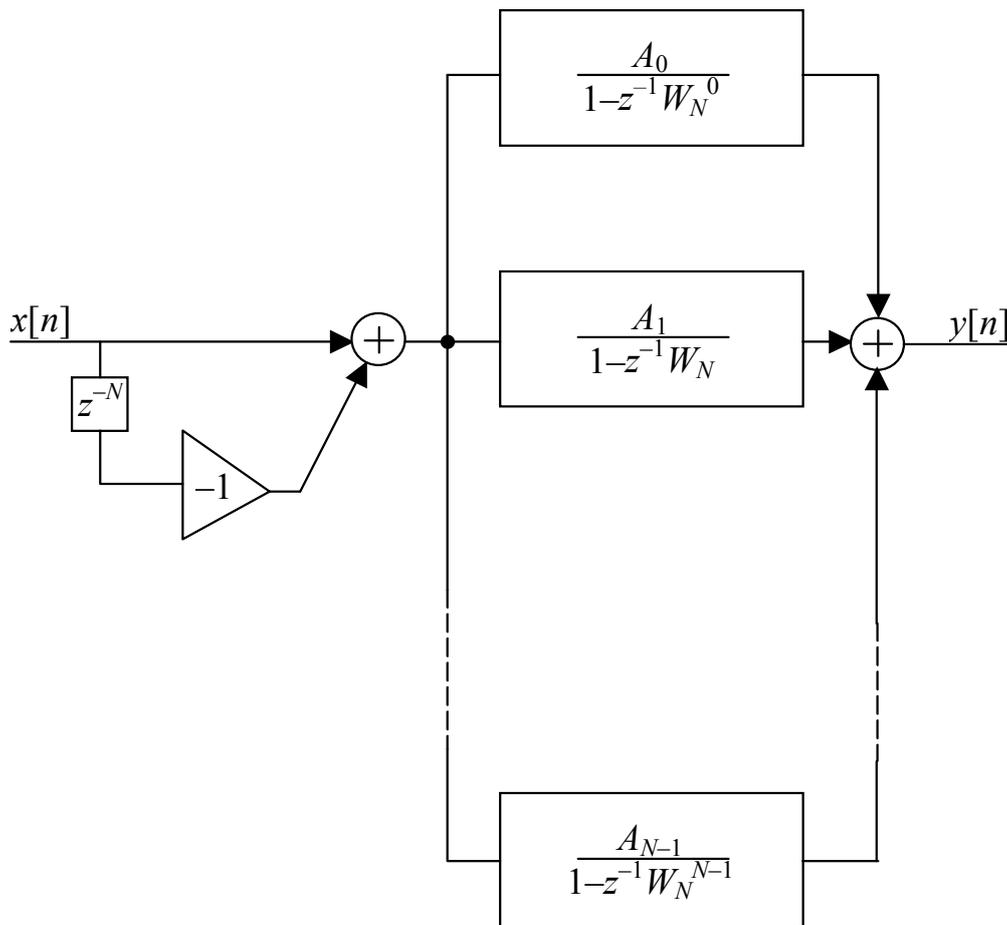


Рис. 6.5. Структура фильтра с частотной выборкой

Пример 6.2. Рассмотрим фильтр «скользящего среднего» с конечной импульсной характеристикой длины $2N$ отсчетов

$$h[n] = 1, \quad n = \overline{0, 2N-1}.$$

Передаточная функция такого фильтра

$$H(z) = \frac{Y(z)}{X(z)} = 1 + z^{-1} + \dots + z^{-2N} = \frac{1 - z^{-2N-1}}{1 - z^{-1}}$$

имеет один простой полюс в точке $z=1$ и $2N+1$ нулей, равномерно размещенных на единичной окружности.

Разностное уравнение данного фильтра имеет вид $y[n] = y[n-1] + x[n] - x[n-2N-1]$ и требует для вычисления одного отсчета выходной последовательности выполнения двух операций сложения/вычитания *независимо* от величины N . Для сравнения: реализация такой же передаточной функции в нерекурсивной форме соответствует разностному уравнению $y[n] = x[n] + x[n-1] + \dots + x[n-2N]$ и требует выполнения $2N$ операций сложения/вычитания. ◀

6.3.3. МЕТОД БЫСТРОЙ СВЕРТКИ

Известно, что воздействие ЛИС-цепи на сигнал исчерпывающе описывается сверткой этого сигнала с импульсной характеристикой цепи. В некоторых случаях в целях уменьшения аппаратурно-временных затрат выгоднее заменить свертку последовательностей перемножением их z -образов. Теоретическим обоснованием такой замены является изоморфизм алгебры l_1 абсолютно суммируемых последовательностей и алгебры Z функций комплексного переменного z , непрерывных на единичной окружности (см. разд. 2). С практической точки зрения эта замена может быть целесообразна благодаря существованию эффективных (быстрых) алгоритмов вычисления ДПФ, называемых алгоритмами *быстрого преобразования Фурье* (БПФ) (Приложение 5).

Очевидно, что z -преобразование *практически* может быть вычислено лишь в конечном множестве точек z -плоскости. Поэтому выполнять свертку косвенным путем, т. е. через умножение z -образов, можно лишь для последовательностей *конечной* длины. Следовательно, этим способом можно реализовать только КИХ-фильтрацию. Метод КИХ-фильтрации последовательностей путем поточечного умножения их ДПФ-спектров на КЧХ цепи, реализуемый на основе БПФ, называется методом *быстрой свертки*.

При реализации метода быстрой свертки необходимо учитывать следующие два обстоятельства. Первое состоит в том, что, как известно, дискретное преобразование Фурье обладает двойствен-

ностью – оно соответствует как последовательностям конечной длины, так и периодическим последовательностям. По этой причине перемножение коэффициентов ДПФ двух последовательностей соответствует не обычной (апериодической), а *циклической* свертке. Убедимся, что это действительно так.

Пусть $\tilde{x}[n]$ и $\tilde{h}[n]$ – периодические последовательности с периодом N . Их циклическая свертка определяется выражением

$$\tilde{y}[n] = \sum_{m=0}^{N-1} \tilde{x}[m] \tilde{h}[n-m].$$

Дискретное преобразование Фурье результирующей последовательности

$$\begin{aligned} \tilde{Y}[k] &= \sum_{n=0}^{N-1} \left(\sum_{m=0}^{N-1} \tilde{x}[m] \tilde{h}[n-m] \right) e^{-j\frac{2\pi}{N}nk} = \\ &= \sum_{m=0}^{N-1} \tilde{x}[m] \left(\sum_{n=0}^{N-1} \tilde{h}[n-m] e^{-j\frac{2\pi}{N}(n-m)k} \right) e^{-j\frac{2\pi}{N}mk} = \tilde{H}[k] \tilde{X}[k]. \end{aligned}$$

Здесь учтен тот факт, что сумма в круглых скобках во второй строке равна $\tilde{H}[k]$ независимо от m в силу *периодичности* суммируемых членов: при различных m суммируются *одни и те же* N слагаемых в разном порядке. Таким образом, видно, что циклическая свертка (или свертка периодических последовательностей) соответствует поточечному произведению ДПФ-спектров последовательностей. При фильтрации же должна выполняться обычная *апериодическая* свертка, определяемая выражением (2.3).

Преодолеть эту трудность можно с помощью следующих рассуждений. z -Образ последовательности длины N есть полином (относительно z^{-1}) степени $N-1$. Полагая, что $x[n]$ и $h[n]$ – последовательности длины N , видим, что произведение их z -образов есть полином степени $2(N-1)$. Но при поточечном перемножении ДПФ-спектров, имеющих по N отсчетов, получается всего N результирующих отсчетов, что соответствует полиному всего лишь $(N-1)$ -й степени. Необходимое условие получения правильного результата умножения двух полиномов состоит в том, чтобы точек вычисления ДПФ «хватило» для представления результата. Иначе говоря, если используется N -точечное ДПФ, то степень *результати-*

рующего полинома должна быть *не выше* $(N - 1)$. Это, в свою очередь, означает, что сумма длин последовательностей $x[n]$ и $h[n]$ (обозначим их L и M) должна удовлетворять очевидному соотношению $M - 1 + L - 1 \leq N - 1$ (или $M + L - 1 \leq N$). Для того чтобы для последовательности $x[n]$ длиной $L < N$ можно было получить N отсчетов ДПФ, следует *перед* вычислением ДПФ последовательность $x[n]$ дополнить $(N - L)$ нулевыми отсчетами (и то же проделать с другой последовательностью). Сказанное иллюстрируется рис. 6.6, где показаны последовательности длиной 22 и 11 отсчетов, а также результат их

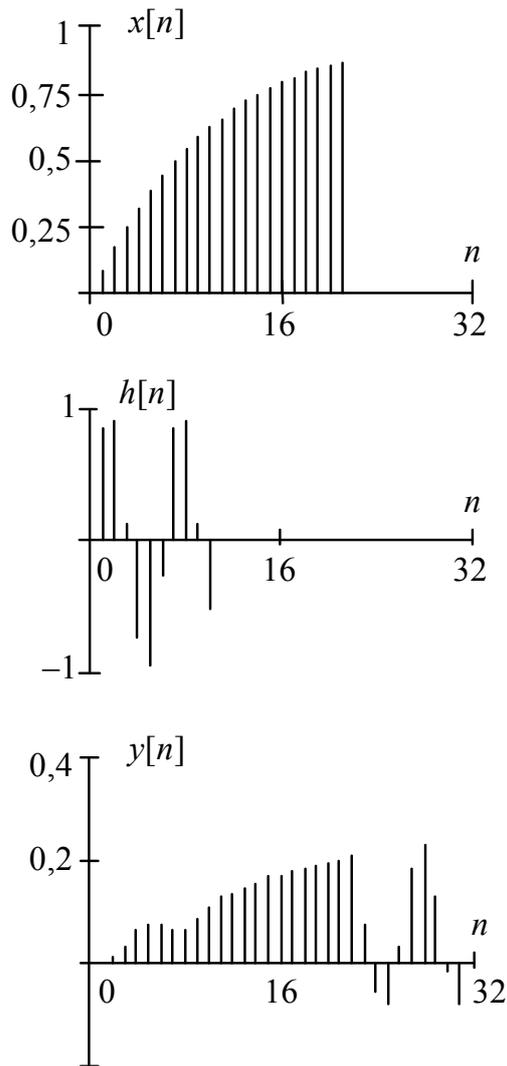


Рис. 6.6. Последовательности длиной 22 и 11 отсчетов, а также результат их циклической свертки

отсчетов, а также результат их циклической свертки, полученный применением 32-точечного БПФ и совпадающий с результатом аperiodической свертки.

Второе обстоятельство, которое должно учитываться при реализации КИХ-фильтрации методом быстрой свертки, относится к фильтрации последовательностей большой (в частности, бесконечной) длины. Если длина входной последовательности велика (сотни тысяч отсчетов и более), что типично для обработки сигналов, применяемых в радиотехнике и связи, то необходимое *основание* БПФ (число вычисляемых отсчетов) оказывается слишком большим, что влечет высокие требования к объему оперативной памяти вычислителя БПФ, а также приводит к большой задержке результирующего сигнала (результат может быть получен не ранее, чем поступит последний отсчет входной последовательности, плюс время, необходимое для вычисления прямого БПФ, умножения и обратного БПФ). Для того чтобы снизить требования к памяти

и уменьшить задержку, применяют так называемое секционирование свертки [1 – 3].

Пусть $h[n]$ – импульсная характеристика фильтра, имеющая длину M , а $x[n]$ – сигнальная последовательность бесконечной длины. Представим $x[n]$ в виде

$$x[n] = \sum_{k=-\infty}^{\infty} x_k[n], \quad \text{где} \quad x_k[n] = \begin{cases} x[n], & kL \leq n < (k+1)L, \\ 0 & \text{в противном случае.} \end{cases}$$

Нетрудно видеть, что таким образом входная последовательность разбивается (*секционируется*) на совокупность неперекрывающихся примыкающих друг к другу сегментов длиной L отсчетов каждый.

В силу *билинейности* свертки (линейности по каждому из операндов)

$$x[n] \otimes h[n] = \left(\sum_{k=-\infty}^{\infty} x_k[n] \right) \otimes h[n] = \sum_{k=-\infty}^{\infty} (x_k[n] \otimes h[n]) = \sum_{k=-\infty}^{\infty} y_k[n].$$

Как видно из полученного выражения, свертка последовательности бесконечной длины с конечной импульсной характеристикой может быть **точно** заменена бесконечной суммой свертки сегментов фиксированной длины с этой же импульсной характеристикой. Каждая частичная свертка требует для своего вычисления $(L + M - 1)$ -точечного БПФ. Учитывая, что L выбирается произвольно, секционирование позволяет реализовать КИХ-фильтрацию длинных (потенциально – *бесконечно длинных*) последовательностей; при этом результат $y_k[n]$ фильтрации сегмента $x_k[n]$ появляется с задержкой, определяемой временем прямого и обратного БПФ и умножения, но эта задержка теперь отсчитывается от момента окончания сегмента, а не всей последовательности.

Заметим, что результат каждой частичной свертки имеет длину $(L + M - 1)$, а следуют секции с относительным сдвигом L . Таким образом, результаты частичных свертки $y_k[n]$ суммируются с перекрытием носителей $kL \leq n < (k+1)L + M - 1$. Этот метод секционирования свертки известен как метод перекрытия с суммированием (*overlap-add method* [1]).

Альтернативный метод перекрытия с накоплением (*overlap-save method* [1]) заключается в том, что перекрываются сегменты

входной последовательности. Последовательность $x[n]$ разбивается на секции $x_k[n]$ длиной $(L + M - 1)$, причем последние $(M - 1)$ отсчетов каждой секции перекрываются с таким же количеством первых отсчетов следующей секции (M – по-прежнему длина импульсной характеристики, L выбирается произвольно). Каждая секция подвергается БПФ с основанием $(L + M - 1)$. Импульсная характеристика длиной M дополняется нулями до основания БПФ. Результат фильтрации (циклической свертки) содержит всего $(L + M - 1)$ отсчетов, из которых последние $(M - 1)$ отсчетов – «ошибочные», не совпадающие с результатом апериодической свертки, а остальные L отсчетов являются «правильными». Суммируются результаты циклических частичных свертки после отбрасывания «ошибочных» отсчетов, т. е. секции выходной последовательности длиной L следуют друг за другом без перекрытия [3].

6.4. СИНТЕЗ БИХ-ФИЛЬТРОВ МЕТОДОМ АНАЛОГО-ЦИФРОВОЙ ТРАНСФОРМАЦИИ

Фильтры с бесконечной импульсной характеристикой имеют по сравнению с КИХ-фильтрами важное преимущество: приемлемая крутизна амплитудно-частотной характеристики КИХ-фильтра достигается обычно при порядке фильтра 100 – 300, в то время как БИХ-фильтр со сравнимыми качествами имеет порядок 10 – 15. Это вполне объяснимо, так как известно, что быстро изменяющиеся сигналы имеют широкий спектр, а благодаря двойственности (дуализму) времени и частоты отсюда следует, что круто изменяющейся функции частоты (АЧХ) должна соответствовать функция времени (импульсная характеристика) большой длительности. Но для КИХ-фильтра порядок – это количество отсчетов импульсной характеристики минус 1, в то время как БИХ-фильтр даже первого порядка имеет импульсную характеристику бесконечной длительности.

Таким образом, в тех случаях, когда вид фазочастотной характеристики не играет определяющей роли для практического применения разрабатываемого фильтра, следует использовать БИХ-фильтр, так как при этом получается существенный выигрыш в быстродействии (или в аппаратных затратах на реализацию) фильтра. То обстоятельство, что не всякий БИХ-фильтр оказывается устойчивым, не представляет такой заметной опасности, как может показаться на первый взгляд. Во-первых, устойчивость может быть проверена (и обеспечена) на этапе проектирования циф-

рового фильтра; во-вторых, характеристики цифровых фильтров не подвержены дрейфу с течением времени, и, следовательно, устойчивый фильтр останется устойчивым в течение всего времени работы (конечно, нужно учитывать возможность выхода аппаратуры из строя в результате *катастрофического отказа*; при этом может возникнуть неустойчивость). Следует также отметить, что в цифровых БИХ-фильтрах могут возникать незатухающие паразитные колебания (так называемые предельные циклы) вследствие своеобразного нарушения устойчивости при округлении дробных чисел; подробнее этот вопрос будет рассмотрен в разд. 7.

Практически все известные методы синтеза цифровых БИХ-фильтров основаны на так называемой *аналого-цифровой трансформации*, т.е. на преобразовании аналогового фильтра с требуемыми характеристиками в цифровой^{*)} фильтр. Это объясняется, во-первых, трудностью решения задачи прямой аппроксимации желаемых характеристик дробно-рациональными передаточными функциями и, во-вторых, наличием развитой теории синтеза аналоговых фильтров и простотой преобразования аналоговых фильтров-прототипов в цифровые фильтры.

Краткий обзор аналоговых фильтров-прототипов, наиболее часто применяемых при синтезе цифровых фильтров, приведен в *Приложении 6*. В этом разделе рассматриваются два наиболее распространенных метода аналого-цифровой трансформации.

6.4.1. МЕТОД ИНВАРИАНТНОСТИ ИМПУЛЬСНОЙ ХАРАКТЕРИСТИКИ

Аналоговые фильтры принято описывать передаточными функциями, которые связаны с импульсными характеристиками преобразованием Лапласа [13]. Преобразование Лапласа связывает аналоговый сигнал $x(t)$ с его образом (*изображением*) в виде функции $X(p)$ комплексного переменного p . Мнимая ось комплексной p -плоскости представляет собой ось частот в описании аналогового сигнала. Аналогичную роль в описании дискретных сигналов играет единичная окружность z -плоскости.

Выше была установлена связь частот в описании аналогового и дискретного сигналов (6.2). Это выражение можно распространить

^{*)} Строго говоря, в результате трансформации получается дискретный фильтр, но поскольку в рамках задачи построения цифрового фильтра не возникает недоразумений, получаемый дискретный фильтр называют цифровым.

с частотных осей на комплексные плоскости, соответствующие описаниям аналоговых и дискретных сигналов при помощи преобразования Лапласа и z -преобразования. Связь комплексных переменных p и z , соответствующая дискретизации аналоговых сигналов, описывается выражениями

$$z = e^{pT_d} \quad (6.12)$$

и

$$p = \frac{1}{T_d} \ln z. \quad (6.13)$$

Заметим, что изменение значения pT_d на $\pm 2\pi$ никак не влияет на значение z . Следовательно, отображение p -плоскости на z -плоскость (рис. 6.7) не является взаимно однозначным и происходит следующим образом: множество точек p -плоскости $\{p = \sigma + j\Omega, \sigma = -\infty, \infty, \Omega T_d \in [-\pi, \pi]\}$, т. е. полоса шириной $2\pi/T_d$, отмеченная на рисунке штриховкой, отображается на всю z -плоскость, и так же отображаются все полосы, получаемые из данной полосы последовательными сдвигами вверх и вниз на $2\pi/T_d$. При

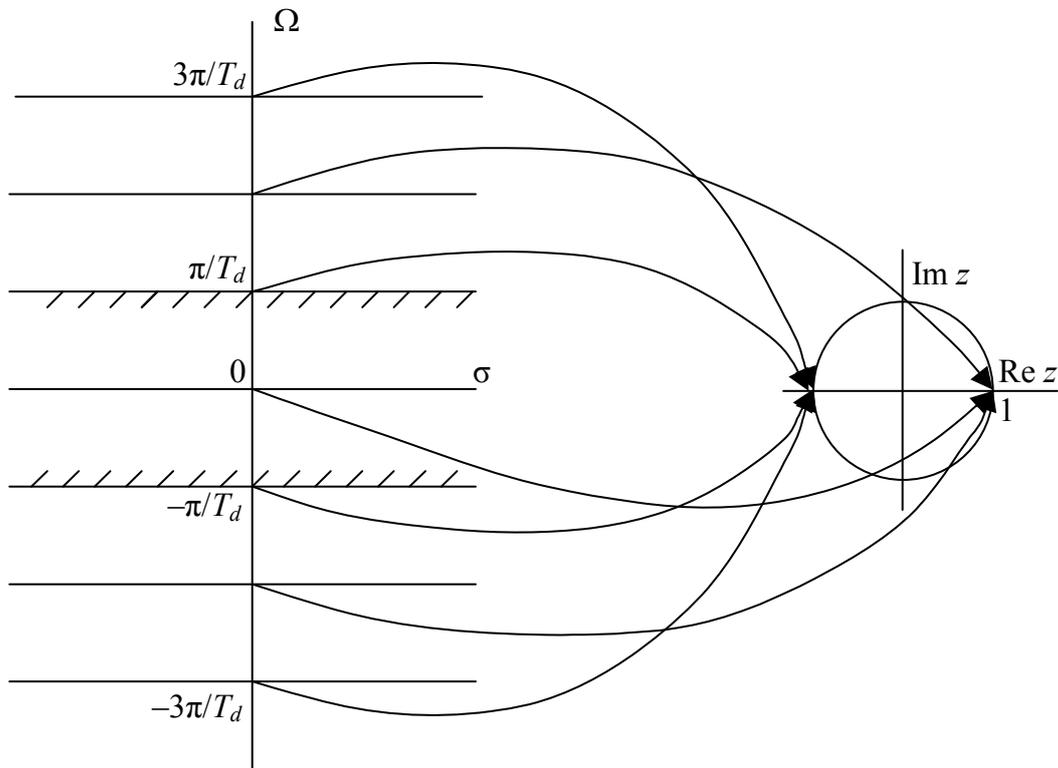


Рис. 6.7. Отображение комплексной p -плоскости на z -плоскость

этом левая половина полосы отображается внутрь единичной окружности z -плоскости, а правая – на всю остальную часть z -плоскости. Точки мнимой оси p -плоскости (оси частот аналогового сигнала) $0, \pm 2\pi/T_d, \pm 4\pi/T_d$ отображаются в единственную точку единичной окружности z -плоскости (оси частот дискретного сигнала) – точку $z=1$. Аналогично точки мнимой оси p -плоскости $\pm\pi/T_d, \pm 3\pi/T_d$ отображаются в единственную точку единичной окружности z -плоскости – точку $z=-1$. Вообще в любую точку единичной окружности z -плоскости отображается бесконечное множество точек мнимой оси p -плоскости. Это объясняет, в частности, наложение спектральных «хвостов» при дискретизации сигналов с нефинитными спектрами, так как «аналоговая» частотная ось отображается на «дискретную» частотную ось *многократно*.

Передаточная функция произвольного аналогового фильтра (с сосредоточенными параметрами) имеет вид дробно-рациональной функции комплексного переменного p . Такая функция может быть представлена в виде суммы дробей

$$H_a(p) = \sum_{k=1}^N \frac{A_k}{(p - p_k)},$$

где $p_k, k = \overline{1, N}$ – полюсы передаточной функции, а коэффициенты A_k определяются из условия равенства числителя передаточной функции $H_a(p)$ и числителя правой части после приведения ее к общему знаменателю. (Здесь мы рассматриваем лишь важный для практики случай, когда степень числителя $H_a(p)$ меньше степени знаменателя, а все корни знаменателя не кратные.)

Ввиду линейности преобразования Лапласа импульсная характеристика такого фильтра имеет вид суммы экспоненциальных

функций непрерывного времени $h_a(t) = \sum_{k=1}^N A_k e^{p_k t} \sigma(t)$, где $\sigma(t)$ – функция Хевисайда, определяемая выражением

$$\sigma(t) = \begin{cases} 1, & t \geq 0, \\ 0 & \text{в противном случае.} \end{cases}$$

Очевидно, для того чтобы импульсная характеристика затухала со временем (т. е. фильтр был устойчивым), необходимо и доста-

точно, чтобы все полюсы были расположены в p -плоскости слева от мнимой оси.

Метод аналого-цифровой трансформации, известный под названием метода *инвариантности импульсной характеристики*, основан на прямом применении теоремы отсчетов (теоремы Котельникова). Рассматривая импульсную характеристику аналогового фильтра-прототипа как функцию времени (сигнал), можно заменить ее последовательностью отсчетов, выбранных с достаточно малым шагом дискретизации T_d .

Результатом дискретизации импульсной характеристики аналогового фильтра будет последовательность

$$h[n] = h_a(nT_d) = \sum_{k=1}^N A_k e^{p_k n T_d} u[n] = \sum_{k=1}^N A_k r_k^n u[n],$$

где

$$r_k = e^{p_k T_d}. \quad (6.14)$$

Из полученного выражения видно, что при дискретизации импульсной характеристики каузального аналогового фильтра с дробно-рациональной передаточной функцией получается сумма каузальных экспоненциальных последовательностей. Таким образом, реализуемому аналоговому фильтру соответствует реализуемый цифровой фильтр (напомним, что импульсная характеристика БИХ-цепи с некрратными полюсами представляет собой сумму экспоненциальных последовательностей). Кроме того, полюсы цифрового фильтра связаны с полюсами фильтра-прототипа простым соотношением $r_k = e^{p_k T_d}$, $k = \overline{1, N}$, вытекающим из (6.12), поэтому устойчивому аналоговому фильтру ($\text{Re}\{p_k\} < 0$) соответствует устойчивый цифровой фильтр того же порядка N (так как $|r_k| = |e^{p_k T_d}| < 1$). Зная полюсы цифрового фильтра, можно сразу записать его передаточную функцию

$$H(z) = \sum_{k=1}^N \frac{A_k}{(1 - r_k z^{-1})},$$

и на этом аналого-цифровая трансформация заканчивается. Полученное выражение передаточной функции соответствует параллельной форме реализации; отсюда нетрудно перейти к каскадной, прямой или канонической формам (см. разд. 3).

Поскольку импульсная характеристика цифрового фильтра есть продукт дискретизации импульсной характеристики аналогового фильтра, то КЧХ цифрового фильтра связана с КЧХ аналогового фильтра соотношением

$$H(e^{j\omega}) = H(e^{j\Omega T_d}) = \frac{1}{T_d} \sum_{k=-\infty}^{\infty} H_a [j(\Omega + k\Omega_d)],$$

где $\Omega_d = 2\pi/T_d$ – частота дискретизации (см. п. 6.1). Очевидно, если КЧХ прототипа не финитна, а это *всегда* так для фильтров конечного порядка, то неизбежно наложение (суммирование) «хвостов» сдвинутых копий $H_a(\cdot)$ и, как следствие, *искажение* формы получаемой КЧХ дискретного фильтра по отношению к КЧХ фильтра-прототипа. Этот эффект ограничивает практическое применение метода инвариантности импульсной характеристики в основном задачами синтеза цифровых фильтров *нижних частот*.

Пример 6.1. Рассмотрим аналоговый фильтр с передаточной функцией

$$H_a(p) = \frac{1/2}{p + \alpha + j\Omega_0} + \frac{1/2}{p + \alpha - j\Omega_0} = \frac{p + \alpha}{(p + \alpha)^2 + \Omega_0^2}.$$

(Амплитудно-частотная характеристика фильтра представлена на рис. 6.8, а нуль-полюсная диаграмма – на рис. 6.9).

Полюсы аналогового фильтра $p_1 = -\alpha - j\Omega_0$ и $p_2 = p_1^* = -\alpha + j\Omega_0$ отобразим в z -плоскость по методу инвариантности импульсной характеристики и получим согласно (6.14) полюсы

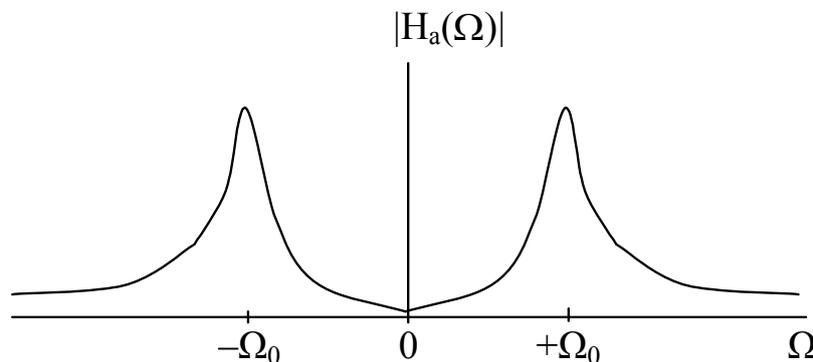


Рис. 6.8. Амплитудно-частотная характеристика фильтра из примера 6.1

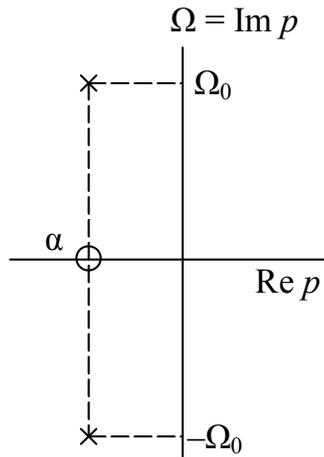


Рис. 6.9. Нуль-полюсная диаграмма аналогового фильтра

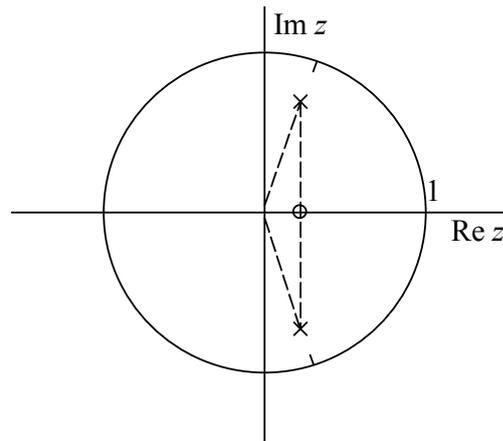


Рис. 6.10. Нуль-полюсная диаграмма дискретного фильтра

цифрового фильтра $r_1 = e^{-\alpha T_d - j\Omega_0 T_d}$ и $r_2 = r_1^* = e^{-\alpha T_d + j\Omega_0 T_d}$. Тогда передаточная функция

$$H(z) = \frac{1/2}{1 - r_1 z^{-1}} + \frac{1/2}{1 + r_2 z^{-1}} =$$

$$= \frac{1 - e^{-\alpha T_d} \cos \Omega_0 T_d z^{-1}}{(1 - e^{-\alpha T_d} e^{-j\Omega_0 T_d} z^{-1})(1 - e^{-\alpha T_d} e^{+j\Omega_0 T_d} z^{-1})},$$

как видно, имеет два полюса и один нуль (рис. 6.10). Это проявление общей закономерности – количество полюсов и нулей при трансформации по методу инвариантности импульсной характеристики сохраняется. Заметим, однако, что расположение нулей передаточной функции цифрового фильтра в z -плоскости нельзя найти таким же простым способом, как расположение полюсов. ◀

6.4.2. МЕТОД БИЛИНЕЙНОГО ПРЕОБРАЗОВАНИЯ

Естественным способом преобразования аналогового фильтра в цифровой была бы прямая подстановка в передаточную функцию $H_a(p)$ фильтра-прототипа выражения (6.13), т. е. замена переменных. Однако такая подстановка дает выражение $H_a(1/T_d \ln z)$, не являющееся дробно-рациональным относительно z^{-1} (что необходимо для реализации передаточной функции цифрового фильтра конечного порядка). Таким образом, точная замена переменных не позволяет получить передаточную функцию дискретного фильтра

из передаточной функции аналогового прототипа. Однако существует приближенная формула замены, сохраняющая дробную рациональность функции. Подставив в (6.13) разложение функции логарифма в ряд, ограниченное первым слагаемым [12], получим

$$p = \frac{2}{T_d} \left(\frac{z-1}{z+1} + \frac{(z-1)^3}{3(z+1)^3} + \frac{(z-1)^5}{5(z+1)^5} + \dots \right) \approx \frac{2}{T_d} \frac{z-1}{z+1}$$

или

$$p = \frac{2}{T_d} \frac{1-z^{-1}}{1+z^{-1}}. \quad (6.15)$$

Это выражение в отличие от (6.13) дробно-рационально относительно z^{-1} , поэтому после его подстановки в дробно-рациональную передаточную функцию $N_a(p)$ аналогового прототипа получается снова дробно-рациональная, а следовательно, реализуемая передаточная функция $H(z)$ цифрового фильтра.

Выясним, как располагаются в z -плоскости полюсы передаточной функции $H(z)$, если полюсы передаточной функции прототипа $N_a(p)$ находятся в левой части комплексной плоскости (иными словами, является ли устойчивым цифровой фильтр, если устойчив фильтр-прототип).

Выразим на основе выражения (6.15) комплексное переменное z через p :

$$z^{-1} p T_d + p T_d = 2 - 2z^{-1},$$

отсюда

$$z = \frac{2 + p T_d}{2 - p T_d}. \quad (6.16)$$

Чтобы выяснить, в какое множество z -плоскости отображается мнимая ось p -плоскости (ось частоты), подставим в это выражение $j\Omega$ вместо p , тогда получим выражение для образа мнимой оси p -плоскости при отображении, описываемом выражением (6.16):

$$z = \frac{2/T_d + j\Omega}{2/T_d - j\Omega}.$$

Числитель и знаменатель этой дроби суть комплексно-сопряженные числа, поэтому модуль дроби равен единице при всех Ω .

Это означает, что мнимая ось p -плоскости отображается преобразованием (6.16) на единичную окружность z -плоскости. Но переменная Ω – это частота, соответствующая описанию аналогового фильтра; роль частотной оси для цифровых цепей играет единичная окружность на z -плоскости (множество точек $e^{j\omega}$ при значениях ω , принимающих значения из интервала от $-\pi$ до π). Заменяя z на $e^{j\omega}$, получим

$$e^{j\omega} = \frac{2/T_d + j\Omega}{2/T_d - j\Omega} = \frac{1 + j\Omega T_d / 2}{1 - j\Omega T_d / 2},$$

тогда

$$\omega = \arg \left\{ \frac{1 + j\Omega T_d / 2}{1 - j\Omega T_d / 2} \right\} = 2 \operatorname{arctg} \frac{\Omega T_d}{2},$$

следовательно, связь «аналоговой» и «цифровой» частот при билинейном преобразовании описывается выражениями

$$\frac{\omega}{2} = \operatorname{arctg} \frac{\Omega T_d}{2}, \quad (6.17)$$

$$\frac{\Omega T_d}{2} = \operatorname{tg} \frac{\omega}{2}. \quad (6.18)$$

Поскольку $-\pi \leq \omega \leq \pi$, а $-\infty < \Omega < \infty$, то нетрудно видеть, что вся аналоговая частотная ось (бесконечной длины) отображается на единичную окружность (длины 2π), причем это отображение *однократно* в отличие от (6.12), и вследствие этого различные участки оси Ω испытывают различное «сжатие» при отображении на ось ω (единичную окружность). Это необходимо учитывать при проектировании цифровых фильтров на этапе определения требований к частотам среза фильтров-прототипов. Способ пересчета требуемых частот цифрового фильтра в характерные частоты прототипа очевиден из рис. 6.11.

Заслуживает внимания вопрос, насколько вредна нелинейная трансформация частотной оси с точки зрения задачи синтеза цифровых фильтров. Очевидно, что при проектировании фильтров с желаемыми АЧХ *кусочно-постоянного* вида указанная нелинейность трансформации частотной оси не влияет на качество цифрового фильтра, так как приводит лишь к необходимости на этапе построения аналогового фильтра-прототипа учитывать последующее изменение характерных частот фильтра (граничных частот)

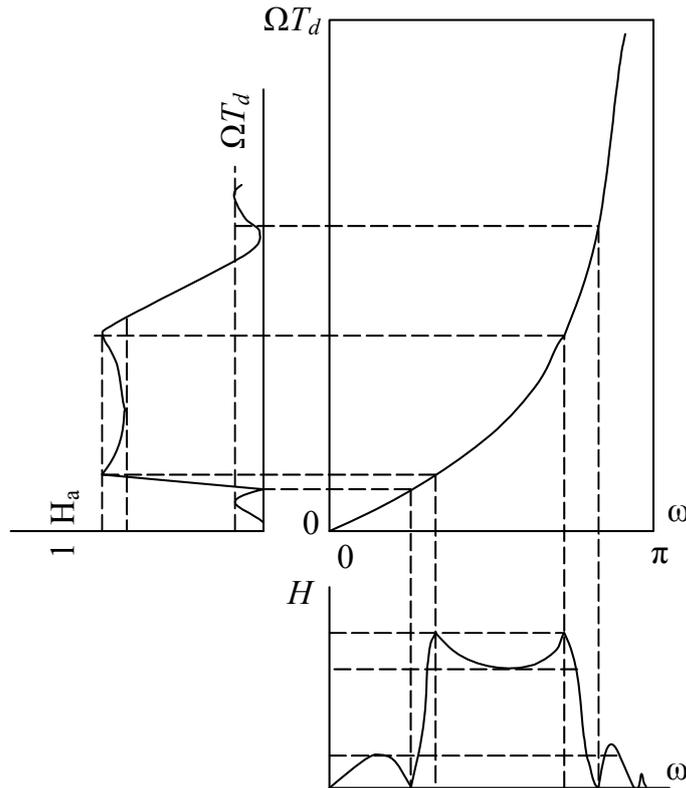


Рис. 6.11 Деформация шкалы частот при билинейном преобразовании

при билинейном преобразовании. Если же требуется построить фильтр, не являющийся типовым (ФНЧ, фильтром верхних частот – ФВЧ, полосовым или режекторным), то в общем случае нелинейность отображения частотной оси приводит к искажению формы АЧХ. Например, интегрирующий аналоговый фильтр имеет амплитудно-частотную характеристику гиперболического вида $1/\Omega$ и при билинейном преобразовании не приводит к интегрирующему цифровому фильтру.

Для того чтобы устойчивый аналоговый фильтр трансформировался в устойчивый же цифровой фильтр, требуется, чтобы при билинейном преобразовании левая полуплоскость p -плоскости отображалась внутрь единичной окружности. Разлагая p на мнимую и вещественную части, для билинейного преобразования (6.16) получим

$$z = \frac{2 + \operatorname{Re}\{pT_d\} + j \operatorname{Im}\{pT_d\}}{2 - \operatorname{Re}\{pT_d\} - j \operatorname{Im}\{pT_d\}}.$$

Поскольку мнимые части числителя и знаменателя одинаковы, то модуль дроби будет меньше 1, если $\operatorname{Re}\{pT_d\} < 0$. Тогда любой полюс функции $H_a(p)$, лежащий *слева* от мнимой оси, отображается в полюс функции $H(z)$, расположенный *внутри* 1-окружности. Это означает, что устойчивый аналоговый фильтр трансформируется преобразованием (6.15) в устойчивый дискретный фильтр.

Итак, установлено, что билинейное преобразование трансформирует устойчивый реализуемый аналоговый фильтр в устойчивый реализуемый цифровой фильтр. При этом вследствие однократности отображения частотной оси на 1-окружность отсутствует наложение «хвостов» КЧХ, что является достоинством билинейного преобразования. К недостаткам следует отнести то, что не сохраняются ни импульсная, ни фазочастотная характеристики фильтра (точнее было бы сказать, что импульсная и фазочастотная характеристики дискретного фильтра могут *сильно отличаться по форме* от соответствующих характеристик прототипа).

Порядок расчета цифрового фильтра (ЦФ) методом билинейного преобразования состоит в следующем:

1. Определение характерных частот ЦФ и пересчет их в частоты аналогового фильтра в соответствии с (6.18).
2. Синтез аналогового фильтра, удовлетворяющего заданным условиям.
3. Подстановка формулы (6.15) билинейного преобразования в выражение $H_a(p)$ передаточной функции фильтра-прототипа.

6.4.3. ЧАСТОТНЫЕ ПРЕОБРАЗОВАНИЯ ЦИФРОВЫХ ФИЛЬТРОВ

Синтез фильтра верхних частот, полосового или режекторного фильтров может быть проведен в два этапа: на первом этапе строится некоторый фильтр нижних частот, а на втором этапе этот ФНЧ преобразуется в фильтр желаемого вида. Такой способ, основанный на специально разработанных для этого *частотных* преобразованиях, часто оказывается более простым, чем непосредственный синтез ЦФ требуемого вида. При помощи частотного преобразования можно также получить фильтр нижних частот из другого ФНЧ.

Пусть имеется передаточная функция цифрового фильтра нижних частот $H(z)$. Передаточную функцию желаемого фильтра обозначим $H_{\text{ж}}(Z)$ и будем рассматривать как функцию другого комплексного переменного Z . Частотные преобразования пред-

ставляют собой отображения комплексной z -плоскости *на себя*, таким образом, Z – это комплексное переменное z после преобразования.

Частотное преобразование должно удовлетворять определенным условиям, а именно: устойчивый фильтр должен остаться устойчивым, реализуемый фильтр – реализуемым, каузальный – каузальным. Это означает, что 1-окружность z -плоскости отображается на 1-окружность Z -плоскости, внутренность единичного круга z -плоскости – на внутренность единичного круга Z -плоскости, а дробно-рациональные функции при этом отображении остаются дробно-рациональными. Требуемыми свойствами обладают преобразования Константиноидиса [2]

$$G(Z^{-1}) = \pm \prod_{k=1}^N \frac{Z^{-1} - \alpha_k}{1 - \alpha_k Z^{-1}} \text{ при } |\alpha_k| < 1.$$

Частотное преобразование в общем виде описывается выражением $H_{\text{ж}}(Z) = H(z)|_{z^{-1}=G(Z^{-1})}$, связывающим желаемую передаточную функцию $H_{\text{ж}}(Z)$ с передаточной функцией исходного фильтра. Чтобы получить желаемую передаточную функцию, следует в передаточную функцию исходного фильтра подставить вместо z^{-1} соответствующее выражение $G(Z^{-1})$.

Рассмотрим для примера преобразование фильтра нижних частот в другой ФНЧ. Примем

$$z^{-1} = G(Z^{-1}) = \frac{Z^{-1} - \alpha}{1 - \alpha Z^{-1}}. \quad (6.19)$$

Рассмотрим отображение частотной оси (единичной окружности) z -плоскости на Z -плоскость. Перепишем выражение (6.19) в форме, не содержащей отрицательных степеней комплексных переменных:

$$z = \frac{Z - \alpha}{1 - \alpha Z}.$$

Если точка Z находится на 1-окружности, то $Z = e^{j\Omega}$, где Ω – частота, соответствующая Z -плоскости. Тогда z равно

$$\frac{e^{j\Omega} - \alpha}{1 - \alpha e^{j\Omega}} = \frac{\cos \Omega + j \sin \Omega - \alpha}{1 - \alpha \cos \Omega - j \alpha \sin \Omega}. \quad (6.20)$$

Легко показать, что модуль этой дроби равен 1 при любом значении Ω . Это значит, что точка z при таком отображении также находится на 1-окружности. При $Z = \alpha$, очевидно, $z = 0$, а это значит, что внутренность 1-круга отображается внутрь 1-круга. Таким образом, данное отображение удовлетворяет всем необходимым требованиям (сохранение дробно-рациональности очевидно).

Частотная ось при таком отображении претерпевает изменения (деформируется). Произведя подстановки $Z = e^{j\Omega}$ и $z = e^{j\omega}$, полу-

чим $e^{-j\Omega} = \frac{\alpha + e^{-j\omega}}{1 + \alpha e^{-j\omega}}$, отсюда

$$\begin{aligned}\Omega &= -\arg\left(\frac{\alpha + e^{-j\omega}}{1 + \alpha e^{-j\omega}}\right) = -\arg\left(\frac{(\alpha + e^{-j\omega})(1 + \alpha e^{-j\omega})^*}{|1 + \alpha e^{-j\omega}|^2}\right) = \\ &= -\arg\left[(\alpha + e^{-j\omega})(1 + \alpha e^{-j\omega})^*\right] = \operatorname{arctg}\left[\frac{(1 - \alpha^2)\sin\omega}{2\alpha + (1 + \alpha^2)\cos\omega}\right].\end{aligned}$$

Связь частот при преобразовании ФНЧ в ФНЧ иллюстрируется графиками на рис. 6.12. Если $\omega_{\text{гр}}$ – граничная частота (частота среза) для исходного ФНЧ, а $\Omega_{\text{гр}}$ – требуемая частота среза после частотного преобразования, то параметр α преобразования определяется выражением

$$\alpha = \frac{\sin[(\omega_{\text{гр}} - \Omega_{\text{гр}})/2]}{\sin[(\omega_{\text{гр}} + \Omega_{\text{гр}})/2]}.$$

Преобразование ФНЧ в фильтр верхних частот обеспечивается выражением

$$z^{-1} = -\frac{Z^{-1} + \alpha}{1 + \alpha Z^{-1}},$$

при этом параметр α преобразования определяется как

$$\alpha = -\frac{\cos[(\Omega_{\text{гр}} + \omega_{\text{гр}})/2]}{\cos[(\Omega_{\text{гр}} - \omega_{\text{гр}})/2]}.$$

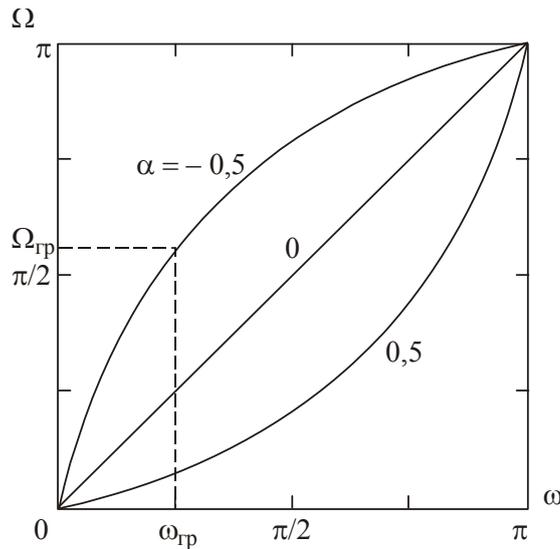


Рис. 6.12. Связь частот при преобразовании ФНЧ в ФНЧ

Преобразование ФНЧ в полосовой (полосно-пропускающий) фильтр

$$z^{-1} = \frac{Z^{-2} - \frac{2\alpha k}{k+1} Z^{-1} + \frac{k-1}{k+1}}{\frac{k-1}{k+1} Z^{-2} - \frac{2\alpha k}{k+1} Z^{-1} + 1}$$

с двумя частотами среза (нижней Ω_1 и верхней Ω_2) характеризуется двумя параметрами, определяемыми выражениями

$$\alpha = -\frac{\cos[(\Omega_2 + \Omega_1)/2]}{\cos[(\Omega_2 - \Omega_1)/2]}, \quad k = \operatorname{ctg}\left(\frac{\Omega_2 - \Omega_1}{2}\right) \operatorname{tg} \frac{\omega_p}{2}.$$

Наконец, режекторный (полосно-заграждающий) фильтр получается из ФНЧ преобразованием

$$z^{-1} = \frac{Z^{-2} - \frac{2\alpha}{1+k} Z^{-1} + \frac{1-k}{1+k}}{\frac{1-k}{1+k} Z^{-2} - \frac{2\alpha}{1+k} Z^{-1} + 1}$$

с параметрами [2]

$$\alpha = \frac{\cos[(\Omega_2 + \Omega_1)/2]}{\cos[(\Omega_2 - \Omega_1)/2]} \quad \text{и} \quad k = \operatorname{tg}\left(\frac{\Omega_2 - \Omega_1}{2}\right) \operatorname{tg} \frac{\omega_p}{2}.$$

Упражнения

1. Докажите, что с точки зрения восстановления аналогового сигнала из АИМ-сигнала форма импульса $d(t)$ не играет роли.
2. Докажите, что при реальной дискретизации (замене аналогового сигнала последовательностью его отсчетов) импульс $d(t)$ должен быть близок к дельта-функции.
3. Докажите, что при нулевой ширине переходной полосы и $1 - \delta_1 > \delta_2$ фильтр нереализуем.
4. Рассчитайте длину секции для реализации метода перекрытия с суммированием, если длина импульсной характеристики равна 47, а основание БПФ составляет 128. Определите требуемое быстродействие вычислителя в умножениях в секунду, учитывая прямое и обратное БПФ, а также поточечное умножение на КЧХ, хранящуюся в памяти (всеми операциями, кроме умножения, следует пренебречь).
5. Сравните количество операций умножения в секунду, полученное в предыдущем упражнении, с аналогичной величиной, рассчитанной для прямой реализации дискретной свертки.
6. Убедитесь, что модуль выражения (6.20) равен 1 при любом значении частоты.

Вопросы для самоконтроля

1. Объясните, почему метод инвариантности импульсной характеристики применяется главным образом для синтеза фильтров нижних частот.
2. Имеет ли значение форма импульса $d(t)$ в выражении (6.6), если он «близок» к дельта-функции? Предложите критерий близости.
3. Чем обеспечивается возможность замены аналоговых сигналов дискретными со спектральной точки зрения?
4. Какую роль играет правильный выбор частоты дискретизации со спектральной точки зрения?
5. Почему перед дискретизацией сигнала с убывающей, но не финитной спектральной плотностью полезно подвергнуть его фильтрации нижних частот?
6. В чем состоит разница между методами синтеза КИХ- и БИХ-фильтров с точки зрения теории аппроксимации?
7. Для чего применяют «окна»?
8. В каких случаях целесообразно применение фильтров на основе частотной выборки?

9. При каких условиях результаты циклической и апериодической свертки совпадают?
10. Как реализовать КИХ-фильтрацию длинных последовательностей через ДПФ?
11. Почему методы аналого-цифровой трансформации применяются для синтеза БИХ-, а не КИХ-фильтров?
12. Назовите преимущества и недостатки методов инвариантности импульсной характеристики и билинейного преобразования.
13. Почему при синтезе ЦФ методом билинейного преобразования отсутствует наложение АЧХ?
14. Неизбежна ли деформация частотной оси при билинейном преобразовании?
15. Что представляют собой частотные преобразования ЦФ и зачем они применяются?



7. РЕАЛИЗАЦИЯ ЦИФРОВЫХ ФИЛЬТРОВ

7.1. РАЗЛИЧИЯ МЕЖДУ ДИСКРЕТНЫМИ И ЦИФРОВЫМИ ЦЕПЯМИ

В предыдущих разделах рассматривались дискретные сигналы и дискретные цепи, т. е. предполагалось, что отсчеты сигналов (последовательностей) принимают значения из непрерывного множества (поля \mathbb{R} вещественных или \mathbb{C} комплексных чисел). То же предполагалось в отношении коэффициентов масштабных звеньев, входящих в состав структурных схем ЛИС-цепей. Такие условия выполняются для дискретных цепей, реализуемых на основе специализированных устройств [9], примером которых могут служить фильтры на ПАВ, упоминавшиеся в разд. 1. Такие цепи, строго говоря, не являются дискретными в полном смысле слова. Реализовать преимущества цифровой обработки сигналов, перечисленные во введении, можно только на основе *цифровых* средств – универсальных программно-управляемых процессоров, а также специализированных устройств, выполненных на цифровой элементной базе и реализующих алгоритмы ЦОС *аппаратным* способом. Промежуточное положение занимают программируемые логические интегральные схемы (ПЛИС), сочетающие черты аппаратного и программного способов реализации (подробнее см. разд. 9).

Цифровые устройства отличаются тем, что сигналы в них неизбежно подвергаются *квантованию* – ограниченная разрядная сетка устройства цифровой обработки приводит к необходимости *округления* чисел. Таким образом, уже на этапе преобразования аналогового сигнала в цифровой происходит его *искажение*. В процессе цифровой обработки сигнал подвергается округлению неоднократно – например, при сложении двух целых чисел фиксированной разрядности результат может иметь разрядность на 1 выше, следовательно, для размещения результата в заданной разрядной сетке

он должен быть сдвинут с потерей значащей цифры (округлен или усечен). Еще более сильное искажение сигнала может произойти при умножении его отсчета на коэффициент масштабного звена – при этом разрядность результата может быть вдвое больше длины разрядной сетки. Так как реализация ЛИС-цепи на практике может потребовать десятков (а для КИХ-фильтров – сотен) умножений и суммирований, то ясно, что их искажающее воздействие на сигнал может быть очень сильным и поэтому заслуживает детального рассмотрения.

Второй важный аспект квантования относится к коэффициентам масштабных звеньев, входящих в состав ЛИС-цепей. При синтезе цифровых фильтров расчеты коэффициентов ведутся с большой точностью, как правило, значительно превышающей точность, достигаемую при реализации. Вследствие неизбежного округления рассчитанных коэффициентов характеристики цифрового фильтра могут существенно отличаться от расчетных. Так, рассчитанный с большой точностью устойчивый фильтр может оказаться неустойчивым, если его полюсы, расположенные близко к единичной окружности, в результате округления коэффициентов переместятся за нее. Кроме того, округление представляет собой нелинейную операцию, вследствие чего состояние устойчивости или неустойчивости цифровой цепи зависит от уровня сигнала в ней: при малых или слишком больших уровнях сигналов цифровые цепи могут становиться неустойчивыми – в них возникают незатухающие колебания, называемые *предельными циклами*.

При рассмотрении преобразования аналогового сигнала в цифровой нельзя не отметить наличия дополнительного источника искажений, связанного с неидеальностью устройства дискретизации (устройства выборки – хранения – УВХ). До сих пор предполагалось, что моменты взятия отсчетов аналогового сигнала определяются согласно теореме отсчетов (Котельникова) и отстоят друг от друга на одну и ту же величину – шаг дискретизации T_d . Любое реальное устройство вследствие влияния внутренних шумов и нестабильностей имеет непостоянные параметры; в данном случае это выражается в случайном изменении величины шага дискретизации при взятии последовательных отсчетов. Хотя эти случайные изменения на практике составляют малую часть от T_d , влияние такого «дрожания отсчетов» на сигнал может быть весьма заметным, и его нельзя не учитывать.

7.2. ДРОЖАНИЕ ОТСЧЕТОВ (ДЖИТТЕР)

Реальная дискретизация сигналов всегда сопровождается так называемым джиттером (дрожанием отсчетов), что влияет на характеристики алгоритмов цифровой обработки сигналов, построенных без учета этого явления. Для ослабления влияния джиттера рекомендуют применять высокостабильные задающие генераторы, управляющие взятием отсчетов в устройствах выборки – хранения. На практике используют кварцевые генераторы с относительной нестабильностью $10^{-8} \dots 10^{-11}$. Однако даже в этих случаях джиттер может оказать значительное влияние на характеристики систем цифровой обработки сигналов.

Рассмотрим обобщенную структурную схему системы дискретной обработки сигналов (рис. 7.1, а). Основными ее элементами являются входной усилитель 1, фильтр нижних частот 2, устройство выборки–хранения 3, генератор синхросигнала 4, а также устройство обработки отсчетов сигнала (УООС) 5, работающее в соответствии с требуемым алгоритмом. Входной усилитель и фильтр образуют линейную часть системы, предназначенную для обеспечения необходимого динамического диапазона

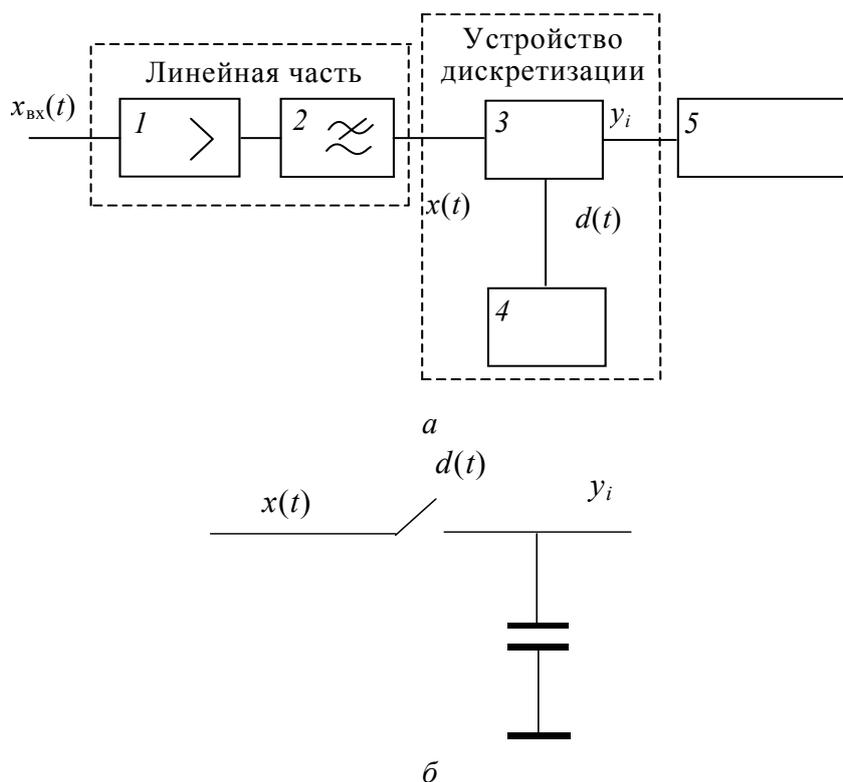


Рис. 7.1. Обобщенная структурная схема устройства дискретизации (а) и упрощенная схема устройства выборки–хранения (б)

и предотвращения наложения «хвостов» спектра при дискретизации входного сигнала, УВХ и генератор синхросигнала вместе образуют устройство дискретизации.

При рассмотрении дрожания отсчетов примем, что УВХ представляет собой идеальный ключ, управляемый генератором синхросигнала, и конденсатор, служащий для запоминания уровня сигнала на время, необходимое для аналого-цифрового преобразования (рис. 7.1, б). Сигнал на входе системы обозначим $x_{\text{вх}}(t)$, на выходе линейной части системы – $x(t)$, значение i -го отсчета на выходе УВХ обозначим y_i , $y_i = x(t_i)$, t_i – момент взятия i -го отсчета. Ключ соединяет конденсатор с выходом фильтра 2 при превышении синхросигналом $d(t)$, поступающим от генератора, порогового уровня h . Синхросигнал представляет собой сумму высокостабильного колебания $S(t)$, вырабатываемого генератором, и стационарного шума $v(t)$, который всегда имеет место в цепях синхронизации. Поэтому момент t_d пересечения порогового уровня h смесью $d(t) = S(t) + v(t)$ смещен относительно момента t_s пересечения этого уровня колебанием $S(t)$ на некоторую случайную величину $\xi = t_d - t_s$ (рис. 7.2).

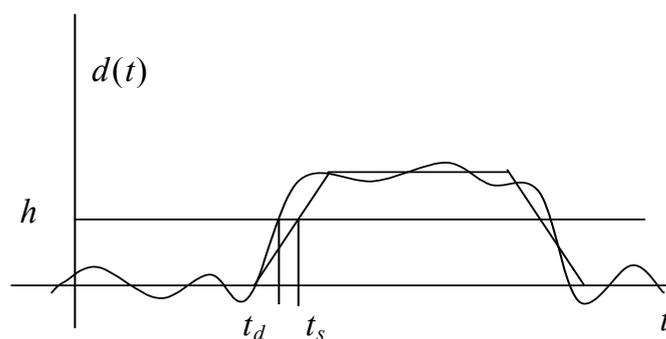


Рис. 7.2. Пересечение синхросигналом порогового уровня

Пусть нестабильность генератора настолько мала, что ее можно не учитывать. Тогда джиттер будет обусловлен только случайной составляющей моментов пересечения синхросигналом уровня h , возникающей из-за наличия шума в цепях синхронизации.

Предположим, что колебание $S(t)$ имеет вид $S(t) = U_s s(t)$, где $s(t)$ – нормированный по амплитуде периодический сигнал, $v(t)$ – дифференцируемый стационарный гауссовский процесс с нулевым средним и дисперсией σ_v^2 ; отношение сигнал/шум в цепях синхро-

низации равно $q_v = U_s / \sigma_v$. Тогда [17] при больших значениях параметра q_v распределение вероятностей случайной величины ξ описывается гауссовским законом с нулевым средним и среднеквадратическим отклонением

$$\sigma_\xi = 1/[q_v |s'(t_s)|],$$

где $s'(t_s)$ – производная нормированного по амплитуде колебания $s(t)$ в момент пересечения им уровня h снизу вверх. В частности, если $h = 0$, $s(t) = \sin(\Omega_d t)$, где $\Omega_d = 2\pi/T_d$ – частота дискретизации, то $\sigma_\xi = 1/(q_v \Omega_d)$. Относительное среднеквадратическое смещение момента дискретизации $\delta_\xi = \sigma_\xi / T_d$ определяется в этом случае выражением $\delta_\xi = 1/(2\pi q_v)$, и при отношении сигнал/шум в цепях синхронизации $q_v = 40$ дБ составляет величину 0,0016, что значительно (на несколько порядков) превосходит относительную нестабильность частоты кварцевого генератора ($10^{-8} \dots 10^{-11}$). Таким образом, дрожание отсчетов даже в системах с высокостабильными генераторами синхросигнала представляет собой заметное явление, которое нельзя игнорировать при расчете характеристик эффективности таких систем. Ниже описана простая модель дрожания отсчетов [18], которая тем не менее позволяет проанализировать существенные элементы искажения сигналов, происходящие вследствие такого дрожания.

Пусть наблюдаемый на выходе линейной части системы процесс $x(t)$ представляет собой сумму полезного сигнала $u(t)$ и стационарного шума $\eta(t)$ с нулевым средним

$$x(t) = u(t) + \eta(t).$$

Полезный сигнал $u(t) = U_m u(t, \lambda)$ может быть квазидетерминированным или случайным. В первом случае U_m – амплитуда сигнала, $u(t, \lambda)$ – нормированный по амплитуде сигнал, $\lambda = \{\lambda_1, \dots, \lambda_M\}$ – вектор параметров сигнала; во втором случае U_m – среднеквадратическое значение сигнала, $u(t, \lambda)$ – реализация нормированного по мощности случайного процесса, $\lambda = \{\lambda_1, \dots, \lambda_M\}$ – вектор параметров процесса. Компоненты вектора параметров λ могут быть частично или полностью неизвестными. Смесь $x(t)$ характеризует-

ся отношением сигнал/шум $q = U_m / \sigma_\eta$, где σ_η – среднеквадратическое отклонение шума.

На выходе устройства дискретизации формируется последовательность из N отсчетов $y_i = x(t_i)$, взятых в моменты времени $t_i, i = \overline{1, N}$; каждый из них представляет собой сумму отсчета $u(t_i) = U_m u(t_i, \lambda)$ полезного сигнала и отсчета $\eta(t_i)$ шумовой составляющей процесса на выходе линейной части системы. При реализации алгоритмов обработки сигналов обычно полагают, что отсчеты берутся в моменты времени $t_{i0} = iT_d$ (T_d – шаг дискретизации), в то время как фактически они берутся в моменты $t_i = t_{i0} + \xi_i, i = \overline{1, N}$, где ξ_i – случайное смещение i -го отсчета, обусловленное его дрожанием.

Положим, что случайные величины $\xi_i, i = \overline{1, N}$, образуют стационарную последовательность с нулевым средним и малым среднеквадратическим отклонением σ_ξ и что функции $u(t)$ и $\eta(t)$ дважды дифференцируемы. Тогда, используя для представления процесса $x(t)$ в окрестности точки t_{i0} ввиду малости σ_ξ три первых члена разложения в ряд Тейлора, для i -го отсчета на выходе устройства дискретизации можно записать выражение

$$\begin{aligned} y_i = x(t_i) &= u(t_{i0} + \xi_i) + \eta(t_{i0} + \xi_i) \approx \\ &\approx u(t_{i0}) + \eta(t_{i0}) + \frac{1}{2} \left[\frac{\partial^2 u(t_{i0})}{\partial t^2} + \frac{\partial^2 \eta(t_{i0})}{\partial t^2} \right] \overline{\xi_i^2} + \left[\frac{\partial u(t_{i0})}{\partial t} + \frac{\partial \eta(t_{i0})}{\partial t} \right] \xi_i + \\ &+ \frac{1}{2} \left[\frac{\partial^2 u(t_{i0})}{\partial t^2} + \frac{\partial^2 \eta(t_{i0})}{\partial t^2} \right] (\xi_i^2 - \overline{\xi_i^2}). \end{aligned} \quad (7.1)$$

В правой части выражения (7.1) первые три слагаемых описывают частотные искажения входного колебания, обусловленные дрожанием отсчетов, остальные два слагаемых – появление добавочного шума.

Частотные искажения входного процесса $x(t)$, возникающие из-за дрожания отсчетов, как видно из выражения (7.1), эквивалентны воздействию на него фильтра с комплексной частотной характеристикой $K(j\Omega) = 1 - \frac{1}{2} \Omega^2 \sigma_\xi^2$, где квадратичный член

возникает вследствие двойного дифференцирования в (7.1), $\Omega = 2\pi f$ – круговая частота в спектральном описании аналогового сигнала; $\sigma_\xi^2 = \overline{\xi_i^2}$ – дисперсия отсчетов случайной последовательности ξ_i , постоянная в силу ее стационарности. Последнее справедливо, если шум $v(t)$ является стационарным, а синхросигнал – строго периодическим. Добавочный шум в отсчетах процесса представлен случайными величинами

$$\begin{aligned} \zeta_i &= \left[\frac{\partial u(t_{i0})}{\partial t} + \frac{\partial \eta(t_{i0})}{\partial t} \right] \xi_i + \frac{1}{2} \left[\frac{\partial^2 u(t_{i0})}{\partial t^2} + \frac{\partial^2 \eta(t_{i0})}{\partial t^2} \right] \left(\xi_i^2 - \overline{\xi_i^2} \right) \approx \\ &\approx \left[\frac{\partial u(t_{i0})}{\partial t} + \frac{\partial \eta(t_{i0})}{\partial t} \right] \xi_i. \end{aligned} \quad (7.2)$$

Очевидно, величины ζ_i и ζ_j при $i \neq j$ являются некоррелированными как между собой, так и с отсчетами шума $\eta(t_i)$. В отсутствие полезного сигнала величины ζ_i , $i = \overline{1, N}$, имеют одинаковое значение дисперсии $\sigma_\zeta^2 \approx \sigma_\eta'^2 \sigma_\xi^2$ (через $\sigma_\eta'^2$ обозначена дисперсия производной $\eta'(t)$ процесса $\eta(t)$). Поэтому их можно рассматривать как отсчеты стационарного шума, имеющего в полосе частот шириной $\Delta F = \frac{1}{2T_d}$ постоянную спектральную плотность мощности

$$N_\zeta = \sigma_\zeta^2 / \Delta F = 2\sigma_\zeta^2 T_d. \quad (7.3)$$

Для случая, когда линейная часть системы имеет П-образную амплитудно-частотную характеристику с граничными частотами F_H и F_B , после подстановки в соотношение (7.3) значения дисперсии σ_ζ^2 , выраженного через параметры входного шума $\eta(t)$ и джиттера ξ_i , получим

$$N_\zeta \approx 8\pi^2 \sigma_\xi^2 T_d \int_{F_H}^{F_B} f^2 N_\eta(f) df, \quad (7.4)$$

где $N_\eta(f)$ – спектральная плотность мощности входного шума.

Появление детерминированного сигнала увеличивает дисперсию добавочного шума, обусловленного дрожанием отсчетов, на величину $\sigma_{ui}^2 = \left(\frac{\partial u(t_{i0})}{\partial t}\right)^2 \sigma_{\xi}^2$, и общая дисперсия i -го отсчета ($i = 1 \dots N$) становится равной

$$\begin{aligned} \sigma_{\zeta_i}^2 &= \left(\frac{\partial u(t_{i0})}{\partial t}\right)^2 \sigma_{\xi}^2 + \sigma_{\eta}^2 \sigma_{\xi}^2 = \\ &= \left[\left(\frac{\partial u(t_{i0})}{\partial t}\right)^2 + 4\pi^2 \int_{F_H}^{F_B} f^2 N_{\eta}(f) df \right] \sigma_{\xi}^2. \end{aligned} \quad (7.5)$$

Как следует из выражения (7.5), наличие детерминированного сигнала на входе приводит к появлению на выходе дополнительной нестационарной составляющей шума, дисперсия которой зависит от производной сигнала в момент взятия отсчетов.

При дважды дифференцируемом стационарном случайном процессе $u(t)$ общая дисперсия i -го отсчета определяется выражением

$$\sigma_{\zeta_i}^2 = 4\pi^2 \sigma_{\xi}^2 \int_{F_H}^{F_B} f^2 [N_{\eta}(f) + N_u(f)] df,$$

где $N_u(f)$ – спектральная плотность мощности сигнала на выходе линейной части системы.

На основе сказанного можно представить устройство дискретизации эквивалентной схемой (рис. 7.3), учитывающей эффект джиттера.

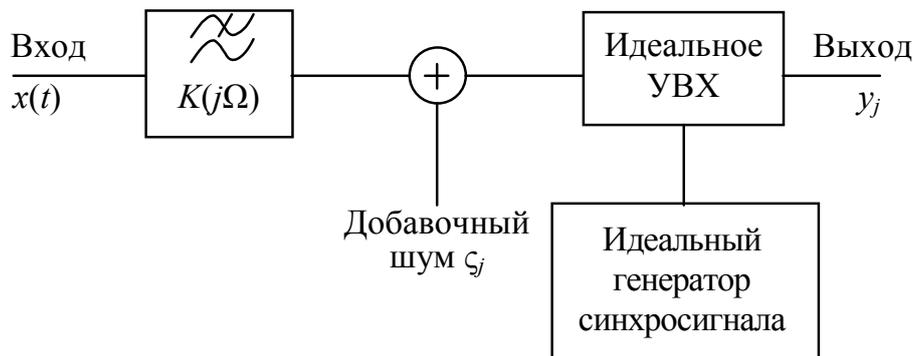


Рис. 7.3. Эквивалентная схема устройства дискретизации с учетом эффекта джиттера

Для характеристики отсчетов наблюдаемого процесса будем использовать отношение сигнал/шум $q_{\Sigma}^2 = \sigma_{\Sigma_c}^2 / \sigma_{\Sigma_{ш}}^2$, где при слу-

чайном входном сигнале $\sigma_{\Sigma_c}^2 = \int_{F_H}^{F_B} N_u(f) |K(j2\pi f)|^2 df$ – дисперсия

последовательности отсчетов сигнальной составляющей процесса $u(t)$ на выходе устройства дискретизации ($N_u(f)$ – спектральная плотность мощности сигнала на выходе линейной части системы, множитель $|K(j2\pi f)|^2$ учитывает эффект *подавления его высокочастотных составляющих* вследствие дрожания отсчетов);

$\sigma_{\Sigma_{ш}}^2 = \int_{F_H}^{F_B} N_{\eta}(f) |K(j2\pi f)|^2 df + \sigma_{\zeta_i}^2$ – дисперсия последовательности отсчетов шумовой составляющей процесса на выходе устрой-

ства дискретизации.

При детерминированном периодическом сигнале $\sigma_{\Sigma_c}^2$ представляет собой среднюю мощность сигнала, а $N_u(f)$ – его спектральную плотность мощности [19], состоящую из взвешенных δ -функций, соответствующих гармоническим составляющим.

Учитывая, что $K(j2\pi f) = 1 - \frac{1}{2}(2\pi f)^2 \sigma_{\xi}^2$, и пренебрегая величинами второго порядка малости, получим

$$q_{\Sigma}^2 \approx \frac{\int_{F_H}^{F_B} N_u(f) df - 4\pi^2 \sigma_{\xi}^2 \int_{F_H}^{F_B} f^2 N_u(f) df}{\int_{F_H}^{F_B} N_{\eta}(f) df + 4\pi^2 \sigma_{\xi}^2 \int_{F_H}^{F_B} f^2 N_u(f) df} =$$

$$= \frac{q^2 - \left(4\pi^2 \sigma_{\xi}^2 \int_{F_H}^{F_B} f^2 N_u(f) df \right) / \left(\int_{F_H}^{F_B} N_{\eta}(f) df \right)}{1 + \left(4\pi^2 \sigma_{\xi}^2 \int_{F_H}^{F_B} f^2 N_u(f) df \right) / \left(\int_{F_H}^{F_B} N_{\eta}(f) df \right)},$$

где q^2 – отношение сигнал/шум на выходе линейной части системы. Видно, что с ростом дисперсии джиттера σ_{ξ}^2 отношение сигнал/шум на выходе устройства дискретизации уменьшается.

Рассмотрим ситуацию, когда и полезный сигнал, и сигнал в цепях синхронизации синусоидальные. Можно показать, что в этом случае

$$q_{\Sigma}^2 = q^2 \left(1 - \frac{T_d^2 f_0^2}{q_v^2} \right) / \left(1 + \frac{q^2 T_d^2 f_0^2}{q_v^2} \right),$$

где f_0 – частота синхросигнала.

Представляет интерес анализ поведения результирующего отношения сигнал/шум при различных соотношениях q^2 и q_v^2 . Видно, что при малом джиттере ($q_v^2 \rightarrow \infty$) результирующее отношение сигнал/шум стремится к величине $q_{\Sigma}^2 = q^2 / \left(1 + \frac{q^2 T_d^2 f_0^2}{q_v^2} \right)$. Можно

выделить два частных случая. При относительно слабом входном сигнале ($q \ll q_v / T_d f_0$) результирующее отношение сигнал/шум определяется входным отношением $q_{\Sigma}^2 \rightarrow q^2$. При большом отношении сигнал/шум на входе ($q \gg q_v / T_d f_0$) результирующее отношение сигнал/шум полностью определяется отношением сигнал/шум в цепях синхронизации и стремится к величине $q_{\Sigma}^2 = q_v^2 / (T_d^2 f_0^2)$. Таким образом, нельзя коренным образом решить проблему джиттера, повышая отношение сигнал/шум только на входе или только в цепях синхронизации, так как, начиная с некоторого момента, результирующее отношение сигнал/шум будет определяться *худшим* из двух указанных отношений. Кроме того, отношение сигнал/шум на выходе устройства дискретизации падает с ростом частоты сигнала вследствие подавления джиттером его высокочастотных составляющих.

7.3. ПРЕДСТАВЛЕНИЕ ЧИСЕЛ В ЦИФРОВЫХ УСТРОЙСТВАХ

В устройствах, выполняющих цифровую обработку сигналов, используются два способа представления чисел, которые называются форматами с фиксированной и с плавающей запятой^{*)}. В формате с *фиксированной запятой* вес (значимость) каждого двоичного разряда считается фиксированным. Например, можно приписать младшему (нулевому) разряду вес $2^0 = 1$, следующему – вес $2^1 = 2$, затем – вес $2^2 = 4$ и т. д. Таким образом получается представление целых чисел. Например, в устройстве с 16 разрядами можно однозначно представить 2^{16} различных чисел от 0 до 65535. Поскольку на практике требуется представлять не только положительные, но и отрицательные числа, один разряд отводят для представления знака: если в этом разряде содержится единица, то число считается отрицательным. Код, в котором отрицательное число образуется из положительного числа с таким же модулем инверсией знакового разряда, называется *прямым* кодом. Например, десятичное число 37 (обозначается 37_{10}) в двоичном прямом коде имеет вид 0000000000100101, а числу -37_{10} соответствует прямой код 1000000000100101. Очевидно, арифметическое устройство, выполняющее операции с числами в прямом коде, должно анализировать знаковые разряды чисел-операндов и в зависимости от их значений выполнять различные действия с остальными разрядами (сложение или вычитание).

Для того чтобы арифметическое устройство могло оперировать с положительными и отрицательными числами одинаковым образом, применяют так называемый *дополнительный код*.

Положительные числа в дополнительном коде представляются обычным способом, при этом старший разряд содержит 0. Например, 16-разрядное двоичное число в дополнительном коде 0000000000010011 соответствует десятичному числу +19. Отрицательные числа представляются в дополнительном коде в соответствии со следующим правилом. Если старший разряд числа содержит 1, число интерпретируется как отрицательное; при этом модуль числа можно найти путем инвертирования всех двоичных разрядов и последующего прибавления к младшему разряду единицы. Так, например, 16-разрядное двоичное число в дополнительном коде 1000110010010011 отрицательное, его модуль находится как

^{*)} В иностранной литературе вместо термина «запятая» употребляется слово «точка»

$0111001101101100+1 = 0111001101101101$, что соответствует десятичному числу 29549, таким образом, исходное число равно -29549 .

Кроме дополнительного кода, используется *обратный* код. Положительные числа в обратном коде представляются так же, как в прямом и дополнительном, а представление отрицательного числа получается вычитанием прямого двоичного кода его модуля из числа, представленного кодом с единицами во всех разрядах, включая знаковый. Например, положительное число 7_{10} имеет обратный код 000000000000111, а отрицательное число -7_{10} представляется обратным кодом $111111111111111-000000000000111=1111111111111000$.

Во многих случаях удобно считать, что все числа являются дробными и не превышают по модулю 1. Тогда двоичная запятая разделяет старший знаковый разряд и все остальные разряды, наибольший из которых имеет вес 2^{-1} , следующий за ним имеет вес 2^{-2} и т.д. При этом наибольшее положительное 16-разрядное число равно $0,111111111111111$, а наименьшее отрицательное – равно $1,000000000000001$. Если реальные сигналы выходят за диапазон представления, их приходится масштабировать (умножать на соответствующий коэффициент).

Формат чисел с *плавающей запятой* предусматривает представление числа в виде двух частей – *мантиссы* M и *порядка* p , причем и мантисса, и порядок могут быть положительными или отрицательными. Значение числа определяется выражением $M \cdot 2^p$. Мантисса приводится к форме дробного числа, находящегося в интервале $1/2 \leq M < 1$ (*нормализуется*). При перемножении двух чисел с плавающей запятой мантиссы операндов перемножаются, а порядки складываются. Если при этом мантисса результата выходит из указанного интервала, то производится ее нормализация с соответствующей коррекцией порядка. При сложении двух чисел с плавающей запятой вначале выравниваются порядки со сдвигом мантиссы меньшего числа вправо (при этом младшие значащие цифры теряются), а затем мантиссы складываются; при необходимости после сложения производится нормализация.

Очевидно, при сложении или умножении чисел в формате с плавающей запятой мантисса результата может иметь большую разрядность, чем мантиссы операндов. При этом происходит *округление* мантиссы или ее *усечение* (отбрасывание младших разрядов). Порядок результата также может иметь значение, выходящее за пределы разрешенного интервала (тогда имеет место *переполнение*). Округление (усечение) результатов приводит к постепенному

накоплению ошибки, в то время как переполнение означает аварийную ситуацию: ошибка при этом составляет значительную величину, соизмеримую с шириной динамического диапазона устройства, поэтому необходимо принимать меры, предотвращающие переполнение.

В формате с фиксированной запятой для дробных чисел переполнение может произойти при сложении, если сумма превысит 1. При умножении таких чисел результат всегда будет меньше любого из операндов, поэтому при многократном умножении результат становится все меньше, при этом относительная погрешность результата, определяемая отношением значимости младшего разряда к результату, возрастает. Относительная погрешность представления чисел в формате с плавающей запятой остается постоянной, так как определяется количеством значащих разрядов мантиссы.

Из сказанного ясно, что формат с плавающей запятой обладает большими возможностями, однако при его использовании требуется больший объем памяти и возрастает сложность арифметического устройства, которое должно выполнять обработку мантисс и порядков операндов по различным алгоритмам, а также обеспечивать выравнивание порядков и нормализацию мантисс.

7.4. ЭФФЕКТЫ КВАНТОВАНИЯ

Квантованием называется процесс получения числового значения заданной разрядности из неквантованного вещественного числа. Вещественные числа в общем случае представляются бесконечными десятичными или двоичными дробями, поэтому квантование сопровождается появлением ошибки (погрешности). В устройствах, реализующих цифровую обработку сигналов, квантование входного сигнала происходит в аналого-цифровом преобразователе (АЦП). Кроме того, округление или усечение промежуточных результатов, неизбежно сопровождающее преобразования сигнала, также является квантованием. Квантованию подвергаются и коэффициенты масштабных звеньев, входящих в состав структурной схемы цепи.

7.4.1. КВАНТОВАНИЕ СИГНАЛОВ

Рассмотрим эффекты квантования сигналов в устройстве, представляющем дробные числа в формате с фиксированной запятой. Отбрасывание «хвоста» бесконечной двоичной дроби, представляющей положительное число, при усечении приводит к погреш-

ности, максимальное значение которой равно весу младшего из удерживаемых разрядов. Например, в 16-разрядном устройстве этот вес составляет 2^{-15} , следовательно, ошибка усечения не может превысить $2^{-15} \approx 0,00003$. При этом ошибка всегда имеет один и тот же знак (усеченное значение не может быть больше исходного). При округлении ошибка может быть как положительной, так и отрицательной, при этом ее модуль не превосходит веса старшего из отброшенных разрядов. В 16-разрядном устройстве этот вес равен 2^{-16} , следовательно, ошибка округления не превосходит по модулю значения $2^{-16} \approx 0,000015$.

Знак ошибки, возникающей при усечении отрицательных чисел, зависит от кода, в котором представлены эти числа. При усечении отрицательного числа, представленного в прямом коде, оно увеличивается (его модуль уменьшается). Усечение числа в дополнительном коде приводит к его уменьшению. Отрицательные числа в обратном коде при усечении увеличиваются.

Округление отрицательных чисел, как и округление положительных чисел, приводит к погрешности, которая может быть положительной или отрицательной и не превосходит по модулю веса старшего из отброшенных разрядов.

При усечении или округлении чисел, представленных в формате с плавающей запятой, изменяется только мантисса. Поэтому абсолютная погрешность округления (усечения) может быть определена с учетом порядка числа. Обозначая через $Q(x)$ результат округления величины x , можно записать для формата с плавающей запятой

$$-2^p \cdot 2^{-(b+1)} < Q(x) - x \leq 2^p \cdot 2^{-(b+1)},$$

где b – номер младшего из удерживаемых разрядов.

Квантование сигнала – принципиально *нелинейная* операция. Ее можно представить как результат воздействия на сигнал нелинейного безынерционного звена со ступенчатой характеристикой, изображенной на рис. 7.4. Если сигнал $x[n]$ представляет собой быстро ме-

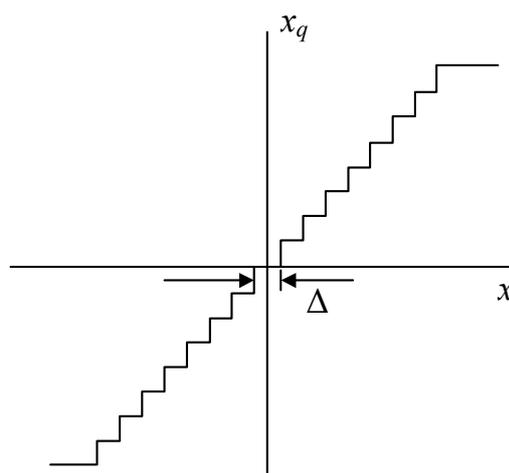


Рис. 7.4. Характеристика квантователя

няющийся (широкополосный) случайный процесс, то можно считать, что квантование эквивалентно сложению дискретного сигнала с шумом, который принято называть *шумом квантования*. Тогда можно нелинейное преобразование сигнала в устройстве квантования представить эквивалентной линейной цепью, изображенной на рис. 7.5.

Если шаг квантования Δ достаточно мал, распределение шума квантования можно считать равномерным с плотностью распределения вероятности, представленной на рис. 7.6. При этом обычно считается, что:

- $\eta[n]$ – стационарная случайная последовательность;
- $\eta[n]$ и $x[n]$ некоррелированы;
- $\eta[n]$ – белый шум.

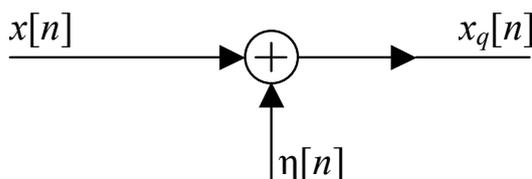


Рис. 7.5. Эквивалентная линейная схема квантователя

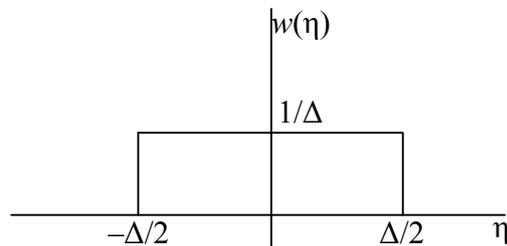


Рис. 7.6. Плотность распределения вероятности шума квантования

Эти предположения тем более обоснованны, чем ближе входной сигнал по своим характеристикам к белому шуму. Такая модель, основанная на замене нелинейного эффекта квантования аддитивным действием шума, называется *статистической моделью* квантования. Качество квантования характеризуют отношением сигнал/шум квантования: чем выше это отношение, тем ближе квантованный (цифровой) сигнал к неквантованному (дискретному). В большинстве практических случаев входное колебание уже представляет собой смесь сигнала с шумом. Таким образом, квантование сигнала приводит к усилению шума, т. е. понижает общее отношение сигнал/шум.

Другой нелинейный эффект, имеющий место в устройствах цифровой обработки сигналов, связан с тем, что значения входной последовательности могут превосходить предельно допустимые для данной цепи (выходить за границы динамического диапазона). При этом происходит нелинейное искажение сигнала типа ограничения. Чтобы избежать подобных искажений, приходится прини-

мать специальные меры, а именно ослаблять сигнал (например, при помощи системы *автоматической регулировки усиления* – АРУ). При этом происходит уменьшение мощности сигнала, а мощность шума квантования остается неизменной, следовательно, уменьшается итоговое отношение сигнал/шум. Очевидно, необходимо уметь правильно выбирать разрядность устройства, чтобы обеспечивать отношение сигнал/шум не ниже заданного.

Рассмотрим шум квантования, считая, что значения сигнала представляются в цифровом устройстве дробными числами с фиксированной запятой, причем устройство имеет b значащих разрядов, т. е. наименьшее значение, представимое в устройстве, равно $\Delta = 2^{-b}$. Тогда мощность шума квантования равна дисперсии случайной величины с распределением вероятности, равномерным в интервале от $-\Delta/2$ до $\Delta/2$:

$$D_{\eta} = \int_{-\Delta/2}^{\Delta/2} \eta^2 \frac{1}{\Delta} d\eta = \frac{\Delta^2}{12} = \frac{2^{-2b}}{12}.$$

Если полезный сигнал представляет собой стационарный случайный процесс (во многих практически важных случаях это предположение оправдано) и дисперсия сигнала равна $D_x = \sigma_x^2$, то отношение сигнал/шум («по мощности») равно

$$q = \frac{\sigma_x^2}{\sigma_{\eta}^2} = \frac{\sigma_x^2}{2^{-2b}} \cdot 12 = 12 \cdot 2^{2b} \sigma_x^2.$$

В логарифмическом представлении (в децибелах) отношение сигнал/шум равно

$$\begin{aligned} q_{\text{дБ}} &= 10 \cdot \lg \frac{\sigma_x^2}{\sigma_{\eta}^2} = 10 \cdot \lg(12) + 2b \cdot 10 \cdot \lg(2) + 10 \cdot \lg(\sigma_x^2) = \\ &= 6,02b + 10,79 + 10 \cdot \lg(\sigma_x^2). \end{aligned}$$

Таким образом, при увеличении разрядности устройства на единицу отношение сигнал/шум возрастает приблизительно на 6 дБ (в 4 раза по мощности).

Умножение сигнала на масштабирующий коэффициент $A < 1$ для согласования динамического диапазона сигнала с динамическим диапазоном устройства приводит к уменьшению мощности

сигнала, поэтому отношение сигнал/шум уменьшается и становится равным

$$q_{\text{дБ}} = 10 \cdot \lg \frac{A^2 \sigma_x^2}{\sigma_\eta^2} = 6,02b + 10,79 + 10 \cdot \lg(\sigma_x^2) + 20 \cdot \lg A.$$

Считая, что входной сигнал – гауссовский стационарный случайный процесс с нулевым средним и дисперсией σ_x^2 , и полагая, что масштабный коэффициент должен обеспечить совпадение интервала « $\pm 3\sigma_x$ » и динамического диапазона устройства^{*)}, примем $A^2 = (9 \cdot \sigma_x^2)^{-1}$, тогда отношение сигнал/шум

$$q_{\text{дБ}} = 10 \cdot \lg \left(\frac{4}{3} \cdot 2^{2b} \right) = 6,02b - 10 \cdot \lg(4/3).$$

Для упрощения можно выбрать ширину динамического диапазона $\pm 3,3\sigma_x$ (с «запасом»), тогда отношение сигнала к шуму квантования составит $\approx 6b$ дБ. Например, 16-разрядному устройству с фиксированной запятой (один разряд является знаковым) при этих условиях соответствует отношение сигнал/шум квантования около 90 дБ, что примерно соответствует стандарту цифровой звукозаписи.

Мы рассмотрели вопрос о выборе разрядности для случая *однократного* квантования, например, при аналого-цифровом преобразовании. При реализации цифровой обработки сигналов квантование всегда осуществляется *многократно*. Поясним сказанное на примере простейшей рекурсивной цепи, описываемой разностным уравнением

$$y[n] = ay[n-1] + x[n].$$

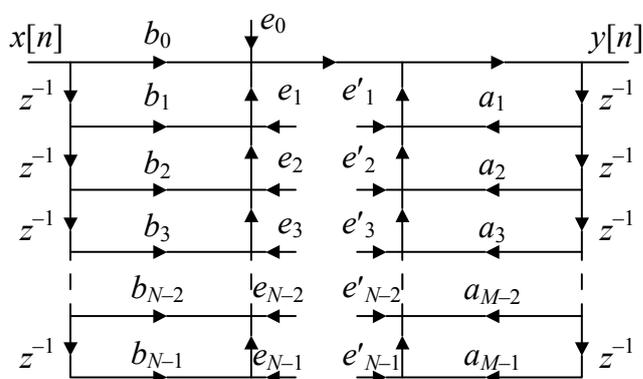
Реализация такого разностного уравнения предполагает выполнение двух операций для вычисления каждого отсчета выходной последовательности: умножения предыдущего отсчета на коэффициент a фильтра и сложения результата с очередным отсчетом входного сигнала. Если предположить для определенности, что отсчеты последовательностей и коэффициент фильтра представлены в одинаковом формате нормализованных дробных чисел с фиксированной запятой, то, очевидно, после каждой из указанных

^{*)} Напомним, что гауссовская случайная величина принадлежит «трехсигмовому интервалу» с вероятностью 0,997.

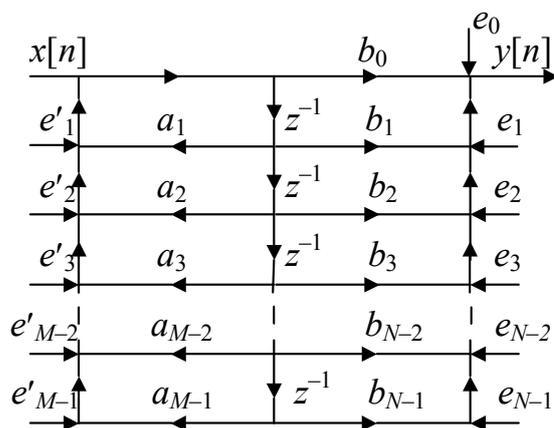
операций требуется нормализация, которая в общем случае сводится к квантованию путем усечения либо округления. Предполагая, что выполняются условия для применения статистической модели квантования, видим, что даже в простейшей рекурсивной цепи существуют две точки, в которых к сигналу добавляется шум квантования.

В общем случае квантование промежуточных результатов (усечение или округление) происходит после выполнения *каждой* из таких операций, как умножение сигнала на масштабный коэффициент или сложение, поэтому для учета влияния эффектов квантования при помощи статистической модели эквивалентные источники шумов квантования должны быть введены в структурную схему цифровой цепи во всех соответствующих точках.

На рис. 7.7 представлен граф ЛИС-цепи, соответствующий прямой форме реализации. Шумы квантования промежуточных



а



б

Рис. 7.7. Эквивалентные источники шумов квантования промежуточных результатов для прямой (а) и канонической (б) форм реализации ЛИС-цепи

результатов обозначены на графе эквивалентными источниками e_0, e_1, \dots, e_N в трансверсальной части цепи и источниками e'_1, e'_2, \dots, e'_M в рекурсивной части. Учитывая, что задержки не изменяют разрядности представления чисел и не сопровождаются округлением или усечением, легко видеть, что действие всех источников эквивалентно действию одного суммарного источника, подключенного к выходу трансверсальной части и ко входу рекурсивной части цепи. Полагая все источники шумов квантования независимыми, можно определить дисперсию шума суммарного источника как сумму дисперсий отдельных источников шумов квантования промежуточных результатов. Характеристики шума квантования могут быть пересчитаны к выходу цепи с учетом амплитудно-частотной характеристики ее рекурсивной части [20].

На рис. 7.8 представлен граф, соответствующий канонической форме реализации ЛИС-цепи. Нетрудно видеть, что действие всех источников e_0, e_1, \dots, e_{N-1} эквивалентно действию одного суммарного источника, подключенного к выходу цепи. Действие всех источников $e'_1, e'_2, \dots, e'_{M-1}$ эквивалентно действию одного суммарного источника, подключенного ко входу цепи, поэтому шум квантования рекурсивной части пересчитывается к выходу с учетом АЧХ всей цепи, где суммируется с шумом квантования трансверсальной части. Таким образом, суммарное действие шумов квантования, сопровождающих реализацию в цифровой форме заданной передаточной функции дискретной цепи, сильно зависит от формы реализации. В большинстве случаев наиболее выгодна с этой точки зрения каскадная форма [20, 21].

7.4.2. КВАНТОВАНИЕ КОЭФФИЦИЕНТОВ ЦИФРОВОГО ФИЛЬТРА

Квантование (округление) коэффициентов цифрового фильтра, неизбежное при его схемной или программной реализации в устройстве ограниченной разрядности, приводит к тому, что реализованным оказывается «не совсем тот» фильтр, который был рассчитан в процессе синтеза. В результате такие определяющие характеристики, как АЧХ, импульсная характеристика, ФЧХ, могут заметно отличаться от ожидаемых. Более того, цифровой фильтр, соответствующий рассчитанному дискретному фильтру, может даже оказаться неустойчивым. Поэтому необходимо «прочитывать» характеристики фильтра с учетом квантования коэф-

фициентов. Анализ эффектов округления коэффициентов в общем виде весьма сложен. В литературе [3] отмечаются два общих подхода к анализу эффектов квантования коэффициентов цифровых фильтров.

Первый подход основан на представлении квантованного коэффициента как суммы коэффициента, определенного с бесконечной точностью, и *случайной* погрешности. Пусть цифровой фильтр с квантованными коэффициентами имеет передаточную функцию

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^{N-1} b_k z^{-k}}{1 - \sum_{r=1}^{M-1} a_r z^{-r}}.$$

Коэффициенты фильтра представляются в виде сумм

$$a_k = \bar{a}_k + \alpha_k, \quad b_k = \bar{b}_k + \beta_k,$$

где \bar{a}_k, \bar{b}_k – точные значения коэффициентов, а α_k, β_k – погрешности квантования, которые считают независимыми случайными величинами, имеющими одинаковую равномерную плотность распределения вероятностей. В [3] показано, что фильтр с квантованными коэффициентами можно представить параллельным соединением «точного» фильтра с передаточной функцией $H_\infty(z)$ и «паразитного» фильтра с передаточной функцией

$$\frac{\beta(z) + \alpha(z)H_\infty(z)}{A_\infty(z)},$$

где $\alpha(z) = \sum_{r=1}^{M-1} \alpha_r z^{-r}$; $\beta(z) = \sum_{k=1}^{N-1} \beta_k z^{-k}$; $A_\infty(z) = 1 - \sum_{r=1}^{M-1} \bar{a}_r z^{-r}$. Пе-

редаточная функция паразитного фильтра является при вышеуказанных предположениях случайной функцией; квадрат ее модуля после интегрирования по единичной окружности и статистического усреднения служит характеристикой качества цифрового фильтра с учетом квантования коэффициентов.

Второй подход [3] основан на оптимальном подборе *квантованных* коэффициентов для каждого фильтра в отдельности, при этом целевая функция, подлежащая минимизации, представляет собой взвешенное среднеквадратическое отклонение получаемой

АЧХ от желаемой. По существу, этот подход представляет собой прямой машинный синтез *цифрового* фильтра, при этом *дискретный* фильтр служит начальным приближением.

Округление различных коэффициентов фильтра в неодинаковой степени сказывается на его характеристиках. Степень влияния коэффициента масштабного звена на расположение полюсов фильтра, которое в наибольшей степени определяет его характеристики, оценивают коэффициентами или *функциями влияния* [20], представляющими собой частные производные полюсов цепи по соответствующим коэффициентам. Рассмотрим определение коэффициентов влияния на примере рекурсивного фильтра второго порядка с передаточной функцией

$$H(z) = \frac{1}{(1 - a_1 z^{-1})(1 - a_2 z^{-1})},$$

соответствующей каскадной форме реализации. Этой передаточной функции соответствуют два полюса $p_1 = a_1$ и $p_2 = a_2$. В данном случае коэффициенты фильтра совпадают с его полюсами, вследствие чего коэффициенты влияния равны $\frac{\partial p_1}{\partial a_1} = \frac{\partial p_2}{\partial a_2} = 1$ и

$\frac{\partial p_2}{\partial a_1} = \frac{\partial p_1}{\partial a_2} = 0$. Та же дискретная цепь может быть реализована в прямой форме с передаточной функцией

$$H(z) = \frac{1}{(1 - a_1 z^{-1})(1 - a_2 z^{-1})} = \frac{1}{1 - c_1 z^{-1} - c_2 z^{-2}},$$

где $c_1 = p_1 + p_2$, $c_2 = -p_1 p_2$. Для определения коэффициентов влияния рассмотрим преобразование вектора коэффициентов в вектор полюсов, которому соответствует дифференциальное выражение

$$\begin{pmatrix} dc_1 \\ dc_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial c_1}{\partial p_1} & \frac{\partial c_1}{\partial p_2} \\ \frac{\partial c_2}{\partial p_1} & \frac{\partial c_2}{\partial p_2} \end{pmatrix} \begin{pmatrix} dp_1 \\ dp_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -p_2 & -p_1 \end{pmatrix} \begin{pmatrix} dp_1 \\ dp_2 \end{pmatrix}.$$

Решая это уравнение относительно dp_1 и dp_2 , получаем

$$\begin{pmatrix} dp_1 \\ dp_2 \end{pmatrix} = \begin{pmatrix} \frac{-p_1}{p_2 - p_1} & \frac{-1}{p_2 - p_1} \\ \frac{p_2}{p_2 - p_1} & \frac{1}{p_2 - p_1} \end{pmatrix} \begin{pmatrix} dc_1 \\ dc_2 \end{pmatrix},$$

где матрица состоит из искомых коэффициентов влияния $\frac{\partial p_1}{\partial c_1} = \frac{p_1}{p_1 - p_2}$, $\frac{\partial p_2}{\partial c_1} = \frac{p_2}{p_2 - p_1}$, $\frac{\partial p_1}{\partial c_2} = \frac{1}{p_1 - p_2}$ и $\frac{\partial p_2}{\partial c_2} = \frac{1}{p_2 - p_1}$. Из полученных выражений видно, что чувствительность полюсов к изменению коэффициентов вследствие округления зависит, вообще говоря, от расположения полюсов конкретной цепи.

Подводя итог, можно сказать, что с точки зрения минимального влияния квантования коэффициентов на расположение полюсов для каждой конкретной цепи существует оптимальная форма реализации, однако общий метод ее выбора в настоящее время не известен. Считается [20], что прямая форма в этом смысле уступает другим формам, в частности каскадной.

7.4.3. ПРЕДЕЛЬНЫЕ ЦИКЛЫ

Пределные циклы представляют собой нелинейный эффект, характерный для цифровых фильтров. Напомним, что операция квантования (округления, усечения) является нелинейной и описывается ступенчатой функцией. Кроме того, вследствие ограниченности разрядной сетки характеристика квантователя имеет горизонтальные участки, соответствующие максимальному и минимальному значениям, представимым в устройстве заданной разрядности. Таким образом, в отличие от дискретных цепей, которые могут быть линейными, цифровые цепи *нелинейны всегда*. Модель линейной цепи, в частности ЛИС-цепи, является, таким образом, идеализацией, допустимой лишь в определенных условиях, например, если уровень сигнала много больше интервала квантования и в то же время много меньше ширины динамического диапазона квантователя. Пределные циклы представляют собой своеобразное нарушение устойчивости цифровых фильтров, причиной которого является принципиально нелинейный характер преобразований сигналов в цифровых фильтрах.

Различают предельные циклы низкого уровня, называемые также эффектом мертвой зоны [2, 3], и предельные циклы переполнения [20].

Предельные циклы *низкого* уровня возникают в результате округления промежуточных результатов вычислений в рекурсивных цепях [2, 3, 20]. Проявляются предельные циклы в форме незатухающих колебаний малой амплитуды, которые имеют место на выходе цифрового фильтра в отсутствие входного сигнала. Во многих практически важных случаях применения цифровой обработки сигналов существование предельных циклов крайне нежелательно. Например, в случае цифровой фильтрации речевого сигнала в системе связи колебания предельных циклов могут проявляться в форме свиста или гудения в паузах между словами и фразами. Точный анализ этого явления в общем случае крайне сложен, поэтому здесь будет рассмотрен лишь простой пример.

Рассмотрим простейшую рекурсивную дискретную цепь, описываемую разностным уравнением

$$y[n] = 0,9 \cdot y[n-1] + x[n].$$

Значения отклика такой цепи на δ -последовательность – импульсной характеристики – для первых десяти неотрицательных значений дискретного времени приведены в таблице (столбец 2). Там же представлены соответствующие значения $y^o[n]$ отклика цифровой цепи, в которой результат умножения округляется с точностью до десятых долей; такая цепь описывается разностным уравнением

$$y^o[n] = \left\langle 0,9 \cdot y^o[n-1] \right\rangle_{0,1} + x[n],$$

где символ $\langle \cdot \rangle_{0,1}$ означает округление до десятых. Разумеется, в реальных цепях округление не бывает таким грубым, кроме того, округляются не десятичные, а двоичные числа, тем не менее причины возникновения предельных циклов проявляются в этом примере достаточно наглядно.

Видно, что в отличие от дискретной цепи в цепи с округлением отклик не стремится к нулю с ростом n , а стабилизируется на уровне 0,5. Объясняется это тем, что нелинейность операции округления приводит к смещению полюса при малых уровнях сигнала в цепи, причем полюс оказывается на единичной окружности (в точке $z = 1$).

**Отклики на δ -последовательность рекурсивных цепей без округления
и с округлением промежуточных результатов**

n	$y[n]$	$y^o[n]$
0	1	1
1	0,9	0,9
2	0,81	0,8
3	0,729	0,7
4	0,6561	0,6
5	0,59049	0,5
6	0,531441	0,5
7	0,4782969	0,5
8	0,43046721	0,5
9	0,387420489	0,5

Если коэффициент разностного уравнения будет не 0,9, а $-0,9$, то отклик становится знакопеременным, т. е. принимает значения 0,5 и $-0,5$ поочередно; полюс при этом оказывается также на единичной окружности (в точке $z = -1$).

С увеличением порядка цепи форма незатухающих колебаний становится более сложной; при некоторых сочетаниях коэффициентов цепи и входных воздействий предельные циклы могут отсутствовать. При повышении разрядности устройства амплитуда колебаний предельных циклов уменьшается. Амплитуда и форма колебаний предельных циклов зависит также от формы реализации цепи (прямой, канонической, каскадной, параллельной и т.д.) Напомним, что анализ предельных циклов в цифровых цепях общего вида крайне затруднен; примеры такого анализа для цепей первого и второго порядков приведены в [2].

Помимо предельных циклов низкого уровня в цифровых цепях могут возникать незатухающие колебания большой амплитуды, вызванные нелинейностью характеристики квантования типа ограничения (насыщения), связанной с переполнением разрядной сетки. Такие колебания представляют собой периодические последовательности, изменяющиеся между предельными отрицательными и положительными значениями, и называются предельными циклами переполнения. Следует отметить, что в некоторых случаях предельные циклы могут быть использованы для генерирования незатухающих периодических колебаний.

7.5. АППАРАТУРНАЯ И ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ЦИФРОВЫХ ФИЛЬТРОВ

Аппаратурной называется реализация цифровой обработки сигналов в специализированном устройстве, построенном на основе соответствующей элементной базы – регистров, триггеров, счетчиков, цифровых сумматоров, умножителей, логических вентилей и т.п. Эти устройства соединяются в соответствии со структурной схемой цифровой цепи, и способ соединения в дальнейшем не изменяется. Такой способ реализации называют «жесткой логикой». Одна и та же дискретная цепь может быть реализована в цифровом устройстве многими способами. При выборе способа реализации учитываются требования к быстродействию, объему памяти, массогабаритным характеристикам и стоимости устройства, надежности и т.п.

Чаще всего аппаратурный способ реализации применяется в тех приложениях цифровой обработки сигналов, где требуется обеспечить обработку в реальном времени – в темпе поступления отсчетов входного сигнала. При этом все операции, необходимые для вычисления одного отсчета выходной последовательности, должны выполняться за время, равное шагу дискретизации аналогового сигнала T_d или интервалу между моментами поступления соседних отсчетов сигнала, являющегося изначально дискретным.

Для определения необходимого быстродействия элементов схемы необходимо исследовать структуру цепи и учесть все операции, которые должны быть выполнены за отведенное время. При этом более предпочтительными являются те способы реализации, которые допускают одновременное (*параллельное*) выполнение разных операций в различных частях схемы. Кроме того, применяется *конвейерный* принцип обработки, предполагающий разбиение операции на несколько шагов, каждый из которых выполняется специализированным устройством, передающим свой результат следующему за ним устройству, также выполняющему часть операции и т.д. Аппаратурный метод используется при реализации широко применяемых алгоритмов ЦОС, в частности алгоритмов быстрого преобразования Фурье, которые воплощаются в больших интегральных схемах (БИС), называемых также специализированными кристаллами ЦОС. Преимуществом аппаратурной реализации цифровой обработки сигналов является высокое быстродействие, а недостатком – невозможность изменения алгоритма работы устройства с жесткой логикой вследствие фиксированных связей между элементами схемы.

Программная реализация цифровой обработки сигналов основана на применении программируемых устройств (универсальных цифровых вычислительных машин, процессоров, микроконтроллеров), работающих под управлением программы, записанной в оперативное или постоянное запоминающее устройство. Если речь идет о реализации цифрового фильтра, то программа реализует алгоритм, который представляет собой не что иное, как разностное уравнение, описывающее фильтр. При этом, как и в случае аппаратной реализации, следует выбирать тот способ организации вычислений, который позволяет достичь наилучших показателей точности, эффективности, надежности, быстродействия и т.п. Программируемые устройства являются вычислителями последовательного действия, с чем связано их относительно низкое быстродействие.

Преимуществом программной реализации является возможность быстрого изменения логики работы устройства путем модификации программы или выбора соответствующей ветви алгоритма, благодаря чему оказывается возможным построение *адаптивных* устройств, меняющих алгоритм работы в зависимости от значений входного сигнала и других факторов. Говоря о сравнительных качествах аппаратной и программной реализации цифровых устройств, не следует забывать о том, что как элементная база специализированных устройств, так и программируемые средства непрерывно совершенствуются. При этом происходит их взаимное «идеологическое» обогащение: в универсальные устройства последовательного действия вводятся специализированные узлы (сопроцессоры), выполняющие трудоемкие операции на аппаратном уровне с высокой эффективностью, в программируемых процессорах реализуется конвейерный принцип, вводятся элементы параллельной обработки и т. п. С другой стороны, специализированные БИС становятся программируемыми за счет организации связей между элементами на основе логических вентилей, управляемых программой, благодаря чему такое устройство приобретает способность к гибкой перестройке алгоритмов, приближающейся к возможностям программируемых процессоров. Подробнее о таких устройствах, называемых программируемыми логическими интегральными схемами, будет сказано в разд. 9.

Упражнения

1. Найдите представление в 16-разрядном дополнительном двоичном коде следующих десятичных чисел: 1; -1; 0; 32767; -32767.

2. Определите десятичные числа, которым соответствуют двоичные числа, представленные в 16-разрядном дополнительном двоичном коде:

- а) 1000000000000001; б) 1000111100001111;
в) 0000111100001111; г) 1111000100010001.

3. Определите десятичные числа, которым соответствуют двоичные числа, представленные в 16-разрядном обратном двоичном коде:

- а) 1000000000000001; б) 1000111100001111;
в) 0000111100001111; г) 1111000100010001.

4. Мантисса и порядок числа занимают вместе две ячейки памяти по 16 разрядов. При этом порядок занимает 7 разрядов, а остальные занимает мантисса. Определите наибольшее и наименьшее десятичные числа, которые могут быть представлены в таком формате.

5. В формате, описанном в предыдущем упражнении, определите мантиссы и порядки чисел $11,25_{10}$ и $-14,875$. Найдите мантиссу и порядок суммы этих чисел.

6. Убедитесь на конкретных примерах, что при усечении отрицательного числа, представленного в прямом и обратном кодах, оно увеличивается (его модуль уменьшается), а усечение отрицательного числа в дополнительном коде приводит к его уменьшению.

7. Найдите плотность распределения вероятностей шума квантования, производимого путем усечения дробных двоичных чисел до 16 разрядов. Рассмотрите отдельно случаи положительных и отрицательных чисел в прямом, обратном и дополнительном кодах.

8. Выведите формулы коэффициентов влияния для параллельной формы реализации рекурсивного фильтра третьего порядка.

9. Определите непосредственными вычислениями форму колебаний предельного цикла в цепи, описываемой разностным уравнением

$$y^o[n] = \left\langle 0,95 \cdot y^o[n-1] \right\rangle_{0,01} + x[n].$$

Вопросы для самоконтроля

1. В чем состоит принципиальное отличие цифровых сигналов от дискретных?
2. Может ли цифровая цепь быть строго линейной?
3. Перечислите причины нелинейности цифровых цепей.
4. Что такое джиттер? В чем состоит причина его возникновения?
5. К каким последствиям приводит «дрожание отсчетов»?
6. Чем отличаются форматы представления чисел с фиксированной и плавающей запятой?
7. Что такое переполнение и в каких случаях оно возникает?
8. В чем особенности прямого, обратного и дополнительного кодов?
9. В каких случаях шум квантования входного сигнала можно считать некоррелированным?
10. Какое распределение имеет шум квантования ?
11. К каким последствиям может привести квантование коэффициентов цифрового фильтра?
12. К чему может привести квантование (округление) промежуточных результатов?
13. При каких условиях обоснована статистическая модель квантования?
14. Что такое предельные циклы и в чем причина их возникновения?
15. В чем состоят преимущества и недостатки аппаратурной и программной реализации цифровой обработки сигналов?



8. ЦИФРОВОЙ СПЕКТРАЛЬНЫЙ АНАЛИЗ И ЕГО ПРИМЕНЕНИЕ

Спектральным анализом принято называть совокупность действий, которые позволяют из наблюдаемой *конечной реализации* случайного процесса вывести некоторую *оценку* его спектральной плотности мощности. Отсюда следует, что методы спектрального анализа применимы только к *стационарным* в широком смысле *эргодическим* процессам. Кроме этих необходимых условий, частные методы спектрального анализа, как правило, предполагают дополнительные ограничения, которым должен удовлетворять анализируемый случайный процесс – только при этом условии полученные результаты могут быть должным образом интерпретированы и, в конечном счете, принесут практическую пользу.

Спектральный анализ имеет долгую историю, связанную с именами И. Ньютона, Ж.Б.Ж. Фурье, Д. Бернулли, Н. Винера, А.Я. Хинчина, А.Н. Колмогорова и др. Особое значение и широкое практическое применение спектральный анализ приобрел в последние десятилетия XX века в связи с интенсивным развитием цифровой обработки сигналов и в особенности с появлением алгоритмов быстрого преобразования Фурье. В настоящее время спектральный анализ на основе цифровой обработки сигналов, называемый цифровым спектральным анализом (ЦСА), находит применение в гидролокации (обнаружение в сигналах гидрофонов периодичностей, связанных с шумоизлучением винтов подводных лодок и надводных кораблей), геофизике (исследование недр Земли с помощью сейсмических датчиков), системах обработки и распознавания речевых сигналов, в экономике (исследование закономерностей сезонных и многолетних колебаний цен и других экономических показателей) и т.п.

Будем обозначать круговую частоту в спектральном описании аналоговых сигналов $\Omega = 2\pi f$, как в п. 6.1. Частота ω , соответствующая дискретным (цифровым) сигналам, связана с ней соотношением $\omega = \Omega T_d$, где T_d – шаг дискретизации, выбираемый в

соответствии с теоремой отсчетов. Если анализируемый сигнал изначально дискретен (что типично, например, для экономических приложений), то следует принять $T_d = 1$. Далее считается, что анализу подлежит дискретный сигнал (последовательность) $x[n]$ вне зависимости от того, какова его природа.

В настоящем разделе кратко рассмотрены основные методы цифрового спектрального анализа. Для более глубокого и подробного изучения методов ЦСА следует обратиться к специальной литературе [22, 23], в том числе к периодическим изданиям.

8.1. КРАТКИЙ ОБЗОР МЕТОДОВ ЦСА

Целью спектрального анализа является нахождение функции $R_x(e^{j\omega})$ – спектральной плотности мощности стационарного случайного процесса $x[n]$ на основании единственной наблюдаемой реализации. Очевидно, что эта функция может быть найдена лишь приближенно, т. е. оценена. Возможны два подхода к этому оцениванию, отличающиеся принципиально. С одной стороны, можно ставить задачу найти некоторую *функцию* частоты, которая в том или ином смысле близка к истинной СПМ. С другой стороны, в некоторых случаях вид искомой функции известен (априорно задан), а оцениванию подлежат неизвестные *параметры* этой функции. В соответствии со сказанным различают *непараметрические* и *параметрические* методы цифрового спектрального анализа. В свою очередь, непараметрические методы делятся на *прямые* и *косвенные* в зависимости от того, оценивается ли непосредственно СПМ или оценка находится для автокорреляционной функции (последовательности), а оценка СПМ вычисляется на основании теоремы Винера – Хинчина. В рамках параметрического подхода чаще всего предполагается, что случайная последовательность порождена некоторой ЛИС-цепью, возбуждаемой дискретным белым шумом, следовательно, СПМ имеет вид квадрата модуля функции, дробно-рациональной относительно $e^{j\omega}$. В соответствии с тем, имеет ли эта *гипотетическая* цепь рекурсивную, трансверсальную или смешанную структуру, различают параметрические методы *авторегрессии (АР)*, *скользящего среднего (СС)* или *авторегрессии – скользящего среднего (АРСС)* [22].

Одной из определяющих характеристик, используемых при сравнении различных методов ЦСА, является разрешающая способность – минимальная разность частот двух гармонических со-

ставляющих, различных в оценке СПМ. Для классических (непараметрических) методов спектрального анализа разрешающая способность в соответствии с принципом неопределенности не может быть меньше величины $\Delta f = 1/T$, где T – длительность наблюдаемой реализации. Параметрические методы предполагают *экстраполяцию* наблюдаемых данных за пределы интервала наблюдения, благодаря чему их разрешающая способность может быть выше, чем у классических методов. В то же время эффективность параметрического подхода сильно зависит от надежности априорной информации о виде СПМ: если функциональный вид истинной СПМ не совпадает с предполагаемым, качество оценки, полученной параметрическим методом, может быть значительно ниже качества непараметрической оценки.

8.2. ПЕРИОДОГРАММНЫЙ МЕТОД ЦСА

Интуитивно понятное определение спектральной плотности мощности аналогового случайного процесса дается формулой [22]

$$R(\Omega) = \mathbf{E} \left\{ \frac{1}{T} \left| \int_0^T x(t) e^{-j\Omega t} dt \right|^2 \right\},$$

где $\mathbf{E}\{\cdot\}$ – символ математического ожидания (усреднения по ансамблю). Заменяя интеграл суммой, реализацию процесса – последовательностью ее отсчетов и учитывая, что длительность реализации $T = NT_d$, после простых преобразований получим

$$R_x(e^{j\omega}) = \mathbf{E} \left\{ \frac{T_d}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j\omega n} \right|^2 \right\}.$$

При использовании дискретного (быстрого) преобразования Фурье вместо функции непрерывной переменной получается последовательность отсчетов

$$R_x[k] = \mathbf{E} \left\{ \frac{T_d}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn} \right|^2 \right\}, \quad k = \overline{0, N-1}. \quad (8.1)$$

Периодограммный метод является прямым непараметрическим методом спектрального анализа и основан на нахождении периодограммы, или выборочного спектра,

$$P_x[k] = \frac{T_d}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} kn} \right|^2, \quad k = \overline{0, N-1}. \quad (8.2)$$

Периодограмма отличается от СПМ (8.1) отсутствием усреднения по ансамблю. Причина этого заключается в том, что практически всегда наблюдается *единственная* реализация случайного процесса, хотя, возможно, и очень длинная. Поэтому периодограмма (8.2) является *несостоятельной* оценкой, так как с ростом длины реализации N дисперсия периодограммной оценки стремится не к нулю, а к квадрату модуля истинной СПМ $R_x[k]$ [22]. По этой причине для практического применения периодограммную оценку (8.2) необходимо модифицировать.

Наиболее просто сделать периодограммную оценку состоятельной можно усреднением нескольких периодограмм, найденных для различных частей (сегментов) наблюдаемой последовательности (метод Бартлетта). Такая замена *ансамбля* реализаций набором отрезков *единственной* реализации правомерна, если процесс обладает свойством эргодичности (см. разд. 4). Вся наблюдаемая последовательность длины N разбивается на m сегментов длины N/m , для каждого (i -го) сегмента вычисляется периодограмма $P_i[k]$, $k = \overline{0, \frac{N}{m}-1}$, а результирующая периодограмма Бартлетта $P_B[k]$ получается арифметическим усреднением:

$$P_B[k] = \frac{1}{m} \sum_{i=1}^m P_i[k], \quad k = \overline{0, \frac{N}{m}-1}. \quad (8.3)$$

На рис. 8.1 представлена реализация узкополосной случайной последовательности длиной 1024 отсчета, на рис. 8.2 – периодограмма, найденная по всей реализации, а на рис. 8.3 – результат



Рис. 8.1. Реализация узкополосной случайной последовательности

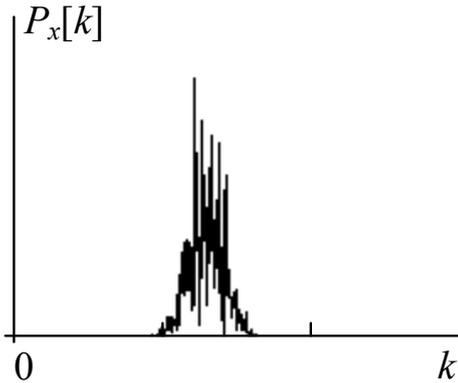


Рис. 8.2. Периодограмма последовательности, показанной на рис. 8.1.

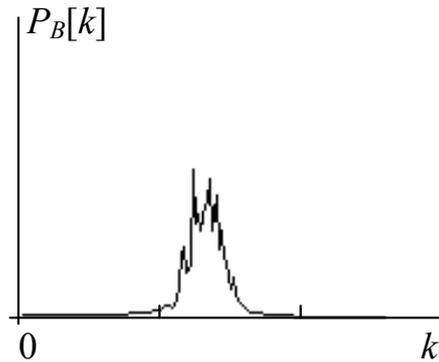


Рис. 8.3. Периодограмма, полученная методом Бартлетта

усреднения периодограмм (8.3), полученных для четырех сегментов длиной по 256 отсчетов.

Следует отметить, что выделение сегмента конечной длины из бесконечно длинной реализации эквивалентно умножению этой реализации на прямоугольное *окно*, вследствие чего периодограмма, полученная на основе этого сегмента, является оценкой не спектральной плотности мощности, а ее свертки с частотной функцией прямоугольного окна и содержит поэтому гиббсовские осцилляции, аналогичные рассмотренным в разд. 6. Для уменьшения осцилляций, которые могут замаскировать реальные подробности СПМ, применяют окна, которые умножаются на фрагменты реализации до вычисления периодограмм по методу Бартлетта. При этом ухудшается разрешающая способность периодограммной оценки.

Другой способ улучшения периодограммной оценки (метод Даньелла) состоит в *сглаживании* периодограммы, найденной по всей последовательности, путем скользящего арифметического усреднения нескольких соседних отсчетов. При этом в простейшем случае каждый k -й отсчет периодограммы (8.2) заменяется сред-

ним арифметическим m отсчетов $\frac{1}{m} \sum_{i=k-(m-1)/2}^{k+(m-1)/2} P_x[i]$. И метод Барт-

летта, и метод Даньелла делают оценку более точной (уменьшается дисперсия), но менее подробной (ухудшается частотное разрешение).

Известны разновидности периодограммного метода, различающиеся алгоритмом усреднения, применяемым окном и т.п. Подробное обсуждение этих разновидностей приведено в [22].

8.3. КОРРЕЛОГРАММНЫЙ МЕТОД ЦСА

Известно, что для стационарных в широком смысле случайных последовательностей справедлива теорема Винера–Хинчина, утверждающая, что СПМ $R_x(e^{j\omega})$ и автокорреляционная последовательность (АКП) $r_x[m]$ связаны парой преобразований Фурье:

$$R_x(e^{j\omega}) = \sum_{m=-\infty}^{\infty} r_x[m] e^{-j\omega m}, \quad (8.4)$$

$$r_x[m] = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_x(e^{j\omega}) e^{j\omega m} d\omega. \quad (8.5)$$

На этой связи основан коррелограммный метод спектрального оценивания – косвенный метод, состоящий в нахождении по наблюдаемой последовательности $x[n]$ оценки $\tilde{r}_x[m]$ ее автокорреляционной последовательности $r_x[m]$ и вычислении оценки СПМ как преобразования Фурье последовательности $\tilde{r}_x[m]$ (на практике вместо преобразования Фурье используют БПФ):

$$\tilde{R}_x[k] = \sum_{m=-\infty}^{\infty} \tilde{r}_x[m] e^{-j\frac{2\pi}{N}km}, \quad k = \overline{0, N-1}. \quad (8.6)$$

Оценка автокорреляционной последовательности (*коррелограмма*) находится путем «усреднения по времени»

$$\tilde{r}_x[m] = \frac{1}{N-m} \sum_{n=0}^{N-m-1} x[n]x[n+m]. \quad (8.7)$$

Следует отметить, что верхний предел суммирования и нормирующий коэффициент зависят от значения переменной m так, что в сумму входят только произведения отсчетов, принадлежащих перекрывающимся частям реализации $x[n]$ и ее сдвинутой копии $x[n+m]$ (рис. 8.4).

Пределы суммирования в (8.7) можно положить бесконечными, но при этом реализацию и ее копию следует умножить на прямоугольные окна вида

$$w[n] = \begin{cases} 1, & n = \overline{0, N-1}, \\ 0 & \text{в противном случае.} \end{cases} \quad (8.8)$$

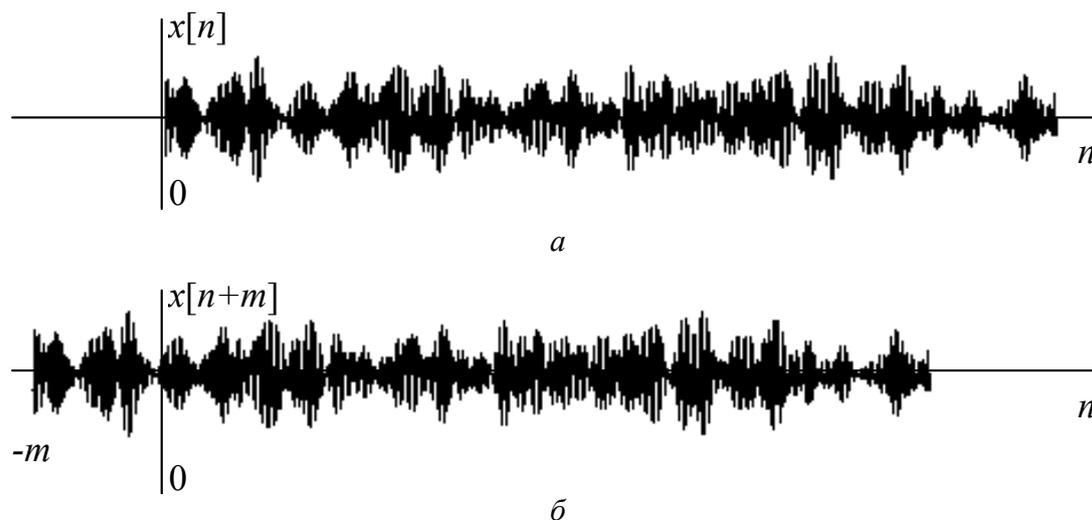


Рис. 8.4. Реализация случайной последовательности (а) и ее сдвинутая копия (б)

Таким образом, по мере увеличения аргумента m область перекрытия уменьшается, а вследствие этого увеличивается дисперсия оценки $\tilde{r}_x[m]$ (рис. 8.5).

Если оценку $\tilde{r}_x[m]$ использовать для вычисления оценки СПМ в соответствии с выражением (8.6), то оценка СПМ оказывается сильно «изрезанной» и мало отличается от периодограммной оценки, найденной по всей реализации. Улучшить коррелограммную оценку СПМ можно путем усечения (обнуления) хвостовой части коррелограммы, которая имеет большую дисперсию.

На рис. 8.6 показана усеченная коррелограмма $r'[m]$, а на рис. 8.7, а – соответствующая ей коррелограммная оценка $R'[k]$ спектральной плотности мощности. На рис. 8.7, б приведена для сравнения коррелограммная оценка СПМ $\tilde{R}_x[k]$, найденная по неусеченной коррелограмме (8.7).

На графике коррелограммной оценки $R'[k]$, показанном на рис. 8.7, а, хорошо видны гиббсовские осцилляции, вызванные прос-



Рис. 8.5. Коррелограмма, найденная по реализации, приведенной на рис. 8.4



Рис. 8.6. Усеченная коррелограмма

тым усечением коррелограммы, эквивалентным применению прямоугольного окна (8.8). В некоторых приложениях такие осцилляции могут замаскировать детали спектральных оценок, представляющие интерес (например, в системе гидроакустического обнаружения такой деталью может быть гармоническая составляющая в спектре шумоизлучения подводной лодки, расположенная поблизости от мощной гармоники шума надводного корабля). Для уменьшения гиббсовских осцилляций применяются различные окна [22], однако, как и в случае периодограммного оценивания, при этом происходит снижение частотного разрешения.

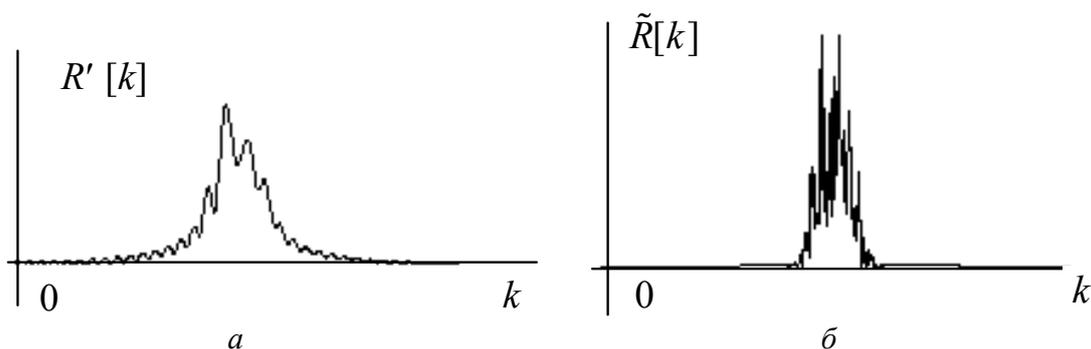


Рис. 8.7. Оценки СПМ, найденные по усеченной (а) и неусеченной (б) коррелограммам

К достоинствам непараметрических методов спектрального анализа следует отнести их робастность (устойчивость к изменениям параметров случайного процесса), а также высокую вычислительную эффективность, обусловленную применением алгоритмов БПФ. Недостатки таких методов – сравнительно низкое частотное разрешение и «эффект просачивания» (так в литературе иногда называют вредное влияние боковых лепестков частотной функции окна) [22].

8.4. ПАРАМЕТРИЧЕСКИЕ МОДЕЛИ СЛУЧАЙНЫХ ПРОЦЕССОВ

Рассмотренные выше непараметрические оценки СПМ были получены при самых общих предположениях о случайном процессе, а именно: случайный процесс полагался стационарным в широком смысле эргодическим. Во многих практически важных случаях имеется некоторая дополнительная (априорная) информация об анализируемом случайном процессе. Например, может быть известно, что СПМ процесса представляет собой довольно гладкую функцию с несколькими выраженными максимумами (в частности, такое предположение оправдано при анализе речевых сигналов, соответствующих гласным звукам) или что СПМ может содержать не только резкие всплески, но и глубокие провалы. В таких случаях применяются методы спектрального анализа, основанные на параметрических моделях случайных процессов.

Самой общей линейной параметрической моделью стационарного случайного процесса является модель [22], описывающая случайный процесс (последовательность) как решение линейного разностного уравнения с постоянными коэффициентами

$$x[n] = \sum_{k=1}^p a_k x[n-k] + \sum_{k=0}^q b_k \eta[n-k] \quad (8.9)$$

при условии, что $\eta[n]$, $n = -\infty, +\infty$, – дискретный белый шум (последовательность независимых случайных величин с одинаковой дисперсией). Выражение (8.9) описывает реакцию на белый шум каузальной ЛИС-цепи с передаточной функцией дробно-рационального вида $H(z) = \frac{B(z)}{A(z)}$, числитель и знаменатель кото-

рой представляют собой полиномы относительно z^{-1} порядков q и p соответственно:

$$B(z) = \sum_{r=0}^q b_r z^{-r}, \quad A(z) = 1 - \sum_{k=1}^p a_k z^{-k}.$$

Известно, что такая цепь имеет бесконечную импульсную характеристику $h[n]$, так что

$$H(z) = \sum_{n=0}^{\infty} h[n] z^{-n}. \quad (8.10)$$

Предполагается, что цепь устойчивая минимально-фазовая, т. е. нули числителя и знаменателя передаточной функции лежат на z -плоскости внутри 1-окружности.

z -Преобразование автокорреляционной последовательности для выходного случайного процесса определяется выражением

$$R_x(z) = R_0(z)H(z)H^*\left(\frac{1}{z^*}\right) = R_0(z)\frac{B(z)B^*(1/z^*)}{A(z)A^*(1/z^*)}, \quad (8.11)$$

где $R_0(z)$ – z -образ АКП входной последовательности, для белого шума $\eta[n]$ постоянный и равный его дисперсии $D_\eta = \sigma_\eta^2$. Тогда (8.11) принимает вид

$$R_x(z) = D_\eta \frac{B(z)B^*(1/z^*)}{A(z)A^*(1/z^*)}. \quad (8.12)$$

Иногда априорная информация о процессе позволяет положить в (8.12) $B(z) = 1$ или $A(z) = 1$. Такие более простые модели получили название моделей авторегрессии и скользящего среднего соответственно. Общая модель (8.9) называется моделью авторегрессии – скользящего среднего. Очевидно, что СПМ процесса $x[n]$, описываемого моделью (8.9) – (8.12), полностью определяется $p + q + 1$ параметрами, а именно p коэффициентами знаменателя, q коэффициентами числителя и дисперсией $D_\eta = \sigma_\eta^2$ (для удобства коэффициенты числителя сокращают на b_0 , компенсируя это изменением D_η).

Параметры модели (8.9) – (8.12) можно определить, если известна автокорреляционная последовательность процесса $x[n]$.

Умножая обе части разностного уравнения (8.9) на $x^*[n - m]$ и выполняя формальное усреднение по ансамблю, получим

$$\begin{aligned} \mathbf{E}\{x[n]x^*[n - m]\} &= \sum_{k=1}^p a_k \mathbf{E}\{x[n - k]x^*[n - m]\} + \\ &+ \sum_{r=0}^q b_r \mathbf{E}\{\eta[n - r]x^*[n - m]\}, \end{aligned}$$

т. е.

$$r_x[m] = \sum_{k=1}^p a_k r_x[m - k] + \sum_{r=0}^q b_r r_{\eta x}[m - r],$$

где взаимная корреляция между входной и выходной последовательностями может быть записана в форме

$$\begin{aligned} r_{\eta x}[i] &= \mathbf{E} \left\{ \eta[n+i] x^*[n] \right\} = \mathbf{E} \left\{ \eta[n+i] \left[\eta^*[n] + \sum_{k=1}^{\infty} h^*[k] \eta^*[n-k] \right] \right\} = \\ &= r_{\eta}[i] + \sum_{k=1}^{\infty} h^*[k] r_{\eta}[i+k]. \end{aligned} \quad (8.13)$$

Здесь использовано выражение для импульсной характеристики формирующей ЛИС-цепи (8.10) с учетом условия $h[0] = 1$. Поскольку автокорреляционная последовательность шума $r_{\eta}[i] = D_{\eta} \delta[i]$, то

$$r_{\eta x}[i] = \begin{cases} D_{\eta} h^*[-i] & \text{при } i \leq 0, \\ 0 & \text{при } i > 0. \end{cases}$$

Отсюда получается окончательное выражение, связывающее параметры АРСС-модели и автокорреляционной последовательности наблюдаемого процесса:

$$\begin{cases} r_x[m] = r_x^*[-m], & m < 0; \\ \sum_{k=1}^p a_k r_x[m-k] + D_{\eta} \sum_{k=m}^q b_k h^*[k-m], & 0 \leq m \leq q; \\ \sum_{k=1}^p a_k r_x[m-k], & m > q. \end{cases} \quad (8.14)$$

Видно, что при $m > q$ значения АКП $r_x[m]$ связаны рекурсивным выражением, зависящим только от коэффициентов знаменателя передаточной функции формирующей ЛИС-цепи, т. е. от параметров авторегрессионной части модели (8.9). На этом основании можно записать *нормальные уравнения Юла–Уолкера* в матричной форме

$$\begin{pmatrix} r_x[q] & r_x[q-1] & \cdot & r_x[q-p+1] \\ r_x[q+1] & r_x[q] & \cdot & r_x[q-p+2] \\ \cdot & \cdot & \cdot & \cdot \\ r_x[q+p-1] & r_x[q+p-2] & \cdot & r_x[q] \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ a_p \end{pmatrix} = \begin{pmatrix} r_x[q+1] \\ r_x[q+2] \\ \cdot \\ r_x[q+p] \end{pmatrix},$$

позволяющие определить параметры a_1, \dots, a_p авторегрессионной части модели по известным значениям АКП для $q - p + 1 \leq m \leq q + p$.

К сожалению, СС-параметры b_0, \dots, b_q модели (8.9) связаны нелинейными уравнениями, содержащими свертки (выражение (8.14) при $0 \leq m \leq q$). Это затрудняет нахождение параметров модели скользящего среднего даже при известной автокорреляционной последовательности процесса.

Чтобы получить уравнения Юла–Уолкера для авторегрессионной модели, следует в (8.14) положить $q = 0$. Тогда из (8.14) получаем уравнения вида

$$\begin{pmatrix} r_x[0] & r_x[-1] & \cdot & r_x[-p] \\ r_x[1] & r_x[0] & \cdot & r_x[-p+1] \\ \cdot & \cdot & \cdot & \cdot \\ r_x[p] & r_x[p-1] & \cdot & r_x[0] \end{pmatrix} \begin{pmatrix} 1 \\ -a_1 \\ \cdot \\ -a_p \end{pmatrix} = \begin{pmatrix} D_\eta \\ 0 \\ \cdot \\ 0 \end{pmatrix}. \quad (8.16)$$

Заметим, что авторегрессионная модель однозначно определяется значениями АКП $r_x[0], r_x[1], \dots, r_x[p]$, поскольку все последующие значения АКП находятся рекурсивно согласно выражению

$$r_x[m] = \sum_{k=1}^p a_k r_x[m-k], \quad (8.17)$$

что следует из (8.14) при $q = 0$. Именно в этом состоит достоинство авторегрессионного метода спектрального анализа по сравнению с классическим коррелограммным методом, в котором используются лишь те значения оценки АКП, которые получены при вычислении коррелограммы, а все остальные полагаются равными нулю. В АР-методе фактически происходит экстраполяция оценки АКП за пределы интервала наблюдения, что повышает разрешающую способность спектрального анализа.

В завершение этого краткого обзора отметим, что выбор параметрической модели для ЦСА определяется априорной информацией о форме спектральной плотности мощности процесса. Так, если ожидаемая СПМ характеризуется наличием пиков, то можно рассчитывать, что подходящей окажется АР-модель; если ожидается наличие глубоких провалов при общей гладкости СПМ, то следует применять СС-модель (передаточная функция формирующей

ЛИС-цепи содержит только нули), если же предполагаемая СПМ может иметь как пики, так и провалы, тогда наиболее оправданной оказывается АРСС-модель. Необходимо помнить, однако, что в вычислительном отношении две последние модели менее удобны, так как требуют решения систем нелинейных уравнений. Кроме того, еще раз отметим, что параметрические методы могут давать очень плохие результаты, если априорные предположения о форме СПМ окажутся ошибочными.

8.5. СПЕКТРАЛЬНОЕ ОЦЕНИВАНИЕ ПО МЕТОДУ АВТОРЕГРЕССИИ

Решив уравнения Юла–Уолкера (8.16) относительно вектора параметров рекурсивной формирующей цепи $(1, -a_1, -a_2, \dots, -a_p)^T$,

можно получить оценку СПМ $R_x(z) = \frac{R_0(z)}{A(z)A^*(1/z^*)}$. При этом

предполагается, что значения автокорреляционной последовательности процесса $x[n]$ точно известны. Реально эти значения могут быть лишь приближенно оценены по наблюдаемой реализации. В настоящее время разработаны многочисленные методы оценивания параметров АР-модели случайной последовательности, которые описаны в литературе [22, 23]. Эти методы отличаются способом организации использования наблюдаемых данных (блочные и последовательные алгоритмы), способом интерпретации результатов (метод максимальной энтропии, метод линейного предсказания), возможностью адаптации к медленным изменениям параметров модели и т.п.

Авторегрессионный процесс описывается разностным уравнением, вытекающим из (8.9) при $b_k = 0$, $k = \overline{1, q}$:

$$x[n] = \sum_{k=1}^p a_k x[n-k] + \eta[n], \quad (8.18)$$

где $\eta[n]$, $n = \overline{-\infty, +\infty}$ – дискретный белый шум. Можно сказать, что очередное значение последовательности $x[n]$ определяется как линейная комбинация p предшествующих значений этой же последовательности и случайной величины $\eta[n]$. Очевидно, если бы этой случайной величины не было, то можно было бы точно опре-

делить (предсказать) значение $x[n]$ согласно (8.18). В действительности можно получить лишь *оценку линейного предсказания*

$$\tilde{x}[n] = \sum_{k=1}^p a_k x[n-k], \quad (8.19)$$

всегда сопровождаемую ошибкой $e[n] = x[n] - \tilde{x}[n]$. Тогда очередное значение последовательности $x[n]$ можно выразить через оценку линейного предсказания и ошибку в виде

$$x[n] = \sum_{k=1}^p a_k x[n-k] + e[n]. \quad (8.20)$$

Это выражение практически совпадает с (8.18), за исключением последнего слагаемого. На самом деле выражение (8.18) описывает процесс на выходе рекурсивного формирующего фильтра с передаточной функцией

$$H^g(z) = 1 / \left(1 - \sum_{k=1}^p a_k z^{-k} \right) = 1 / A(z),$$

возбуждаемого дискретным белым шумом. Выражение (8.20), переписанное в виде

$$e[n] = x[n] - \sum_{k=1}^p a_k x[n-k], \quad (8.21)$$

описывает процесс на выходе трансверсального фильтра с передаточной функцией $H^i(z) = 1 - \sum_{k=1}^p a_k z^{-k} = A(z)$, являющегося обрат-

ным (инверсным) по отношению формирующему фильтру, когда на вход обратного фильтра воздействует процесс $x[n]$. Очевидно, что при точном соответствии фильтров ошибка $e[n]$ представляет собой входной белый шум модели формирования авторегрессионного процесса. При этом дисперсия (мощность) ошибки будет минимальной. Таким образом, задача оценивания коэффициентов АР-модели сводится к такой настройке коэффициентов трансверсального обратного фильтра, когда процесс на его выходе имеет наименьшую дисперсию и является белым шумом («отбеливающий» фильтр) (рис. 8.8).

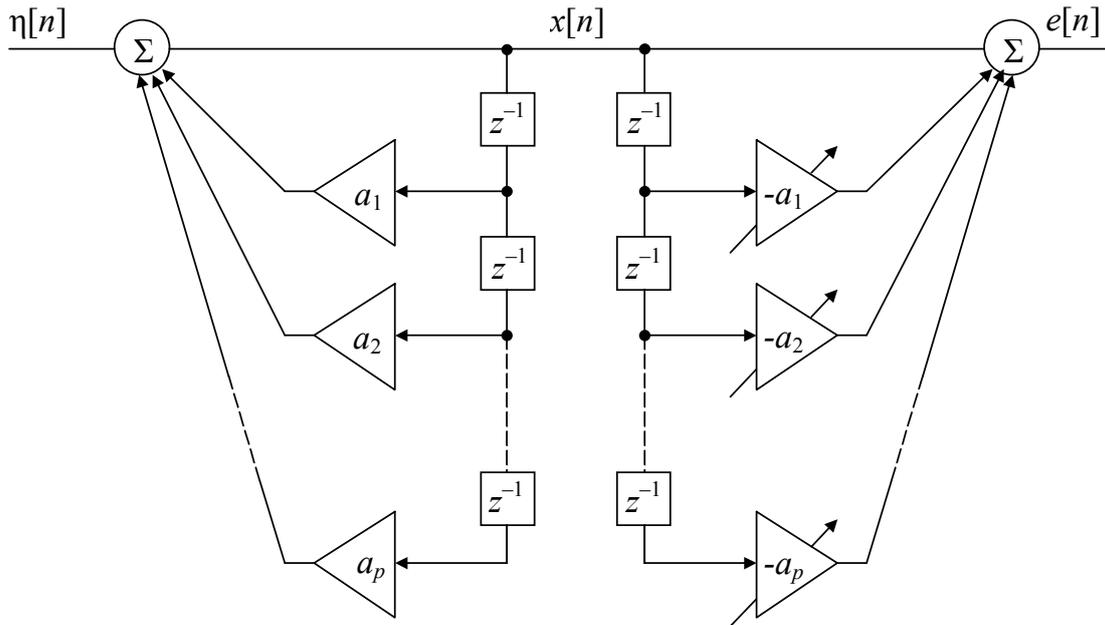


Рис. 8.8. Оценивание параметров АР-модели как настройка отбеливающего фильтра

Следует отметить, что при практическом использовании метода линейного предсказания, как правило, порядок формирующей цепи точно не известен. С другой стороны, очевидно, что увеличение порядка отбеливающего фильтра сверх необходимого (p) не приведет к дальнейшему уменьшению дисперсии. Это обстоятельство на практике может служить критерием правильности выбора порядка АР-модели. Стоит заметить, что отбеливающий фильтр является минимально-фазовым (почему?).

В некоторых случаях рассматривают антикаузальный вариант отбеливающего фильтра, последовательность на выходе которого называется «ошибкой предсказания назад». Этот фильтр имеет такую же структуру, как и фильтр «предсказания вперед», описываемый выражением (8.21), а его коэффициенты a_k^b , $k = \overline{1, p}$ являются комплексно-сопряженными по отношению к коэффициентам, входящим в (8.21), т. е. $a_k^b = (a_k)^*$. Пару фильтров, формирующих ошибки предсказания вперед и назад, можно представить схемой, показанной на рис. 8.9.

Коэффициенты линейного предсказания вперед называют также коэффициентами отражения (в соответствии с их физической интерпретацией в задачах обработки речевых сигналов и сейсмических данных), в этих случаях для них используют обозначения

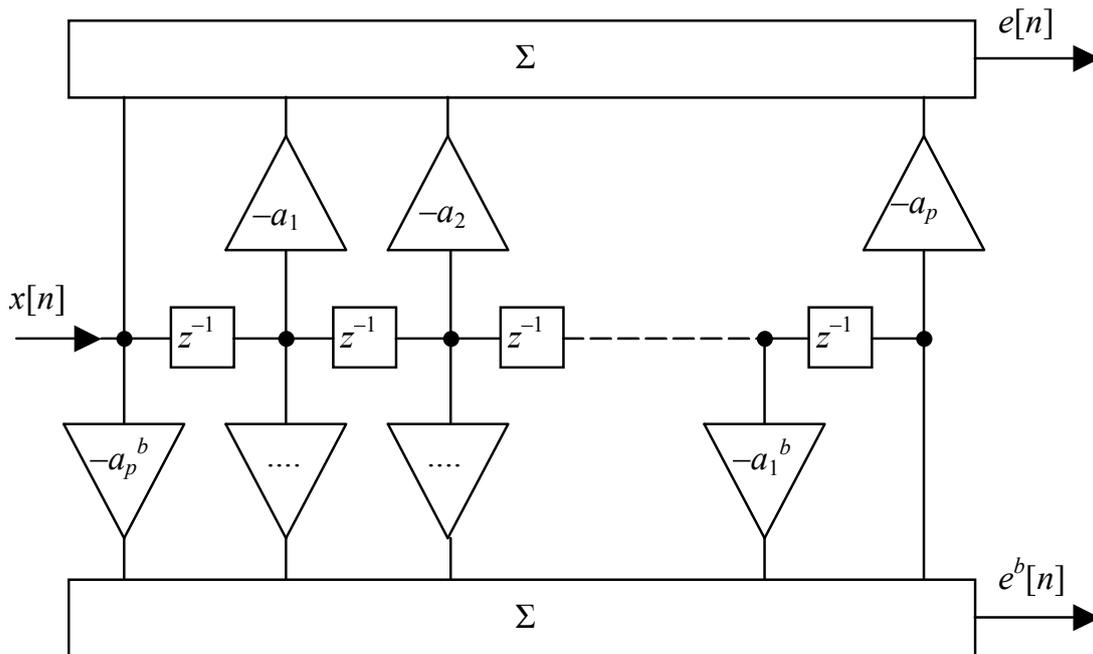


Рис. 8.9. Формирование ошибок предсказания «вперед» и «назад»

k_m , $m = \overline{1, p}$. Взятые со знаком минус, эти же коэффициенты называются частными коэффициентами корреляции [22].

Упражнения

1. Объясните причину снижения разрешающей способности периодограммной и коррелограммной оценок, обусловленного применением окон.
2. Объясните причину увеличения дисперсии коррелограммной оценки (8.7) при увеличении значения переменной m .
3. Покажите, что уравнения (8.16) получаются из (8.14) при $q = 0$.

Вопросы для самоконтроля

1. Почему периодограмма последовательности не применяется в качестве самостоятельной оценки спектральной плотности мощности?
2. В чем состоит смысл методов Бартлетта и Даньелла?
3. Почему АР-модель применяется в тех случаях, когда ожидаемая СПМ характеризуется наличием пиков?

4. Почему СС-модель применяется в тех случаях, когда предполагается наличие глубоких провалов СПМ?

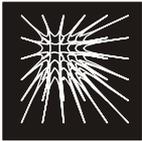
5. Когда целесообразно использовать для спектрального анализа АРСС-модель?

6. Когда целесообразно использовать для спектрального анализа непараметрические методы?

7. Чем объясняется более высокое частотное разрешение параметрических методов спектрального анализа?

8. Что такое эффект просачивания, и как его ослабить? Какая характеристика спектральной оценки при этом ухудшается и почему?

9. Что представляет собой метод линейного предсказания? В чем состоят его преимущества?



9. СИГНАЛЬНЫЕ ПРОЦЕССОРЫ И ИХ ПРИМЕНЕНИЕ

Реализация алгоритмов цифровой обработки сигналов, как отмечалось ранее, возможна на различной элементной базе. Выбор конкретного воплощения алгоритма цифровой обработки сигналов производится разработчиком с учетом различных показателей, к которым относятся стоимость, массогабаритные характеристики, энергопотребление, быстродействие и т.п. В каждом отдельном случае один или несколько показателей играют наиболее важную роль в выборе способа реализации. Например, в системах *подвижной радиосвязи* главными показателями являются быстродействие (обработка должна выполняться в реальном времени) и массогабаритные характеристики, при этом желательно обеспечить малое энергопотребление и умеренную цену мобильной станции. Устройство обработки сигналов в таких системах работает по жестким алгоритмам, которые не изменяются в течение всего срока эксплуатации изделия; точность представления данных (разрядность) должна быть достаточна для обеспечения комфортности восприятия речи (разборчивости и возможности узнавания собеседника) и является поэтому сравнительно невысокой. В системах обработки геофизической информации, напротив, обработка в реальном времени не требуется, массогабаритные характеристики не играют доминирующей роли, однако на передний план выступают точность представления данных и гибкость реализуемых алгоритмов. Поэтому в таких системах обработка сигналов реализуется обычно на универсальных цифровых вычислительных машинах. В системах реального времени (например, в подвижной радиосвязи), когда на обработку отсчета сигнала отводится временной интервал, равный шагу дискретизации, как правило, используются специализированные цифровые устройства, называемые *цифровыми сигнальными процессорами (ЦСП)*.

9.1. АРХИТЕКТУРНЫЕ ОСОБЕННОСТИ СИГНАЛЬНЫХ ПРОЦЕССОРОВ

Первым микропроцессором был выпущенный в 1971 году фирмой Intel прибор i4004. Этот 4-разрядный микропроцессор был разработан для использования в микрокалькуляторах, но результат оказался неизмеримо более впечатляющим. Фактически появление этой микросхемы вызвало научно-техническую революцию, которая не закончилась и по сей день.

Микропроцессоры общего назначения применяются в универсальных устройствах (таких, например, как персональные компьютеры). Архитектура таких микропроцессоров имеет следующие характерные особенности:

- наличие развитой системы команд, включающей команды условного и безусловного ветвления;
- высокая разрядность и возможность обработки данных различных типов, включая данные с плавающей запятой и повышенной точностью;
- наличие стековой памяти достаточно большой глубины, а также системы обработки прерываний;
- наличие регистровой памяти (блока регистров общего назначения);
- развитая система адресации;
- наличие прямого доступа к памяти, внутреннего математического сопроцессора (фактически представляющего собой самостоятельный процессор, специализированный для выполнения трудоемких операций умножения и деления чисел с плавающей запятой) и т.п.

Чаще всего микропроцессоры общего назначения имеют так называемую *принстонскую (фон-неймановскую)* архитектуру, которая характеризуется общей памятью для программ и данных. *Гарвардская* архитектура предусматривает для программ и данных отдельную память (и разные шины). На рис. 9.1 и 9.2 показаны упрощенные структурные схемы, соответствующие принстонской и гарвардской архитектурам. В принстонской архитектуре имеется

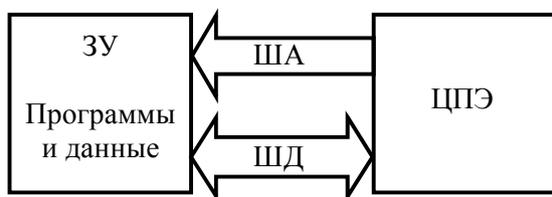


Рис. 9.1. Принстонская архитектура

одно запоминающее устройство (ЗУ), и *центральный процессорный элемент (ЦПЭ)* обменивается с ним информацией по *шине* (магистральной), которая фактически содержит шину *данных (ШД)*, шину

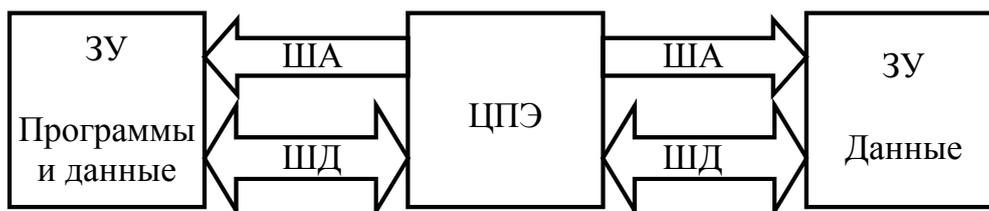


Рис. 9.2. Гарвардская архитектура

адресов (ША) и шину команд, не показанную на рисунке. Следует заметить, что когда говорят об *одном* запоминающем устройстве в принстонской архитектуре, то подразумевают одно общее *адресное пространство*, т. е. множество *адресов* элементов памяти, к которым можно обратиться при помощи одной адресной шины. Физически же адреса, принадлежащие этому адресному пространству, могут быть распределены между оперативным запоминающим устройством (ОЗУ), постоянным запоминающим устройством (ПЗУ) и регистрами внешних (периферийных) устройств. В устройстве принстонской архитектуры команды и данные (операнды) могут передаваться по единственной шине данных лишь *поочередно*, что приводит к значительным временным затратам, так как за время одного *командного цикла* в общем случае производятся выборка команды, выборка двух операндов, выполнение команды и запись в память полученного результата. Гарвардская архитектура обеспечивает потенциально большее быстродействие за счет возможности *одновременной* пересылки команд и данных по *разным* шинам.

Развитие микропроцессоров привело к появлению новых устройств и систем, которые без них просто не могли бы быть реализованы, таких, например, как системы мобильной связи. Это, в свою очередь, стимулировало появление специализированных сигнальных процессоров, которые ценой уменьшения степени универсальности достигают в некоторых отношениях предельных показателей. Часто сигнальные процессоры имеют кроме узлов, традиционных для микропроцессоров, также встроенные элементы сопряжения с аналоговыми устройствами (в частности, аналого-цифровые и цифроаналоговые преобразователи), что позволяет выполнять *цифровую обработку аналоговых сигналов* при помощи однокристалльного устройства.

Специализация сигнальных процессоров определяет следующие особенности их архитектуры :

- сокращенная система команд, ориентированная на выполнение типовых алгоритмов ЦОС;
- невысокая разрядность и фиксированный тип данных;

– отдельная память для данных и программ, позволяющая производить чтение команд и данных *одновременно* по разным шинам;

– наличие циклической адресации, наиболее удобной при реализации алгоритмов ЦОС;

– конвейерная организация выборки и исполнения команд и т.д.

В настоящее время выпускается большое количество сигнальных процессоров различных фирм-производителей, среди которых наиболее известны Motorola, Texas Instruments и Analog Devices. На одном кристалле ЦСП содержится от 30 000 до 50 000 и более транзисторов. Подробное описание устройства, системы команд и особенностей программирования одного сигнального процессора занимает многие сотни страниц (см., например, [24, 25]). В рамках одного раздела невозможно рассмотреть подробно даже один сигнальный процессор; кроме того, развитие микропроцессорной техники происходит такими темпами, что содержание раздела неизбежно устареет к моменту выхода учебника в свет. Поэтому здесь кратко рассматриваются лишь *основные* архитектурные принципы построения сигнальных процессоров, которые, с одной стороны, меняются не так быстро, а с другой – их понимание необходимо для самостоятельного изучения архитектуры конкретных процессоров по технической документации фирм-производителей, которая доступна, например, в сети Internet.

Рассмотрим основные понятия, относящиеся к архитектурным особенностям сигнальных процессоров, отличающих их от универсальных процессоров широкого применения.

Базовая операция ЦСП. Одной из наиболее распространенных операций в цифровой обработке сигналов является дискретная свертка, представляющая собой многократно выполняемое умножение двух чисел и сложение результата с уже накопленной суммой (*умножение с накоплением*). Поэтому в состав большинства ЦСП входит узел, называемый *суммирующим множителем* или *накапливающим сумматором* и реализующий указанную операцию за один командный цикл (период времени, необходимый для выполнения короткой команды).

Модифицированная гарвардская архитектура, характерная для многих сигнальных процессоров, предполагает наличие нескольких одновременно адресуемых блоков памяти. Например, процессоры семейства DSP56000 фирмы Motorola имеют на кристалле два блока *памяти данных* (X-память и Y-память) и один блок *памяти программ* (P-память). Соответственно имеются три адресные ши-

ны – шина адресов памяти X, шина адресов памяти Y, шина адресов памяти P, а также три шины данных, которые могут действовать одновременно и независимо друг от друга. Таким образом, в течение одного командного цикла могут быть параллельно осуществлены выборка двух операндов (из X-памяти и Y-памяти) и команды из P-памяти, умножение операндов и суммирование с уже накопленным результатом.

Конвейер команд. Чтение (выборка) команд из памяти и их исполнение осуществляются за несколько шагов, называемых *машинными* циклами. Этот процесс происходит под управлением *программного контроллера*, который представляет собой цифровой автомат, выполняющий формирование адреса и выдачу его на адресную шину P-памяти, прием кода команды с шины данных P-памяти, декодирование кода команды и формирование управляющих сигналов для арифметико-логического устройства (АЛУ). Эти действия происходят не одновременно, а в течение нескольких тактовых интервалов, так что исполнение одной команды состоит из нескольких *фаз*, относительно независимых друг от друга. Используя то обстоятельство, что порядок следования команд в большинстве случаев predetermined программистом и соответствует естественному порядку адресов P-памяти, можно совместить во времени различные фазы исполнения нескольких команд. Например, когда n -я команда подвергается декодированию, $(n-1)$ -я команда исполняется, а $(n+1)$ -я считывается из памяти. Такой *конвейерный* принцип исполнения команд позволяет существенно повысить быстродействие сигнальных процессоров, хотя требует от программиста внимательности при составлении программы. Нельзя, например, использовать в команде данные, которые формируются предшествующей командой, так как к моменту использования они могут быть еще не готовы (фаза исполнения кода предыдущей команды совпадает по времени с фазой выборки текущей команды).

Циклическая (модульная) адресация. При реализации линейных алгоритмов цифровой фильтрации, которые наиболее часто применяются в устройствах широкого применения, работающих в реальном времени, для вычисления одного выходного значения используются несколько соседних значений входного сигнала и – в случае БИХ-фильтрации – несколько предшествующих значений выходной последовательности. Для того чтобы оптимизировать последовательную выборку данных, применяется *кольцевой буфер*, организованный в памяти данных при помощи *модульной адресации*. Данные, используемые в качестве операндов, выбираются из буфера с использованием *указателя*, который всегда равен адресу

элемента, подлежащего выборке. На каждом шаге к указателю прибавляется (или вычитается) единица, но сложение/вычитание производится *по модулю*, определяемому порядком фильтра. На рис. 9.3, *а* показан участок памяти, в котором хранятся значения последовательности $x[n]$ при реализации КИХ-фильтра 8-го порядка [26]. Для вычисления одного отсчета выходной последовательности $y[n]$ по формуле дискретной свертки используются 8 значений входной последовательности от $x[n]$ до $x[n-7]$. После того как значение $y[n]$ найдено, число $x[n-7]$, хранящееся по адресу 10025, больше никогда не понадобится, и на его место можно записать очередное входное значение, которое теперь будет самым «новым» и должно быть обозначено $x[n]$. Указатель становится равен адресу 10025, и на следующем шаге именно с него начинается вычисление свертки. Значение 0,7367, хранящееся по адресу 10024, теперь получает обозначение $x[n-1]$, но хранится по тому же адресу. Аналогично и все остальные значения входной последовательности хранятся там же, где они находились на предыдущем шаге, меняется лишь указатель – число, которое указывает, с какого адреса начинается выборка совокупности отсчетов (новое состояние буфера показано на рис. 9.3, *б*). Значение указателя при этом увеличивается на 1, в рассматриваемом примере оно становится равным 10025. Через три шага указатель станет равным 10028, а затем 10021 (смещение относительно начального адреса 10021 станет равным 8, но поскольку сложение реализуется по модулю 8, то *эффективное* значение смещения будет равно 0). Таким образом, указатель при последовательном обращении к буферу будет бесконечно циркулировать в выделенном диапазоне адресов.

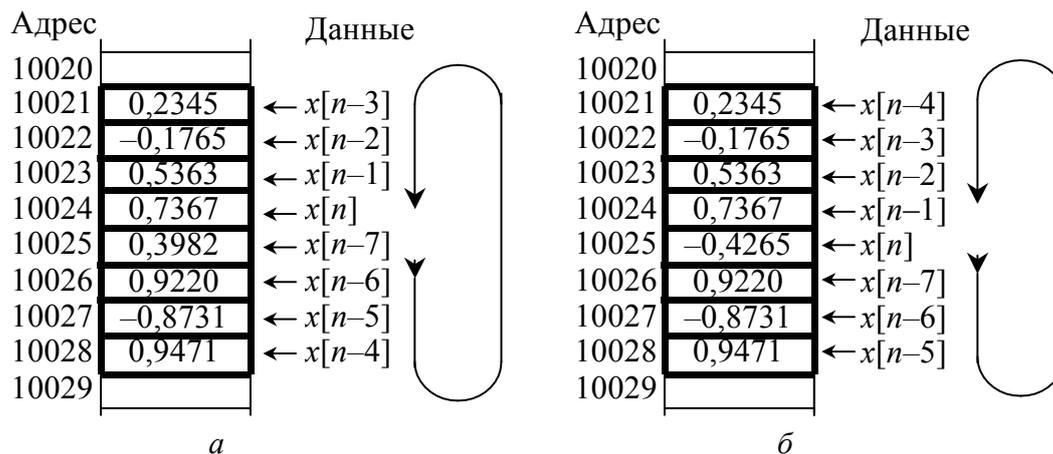


Рис. 9.3. Содержимое кольцевого буфера на двух последовательных шагах (*а*, *б*) обработки сигнала

Легко видеть, что циклическая адресация дает существенное преимущество в быстродействии по сравнению с обычной индексной адресацией, так как не требуется переписывать значения входной последовательности; вместо этого на каждом шаге обновляется лишь один отсчет входных данных и модифицируется указатель. Отметим, что запись новых значений входного сигнала в ячейки кольцевого буфера может производиться без участия ЦПЭ путем прямого доступа к памяти (ПДП) внешнего устройства (например, АЦП) или за счет дополнительных аппаратных средств сигнального процессора.

Бит-реверсивная адресация. Реализация цифровых фильтров методом быстрой свертки предполагает применение алгоритма быстрого преобразования Фурье. Чаще всего используют алгоритмы по основанию 2^m с прореживанием по частоте или по времени. В обоих случаях входные или выходные данные оказываются переставлены в определенном порядке (см. Приложение 4). Для ускорения выборки (или записи) таких данных в сигнальных процессорах применяют так называемую бит-реверсивную адресацию. Суть ее состоит в перестановке в обратном порядке разрядов двоичного числа, указывающего номер отсчета соответствующей последовательности. В таблице приведены номера восьми отсчетов последовательности и результаты их бит-реверсии [8]. Видно, что бит-реверсия номеров отсчетов привела к их перестановке в соответствии с алгоритмом 8-точечного БПФ с прореживанием по времени.

Номера восьми отсчетов последовательности и результат их бит-реверсии

Порядок расположения отсчетов	Номера отсчетов							
	0	1	2	3	4	5	6	7
Прямой в десятичной системе	0	1	2	3	4	5	6	7
Прямой в двоичной системе	000	001	010	011	100	101	110	111
Бит-реверсивный в двоичной системе	000	100	010	110	001	101	011	111
Бит-реверсивный в десятичной системе	0	4	2	6	1	5	3	7

Теневые регистры. Дополнительным средством ускорения работы ЦСП является применение теневых регистров. Так называются ячейки памяти, которые могут обмениваться содержимым с внутренними регистрами ЦПЭ за один тактовый интервал, что ускоряет сохранение состояния процессора при переходе к подпро-

граммам обработки прерываний и восстановление после завершения подпрограмм [26].

Хост-интерфейс. Нередко ЦСП используются совместно с другими процессорами в составе мультипроцессорной системы, играя в этом случае роль подчиненного устройства по отношению к ведущему устройству, которое принято называть хост-машиной (хост-процессором). Хост-процессором может быть другой сигнальный процессор, универсальный компьютер, специализированный прибор цифровой обработки сигналов или память прямого доступа [8]. Сопряжение с хост-машиной осуществляется при помощи соответствующего интерфейсного устройства – порта. Порт хост-интерфейса содержит два комплекта регистров, один из которых отображается на адресное пространство памяти данных сигнального процессора, а другой доступен хост-процессору как набор ячеек памяти из его собственного адресного пространства. Порт хост-интерфейса является *асинхронным*, т. е. позволяет хост-машине обращаться к нему независимо от выполняемых сигнальным процессором действий.

Аналоговый интерфейс. Некоторые сигнальные процессоры содержат на кристалле средства сопряжения с аналоговыми устройствами, такие как аналого-цифровой (АЦП) и цифро-аналоговый (ЦАП) преобразователи, предварительный («антиэлайсинговый»^{*)} и интерполирующий фильтры нижних частот и т.п. Например, приборы ADSP-21msp58/59 фирмы Analog Devices имеют в своем составе два эллиптических фильтра нижних частот 10-го порядка (предварительный и интерполирующий) с граничной частотой полосы пропускания 3,7 кГц, два эллиптических фильтра верхних частот 4-го порядка, аналоговый сглаживающий фильтр и т.п. Такие ЦСП ориентированы на использование в телефонных каналах связи. Управление аналоговым интерфейсом осуществляется при помощи регистров, отображаемых на адресное пространство ЦСП.

9.2. СРЕДСТВА РАЗРАБОТКИ ПРИЛОЖЕНИЙ ДЛЯ СИГНАЛЬНЫХ ПРОЦЕССОРОВ

Цифровой сигнальный процессор представляет собой сложное программно-управляемое устройство. Успех разработчика аппара-

^{*)} От aliasing (англ.) – наложение; так называют ФНЧ, ограничивающий спектр сигнала перед дискретизацией во избежание наложения спектральных «хвостов».

туры определяется не только внутренней архитектурой выбранного ЦСП, но и наличием инструментария, позволяющего реализовать потенциальные возможности архитектуры в конкретном устройстве. Поэтому фирмы-производители выпускают кроме ЦСП также инженерные средства создания приложений на их основе. Как правило, они включают следующие компоненты.

С-компилятор. Составление программы для сигнального процессора на его внутреннем (машинном) языке представляет собой весьма утомительный процесс, требующий высокой квалификации и особого внимания. Для облегчения этой работы создаются программные средства, называемые компиляторами и выполняющие перевод программы, написанной на языке программирования высокого уровня, на машинный язык конкретного процессора. Среди языков высокого уровня язык С является наиболее удобным для программирования ЦСП, так как он предоставляет программисту возможность обращаться к программно-доступным элементам архитектуры. Поэтому С-компилятор является незаменимым средством разработки программ для ЦСП.

Ассемблер. Язык ассемблера служит промежуточным звеном между языком внутреннего представления ЦСП (машинными кодами) и языком высокого уровня. С одной стороны, язык ассемблера ориентирован на архитектуру конкретного процессора и позволяет реализовать все его потенциальные возможности; с другой стороны, удобные для запоминания (мнемонические) обозначения команд облегчают кодирование программ. Программируя на языке ассемблера, можно не заботиться о распределении памяти, назначении физических адресов операндов. Кроме того, язык ассемблера, как правило, содержит *макросредства*, позволяющие оформлять повторяющиеся последовательности действий в виде макрокоманд. Эффективность получаемых программ близка к потенциально достижимой, чего нельзя сказать о программах, полученных с помощью С-компилятора. В любом случае знание языка ассемблера является необходимым условием для разработки устройств ЦОС с применением цифровых сигнальных процессоров.

Преобразование программы на языке ассемблера в исполняемый код осуществляется в два этапа. На первом этапе исходная программа транслируется программой-*ассемблером* в объектный код. Для получения исполняемого кода используется компоновщик (заметим, что некоторые ассемблеры позволяют получить сразу исполняемый код [8]).

Компоновщик. Назначение компоновщика (построителя, редактора связей) состоит в построении исполняемой программы на ос-

нове *объектного* кода, созданного ассемблером, а также библиотечных модулей. Возможно построение исполняемого кода из нескольких независимо разработанных и оттранслированных программных модулей. Объектные коды, полученные при ассемблировании, не привязаны к конкретным физическим адресам и являются перемещаемыми. Компоновщик назначает соответствие между относительными адресами объектных кодов и физическими (абсолютными) адресами памяти и обеспечивает перекрестные ссылки между различными модулями. Получаемый исполняемый (абсолютный) код может быть загружен в оперативную память для исполнения или записан в ПЗУ при помощи программатора.

Отладчик. Отладчик – это программное средство, позволяющее выполнять программу в режиме, удобном для поиска и устранения ошибок. Некоторые процессоры обеспечивают *аппаратную* поддержку отладочных режимов (контроль регистров, трассировку, контроль точек останова и т.п.) [6].

Симулятор. Симулятор представляет собой программу, *имитирующую* архитектуру сигнального процессора и позволяющую смоделировать его поведение при исполнении разработанных программ. Наличие симулятора позволяет разработать и полностью отладить программу до ее перенесения в память ЦСП, что дает возможность избежать материальных и временных потерь из-за ошибочных решений.

Отладочный модуль. Фирмы-производители выпускают отладочные модули (например, EVM56K фирмы Motorola [6] или TMS320C6201 фирмы Texas Instruments [25]), включающие сигнальные процессоры и вспомогательные устройства и позволяющие выполнять отлаживаемые программы в реальном масштабе времени.

9.3. АЛЬТЕРНАТИВА СИГНАЛЬНЫМ ПРОЦЕССОРАМ – ПРОГРАММИРУЕМЫЕ ЛОГИЧЕСКИЕ ИНТЕГРАЛЬНЫЕ СХЕМЫ

В последнее время в устройствах цифровой обработки сигналов широкое распространение получили *программируемые логические интегральные схемы*. ПЛИС представляет собой интегральную схему сверхвысокой степени интеграции, содержащую на кристалле порядка одного миллиона логических вентилях, которые могут быть *программным* путем соединены в логическую структуру, реализующую заданный алгоритм цифровой обработки сигналов *аппаратным способом*. Таким образом, ПЛИС сочетают в себе

преимущества аппаратной реализации алгоритмов (главное из которых – быстродействие) с достоинствами программируемых устройств.

Классификацию ПЛИС по виду, количеству и способам соединения (коммутации) элементов можно найти в [27].

Программируемые логические матрицы (ПЛИМ) (в зарубежной литературе называемые FPLA – Field Programmable Logic Array или FPLS – Field Programmable Logic Sequencer) – наиболее традиционный вид ПЛИС – представляют собой массивы логических элементов типа «И» и «ИЛИ». Дальнейшее развитие эта архитектура получила в микросхемах *программируемой матричной логики (ПМЛ)* (PAL – Programmable Array Logic), имеющих программируемую матрицу элементов «И» и фиксированную матрицу элементов «ИЛИ», а иногда – только программируемую матрицу элементов «И». Следующая ступень в развитии этого типа ПЛИС – программируемая макрологика, для которой характерно наличие единственной матрицы элементов «И-НЕ» или «ИЛИ-НЕ», позволяющей реализовать сложные логические функции за счет многочисленных обратных связей. Все эти типы ПЛИС в настоящее время морально устарели [27] и применяются в основном для реализации в устройствах тех функций, для которых отсутствуют серийные микросхемы средней степени интеграции.

Программируемые коммутируемые матричные блоки (ПКМБ) содержат несколько матричных логических блоков (МЛБ), соединенных коммутационной матрицей. В состав МЛБ входят программируемые матрицы элементов «И», фиксированные матрицы элементов «ИЛИ» и макроячейки. Степень интеграции таких ПЛИС доходит до 10000 эквивалентных вентилях. В зарубежной литературе они имеют название CPLD – Complex Programmable Logic Devices.

Программируемые вентилярные матрицы (ПВМ) (FPGA – Field Programmable Gate Array) состоят из логических блоков и программируемых коммутирующих путей. Логический блок может включать таблицы перекодировки, мультиплексоры, триггеры, элементы управления и т. п. Таким образом, каждый блок может реализовать операции типа перемножения, умножения на константу, задержки, а наличие во многих ПЛИС этого типа реконфигурируемых модулей памяти делает их удобным средством реализации алгоритмов цифровой обработки сигналов, особенно цифровой фильтрации.

Дальнейшее развитие архитектур ПЛИС идет в направлении введения двух уровней матриц соединений: локальных, которые

соединяют логические элементы в логические блоки, и глобальных, соединяющих блоки между собой и с портами ввода-вывода. Такие ПЛИС пригодны для реализации «систем на кристалле» (System on Chip, SOC), представляющих серьезную альтернативу сигнальным процессорам.

Что особенно важно, фирмы-производители ПЛИС выпускают программное обеспечение, позволяющее выполнять сквозное проектирование устройств цифровой обработки сигналов. Примером может служить система проектирования MAX+PLUS фирмы ALTERA [27]). Система проектирования версии MAX+PLUS II включает *графический редактор* для задания электрической принципиальной схемы, *текстовый редактор* для описания проекта, *сигнальный редактор* для задания временных диаграмм, *поуровневый планировщик* для размещения узлов проектируемой системы по логическим блокам. Для испытания и отладки полученной конфигурации служат *симулятор*, *сигнальный редактор* и *временной анализатор*. *Компилятор* выполняет логический синтез, трассировку соединений и формирует файл программирования или загрузки. *Программатор* выполняет с помощью соответствующего оборудования программирование ПЛИС, обеспечивая реализацию разработанного проекта на кристалле. В целом подобное программное обеспечение представляет собой незаменимый инструмент автоматизированного проектирования систем цифровой обработки сигналов на элементной базе ПЛИС.

Вопросы для самоконтроля

1. Перечислите основные отличия сигнальных процессоров от микропроцессоров широкого применения.
2. Что такое кольцевой буфер и для чего он используется?
3. Для чего применяется бит-реверсивная адресация?
4. В чем состоит отличие принстонской и гарвардской архитектур? Какая из них используется для построения сигнальных процессоров и почему?
5. Что такое конвейерный принцип обработки и как он реализуется в сигнальных процессорах?
6. Что такое ассемблер? Что такое язык ассемблера и какое положение он занимает в иерархии языков программирования?
7. Что такое ПЛИС?
8. Каковы преимущества ПЛИС перед сигнальными процессорами?



10. ЦИФРОВАЯ ОБРАБОТКА ВРЕМЕННЫХ И ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ СИГНАЛОВ В СИСТЕМАХ ПОДВИЖНОЙ РАДИОСВЯЗИ

Цифровая обработка сигналов благодаря своим неоспоримым преимуществам перед аналоговой обработкой все шире распространяется в практике средств связи. В одном разделе невозможно даже кратко рассмотреть все применения ЦОС в системах подвижной связи, поэтому рассматриваются лишь некоторые операции обработки сигналов, реализуемые в цифровой форме.

10.1. ЦИФРОВАЯ СОГЛАСОВАННАЯ ФИЛЬТРАЦИЯ

Одна из самых распространенных операций в системах радиолокации и связи – согласованная фильтрация, предназначенная для максимального повышения отношения сигнал/шум с целью надежного обнаружения сигнала [28]. Сигнал может быть простым, как радиоимпульс с прямоугольной огибающей, или сложным, как, например, радиоимпульс с внутренней частотной модуляцией по линейному закону (ЛЧМ-сигнал) или псевдослучайная последовательность с частотной или фазовой манипуляцией элементарных посылок. Сложные сигналы обладают способностью к сжатию, связанному с большой базой $B = \tau \Delta f$, где τ – длительность сигнала, Δf – ширина его спектра. Реализовать эту способность можно только путем согласованной фильтрации (или эквивалентной ей корреляционной обработки).

Известно [13], что фильтр, согласованный с аналоговым сигналом $s(t)$, имеет импульсную характеристику вида $h(t) = s(t_0 - t)$, где t_0 – задержка, вводимая для обеспечения каузальности фильтра. Согласованный фильтр обеспечивает на выходе наибольшее (среди всех возможных ЛИС-фильтров) значение отношения сиг-

нал/шум в момент времени t_0 при условии, что на входе действует аддитивная смесь данного сигнала $s(t)$ с белым шумом. Со спектральной точки зрения согласованный фильтр описывается комплексной частотной характеристикой $H(j\Omega) = S^*(j\Omega)e^{-j\Omega t_0}$, где $S^*(j\Omega)$ – функция, комплексно-сопряженная по отношению к спектральной плотности $S(j\Omega)$ сигнала $s(t)$. Форма временного отклика согласованного фильтра на произвольный детерминированный сигнал повторяет форму взаимной корреляционной функции этого сигнала с сигналом, по отношению к которому данный фильтр является согласованным. Таким образом, если фильтр согласован с входным сигналом, его отклик по форме повторяет автокорреляционную функцию (АКФ) этого сигнала. Сложные сигналы, применяемые в системах связи и локации, обычно имеют «игольчатую» АКФ, основной пик которой имеет ширину примерно в B раз меньше длительности сигнала, что повышает точность измерения дальности и разрешающую способность по дальности (в радиолокации), обеспечивает высокую помехоустойчивость и облегчает синхронизацию (в системах связи).

Рассмотрим случай, когда аналоговый видеосигнал конечной длительности и известной формы, наблюдаемый на фоне белого шума, подвергается дискретизации и последующей согласованной фильтрации. Перед дискретизацией наблюдаемая смесь подвергается НЧ-фильтрации для уменьшения (в идеале – исключения) искажения сигнала вследствие наложения спектральных «хвостов». Если предположить, что спектр сигнала финитен, то фильтрация все же целесообразна для повышения отношения сигнал/шум перед аналого-цифровым преобразованием, что улучшает использование динамического диапазона АЦП. Предположим, что ФНЧ имеет П-образную АЧХ с полосой пропускания $(-B, B)$, а частота

дискретизации $\Omega_d = \frac{2\pi}{T_d} = 2B$. Шум на входе белый со спектральной плотностью N_0 , поэтому шум на выходе ФНЧ имеет дисперсию

$$\sigma^2 = \frac{1}{2\pi} \int_{-B}^B N_0 d\Omega = \frac{BN_0}{\pi}.$$

После дискретизации смесь сигнала с шумом поступает на КИХ-фильтр с импульсной характеристикой $h[n]$, $n = \overline{0, N-1}$. Найдем вид импульсной характеристики фильтра, который обеспечивает максимальное отношение сигнал/шум

при воздействии на вход смеси сигнала $x[n]$, $n = \overline{0, N-1}$ с шумом. Для этого запишем значение выходного сигнала КИХ-фильтра в момент окончания входного сигнала, т.е. при $n = N-1$, как свертку

$$y[N-1] = \sum_{k=0}^{N-1} h[k]x[N-1-k].$$

Шумовые отсчеты на входе КИХ-фильтра являются некоррелированными, так как автокорреляционная функция шума, белого в полосе $(-B, B)$, проходит через нуль в моменты взятия отсчетов, поэтому дисперсия шума на выходе КИХ-фильтра находится как дисперсия взвешенной суммы некоррелированных случайных величин с одинаковой дисперсией:

$$\sigma_{\text{вых}}^2 = \sum_{k=0}^{N-1} h^2[k]\sigma^2 = \sigma^2 \sum_{k=0}^{N-1} h^2[k] = \frac{BN_0}{\pi} \sum_{k=0}^{N-1} h^2[k].$$

Отношение сигнал/шум (по мощности) на выходе КИХ-фильтра

$$\eta = \frac{\left(\sum_{k=0}^{N-1} h[k]x[N-1-k] \right)^2}{\frac{BN_0}{\pi} \sum_{k=0}^{N-1} h^2[k]}$$

будет максимальным для фильтра, импульсная характеристика которого удовлетворяет системе уравнений, полученной приравнованием нулю производных η по всем $h[n]$, $n = \overline{0, N-1}$:

$$\left\{ \left(\sum_{k=0}^{N-1} h[k]x[N-1-k] \right) x[N-1-m] \sum_{k=0}^{N-1} h^2[k] - \right. \\ \left. -h[m] \left(\sum_{k=0}^{N-1} h[k]x[N-1-k] \right)^2 = 0, \quad m = \overline{0, N-1} \right\}.$$

Видно, что при $h[k] = x[N-1-k]$ эти уравнения обращаются в равенства, поэтому согласованный фильтр, для которого выполне-

но условие $h[k] = x[N-1-k]$, $k = \overline{0, N-1}$, обеспечивает максимальное отношение сигнал/шум, равное

$$\eta_{\max} = \frac{\left(\sum_{k=0}^{N-1} h^2[k] \right)^2}{\frac{BN_0}{\pi} \sum_{k=0}^{N-1} h^2[k]} = \frac{\sum_{k=0}^{N-1} h^2[k]}{\frac{BN_0}{\pi}} = \frac{\|x\|_2^2}{\sigma^2},$$

где $\|x\|_2^2$ – энергия сигнала на входе согласованного фильтра. Если предположить, что согласованной фильтрации подвергается дискретный импульс с прямоугольной огибающей в смеси с шумом, то, очевидно, согласованный фильтр повышает отношение сигнал/шум в N раз (сравните с аналоговым согласованным фильтром!).

Следует отметить, что на практике, как правило, сигналы имеют большую длительность, поэтому с вычислительной точки зрения наиболее эффективна реализация согласованного фильтра методом быстрой свертки [28]. Заметим также, что для *любого* дискретного сигнала конечной длительности согласованный фильтр реализуем, чего нельзя сказать об аналоговых согласованных фильтрах.

10.2. ИНТЕРПОЛЯЦИЯ И ДЕЦИМАЦИЯ ДИСКРЕТНЫХ СИГНАЛОВ

Интерполяция и децимация (иначе – повышение и понижение частоты следования отсчетов дискретных сигналов) широко применяются в системах цифровой связи. Это вызвано необходимостью сопряжения систем обработки и передачи цифровых сигналов, работающих с различными тактовыми частотами [5].

Интерполировать можно любой дискретный сигнал. Основанием для такого утверждения служит то обстоятельство, что по дискретному сигналу, полученному дискретизацией аналогового сигнала с соблюдением условий теоремы отсчетов, можно в принципе *точно* восстановить исходный аналоговый сигнал. Этот восстановленный аналоговый сигнал может быть снова продискретизирован с повышенной частотой дискретизации. Децимация, или понижение частоты, может быть выполнена без потери информации только в том случае, если первоначальная дискретизация была

выполнена «с запасом» и частота повторной дискретизации не меньше удвоенной верхней частоты аналогового сигнала.

Интерполяция может быть представлена в виде процедуры, выполняемой в два этапа (рис. 10.1): сначала между отсчетами сигнала вставляются нулевые отсчеты, а затем полученный сигнал пропускается через интерполирующий фильтр. Устройство, увеличивающее частоту отсчетов в целое число m раз путем вставления между отсчетами сигнала $(m-1)$ нулевых отсчетов, называется

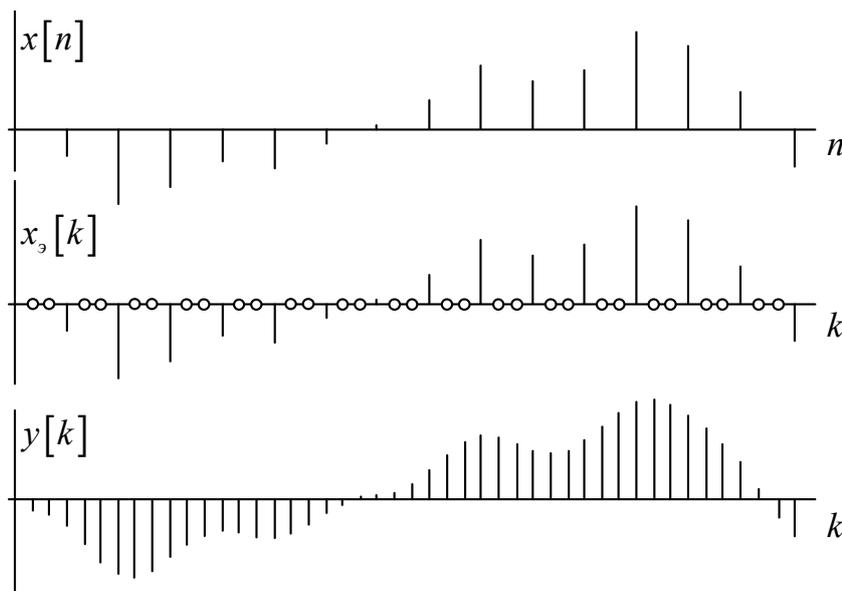


Рис. 10.1. Интерполяция последовательности

экспандером частоты дискретизации. На рис. 10.2 экспандер обозначен символом $m \uparrow$. Предположим, что исходным является дискретный сигнал $x[n], n = \overline{-\infty, \infty}$ с z -образом $X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n}$.

Сигнал на выходе экспандера $x_3[k], k = \overline{-\infty, \infty}$ связан с исходным сигналом следующим условием:

$$x_3[k] = \begin{cases} x[n] & \text{при } k = mn, \\ 0 & \text{в противном случае.} \end{cases}$$

z -Образ выходного сигнала, очевидно, равен

$$X_3(z) = \sum_{k=-\infty}^{\infty} x_3[k]z^{-k} = \sum_{n=-\infty}^{\infty} x[n]z^{-mn} = X(z^m),$$

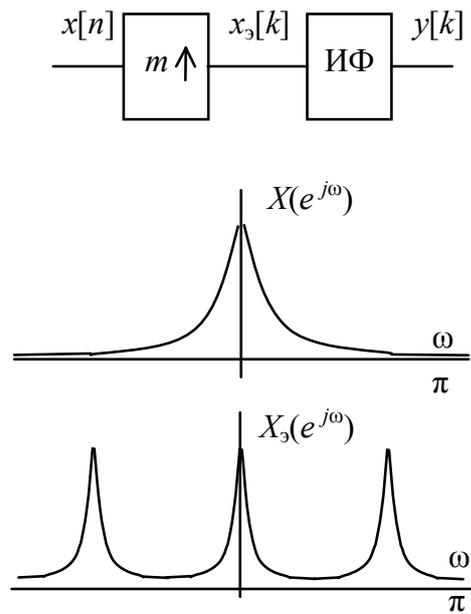


Рис. 10.2. Изменение спектра сигнала при интерполяции

откуда следует, что его спектральная плотность $X_3(e^{j\omega}) = X(e^{jm\omega})$. Если исходная спектральная плотность имела период 2π , то период спектральной плотности выходного сигнала экспандера будет равен $2\pi/m$ (рис. 10.2). Если теперь на такой сигнал воздействовать дискретным фильтром нижних частот с П-образной АЧХ и частотой среза π/m , то, очевидно, получится дискретный сигнал, эквивалентный результату дискретизации аналогового сигнала с частотой дискретизации, в m раз превышающей частоту Найквиста, т. е. именно тот сигнал, который отвечает введенному понятию интерполяции. Следует заметить, что

фильтр с П-образной АЧХ нереализуем, поэтому реальная интерполяция всегда будет неточной.

Интерполяция может быть также выполнена с применением дискретного преобразования Фурье. Предположим, что имеется последовательность $x[n], n = 0, N-1$ конечной длины N . Дискретное преобразование от нее дает последовательность спектральных отсчетов такой же длины. Если ее дополнить $(m-1)N$ нулевыми отсчетами, то получится последовательность длины mN , которая при обратном ДПФ дает временную последовательность длины $M = mN$. Нетрудно видеть, что полученная последовательность удовлетворяет введенному выше понятию интерполяции и соответствует дискретизации аналогового сигнала с частотой, в m раз превышающей частоту Найквиста. Если исходная последовательность имеет большую длительность, ее можно разбить на секции приемлемой длины и интерполяцию производить для каждой секции в отдельности. Очевидно, таким способом можно выполнять интерполяцию с любым рациональным коэффициентом M/N , а не только с целым.

Децимация сигнала в простейшем случае выполняется в устройстве, называемом *компрессором* частоты дискретизации, путем отбрасывания части отсчетов. При децимации в m раз из группы, содержащей m отсчетов, оставляется один, а $(m-1)$ отсчетов ис-

ключаются. Отсчеты результирующего сигнала следуют с частотой, в m раз меньшей первоначальной. Ясно, что результат получается в точности такой же, как если бы частота дискретизации сразу была выбрана в m раз меньше. Поэтому очевидно, что децимация не сопровождается потерей информации только при условии, что вторичная частота следования удовлетворяет условиям теоремы отсчетов, иначе неизбежны искажения вследствие наложения спектральных «хвостов».

Подробное обсуждение интерполяции и децимации можно найти в [5].

10.3. ПЕРЕНОС СПЕКТРА ДИСКРЕТНОГО СИГНАЛА

Перенос спектра сигнала широко используется в цифровых системах связи [5, 6] с частотным разделением каналов (ЧРК). Для переноса спектра вещественного сигнала на величину γ вправо по оси частот необходимо умножить этот сигнал на комплексную экспоненциальную последовательность $\exp(j\gamma n)$, $n = -\infty, \infty$.

Действительно, если сигнал $x[n]$ имеет спектральную плотность $X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}$, то спектральная плотность произведения $x[n]\exp(j\gamma n)$, $n = -\infty, \infty$ равна

$$Y(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{j\gamma n}e^{-j\omega n} = \sum_{n=-\infty}^{\infty} x[n]e^{-j(\omega-\gamma)n} = X(e^{j(\omega-\gamma)}),$$

что и означает сдвиг спектральной плотности сигнала вправо на γ .

В реальных устройствах сигналы принимают только вещественные значения, поэтому комплексную последовательность $x[n] = x_r[n] + jx_i[n]$ представляют двумя вещественными последовательностями $x_r[n]$ и $x_i[n]$. Устройства обработки комплексных последовательностей являются фактически

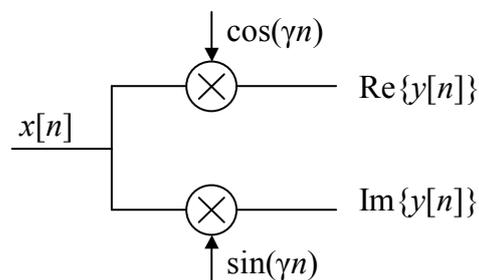


Рис. 10.3. Умножение вещественной последовательности на комплексную экспоненту

двухканальными, причем обработка в каналах должна быть согласованной в соответствии с правилами комплексной арифметики.

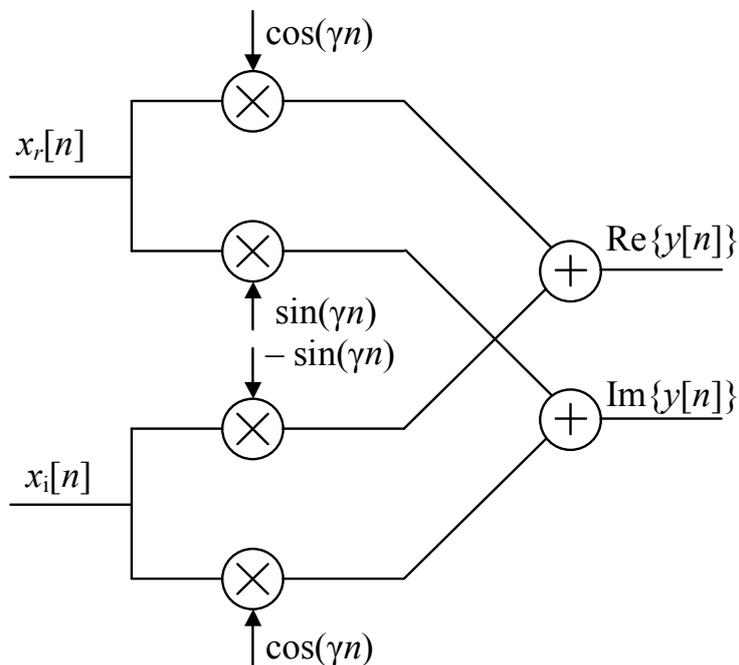


Рис. 10.4. Перенос спектра комплексной последовательности

Умножение вещественной последовательности $x[n]$ на комплексную экспоненциальную последовательность выполняется схемой, представленной на рис. 10.3. На рис. 10.4 показана цепь, выполняющая перенос спектра комплексной последовательности $x[n] = x_r[n] + jx_i[n]$ на величину γ .

10.4. ФОРМИРОВАНИЕ СИГНАЛА С ОДНОЙ БОКОВОЙ ПОЛОСОЙ

Известно, что спектральная плотность вещественного сигнала является сопряженно-симметричной, т. е. имеет четную вещественную и нечетную мнимую части. Вследствие этого сигнал может быть восстановлен по одной *боковой полосе*, например, по половине его спектра, лежащей в области положительных частот. Эта симметрия спектра, позволяющая вдвое сузить полосу частот сигнала без потери информации, используется в системах связи с частотным разделением для увеличения числа каналов.

Один из способов формирования сигнала с одной боковой полосой (ОБП-сигнала) реализуется устройством, представленным на

рис. 10.5, а. Вначале спектр сигнала переносится влево на величину γ , равную полусумме верхней и нижней граничных частот спектра. Затем производится обработка полученного сигнала фильтром нижних частот с соответствующей полосой пропускания. Последней операцией является перенос спектра вправо на величину γ .

Более простой способ формирования ОБП-сигнала реализуется с использованием дискретного преобразователя Гильберта ПГ (Приложение 7) устройством, представленным на рис. 10.5, б.

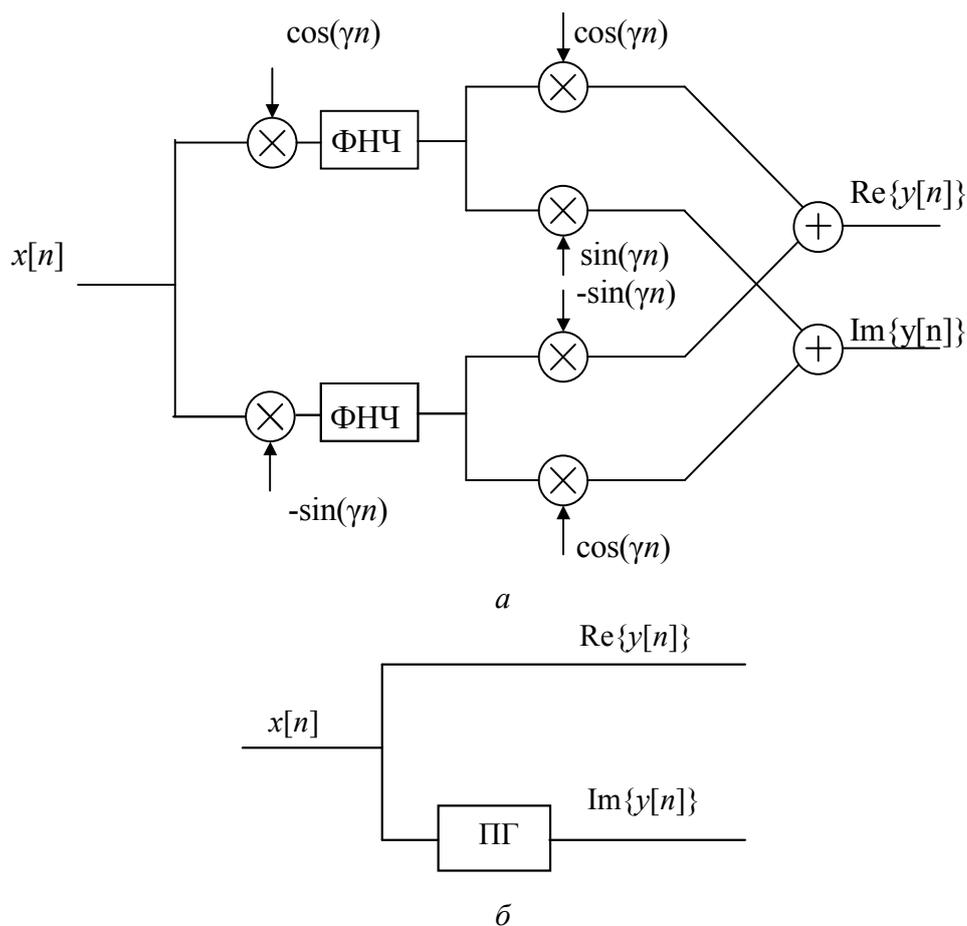


Рис. 10.5. Формирование ОБП-сигнала

10.5. ЦИФРОВАЯ ДЕМОДУЛЯЦИЯ СИГНАЛОВ

В системах радиолокации, радионавигации и связи цифровая обработка сигналов все чаще применяется для выполнения функций, традиционных для аналоговых радиотехнических устройств. Рассмотрим один из вариантов цифрового устройства демодуляции амплитудно-модулированного (АМ) сигнала [29].

Демодуляция АМ-сигнала состоит в формировании колебания, пропорционального его огибающей, поэтому первый этап обработки представляет собой синхронное детектирование квадратурных компонент АМ-сигнала и может быть выполнен в аналоговой форме (рис. 10.6).

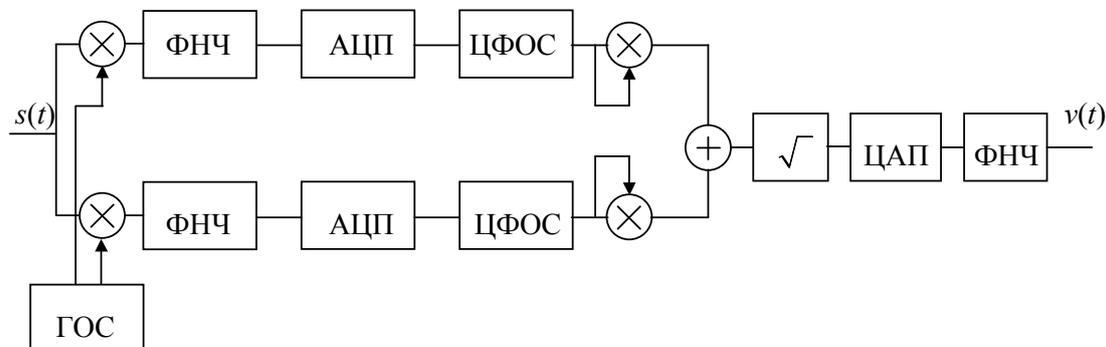


Рис. 10.6. Демодуляция амплитудно-модулированного сигнала

Генератор опорных сигналов (ГОС) вырабатывает два гармонических колебания с разностью фаз 90° , каждое из которых после умножения на входной АМ-сигнал $s(t)$ обрабатывается аналоговым фильтром нижних частот, где происходит подавление суммарной частоты (напомним, что при умножении двух гармонических колебаний появляются комбинационные частоты, из которых интерес в данном случае представляют разностные). Квадратурные компоненты сигнала поступают на аналого-цифровой преобразователь, а затем – на цифровые фильтры основной селекции ЦФОС, полоса пропускания которых определяется шириной спектра сообщения. Квадратурные последовательности $U_c[n]$ и $U_s[n]$ преобразуются согласно выражению $U[n] = \sqrt{U_c^2[n] + U_s^2[n]}$, результат подвергается цифроаналоговому преобразованию, сглаживается в ФНЧ, и выделенная огибающая $v(t)$ передается получателю.

10.6. ЦИФРОВАЯ ОБРАБОТКА ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ СИГНАЛОВ

Пространственно-временной аналоговый сигнал $s(\mathbf{x}, t)$ представляет собой функцию четырех непрерывных аргументов, один из которых – время, а три остальных – пространственные координаты, составляющие координатный вектор $\mathbf{x} = (x_1, x_2, x_3)^T$. Сигнал

$s(\mathbf{x}, t)$ можно представить суперпозицией плоских волн при помощи четырехмерного обратного преобразования Фурье [16]

$$s(\mathbf{x}, t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\mathbf{k}, \Omega) \exp[j(\Omega t - \mathbf{k}^T \mathbf{x})] d\mathbf{k} d\Omega. \quad (10.1)$$

Здесь $\mathbf{k} = (k_1, k_2, k_3)^T$ – вектор-столбец, называемый волновым вектором и представляющий собой пространственный аналог временной частоты. Компоненты вектора \mathbf{k} показывают количество волн, укладываемых на отрезках длины 2π вдоль каждого из трех взаимно перпендикулярных направлений в трехмерном пространстве, так же, как круговая частота Ω равна числу периодов, укладываемых на временном отрезке длительностью 2π . Спектральная плотность пространственно-временного сигнала определяется прямым преобразованием Фурье

$$S(\mathbf{k}, \Omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\mathbf{x}, t) \exp[-j(\Omega t - \mathbf{k}^T \mathbf{x})] d\mathbf{x} dt. \quad (10.2)$$

Точка (\mathbf{k}_0, Ω_0) в частотном пространстве (\mathbf{k}, Ω) соответствует набору пространственно-временных частот плоской элементарной волны $e(\mathbf{x}, t) = \exp[j(\Omega_0 t - \mathbf{k}_0^T \mathbf{x})]$, которая имеет спектральную плотность $E(\mathbf{k}, \Omega) = (2\pi)^4 \delta(\mathbf{k} - \mathbf{k}_0) \delta(\Omega - \Omega_0)$, в чем легко убедиться, подставив $E(\mathbf{k}, \Omega)$ в выражение (10.1).

Линейный инвариантный к сдвигу (в пространстве и во времени) фильтр описывается импульсной характеристикой $h(\mathbf{x}, t)$, поэтому отклик фильтра на сигнал $s(\mathbf{x}, t)$ определяется в пространственно-временной области сверткой

$$f(\mathbf{x}, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\mathbf{x} - \xi, t - \tau) s(\xi, \tau) d\xi d\tau, \quad (10.3)$$

а в частотной области – произведением $F(\mathbf{k}, \Omega) = H(\mathbf{k}, \Omega) S(\mathbf{k}, \Omega)$. На практике чаще всего требуется из всех пространственно-временных сигналов выделить сигналы, пришедшие с направления, определяемого вектором \mathbf{k}_0 , и содержащие частоты, близкие

к Ω_0 , что соответствует полосовому фильтру с центральной частотой (\mathbf{k}_0, Ω_0) .

Из теории и практики обработки одномерных (в частности, временны́х) сигналов известно, что далеко не всякое преобразование сигналов может быть реализовано. Одно из ограничений такого рода заключается в требовании каузальности цепи. Второе – не всегда осознаваемое – ограничение состоит в неизбежной конечности времени обработки. При рассмотрении пространственно-временных сигналов следует отдавать себе отчет в том, что для наблюдения может быть доступна только часть сигнала, ограниченная как временными, так и пространственными рамками – размерами антенной системы. Если антенна к тому же представляет собой *фазированную антенную решетку* (ФАР), то применение такой антенны для приема пространственно-временного сигнала соответствует его *дискретизации по пространству*. Антенные решетки широко применяются в радиотехнике, гидролокации, в системах сейсмической разведки и т.п.

Предположим, что антенна состоит из N точечных элементов (датчиков), расположенных в пространстве так, что i -й датчик находится в точке $\mathbf{x}_i, i = 0, N-1$. Сигнал на выходе i -го датчика $r_i(t) = s(\mathbf{x}_i, t)$ представляет собой одномерное сечение (пространственный отсчет) пространственно-временного сигнала. Простейший способ формирования выходного сигнала антенной решетки состоит во взвешенном суммировании задержанных сигналов отдельных элементов решетки $f_p(t) = \sum_{i=0}^{N-1} w_i r_i(t - \tau_i)$, где весовой коэффициент и задержка для i -го датчика обозначены соответственно w_i и τ_i .

Отклик решетки на плоскую волну $e(\mathbf{x}, t) = \exp[j(\Omega t - \mathbf{k}^T \mathbf{x})]$ описывается выражением

$$\begin{aligned} f_p(t) &= \sum_{i=0}^{N-1} w_i r_i(t - \tau_i) = \sum_{i=0}^{N-1} w_i e(\mathbf{x}_i, t - \tau_i) = \\ &= \sum_{i=0}^{N-1} w_i \exp\left\{j\left[\Omega(t - \tau_i) - \mathbf{k}^T \mathbf{x}_i\right]\right\}. \end{aligned} \quad (10.4)$$

Задавая задержки $\tau_i = -\frac{1}{\Omega_0} \mathbf{k}_0^T \mathbf{x}_i$, где \mathbf{k}_0 и Ω_0 – волновой вектор и частота элементарной плоской волны, на которую настроена решетка, получим

$$\begin{aligned} f_p(t) &= \sum_{i=0}^{N-1} w_i \exp \left\{ j\Omega t + j \frac{\Omega}{\Omega_0} \mathbf{k}_0^T \mathbf{x}_i - j\mathbf{k}^T \mathbf{x}_i \right\} = \\ &= \exp(j\Omega t) \sum_{i=0}^{N-1} w_i \exp \left[-j \left(\mathbf{k} - \frac{\Omega}{\Omega_0} \mathbf{k}_0 \right)^T \mathbf{x}_i \right]. \end{aligned} \quad (10.5)$$

Можно записать сигнал на выходе решетки в виде

$$f_p(t) = W \left(\mathbf{k} - \frac{\Omega}{\Omega_0} \mathbf{k}_0 \right) \exp(j\Omega t),$$

где функция

$$W(\mathbf{k}) = \sum_{i=0}^{N-1} w_i \exp(-j\mathbf{k}^T \mathbf{x}_i) \quad (10.6)$$

представляет собой диаграмму направленности антенной решетки. Функция $W(\mathbf{k} - \Omega/\Omega_0 \mathbf{k}_0)$ показывает ослабление плоской волны с частотой Ω и значением волнового вектора \mathbf{k} в антенной решетке, настроенной на прием плоской волны, определяемой вектором \mathbf{k}_0 и частотой Ω_0 . Очевидно, (10.6) представляет собой преобразование Фурье для пространственной последовательности (пространственного дискретного сигнала), состоящей из отсчетов w_i , размещенных в пространстве в точках \mathbf{x}_i , $i = \overline{0, N-1}$.

В общем случае пространственно-временной сигнал представляет собой суперпозицию плоских волн

$$s(\mathbf{x}, t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\mathbf{k}, \Omega) \exp \left[j(\Omega t - \mathbf{k}^T \mathbf{x}) \right] d\mathbf{k} d\Omega,$$

поэтому выражение для выходного сигнала можно записать в виде

$$f_p(t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\mathbf{k}, \Omega) W \left(\mathbf{k} - \frac{\Omega}{\Omega_0} \mathbf{k}_0 \right) \exp(j\Omega t) d\mathbf{k} d\Omega.$$

Это выражение описывает *пространственно-временную фильтрацию* сигнала четырехмерным фильтром, обладающим избира-

тельностью по трем компонентам волнового вектора и временно, ω частоте и описываемым функцией

$$H(\mathbf{k}, \Omega) = W\left(\mathbf{k} - \frac{\Omega}{\Omega_0} \mathbf{k}_0\right).$$

Рассмотрим простейший случай решетки, состоящей из датчиков, размещенных равномерно (с шагом D) вдоль одной из координатных осей, например, так, что $\mathbf{x}_i = (iD, 0, 0)^T, i = \overline{0, N-1}$, при равных весах элементов $w_i = 1, i = \overline{0, N-1}$. В этом случае

$$\begin{aligned} W(\mathbf{k}) &= \sum_{i=0}^{N-1} \exp(-j\mathbf{k}^T \mathbf{x}_i) = \sum_{i=0}^{N-1} \exp(-jk_1 iD) = \\ &= \frac{\sin(Nk_1 D/2)}{\sin(k_1 D/2)} \exp\left[\frac{-j(N-1)k_1 D}{2}\right]. \end{aligned} \quad (10.7)$$

Диаграмма направленности такой решетки, построенной по методу взвешенного сложения с задержкой, состоящей из 12 элементов, представлена на рис. 10.7. Видно, что она имеет значительные боковые лепестки. Для ослабления боковых лепестков можно использовать в качестве весовых множителей отсчеты какого-либо окна (см. п. 6.3.1), при этом происходит расширение главного лепестка диаграммы направленности.

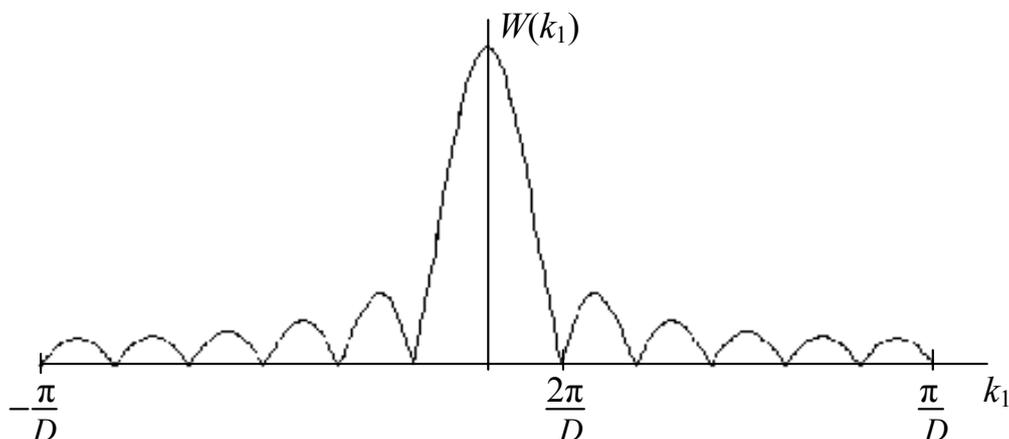


Рис. 10.7. Диаграмма направленности решетки из 12 элементов

Решетка более общего вида получается, если предположить, что каждый элемент решетки представляет собой временной фильтр с импульсной характеристикой $w_i(t)$, или, что эквивалент-

но, с комплексной частотной характеристикой $W_i(j\Omega)$. Тогда сиг-

нал на выходе решетки $f_p(t) = \sum_{i=0}^{N-1} q_i(t - \tau_i)$, где $q_i(t) = w_i(t) * r_i(t)$;

символ $*$ обозначает непрерывную свертку (интеграл Дюамеля). В частотной области этот сигнал описывается выражением

$$F_p(\Omega) = \sum_{i=0}^{N-1} W_i(\Omega) R_i(\Omega) \exp(-j\Omega\tau_i). \quad (10.8)$$

В этом случае пространственно-временная комплексная частотная характеристика принимает вид

$$H(\mathbf{k}, \Omega) = \sum_{i=0}^{N-1} W_i(j\Omega) \exp \left[-j \left(\mathbf{k} - \frac{\Omega}{\Omega_0} \mathbf{k}_0 \right)^T \mathbf{x}_i \right].$$

До сих пор рассматривалась обработка пространственно-временных сигналов, дискретизированных по пространственным координатам за счет использования антенной *решетки*. Теперь будут затронуты эффекты, связанные с дискретизацией по времени. Заменяя время t , а также задержки, используемые при формировании диаграммы направленности, на величины, кратные временно-му шагу дискретизации T_d , получим для решетки, построенной по методу взвешенного сложения с задержкой, выражение

$$f_d(nT_d) = \sum_{i=0}^{N-1} w_i r_i(nT_d - n_i T_d). \quad (10.9)$$

Вместо идеальных задержек $\tau_i = -\frac{1}{\Omega_0} \mathbf{k}_0^T \mathbf{x}_i$ в такой решетке реализуются задержки $n_i T_d$, что эквивалентно введению дополнительных задержек – ошибок $\Delta\tau_i = n_i T_d - \tau_i$, вследствие чего выражение (10.5) принимает форму

$$f_d(nT_d) = \exp(j\Omega n T_d) \sum_{i=0}^{N-1} w_i \exp \left[-j \left(\mathbf{k} - \frac{\Omega}{\Omega_0} \mathbf{k}_0 \right)^T \mathbf{x}_i \right] \exp(-j\Omega \Delta\tau_i).$$

Отсюда видно, что дискретизация пространственно-временного сигнала по *времени* приводит к искажению пространственно-временной комплексной частотной характеристики, а следовательно, и *диаграммы направленности*.

Уменьшить это искажение можно искусственным повышением частоты дискретизации сигнала (т. е. путем его интерполяции). При этом уменьшается T_d , а следовательно, и ошибки $\Delta\tau_i = n_i T_d - \tau_i$ могут быть уменьшены [16].

10.7. ЦИФРОВАЯ ОБРАБОТКА РЕЧЕВЫХ СИГНАЛОВ

Речевые сигналы являются предметом неослабевающего внимания инженеров и исследователей на протяжении нескольких последних десятилетий. Этот интерес обусловлен естественной ролью речи как способа передачи информации от человека к человеку, вследствие чего огромное количество сигналов, передаваемых по каналам связи, – это именно речевые сигналы. Кроме того, последние десятилетия XX века отмечены бурным развитием исследований по искусственному интеллекту, а одним из проявлений интеллекта, по общему мнению, является способность воспринимать речевые сигналы. Поэтому разработка систем, способных имитировать речевое общение, занимает одно из центральных мест в работах по созданию искусственного интеллекта. Ввиду сложности речевых сигналов алгоритмы их обработки также оказываются достаточно сложными, поэтому многие из них могут быть реализованы только в цифровом виде. Таким образом, обработка речевых сигналов является той областью, где преимущества цифровой техники перед аналоговой проявляются особенно отчетливо [30].

Термином «цифровая обработка речевых сигналов» принято обозначать совокупность различных теоретических исследований и практических методов, объединенных общим предметом – речевыми сигналами – и охватывающих различные задачи, включая цифровую *передачу и хранение* речевых сигналов, *синтез* искусственных речевых сигналов, *верификацию* и *идентификацию* дикторов, *распознавание* речи, *улучшение* качества речевых сигналов и т.п.

10.7.1. МОДЕЛИ РЕЧЕВЫХ СИГНАЛОВ

Речевые сигналы отличаются сложностью, случайностью и нестационарностью, что обусловлено их естественным происхождением. Для построения эффективных алгоритмов обработки необходимо иметь соответствующее математическое описание (модель) сигнала. Первые исследования речевых сигналов сводились к попыткам копирования естественных органов слуха человека и животных и опирались главным образом на изучение механизмов образования речи [30]. На рис. 10.8 представлена уп-

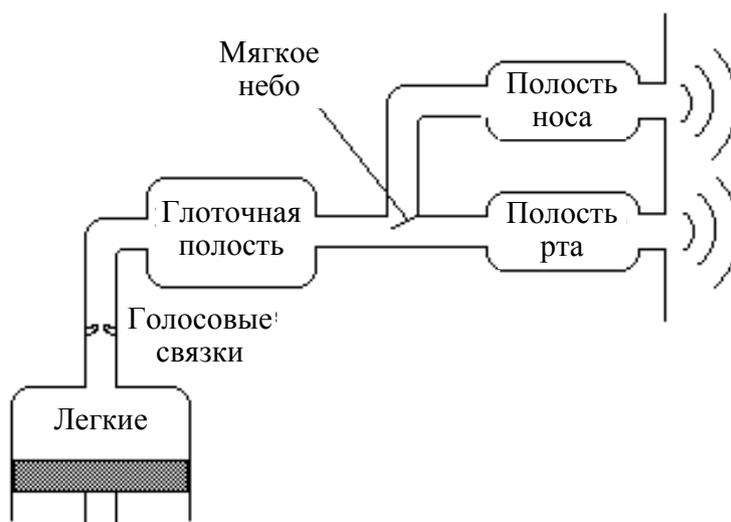


Рис. 10.8. Упрощенная физическая модель механизма речеобразования

рошенная физическая модель [30] механизма речеобразования, включающего легкие, выталкивающие воздух под действием сокращающихся мышц, голосовые связки, вибрирующие под воздействием проходящего через них воздушного потока, глоточную полость и полость рта, выполняющие функции объемного резонатора, и мягкое нёбо (нёбную занавеску), изменяющее объем и форму резонатора за счет полости носа. Обилие деталей голосового тракта и сложность их движения (*артикуляции*) в процессе речеобразования объясняют сложность речевых сигналов.

Существуют разнообразные подходы к описанию речи и, в частности, к членению ее на элементарные единицы. Согласно одному из подходов считается, что звуки речи делятся на три основные группы. *Вокализованные* (а, о, ...) звуки образуются при проталкивании воздуха через напряженные голосовые связки, которые модулируют звуковой поток последовательностью импульсов, возбуждающих объемный резонатор (голосовой тракт). *Фрикативные* или невокализованные звуки образуются при проталкивании воздуха через сужение голосового тракта в начале (х), середине (ш) или в конце (с, ф). При образовании *взрывных* звуков происходит полное смыкание тракта и повышение давления воздуха с последующим резким высвобождением воздушного потока (п, т, к, ч).

Такое деление звуков речи является весьма приблизительным и условным, так как существуют звуки, являющиеся одновременно фрикативными и вокализованными (в, ж), вокализованными и взрывными (г, д). Наиболее известным в настоящее время является

фонемное описание звуков речи, которое предполагает деление звучащей речи на *фонемы* – минимально различимые единицы. В последнее время преобладает точка зрения, согласно которой деление речи на фонемы не учитывает динамики живой речи, и предлагается делить речь на *фоноиды* или *аллофоны*, которые можно в первом приближении определить как фонемы, изменяющиеся под влиянием окружающих фонем.

В дальнейшем будет рассматриваться упрощенная модель речеобразования [30], предполагающая, что звуки речи различаются по механизму возбуждения (импульсная последовательность для вокализованных звуков и шум для фрикативных), а также по частотным свойствам голосового тракта. Произнесение вокализованных

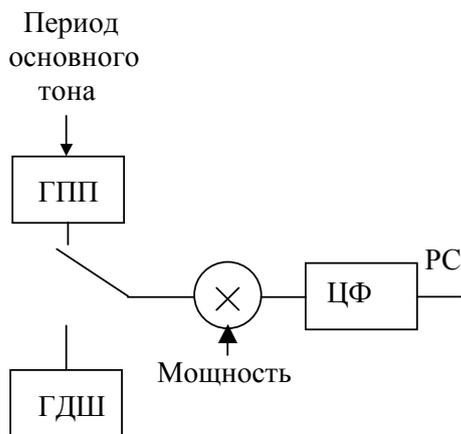


Рис. 10.9. Упрощенная математическая модель механизма речеобразования

звуков характеризуется выраженными резонансами голосового тракта, которые называются *формантами*. Такой предельно упрощенной модели соответствует эквивалентная схема, приведенная на рис. 10.9. Голосовой тракт, представленный в этой схеме цифровым фильтром с изменяемыми параметрами, возбуждается либо генератором периодической последовательности (ГПП), либо генератором дискретного шума (ГДШ). На генератор периодической последовательности подается внешний управляющий сигнал,

задающий период последовательности – период *основного тона*. Выходной сигнал ГПП представляет собой последовательность «единичных импульсов», следующих с частотой основного тона, которая медленно меняется со временем. Генератор дискретного шума вырабатывает стационарную случайную последовательность с независимыми отсчетами (дискретный белый шум); изменение мощности возбуждающего сигнала обеспечивается умножением его на соответствующий коэффициент.

Структуру речевого сигнала удобно анализировать с помощью *спектрограммы* [28], представляющей собой двумерный график, по горизонтальной оси которого откладывается время, по вертикальной – частота, а интенсивность показывается степенью затемнения (или, наоборот, яркости). На рис. 10.10 приведена осциллограмма (временной график) речевого сигнала, а на рис. 10.11 –



Рис. 10.10. Типичная реализация речевого сигнала



Рис. 10.11. Спектрограмма сигнала, показанного на рис. 10.10

соответствующая ему спектрограмма, на которой темные участки соответствуют более интенсивному сигналу.

10.7.2. КРАТКОВРЕМЕННОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Необходимость анализа нестационарных сигналов, в частности с использованием спектрограмм, привела к необходимости введения характеристики, учитывающей зависимость спектрального состава сигнала от времени. Таким образом возникло понятие спектра, зависящего от времени. Как известно, Фурье-спектр (или спектральная плотность) характеризует представление сигнала в виде суммы (суперпозиции) гармонических составляющих, которые имеют бесконечную протяженность, поэтому базис Фурье является подходящим для анализа *стационарных* процессов на *бесконечной* временной оси. *Кратковременное преобразование Фурье* [28], предназначенное для анализа нестационарных процессов, представляет собой функцию двух переменных n и ω и определяется выражением

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} w[n-m]x[m]e^{-j\omega m}, \quad (10.10)$$

где $w[m]$ – весовая последовательность, или временное окно, «вырезающее» участок сигнала. При фиксированном n можно рассматривать (10.10) как функцию частоты ω – обычное преобразование Фурье последовательности $w[n-m]x[m]$, пред-

ставляющей собой сигнал, умноженный на окно, инвертированное во времени и сдвинутое на n отсчетов вправо (рис. 10.12).

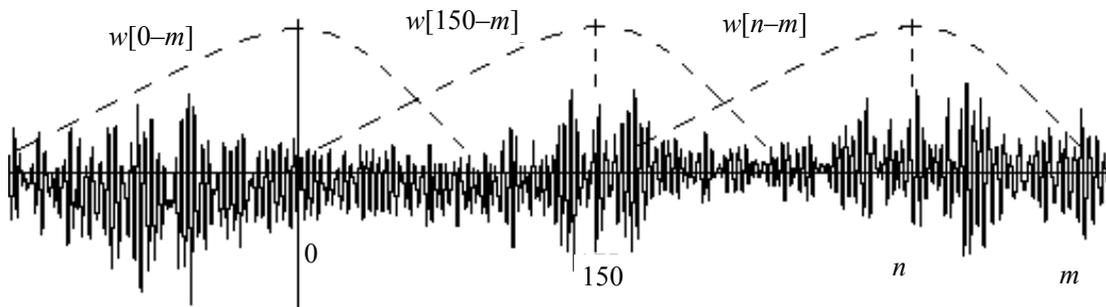


Рис. 10.12. Применение окна при вычислении кратковременного преобразования Фурье

При фиксированном ω выражение (10.10) представляет собой свертку окна $w[m]$ и последовательности $x[m]e^{-j\omega m}$. Эта последовательность – сигнал, сдвинутый влево по оси частот на ω (см. п. 10.4). Учитывая, что последующая свертка с окном представляет собой фильтрацию нижних частот, приходим к заключению, что выражение (10.10) при фиксированном ω описывает результат синхронного детектирования (огibaющую) сигнала, получаемого при полосовой фильтрации анализируемой последовательности фильтром, полоса пропускания которого сосредоточена около ω . Таким образом, вычисление кратковременного преобразования Фурье для набора частот соответствует нахождению огibaющих откликов на данный сигнал *гребенки* полосовых фильтров, т. е. реализации *цифрового анализатора спектра*. Спектрограмма, показанная на рис. 10.11, отображает набор таких откликов, расположенных друг над другом в порядке возрастания частот, при этом степень черноты показывает мощность отклика фильтров спектроанализатора.

Очевидно, что степень детальности спектрограммы определяется выбранным окном. Действительно, если окно короткое, то спектрограмма будет хорошо «отслеживать» временные изменения сигнала, но короткой импульсной характеристике соответствует широкая полоса пропускания фильтра, поэтому такая *широкополосная* спектрограмма будет иметь плохое разрешение по частоте. Наоборот, длинному окну соответствует узкополосный фильтр, поэтому *узкополосная* спектрограмма имеет высокую разрешающую способность по частоте, но низкую по времени. При выборе размера окна необходимо руководствоваться физическими сооб-

ражениями, вытекающими из содержания решаемой задачи. Важно также правильно выбрать форму окна, так как наложение боковых лепестков частотной функции окна ухудшает качество спектрограммы. Спектрограмма, представленная на рис. 10.11, получена с применением окна Блэкмана для набора из 256 фильтров, перекрывающих полосу частот от 0 до 10 кГц.

Анализ речевых сигналов показывает, что такие сигналы обладают большой *избыточностью* с точки зрения теории связи. Участки временного графика речевого сигнала, соответствующие вокализованным звукам, имеют вид, характерный для узкополосных сигналов, где явно видна медленно меняющаяся огибающая и заполнение с практически постоянной частотой. Участки, соответствующие невокализованным, в частности фрикативным, звукам (фонемам), имеют шумоподобный (широкополосный) характер. Поэтому уже в середине XX века инженеры и исследователи обратили внимание на возможность сокращения избыточности речевых сигналов для повышения эффективности систем связи. Интуитивно ясно, что для передачи медленно меняющейся огибающей сигнал можно дискретизировать с частотой, во много раз меньшей частоты Найквиста. На приемном конце канала связи восстановленной по отсчетам огибающей можно промодулировать гармоническую несущую и тем самым восстановить сигнал. Таким образом, в принципе можно передавать по каналу связи сравнительно редкие отсчеты огибающей сигнала и значения медленно меняющейся частоты заполнения. Эти и другие соображения привели к созданию устройств сжатия речевых сигналов, называемых вокодерами^{*)}.

10.7.3. ПРИНЦИП ДЕЙСТВИЯ ВОКОДЕРА

Вокодер представляет собой устройство, позволяющее резко уменьшить пропускную способность канала, необходимую для передачи речевого сигнала, на основе сокращения его избыточности.

В *полосном* вокодере, анализирующая часть которого показана на рис. 10.13 [3], сигнал обрабатывается гребенкой полосовых фильтров (ПФ), каждый из которых реализует кратковременное преобразование Фурье. Полосы пропускания фильтров выбираются таким образом, чтобы они не перекрывались и в то же время в совокупности занимали всю полосу частот сигнала. Отклики фильтров, представляющие собой узкополосные сигналы, детектируются детектором (Д) и фильтруются (сглаживаются) с целью вы-

^{*)} От английских слов voice – голос и coder – устройство кодирования

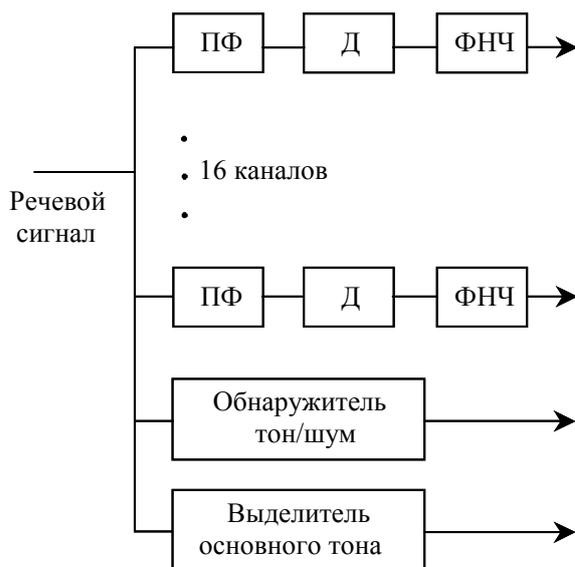


Рис. 10.13. Анализирующая часть полосного вокодера

деления огибающих. Поскольку огибающие таких узкополосных сигналов меняются медленно, то их дискретизация производится с частотой, значительно меньшей частоты Найквиста для исходного сигнала.

Кроме гребенки полосовых фильтров, вокодер содержит обнаружитель тон/шум – устройство, определяющее *тип возбуждения* для текущего отрезка речевого сигнала. Если сигнал признается вокализованным, то принимается решение «тон», иначе – «шум». Для вокализованного сигнала производится выделение *основного тона*, т. е. определяется период возбуждающей импульсной последовательности.

Отсчеты огибающих вместе с признаком тон/шум и значением периода основного тона передаются по каналу связи, а на приемном конце канала в синтезирующей части вокодера (рис. 10.14) происходит *синтез* искусственного речевого сигнала. Пропускная способность канала, необходимая для передачи речевого сигнала

деления огибающих вместе с признаком тон/шум и значением периода основного тона передаются по каналу связи, а на приемном конце канала в синтезирующей части вокодера (рис. 10.14) происходит *синтез* искусственного речевого сигнала. Пропускная способность канала, необходимая для передачи речевого сигнала

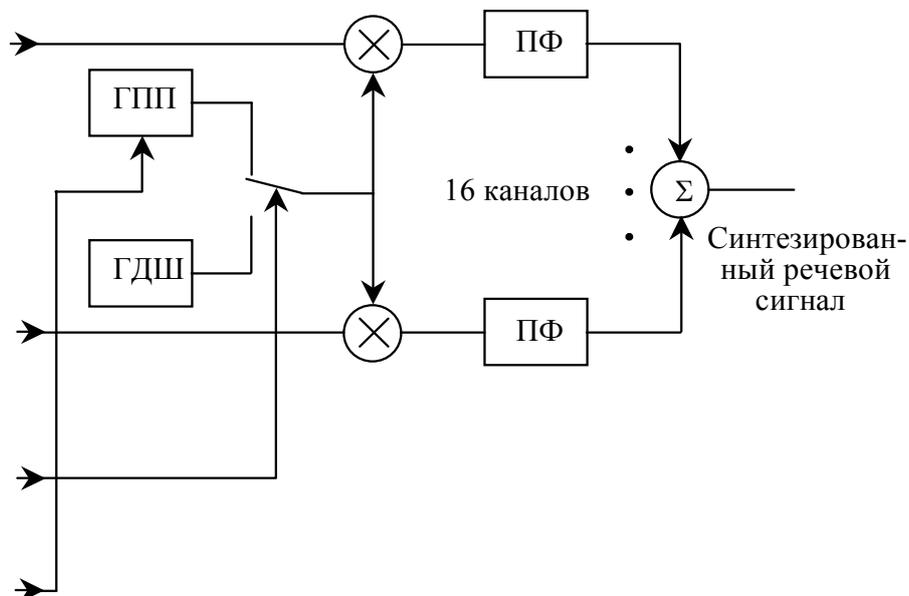


Рис. 10.14. Синтезирующая часть полосного вокодера

с использованием полосного вокодера, составляет от 2400 до 9600 бит/с [3]. Для сравнения: пропускная способность канала, необходимого для передачи речевого сигнала с полосой частот 4 кГц, дискретизированного согласно теореме отсчетов с частотой 8 кГц и квантованного на 256 уровней, составляет 64000 бит/с.

10.7.4. ГОМОМОРФНАЯ ОБРАБОТКА РЕЧЕВЫХ СИГНАЛОВ

Гомоморфными принято называть системы и методы, использующие отображение пространства сигналов в некоторое другое пространство, в котором обработка оказывается более простой, причем отображение должно представлять собой гомоморфизм, а еще лучше – изоморфизм [2]. На практике используется главным образом отображение, преобразующее алгебру абсолютно суммируемых последовательностей l_1 с дискретной сверткой в другое пространство сигналов, где свертке двух последовательностей из l_1 соответствует сложение образов этих последовательностей (см. разд. 2). Такое отображение представляет собой композицию (последовательное применение) z -преобразования и логарифма (рис. 10.15). Исходная последовательность x может представлять собой продукт свертки полезного сигнала s и мешающей последовательности (например, импульсной характеристики искажающей цепи) h . Образ x' такой последовательности при гомоморфном отображении представляет собой сумму образов $x' = s' + h'$ сигнальной и мешающей последовательностей. Целесообразность такого преобразования объясняется тем, что методы оптимальной обработки сигналов наиболее глубоко разработаны для случая аддитивного взаимодействия сигналов и помех. Таким образом, сложное нелинейное взаимодействие последовательностей, выполняемое путем свертки, заменяется аддитивным взаимодействием образов, а выделение полезного сигнала из аддитивной смеси сводится к *линейной фильтрации*.

Если гомоморфное отображение $H\{\cdot\}$ является обратимым, то после линейной обработки оператором $L\{\cdot\}$ путем обратного преобразования $H^{-1}\{\cdot\}$ можно получить результат обработки в виде последовательности \tilde{s} , близкой к s (рис. 10.16).

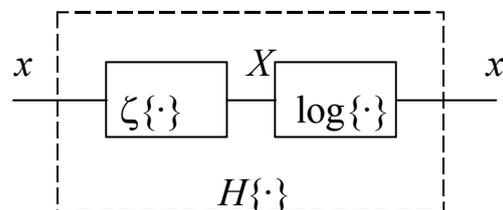


Рис. 10.15. Гомоморфное преобразование

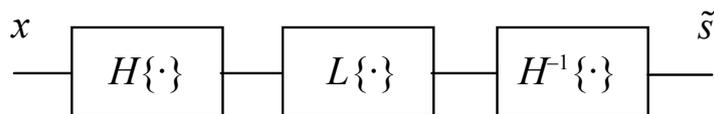


Рис. 10.16. Принцип гомоморфной фильтрации

При практическом применении гомоморфных систем вместо z -преобразования используется дискретное преобразование Фурье, поэтому результат преобразования $H\{\cdot\}$ в соответствии с рис. 10.15 представляет собой последовательность отсчетов, упорядоченных по частоте. Во многих случаях удобнее иметь дело с последовательностью, упорядоченной по времени, поэтому было предложено [31] дополнить преобразование $H\{\cdot\}$ обратным ДПФ. В результате такого преобразования получается последовательность отсчетов, которая носит название *кепстр* (слово получено перестановкой букв слова *спектр*). Таким образом, *кепстр* последовательности x описывается выражением $\hat{x} = F^{-1}\{\ln[F(x)]\}$, где $F(\cdot)$ и $F^{-1}\{\cdot\}$ – соответственно прямое и обратное ДПФ, и представляет собой последовательность отсчетов, упорядоченную по переменной, аналогичной времени, но это особое, «кепстральное» время. Поскольку результат ДПФ произвольной последовательности в общем случае выражается комплексными коэффициентами, логарифм, входящий в гомоморфное преобразование, понимается как комплексный логарифм, представляющий собой многозначную (многолиственную) функцию комплексного переменного. Поэтому неоднозначность определения фазового спектра сигнала с точностью до $\pm 2\pi k$, которая при обычной обработке не играет никакой роли и не учитывается при построении алгоритмов, в данном случае приводит к неоднозначности определения *кепстра*, которую приходится учитывать. Чаще всего применяют алгоритмы *развертывания фазы* [2], приводящие к однозначному определению фазового спектра на основе предположения о его непрерывности. В таком случае комплексный *кепстр* оказывается определенным единственным образом и существует обратное преобразование, описываемое выражением $x = F^{-1}\{\ln[F(\hat{x})]\}$. В некоторых случаях, когда фазовый спектр не имеет существенного значения с точки зрения решаемой задачи, вместо комплексного *кепстра* используют *кепстр мощности*, определяемый выражением $x = F^{-1}\{\ln|F(x)|\}$. Следует заметить, что такое преобразование не имеет обратного в силу необратимости потери информации о фазе.

Гомоморфная обработка речевых сигналов базируется на предположении, что речевой сигнал на коротких временных интервалах представляет собой *свертку* сигнала возбуждения с импульсной характеристикой голосового тракта. Поэтому гомоморфное преобразование можно применить для *разделения* этих составляющих линейными методами. В частности, передаточная функция голосового тракта хорошо аппроксимируется дробно-рациональной функцией, поэтому кепстр импульсной характеристики довольно быстро убывает [2]. Кепстр периодической импульсной последовательности (сигнала возбуждения) *периодичен* с тем же периодом. Учитывая эти различия, составляющие h и v , которые соответствуют импульсной характеристике голосового тракта и сигналу возбуждения, можно разделить при помощи кепстральных окон, схожих с фильтрами нижних и верхних частот (рис. 10.17).

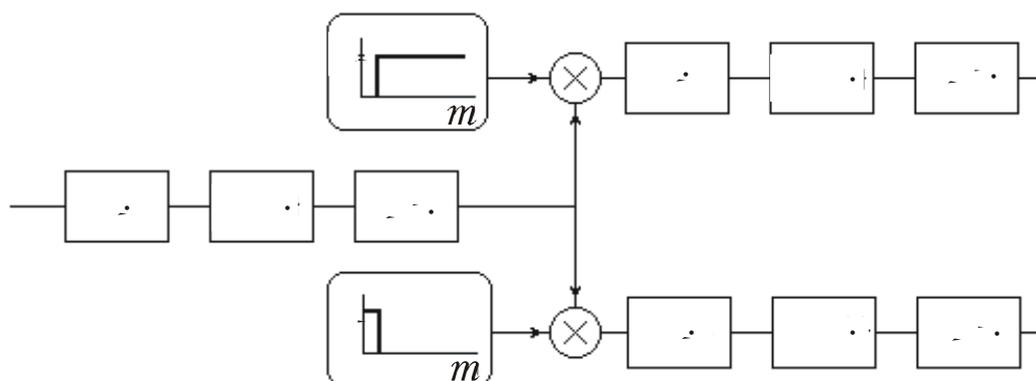


Рис. 10.17. Структура анализирующей части гомоморфного вокодера

Такое разделение используется в описанном в [3] *гомоморфном* вокодере. По каналу связи передаются первые 30 отсчетов кепстра, содержащие информацию о голосовом тракте, а также период основного тона, мощность сигнала и признак тон/шум, которые оцениваются по периодической части кепстра. На приемном конце канала связи для восстановления речевого сигнала импульсная характеристика, восстановленная по первым отсчетам кепстра, свертывается с синтезированным сигналом возбуждения.

10. 8. ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ В СИСТЕМАХ СОТОВОЙ СВЯЗИ

Сотовая связь представляет собой яркий пример технической отрасли, которая претерпела революционные изменения в связи с применением цифровой обработки сигналов. Идея сотовой связи

была высказана в 1971 году [32], и первые сотовые системы были аналоговыми (например, AMPS – *Advanced Mobile Phone Service*). В них по необходимости использовался принцип частотного разделения каналов (ЧРК, FDMA – *Frequency Division Multiple Access*), слабой стороной которого является нерациональное использование выделенной полосы частот. В таких системах различным абонентам предоставляются участки частотного диапазона, которые используются неэффективно вследствие уже упоминавшейся избыточности речевых сигналов, поэтому аналоговые системы сотовой связи быстро достигли предела своих возможностей. Появление цифровых сотовых систем, использующих наряду с ЧРК также принципы временного разделения каналов (ВРК, TDMA – *Time Division Multiple Access*) и кодового разделения каналов (КРК, CDMA – *Code Division Multiple Access*), сделало сотовую связь более доступной и вызвало ее широкое распространение.

Говоря коротко, система сотовой связи включает сеть стационарных базовых станций, покрывающих своими зонами («сотами») обслуживаемую территорию, и множество подвижных станций (абонентских аппаратов), а также центр коммутации, обеспечивающий переключение подвижной станции от одной базовой станции к другой при перемещении или ухудшении качества связи. Ниже будут рассмотрены некоторые вопросы применения цифровой обработки сигналов в подвижных станциях сотовой связи стандарта GSM (*Global System for Mobile communications*).

Упрощенная структурная схема подвижной станции [32] показана на рис. 10.18. Легко видеть, что начиная с аналого-цифрового преобразователя и вплоть до цифроаналогового преобразователя обработка сигнала происходит в цифровой форме. Хотя генератор, гетеродин и смеситель представляют собой аналоговые устройства, методы модуляции и демодуляции, применяемые в таких системах, специфичны для передачи именно цифровых сигналов, поэтому в литературе они издавна именуется методами цифровой передачи сигналов [34].

Аналого-цифровой преобразователь выполняет дискретизацию входного сигнала с частотой 8000 Гц (полоса частот для телефонной связи принята от 300 до 3400 Гц) и квантование с разрядностью 8. Таким образом, поток информации составляет 64 Кбит/с. Как уже говорилось, речевой сигнал обладает большой избыточностью, связанной с его природой. Используя это свойство речевого сигнала, можно передавать не сам сигнал, а некоторые его параметры, существенно уменьшая необходимую пропускную способность канала. Кодер речи (вокодер) и осуществляет такое преобра-

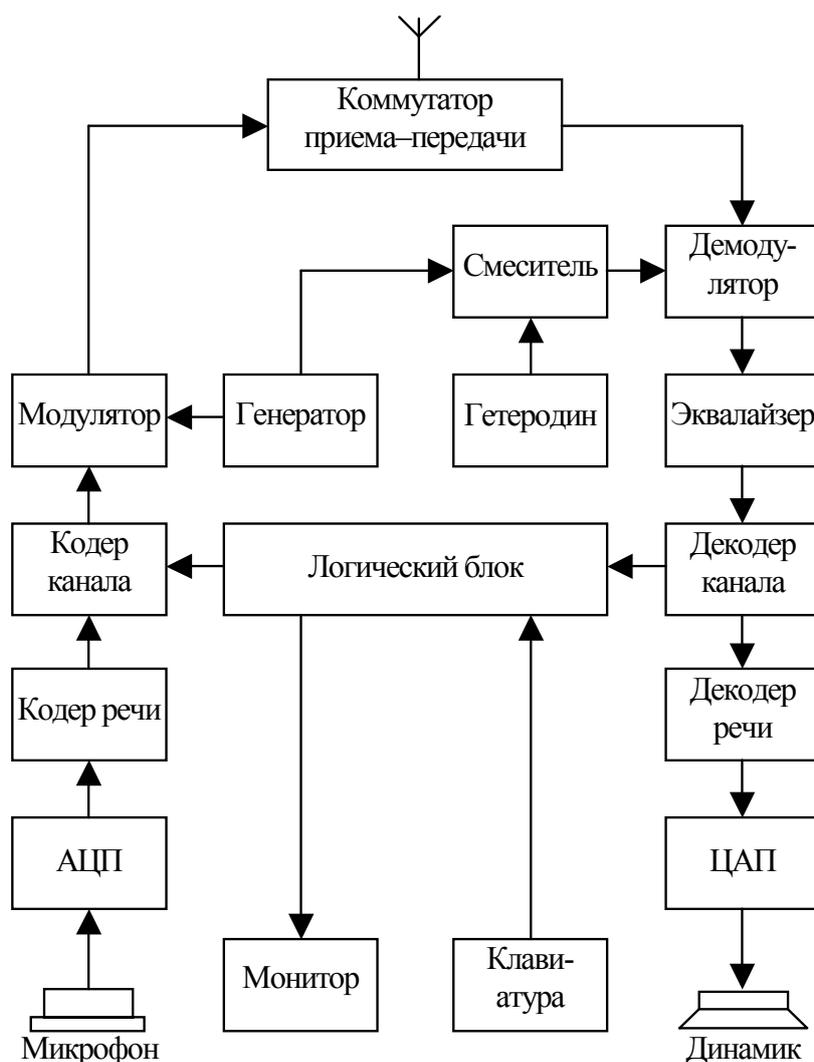


Рис. 10.18. Структурная схема абонентского аппарата сотовой связи

зование. Кодер канала, наоборот, вводит в передаваемый сигнал некоторую избыточность, но это избыточность особого рода, предназначенная для контроля и устранения возможных ошибок при передаче сигналов по каналу (так называемое помехоустойчивое кодирование [34]). Помимо этого кодер канала вводит в передаваемый сигнал служебную информацию, поступающую от логического блока. Декодер канала выполняет обратное преобразование, т. е. обнаруживает и исправляет ошибки, а также выделяет служебную информацию и передает ее в логический блок. Декодер речи восстанавливает естественную форму речевого сигнала, а ЦАП преобразует этот цифровой сигнал в аналоговый для подачи на звуковоспроизводящее устройство (динамик). Эквалайзер, представляющий собой адаптивный цифровой фильтр для компенсации

искажений, связанных с многолучевым распространением сигналов и переотражениями, неизбежными в городских условиях, не входит в стандарт GSM и может отсутствовать в конкретных реализациях подвижной станции.

Рассмотрим более подробно принцип работы кодера и декодера речи (часто их рассматривают как одно устройство и называют словом «кодек»). В основе применяемых в настоящее время кодеков лежит метод линейного предсказания.

Линейное предсказание является одним из самых эффективных методов анализа речевых сигналов [33], широко применяемых в современных системах мобильной связи [32]. Принцип линейного предсказания состоит в том, что очередной отсчет речевого сигнала аппроксимируется линейной комбинацией (взвешенной суммой) некоторого количества предшествующих отсчетов. Коэффициенты взвешивания подстраиваются таким образом, чтобы сделать аппроксимацию наиболее точной (минимизировать ошибку предсказания). Коэффициенты линейного предсказания (КЛП) могут передаваться по каналу связи вместе с ошибкой предсказания с тем, чтобы на приемном конце канала восстановить речевой сигнал.

Метод линейного предсказания основан на представлении речевого сигнала в виде отклика на сигнал возбуждения линейного рекурсивного фильтра с передаточной функцией

$$H(z) = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}}. \quad (10.11)$$

Сигнал $s[n]$ на выходе такого фильтра связан с входным сигналом $v[n]$ разностным уравнением $s[n] = Gv[n] + \sum_{k=1}^p a_k s[n-k]$.

Таким образом, на участках между импульсами возбуждения, когда $v[n] = 0$, выходной сигнал формируется как линейная комбинация собственных предшествующих значений.

Ошибка предсказания определяется как $e[n] = s[n] - \sum_{k=1}^p a_k s[n-k]$, что соответствует отклику на речевой сигнал фильтра с передаточной функцией $A(z) = 1 - \sum_{k=1}^p a_k z^{-k}$. Этот фильтр-анализатор (фильтр погрешности предсказания – ФПП) является обратным

(инверсным) по отношению к фильтру (10.11), описывающему голосовой тракт, и его коэффициенты выбираются (подстраиваются) таким образом, чтобы при речевом сигнале на входе минимизировать мощность выходного сигнала. Поскольку фильтр погрешности предсказания является КИХ-фильтром, задача минимизации ошибки предсказания сводится к решению системы линейных уравнений, получаемых путем приравнивания нулю частных производных дисперсии по всем коэффициентам. Величина p выбирается исходя из условия компромисса между пропускной способностью канала и качеством восстановленной речи и на практике имеет значение порядка 10.

Для решения уравнений применяются в основном два метода – *автокорреляционный*, основанный на оценке автокорреляционной матрицы, получаемой при сдвигах одной реализации определенной длины, и *ковариационный*, при котором отрезки равной длины выбираются из разных участков одной более длинной реализации. Соответственно в первом случае оценочная матрица является трёхдиагональной (содержащей на каждой диагонали одинаковые элементы), а во втором – симметричной, но не трёхдиагональной. Коэффициенты линейного предсказания находятся итерационной процедурой Дарбина для автокорреляционного метода и методом разложения Холецкого – для ковариационного [22].

Получаемый на выходе ФПП сигнал является (с точностью до погрешностей оценки коэффициентов) сигналом возбуждения голосового тракта. Упрощенная схема кодека на основе метода линейного предсказания приведена на рис. 10.19. По каналу связи передаются набор коэффициентов линейного предсказания и ошибка предсказания, которая на приемном конце используется в качестве сигнала возбуждения для синтезирующего фильтра с передаточной функцией (10.11).

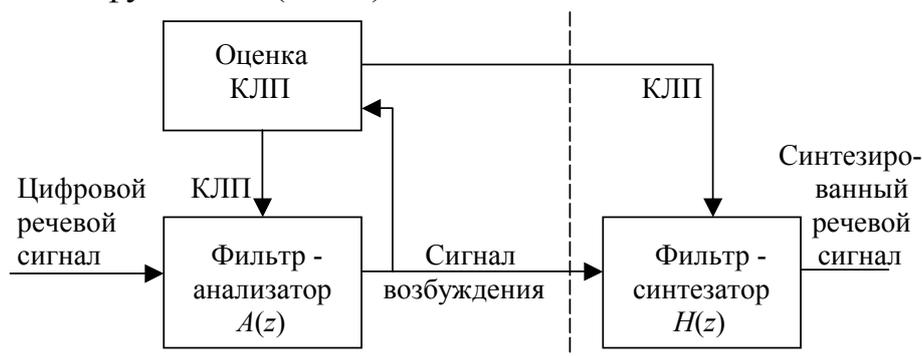


Рис. 10.19. Упрощенная структура кодека на основе метода линейного предсказания

На практике в дополнение к описанному *кратковременному* линейному предсказанию используют еще так называемое *долговременное* линейное предсказание [32]. Дело в том, что описанные преобразования все же не полностью устраняют избыточность речевого сигнала и ошибка предсказания требует для кодирования с целью передачи слишком большого числа бит. Поэтому для снижения загрузки каналов связи (ценой увеличения вычислительных затрат на передающем и приемном концах) ошибка предсказания кратковременного линейного предсказателя подается на фильтр долговременного предсказания с полиномиальной передаточной

функцией $P(z) = 1 - \sum_{k=-M_1}^{M_2} g_k z^{-(d+k)}$, где g_k – коэффициенты *дол-*

говременного предсказания; M_1 и M_2 – константы, определяющие порядок предсказания (обычно небольшой), а временная задержка d соответствует периоду основного тона. При $M_1 = M_2 = 0$ фильтр определяется единственным коэффициентом предсказания и задержкой $P(z) = 1 - gz^{-d}$. Когда на вход этого фильтра (*анализатора*) поступает ошибка кратковременного предсказания $e[n]$, на его выходе получается ошибка (остаток) долговременного предсказания, близкая к белому шуму $r[n] = e[n] - ge[n-d]$. Параметры g и d определяются из условия минимума среднеквадратической ошибки долговременного предсказания на некотором временном интервале, т. е. минимума мощности $r[n]$.

Синтезирующий «долговременный» фильтр является обратным по отношению к анализатору и имеет передаточную функцию

$$D(z) = \frac{B}{1 - gz^{-d}}, \text{ где } B \text{ – коэффициент усиления.}$$

Сигнал с выхода фильтра долговременного предсказания $r[n]$, аппроксимирующий (с точностью до ошибок оценивания) сигнал возбуждения, не передается по каналу, так как это привело бы к его чрезмерной нагрузке. Вместо этого передаются его *параметры*, используемые в декодере для формирования сигнала возбуждения по определенным алгоритмам. В ранних кодеках применялось формирование сигнала возбуждения на основе признака тон-шум; в последнее время используются алгоритмы многоимпульсного возбуждения, не требующие классификации участка речевого сигнала по признаку вокализованности.

На рис. 10.20 представлена структурная схема [32] кодека стандарта GSM. В блоке предварительной обработки (БПО) выполняется цифровая фильтрация (предыскажение) с целью подчеркивания верхних частот речевого сигнала $s[n]$, так как естественный речевой сигнал характеризуется быстрым убыванием спектральной плотности, вследствие чего динамический диапазон канала используется нерационально. После этого сигнал разбивается на сегменты по 160 отсчетов (20 мс) и каждый сегмент умножается на окно Хэмминга. Далее для каждого сегмента оцениваются 8 коэффициентов частичной корреляции $k_j, j = \overline{1,8}$, однозначно связанных с КЛП, которые преобразуются в так называемые логарифмические отношения площадей $l_j = \lg[(1 - k_j)/(1 + k_j)]$, причем логарифмы вычисляются на основе кусочно-линейной аппроксимации. Это приводит к дополнительным преимуществам, так как эти величины по своим статистическим свойствам требуют меньшего числа бит для кодирования, чем непосредственно КЛП. По остатку предсказания $e[n]$ оцениваются параметры долговременного предсказания, при этом каждый сегмент разбивается на 4 подсегмента по 40 отсчетов и параметры g и d определяются для каждого подсегмента.

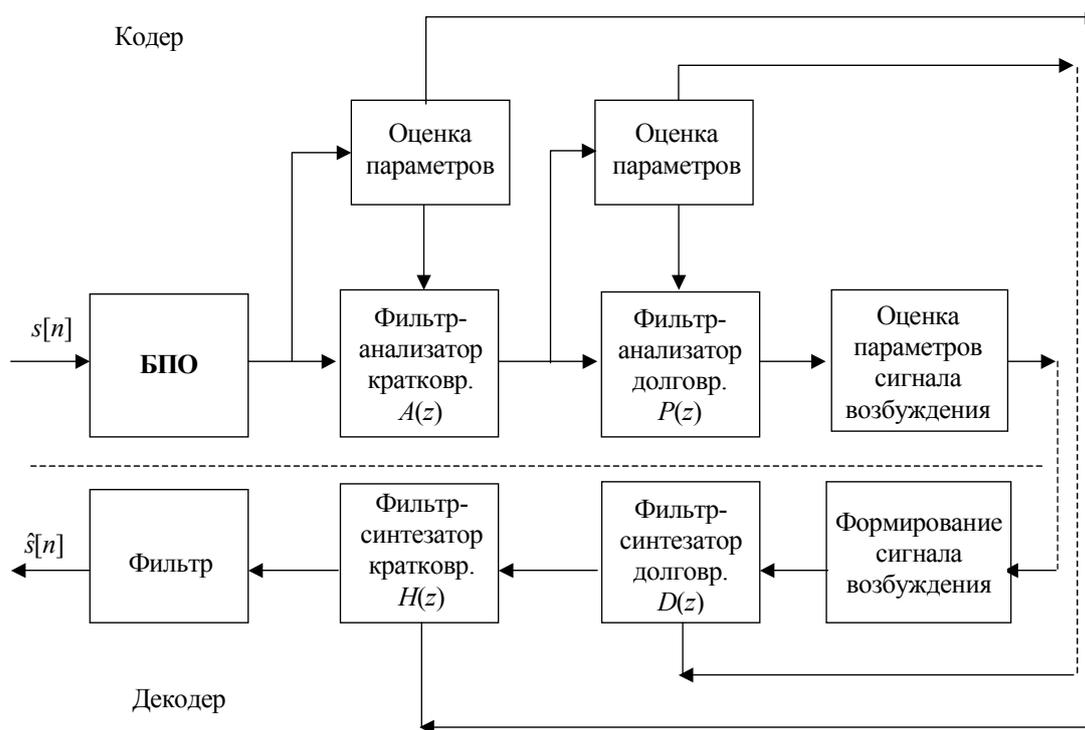


Рис. 10.20. Кодеки стандарта GSM

та. По остатку долговременного предсказания $r[n]$ после его сглаживания вычисляются параметры сигнала возбуждения для каждого подсегмента. Делается это следующим образом [32]. Последний (40-й) отсчет подсегмента отбрасывается, а оставшаяся часть разбивается на три подпоследовательности, первая из которых включает отсчеты с номерами 1,4,...,37, вторая – с номерами 2,5,...,38 и последняя – с номерами 3,6,...,39. Отсчеты подпоследовательности с максимальной энергией оставляются в сигнале возбуждения, остальные заменяются нулями. Значения отсчетов нормируются по отношению к максимальному, нормированные значения кодируются тремя битами каждое, а максимальное значение кодируется шестью битами по логарифмической шкале. Номер выбранной подпоследовательности кодируется двумя битами. Всего для одного 20-миллисекундного интервала передаются 260 бит, причем количество бит для кодирования различных параметров определяется согласно таблице.

В декодере происходит восстановление сигнала возбуждения, который последовательно обрабатывается фильтрами-синтезаторами долговременного и кратковременного предсказаний, а затем фильтром, компенсирующим предыскажение спектра речевого сигнала, в результате чего получается синтезированный сигнал $\hat{s}[n]$, близкий по качеству восприятия к исходному речевому сигналу $s[n]$.

Передаваемые параметры речевых сигналов и число бит, используемое для их кодирования

Передаваемые параметры	Число бит	Примечание
Параметры фильтра кратковременного линейного предсказания (логарифмические отношения площадей) $l_j, j = \overline{1,8}$	36	l_1, l_2 – по 6 бит l_3, l_4 – по 5 бит l_5, l_6 – по 4 бита l_7, l_8 – по 3 бита
Параметры фильтра долговременного линейного предсказания g и d для каждого подсегмента	36	g – 2 бита d – 7 бит
Параметры сигнала возбуждения: номер последовательности N , максимальная амплитуда v , нормированные амплитуды $b_i, i = \overline{1,13}$, для каждого подсегмента	188	N – 2 бита v – 6 бит b_i – 3 бита
Всего за 20 мс	260	–

Стандарт D-AMPS (*Digital AMPS*), представляющий собой альтернативу стандарту GSM, отличается, во-первых, тем, что для каждого 20-миллисекундного фрагмента из 160 отсчетов вычисляются 10 коэффициентов частичной корреляции, которые и кодируются без каких-либо преобразований. Во-вторых, для формирования сигнала возбуждения используются так называемые *кодовые книги*, представляющие собой наборы двоичных данных, хранящихся в запоминающем устройстве. Каждая из кодовых книг содержит 128 кодовых векторов длиной 40 отсчетов. Эти векторы не являются линейно-независимыми, а принадлежат 7-мерному подпространству 40-мерного пространства. Подпространство натянуто на 7 базисных векторов, а коэффициенты разложения всех 128 векторов в этом базисе равны 1 и -1 . Таким образом, в каждой кодовой книге фактически хранятся 7 векторов из 40 элементов и 128 наборов по 7 двоичных (однобитовых) коэффициентов. Параметрами сигнала возбуждения согласно этому стандарту являются номера l_1 и l_2 векторов из первой и второй кодовых книг и соответствующие коэффициенты усиления.

ВОПРОСЫ ДЛЯ САМОКОНТРОЛЯ

1. Какую форму имеет отклик согласованного фильтра на сигнал, с которым он согласован? на другой сигнал?
2. Как найти дисперсию шума на выходе согласованного фильтра?
3. Всегда ли реализуем цифровой согласованный фильтр?
4. Всегда ли можно выполнить интерполяцию и децимацию сигналов?
5. Что происходит со спектрами сигналов при интерполяции и децимации?
6. Зачем применяют интерполяцию и децимацию?
7. Зачем применяют перенос спектра дискретных сигналов?
8. Что такое ОБП-сигнал? В чем его преимущества?
9. Влияет ли временная дискретизация сигналов в антенных решетках на диаграмму направленности?
10. Как уменьшить искажение диаграммы направленности, возникающее вследствие временной дискретизации сигналов?
11. Что такое основной тон?
12. Что такое кратковременное преобразование Фурье, и как оно связано с «обычным» преобразованием Фурье?
13. В чем состоит принцип действия полосного вокодера?
14. В чем состоит принцип действия гомоморфного вокодера?

15. За счет чего достигается повышение эффективности передачи информации в системах с вокодерами?
16. Что такое кодек?
17. Какие требования предъявляются к каналу передачи речевых сигналов в системах мобильной связи?
18. В чем состоит суть метода линейного предсказания?
19. Что такое «долговременное предсказание», и для чего оно применяется?
20. По какому критерию настраиваются анализирующие фильтры кратковременного и долговременного линейных предсказаний?
21. Какая информация передается по каналу связи в системах стандарта GSM?
22. Что такое «кодовая книга»?
23. Что такое «предыскажение», и зачем оно применяется?

ЗАКЛЮЧЕНИЕ

В рамках учебника невозможно охватить все стороны цифровой обработки сигналов, которая переживает этап бурного развития. Постоянно совершенствуется элементная база цифровых устройств, развиваются теоретические модели и методы анализа и синтеза алгоритмов цифровой обработки сигналов. Нет сомнений, что и в обозримом будущем это положение сохранится, поэтому инженер, работающий в области цифровой связи, должен постоянно следить за научным и техническим прогрессом в этой области, заниматься непрерывным самообразованием, используя для этого научно-технические журналы, сеть Internet и техническую литературу.

ОСНОВНЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ФУНКЦИЙ КОМПЛЕКСНОГО ПЕРЕМЕННОГО

Следующие три определения аналитической функции эквивалентны:

1. Функция $f(z)$ комплексного переменного z называется аналитической в области D комплексной плоскости C , если она в этой области дифференцируема по z .

2. Задание комплексной функции $f(z)$ комплексного переменного $z = x + iy$ эквивалентно заданию двух вещественных функций $u = \varphi(x, y)$, $v = \psi(x, y)$ с областью определения $\{x + iy \in D\}$. При этом подходе функция $f(z)$ аналитична в D в том и только в том случае, если выполняется условие Коши – Римана

$$\frac{d\varphi}{dx} = \frac{d\psi}{dy}; \quad \frac{d\varphi}{dy} = -\frac{d\psi}{dx}$$

(предполагается, что все частные производные существуют).

3. Функция $f(z)$ комплексного переменного z называется аналитической в области D , если в каждой точке $z_0 \in D$ она может быть представлена степенным рядом

$$f(z) = a_0 + a_1(z - z_0) + \dots + a_n(z - z_0)^n + \dots$$

Отсюда следует, что аналитическая функция бесконечно дифференцируема в области D .

Примеры

1. Функция z^2 аналитична во всей плоскости C .

2. Функция $w = x - iy$ не является аналитической, так как ее нельзя записать, как $w(z)$, а следовательно, и дифференцировать по z (хотя нет сомнений в том, что w является функцией от $z = x + iy$).

3. Функция $1/z$ аналитична во всей плоскости C , за исключением точки $z = 0$, в которой нарушается непрерывность, а значит, и аналитичность.

Точки нарушения аналитичности (особые точки) играют важную роль в теории функций комплексного переменного. Если

$f(z)$ – функция, аналитическая в кольце $0 < |z - z_0| < \rho$, охватывающем особую точку z_0 , то она разлагается в этом кольце в ряд Лорана

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n .$$

Коэффициенты ряда Лорана определяют тип особой точки в соответствии со следующими утверждениями:

а) если ряд Лорана не содержит отрицательных степеней ($a_n = 0 \quad \forall n < 0$), то z_0 – *устраняемая особая точка*, т.е. при замене $f(z_0) = \lim_{z \rightarrow z_0} f(z) = a_0$ функция становится аналитической во

всем круге $|z - z_0| < \rho$;

б) если ряд Лорана содержит лишь конечное число m отрицательных степеней ($a_n = 0 \quad \forall n < -m$), то z_0 – *полюс* (кратности m), при этом $\lim_{z \rightarrow z_0} f(z) = \infty$;

в) если ряд Лорана содержит бесконечное число отрицательных степеней, то z_0 – *существенно особая точка* и $\lim_{z \rightarrow z_0} f(z)$ не существует.

Коэффициент a_{-1} ряда Лорана называется *вычетом* функции $f(z)$ в точке z_0 :

$$a_{-1} = \operatorname{res}_{z=z_0} f(z) = \frac{1}{2\pi i} \int_{\gamma} f(z) dz ,$$

где γ – контур, охватывающий точку z_0 (и не охватывающий других особых точек). Вычеты используются для нахождения интегралов в соответствии с формулой

$$\int_{\Gamma} f(z) dz = 2\pi i \sum_{k=1}^n \operatorname{res}_{z=z_k} f(z) ,$$

где Γ – контур, охватывающий n особых точек, вычет в m -кратном полюсе

$$\operatorname{res}_{z=z_k} f(z) = \frac{1}{(m-1)!} \lim_{z \rightarrow z_k} \left\{ \frac{d^{m-1}}{dz^{m-1}} (z - z_k)^m f(z) \right\}.$$

Простейший класс аналитических функций составляют целые функции, представимые (на всей комплексной плоскости) сходящимися степенными рядами, например

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots,$$

$$\sin z = z - \frac{z^3}{3!} + \dots + (-1)^n \frac{z^{2n+1}}{(2n+1)!} + \dots,$$

$$\cos z = 1 - \frac{z^2}{2!} + \dots + (-1)^n \frac{z^{2n}}{(2n)!} + \dots.$$

К целым функциям относятся также все полиномы от z .

Очень важную роль в цифровой обработке сигналов играют дробно-рациональные функции, имеющие вид $F(z) = \frac{B(z)}{A(z)}$, где

$A(z)$ и $B(z)$ – полиномы степеней M и N соответственно. Корни уравнения $A(z) = 0$, которых насчитывается ровно M , и которые, в частности, могут совпадать, называются нулями полинома $A(z)$; они, очевидно, являются полюсами $F(z)$. Других особенностей $F(z)$ не имеет.

Если $F(z) = \frac{\varphi(z)}{\psi(z)}$, причем $\varphi(z)$ и $\psi(z)$ – аналитические в ок-

рестности z_0 , и z_0 есть простой (однократный) нуль $\psi(z)$, то

$$\operatorname{res}_{z=z_0} \left[\frac{\varphi(z)}{\psi(z)} \right] = \frac{\varphi(z_0)}{\psi'(z_0)}.$$

СВОЙСТВА Z-ПРЕОБРАЗОВАНИЯ

Преобразование Лапласа широко применяется в аналоговой теории цепей и сигналов благодаря тому, что дифференциальному уравнению, связывающему функции времени, соответствует алгебраическое уравнение для их лапласовских изображений. Роль z -преобразования в анализе дискретных сигналов и цепей совершенно аналогична. В большинстве случаев интерес представляют z -преобразования, приводящие к функциям дробно-рационального вида, что соответствует реализуемым дискретным ЛИС-цепям конечного порядка (см. разд. 3). Поэтому здесь приводятся свойства z -преобразования, справедливые для дробно-рациональных функций. Не следует, однако, забывать, о существовании абсолютно суммируемых последовательностей, имеющих z -преобразования вида, отличного от дробно-рационального (см. пример 2.9).

1. Области сходимости

Область сходимости z -преобразования каузальной *быстро убывающей* последовательности ограничена изнутри окружностью, проходящей через полюс с максимальным модулем, а извне – окружностью бесконечно большого радиуса. Предполагая, что все полюсы z -образа – простые, запишем последовательность в виде (см. п. 3.5)

$$x[n] = \sum_{k=1}^N C_k (d_k)^n, \quad n \geq 0,$$

где $d_k, k = \overline{1, N}$ – полюсы; $C_k, k = \overline{1, N}$ – «амплитудные» коэффициенты. Очевидно, для того чтобы z -преобразование

$$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n} = \sum_{n=0}^{\infty} \left(\sum_{k=1}^N C_k (d_k)^n \right) z^{-n} = \sum_{k=1}^N C_k \left(\sum_{n=0}^{\infty} (d_k z^{-1})^n \right)$$

было ограниченным, должно выполняться условие $|d_k| < |z|$; при этом все экспоненты затухают при $n \rightarrow \infty$. Доказательство соответствующего утверждения для некрратных и комплексно-сопряженных полюсов аналогично. Также можно показать, что для *антикаузальной* последовательности область сходимости ограни-

чена *снаружи* окружностью, проходящей через полюс с *минимальным* модулем. В самом деле, для антикаузальной последовательности z -преобразование

$$\begin{aligned} X(z) &= \sum_{n=-\infty}^0 x[n]z^{-n} = \sum_{n=-\infty}^0 \left(\sum_{k=1}^N C_k (d_k)^n \right) z^{-n} = \\ &= \sum_{k=1}^N C_k \left(\sum_{n=-\infty}^0 (d_k z^{-1})^n \right) \end{aligned}$$

ограничено, только если $(d_k z^{-1})^n$ убывает при $n \rightarrow -\infty$, а это имеет место лишь при $|z| < |d_k| \quad \forall k$.

Рассматривая некаузальную последовательность как сумму каузальной и антикаузальной последовательностей, приходим к выводу, что областью сходимости в этом случае является множество точек z -плоскости, в которых конечны суммы для отрицательных и для положительных n

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n} = \sum_{n=0}^{\infty} x[n]z^{-n} + \sum_{n=-\infty}^{-1} x[n]z^{-n}.$$

Таким образом, z -преобразование некаузальной последовательности сходится в тех точках z -плоскости, где сходятся z -преобразования каузальной и антикаузальной частей последовательности, т. е. на множестве, представляющем собой *пересечение* областей сходимости – открытое кольцо, ограниченное двумя окружностями, проходящими через два полюса (рис. П2.1). При этом все полюсы, лежащие вне кольца, относятся к антикаузальной части последовательности, а все полюсы внутри кольца – к каузальной части. Обозначив радиусы указанных окружностей, ограничивающих области сходимости для антикаузальной и каузальной частей последовательности, соответственно R_{x-} и R_{x+} , можно записать для области сходимости некаузальной последовательности $R_{x-} < |z| < R_{x+}$. Если $R_{x-} > R_{x+}$, то z -преобразование некаузальной последовательности не сходится нигде. Одному и тому же набору полюсов могут соответствовать разные последовательности в соответствии с различным выбором области сходимости. Например, для нуль-полюсной диаграммы, представленной на рис. П2.2, существуют три варианта выбора области сходимости z -преобразова-

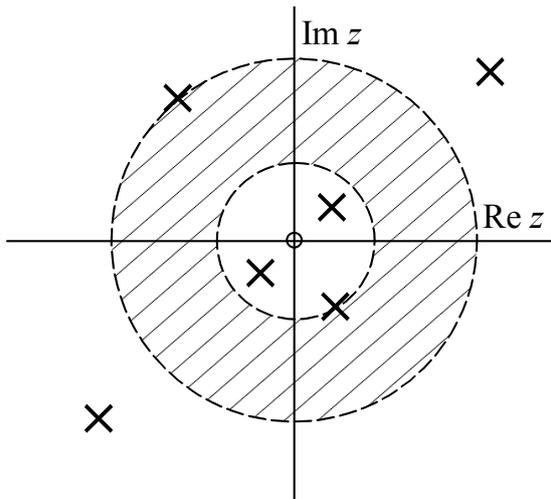


Рис. П2.1

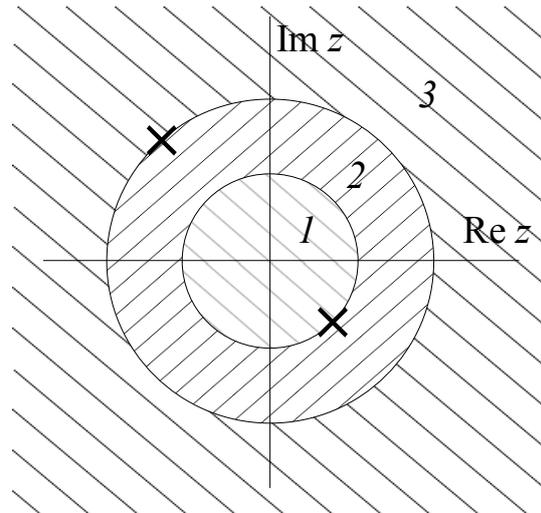


Рис. П2.2

ния и соответственно три возможные последовательности. Область 1 соответствует антикаузальной, область 2 – некаузальной, а область 3 – каузальной последовательностям.

2. Обратное z-преобразование

Согласно теореме Коши

$$\frac{1}{2\pi i} \oint_C^{\leftarrow} z^{k-1} dz = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0, \end{cases} \quad (\text{П2.1})$$

где интеграл берется по контуру, охватывающему начало координат z -плоскости, против часовой стрелки. z -Преобразование определяется выражением

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

Умножим обе части на z^{k-1} и проинтегрируем их по контуру, лежащему в области сходимости $X(z)$:

$$\frac{1}{2\pi i} \oint_C^{\leftarrow} X(z)z^{k-1} dz = \frac{1}{2\pi i} \oint_C^{\leftarrow} \sum_{n=-\infty}^{\infty} x[n]z^{-n+k-1} dz.$$

Меняя местами сумму и интеграл в правой части выражения (это можно делать, если ряд сходится!), получим

$$\frac{1}{2\pi i} \oint_C^{\leftarrow} X(z)z^{k-1} dz = \sum_{n=-\infty}^{\infty} x[n] \frac{1}{2\pi i} \oint_C^{\leftarrow} z^{-n+k-1} dz,$$

откуда следует, согласно (П 2.1),

$$\frac{1}{2\pi i} \oint_C^{\leftarrow} X(z)z^{k-1} dz = x[k].$$

Таким образом, обратное z -преобразование определяется выражением

$$x[n] = \frac{1}{2\pi i} \oint_C^{\leftarrow} X(z)z^{n-1} dz, \quad (\text{П2.2})$$

справедливым как для положительных, так и для отрицательных n .

Для дробно-рациональных функций контурный интеграл, входящий в обратное z -преобразование, можно находить с использованием вычетов [2], так как

$$x[n] = \frac{1}{2\pi i} \oint_C^{\leftarrow} X(z)z^{n-1} dz = \sum_{k=1}^N \operatorname{res}_{z=z_k} \left\{ X(z)z^{n-1} \right\},$$

где суммируются вычеты в N полюсах функции $X(z)z^{n-1}$ внутри контура интегрирования. Предположим, что $X(z)z^{n-1}$ – рациональная функция от z , имеющая в точке z_0 полюс кратности s .

Тогда можно записать $X(z)z^{n-1} = \frac{\psi(z)}{(z-z_0)^s}$, при этом функция $\psi(z)$ полюсов в точке z_0 , очевидно, не имеет. Вычет функции $X(z)z^{n-1}$ в точке $z = z_0$ определяется формулой

$$\operatorname{res}_{z=z_0} \{X(z)z^{n-1}\} = \frac{1}{(s-1)!} \left. \frac{d^{s-1}\psi(z)}{dz^{s-1}} \right|_{z=z_0}.$$

В частности, если полюс простой, то

$$\operatorname{res}_{z=z_0} \{X(z)z^{n-1}\} = \psi(z_0).$$

Вычисление обратного z -преобразования по (П 2.2) часто бывает слишком трудоемким, поэтому в частных случаях применяют следующие приемы, упрощающие вычисления [2].

Разложение в степенной ряд. Если функция переменного z разлагается в степенной ряд по z , то коэффициент этого ряда при z^{-n} есть не что иное, как значение $x[n]$. Например, для функции $X(z) = \ln(1 + az^{-1})$, где $|a| < |z|$, можно записать разложение

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

Сравнивая полученное выражение с формулой z -преобразования, заключаем, что $x[n] = \begin{cases} (-1)^{n+1} \frac{a^n}{n}, & n \geq 1, \\ 0, & n \leq 0. \end{cases}$

Деление числителя дробно-рациональной функции на знаменатель, рассмотренное в п. 3.3, и *разложение* дробно-рациональной функции в сумму элементарных дробей, применяемое в п. 3.5 для анализа строения импульсной характеристики ЛИС-цепи конечного порядка, также могут быть использованы как альтернативные способы нахождения обратного z -преобразования.

3. Линейность z -преобразования

Линейность z -преобразования означает, что z -образом линейной комбинации последовательностей является соответствующая линейная комбинация их z -образов. Обозначая z -образ последовательности $x[n]$ как $\zeta\{x[n]\}$, можно записать

$$\zeta\{ax[n] + by[n]\} = a\zeta\{x[n]\} + b\zeta\{y[n]\}.$$

Это свойство практически очевидно. Следует только отметить особенность, связанную с областью сходимости результирующего z -образа, которая в общем случае *содержит* пересечение областей сходимости z -преобразований исходных последовательностей, но может быть *больше* него, если некоторые полюсы одного слагаемого компенсируются нулями другого. Например, для z -преобразований последовательностей $a^n u[n]$ и $a^n u[n-1]$ область сходимости определяется условием $|z| > |a|$, а для z -преобразования их разности областью сходимости является вся z -плоскость.

4. Сдвиг последовательности

Если z -образом последовательности $x[n]$ является функция $X(z)$ в области сходимости $R_{x-} < |z| < R_{x+}$, то для последовательности $x[n + n_0]$, сдвинутой на целое число шагов, z -образ имеет вид

$$\zeta\{x[n + n_0]\} = \sum_{n=-\infty}^{\infty} x[n + n_0]z^{-n} = \sum_{k=-\infty}^{\infty} x[k]z^{-k+n_0} = z^{n_0} X(z),$$

при этом область сходимости остается неизменной, если не учитывать точки $z = 0$ и $z = \infty$. Очевидно, что сдвиг на n_0 шагов при $n_0 > 0$ добавляет к нуль-полюсной диаграмме z -образа последовательности нуль кратности n_0 в точке $z = 0$ и полюс такой же кратности в точке $z = \infty$, а при $n_0 < 0$, наоборот, нуль кратности n_0 в точке $z = \infty$ и полюс такой же кратности в точке $z = 0$.

5. Умножение последовательности на экспоненту

Если z -образом последовательности $x[n]$ является функция $X(z)$ в области сходимости $R_{x-} < |z| < R_{x+}$, то для последовательности $x[n]$, умноженной на экспоненциальную последовательность a^n , где a – комплексное число, z -образ

$$\zeta\{a^n x[n]\} = \sum_{n=-\infty}^{\infty} a^n x[n]z^{-n} = \sum_{n=-\infty}^{\infty} x[n](a^{-1}z)^{-n} = X(a^{-1}z)$$

в области сходимости $|a|R_{x-} < |z| < |a|R_{x+}$. Вообще, если функция $X(z)$ имеет полюс или нуль при $z = z_0$, то очевидно, что $X(a^{-1}z)$ имеет полюс или нуль в точке $z = az_0$. Заметим, что при вещественном a происходит масштабное расширение или сжатие нуль-полюсной диаграммы; при комплексном a с единичным модулем происходит поворот вокруг начала координат; в общем случае преобразование нуль-полюсной диаграммы представляет собой комбинацию поворота и масштабного преобразования.

6. Умножение последовательности на линейную последовательность

Предположим, что функция $X(z)$ в области сходимости $R_{x-} < |z| < R_{x+}$ является z -образом последовательности $x[n]$. Продифференцируем $X(z)$ по z :

$$\begin{aligned} \frac{dX(z)}{dz} &= \frac{d}{dz} \left\{ \sum_{n=-\infty}^{\infty} x[n]z^{-n} \right\} = \sum_{n=-\infty}^{\infty} x[n] \frac{d}{dz} (z^{-n}) = \\ &= \sum_{n=-\infty}^{\infty} x[n] \frac{d}{dz} (z^{-n}) = \sum_{n=-\infty}^{\infty} (-n)x[n]z^{-n-1}. \end{aligned}$$

Умножив обе части выражения на z , получаем

$$\zeta \{nx[n]\} = -z \frac{dX(z)}{dz}.$$

Область сходимости при этом не изменяется.

7. Переход к комплексно-сопряженной последовательности

z -Преобразование последовательности $x^*[n]$, комплексно-сопряженной по отношению к $x[n]$, имеет вид

$$\zeta \{x^*[n]\} = \sum_{n=-\infty}^{\infty} x^*[n]z^{-n} = \left(\sum_{n=-\infty}^{\infty} x[n](z^*)^{-n} \right)^* = X^*(z^*),$$

область сходимости $R_{x-} < |z| < R_{x+}$.

8. Обращение последовательности во времени

z -Преобразование последовательности $x[-n]$ имеет вид

$$\zeta \{x[-n]\} = \sum_{n=-\infty}^{\infty} x[-n]z^{-n} = \sum_{k=-\infty}^{\infty} x[k]z^k = \sum_{k=-\infty}^{\infty} x[k](z^{-1})^{-k} = X(z^{-1}).$$

Если функция $X(z)$ имеет полюс в точке $z = z_0$, то очевидно, что $X(z^{-1})$ имеет полюс в точке $z = z_0^{-1}$. Поэтому при переходе от последовательности $x[n]$ к $x[-n]$ нуль-полюсная диаграмма z -образа претерпевает преобразование $z \rightarrow 1/z$, т. е. «отражение» относительно 1-окружности. Область сходимости после преобразования также изменяется и приобретает вид $1/R_{x+} < |z| < 1/R_{x-}$.

9. Свертка последовательностей

Основное свойство z -преобразования, благодаря которому оно применяется в анализе дискретных сигналов и цепей, описывается выражением

$$\zeta \{x[n] \otimes y[n]\} = X(z)Y(z),$$

где символом \otimes обозначена дискретная свертка последовательностей. Область сходимости результата содержит пересечение областей сходимости z -образов сомножителей, но может быть шире его при взаимной компенсации нулей и полюсов. Доказательство этого свойства изложено в разд. 2.

10. «Комплексная свертка» z -образов

Дуальное («симметричное») свойство z -преобразования соответствует умножению последовательностей и описывается выражением

$$\zeta \{x[n]y[n]\} = \frac{1}{2\pi i} \oint_C^{\leftarrow} X(v)Y(z/v)v^{-1}dv,$$

где C – замкнутый контур, лежащий в пересечении областей сходимости $X(z)$ и $Y(z)$. Область сходимости результата содержит область, определяемую выражением $R_{x-}R_{y-} < |z| < R_{x+}R_{y+}$, но снова может быть шире. Доказательство приведено в [2].

Некоторые свойства z -преобразования сведены в таблицу для удобства использования.

Основные свойства z -преобразования

Последовательность	z -Преобразование	Область сходимости
$x[n]$	$X(z)$	$R_{x-} < z < R_{x+}$
$y[n]$	$Y(z)$	$R_{y-} < z < R_{y+}$
$ax[n] + by[n]$	$aX(z) + bY(z)$	$\max[R_{x-}, R_{y-}] < z < \min[R_{x+}, R_{y+}]$
$x[n + n_0]$	$z^{n_0} X(z)$	$R_{x-} < z < R_{x+}$
$a^n x[n]$	$X(a^{-1}z)$	$ a R_{x-} < z < a R_{x+}$

Окончание таблицы

Последовательность	z -Преобразование	Область сходимости
$nx[n]$	$-z \frac{dX(z)}{dz}$	$R_{x-} < z < R_{x+}$
$x^*[n]$	$X^*(z^*)$	$R_{x-} < z < R_{x+}$
$x[-n]$	$X(1/z)$	$1/R_{x+} < z < 1/R_{x-}$
$\operatorname{Re}\{x[n]\}$	$\frac{1}{2} [X(z) + X^*(z^*)]$	$R_{x-} < z < R_{x+}$
$\operatorname{Im}\{x[n]\}$	$\frac{1}{2i} [X(z) - X^*(z^*)]$	$R_{x-} < z < R_{x+}$
$x[n] \otimes y[n]$	$X(z)Y(z)$	$\max[R_{x-}, R_{y-}] < z < \min[R_{x+}, R_{y+}]$
$x[n] \cdot y[n]$	$\frac{1}{2\pi i} \oint_C^{\leftarrow} X(v)Y(z/v)v^{-1}dv$	$R_{x-}R_{y-} < z < R_{x+}R_{y+}$

ПРЯМОЕ И ОБРАТНОЕ ДПФ

Прямое ДПФ последовательности длины N определяется выражением (2.9)

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}, \quad k = \overline{0, N-1}.$$

Умножим обе части выражения на $e^{j\frac{2\pi}{N}km}$, $m = \overline{0, N-1}$, и просуммируем по $k = \overline{0, N-1}$:

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

Выражение в фигурных скобках можно рассматривать как сумму N векторов единичной длины на комплексной плоскости. Нетрудно видеть, что при $m = n$ все эти векторы направлены вдоль вещественной оси и их сумма равна N . При $m \neq n$ каждый вектор отстоит от соседних векторов на одинаковый угол и в силу симметрии сумма равна 0. Таким образом,

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

где δ_{mn} – символ Кронекера, определяемый выражением

$$\delta_{mn} = \begin{cases} 1, & m = n, \\ 0, & m \neq n. \end{cases}$$

Поэтому

$$\sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi}{N}mk} = \sum_{n=0}^{N-1} x[n] N \delta_{mn} = x[m]N.$$

Заменяя m на n , получаем формулу (2.10) обратного ДПФ

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi}{N}kn}.$$

СТРУКТУРА КИХ-ЦЕПИ НА ОСНОВЕ ИНТЕРПОЛЯЦИОННОЙ ФОРМУЛЫ ЛАГРАНЖА

Интерполяционная формула Лагранжа предназначена для построения полинома, который в точках z -плоскости, называемых узлами интерполяции, принимает заданные значения. Пусть на z -плоскости даны N несовпадающих точек z_n , $n = \overline{0, N-1}$, и заданы значения функции H_n в этих точках. Существует полином степени $N-1$, называемый полиномом Лагранжа и определяемый выражением [3]

$$H(z) = \prod_{n=0}^{N-1} (1 - z^{-1}z_n) \sum_{m=0}^{N-1} \frac{H_m}{(1 - z^{-1}z_m) \prod_{\substack{k=0 \\ k \neq m}}^{N-1} (1 - z_m^{-1}z_k)}.$$

Как видно из формулы, функция $H(z)$ полностью определяется значениями H_n в точках z_n и самими этими точками z_n , $n = \overline{0, N-1}$. Перепишем это выражение в виде

$$H(z) = \prod_{n=0}^{N-1} (1 - z^{-1}z_n) \sum_{m=0}^{N-1} \frac{A_m}{(1 - z^{-1}z_m)}, \quad (\text{П4.1})$$

где

$$A_m = \frac{H_m}{\prod_{\substack{k=0 \\ k \neq m}}^{N-1} (1 - z_m^{-1}z_k)} \quad (\text{П4.2})$$

– константа, не зависящая от z .

В выражении (П4.1) первый сомножитель как произведение двучленов равен полиному относительно z^{-1} степени N ; второй сомножитель (сумма дробей) имеет в качестве общего знаменателя этот самый полином. Если дроби привести к общему знаменателю и просуммировать, то в числителе получится полином степени $N-1$; после сокращения оказывается, что этот полином и есть

$H(z)$. Таким образом, видно, что выражение (П4.1) определяет полином порядка $N-1$.

Рассмотрим поведение полинома (П4.1) в точке z_k при некотором k . При подстановке z_k в (П4.1) легко видеть, что все слагаемые во втором сомножителе конечны, за исключением одного, а именно слагаемого $A_k / (1 - z_k^{-1} z_k)$, поэтому всеми слагаемыми, кроме k -го, можно пренебречь. Запишем с учетом (П4.2)

$$H(z_k) = \prod_{n=0}^{N-1} (1 - z_k^{-1} z_n) \frac{A_k}{(1 - z_k^{-1} z_k)} = \prod_{\substack{n=0 \\ n \neq k}}^{N-1} (1 - z_k^{-1} z_n) A_k = H_k.$$

Таким образом, формула Лагранжа действительно представляет собой полином порядка $N-1$, принимающий в N данных точках z -плоскости заданные значения.

Выражение (П4.1) дробно-рационально, следовательно, оно может рассматриваться в качестве передаточной функции ЛИС-цепи. Очевидно, эта цепь представляет собой каскадное соединение трансверсальной цепи (первый сомножитель в (П4.2)) и рекурсивной цепи (второй сомножитель). Трансверсальная цепь, в свою очередь, образована каскадным соединением трансверсальных звеньев 1-го порядка с нулями z_n , $n = \overline{0, N-1}$; рекурсивная часть цепи – это параллельное соединение простейших рекурсивных звеньев с полюсами z_n , $n = \overline{0, N-1}$, и масштабными множителями A_n , определяемыми согласно выражению (П4.2).

На первый взгляд цепь, реализованная в соответствии с формулой Лагранжа, должна иметь бесконечную импульсную характеристику как цепь с рекурсивными звеньями. На самом деле каждому полюсу рекурсивной части соответствует компенсирующий нуль в трансверсальной части схемы, и в действительности это – КИХ-цепь. (Это естественно, так как ранее было показано, что выражение (П4.1) представляет собой полином.)

БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Существует класс алгоритмов вычисления дискретного преобразования Фурье, называемых алгоритмами быстрого преобразования Фурье. Такие алгоритмы используют симметрию ДПФ для группирования данных с целью экономии числа операций. Здесь рассматривается лишь один из наиболее простых алгоритмов БПФ – алгоритм с прореживанием по времени [3].

Рассмотрим ДПФ последовательности $x[n]$ по основанию $N = 2^r$, где r – целое число:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} kn}, \quad k = \overline{0, N-1}. \quad (\text{П.5.1})$$

При рассмотрении алгоритмов БПФ общепринятым является обозначение комплексного множителя $W_N = e^{-j \frac{2\pi}{N}}$. С учетом этого

$$X[k] = \sum_{n=0}^{N-1} x[n] W_N^{kn}, \quad k = \overline{0, N-1}. \quad (\text{П.5.2})$$

Заметим, что W_N^{kn} при фиксированном k можно рассматривать как последовательность $\tilde{y}[n] = W_N^{kn}$, $n = \overline{-\infty, \infty}$, периодичную с периодом N/k .

Аналогично при фиксированном n можно рассматривать W_N^{kn} как последовательность $\tilde{y}[k] = W_N^{kn}$, $k = \overline{-\infty, \infty}$, N/n -периодичную по k .

Разобьем сумму, входящую в (П.5.1), на сумму всех слагаемых с четными номерами и сумму всех слагаемых с нечетными номерами:

$$X[k] = \underbrace{\sum_{n=0}^{N-2} x[n] W_N^{kn}}_{\text{(четные)}} + \underbrace{\sum_{n=1}^{N-1} x[n] W_N^{kn}}_{\text{(нечетные)}}, \quad k = \overline{0, N-1}.$$

Введем новый индекс $v = \overline{0, N/2 - 1}$ такой, что $n = 2v$ для четных n , а для нечетных $n = 2v + 1$. Тогда можно записать

$$X[k] = \sum_{v=0}^{\frac{N}{2}-1} x_{10}[v] W_N^{2vk} + \sum_{v=0}^{\frac{N}{2}-1} x_{11}[v] W_N^{(2v+1)k},$$

где $x_{10}[v] = x[2v] = x[n]$ – последовательность отсчетов из $x[n]$ с четными номерами n , а $x_{11}[v] = x[2v + 1] = x[n]$ – последовательность отсчетов с нечетными номерами. Таким образом, последовательность $x[n]$ представляет собой две *подпоследовательности*, как бы «вдвинутые» друг в друга.

Рассмотрим 1-ю сумму

$$\begin{aligned} \sum_{v=0}^{\frac{N}{2}-1} x_{10}[v] W_N^{2vk} &= \sum_{v=0}^{\frac{N}{2}-1} x_{10}[v] e^{-j \frac{2\pi}{N} 2vk} = \sum_{v=0}^{\frac{N}{2}-1} x_{10}[v] e^{-j \frac{2\pi}{N/2} vk} = \\ &= \sum_{v=0}^{\frac{N}{2}-1} x_{10}[v] W_{\frac{N}{2}}^{vk}. \end{aligned}$$

Это не что иное, как $N/2$ -точечное ДПФ подпоследовательности отсчетов с четными номерами исходной последовательности.

Аналогично 2-я сумма представляет собой $N/2$ -точечное ДПФ подпоследовательности отсчетов с нечетными номерами исходной последовательности:

$$\begin{aligned} \sum_{v=0}^{\frac{N}{2}-1} x_{11}[v] W_N^{(2v+1)k} &= W_N^k \sum_{v=0}^{\frac{N}{2}-1} x_{11}[v] e^{-j \frac{2\pi}{N} 2vk} = \\ &= W_N^k \sum_{v=0}^{\frac{N}{2}-1} x_{11}[v] e^{-j \frac{2\pi}{N/2} vk} = W_N^k \sum_{v=0}^{\frac{N}{2}-1} x_{11}[v] W_{\frac{N}{2}}^{vk}. \end{aligned}$$

Таким образом, ДПФ последовательности $x[n]$ может быть выражено через ДПФ двух подпоследовательностей вдвое меньшей длины:

$$X[k] = X_{10}[k] + W_N^k X_{11}[k], \text{ если } k = \overline{0, N/2 - 1}.$$

Получено выражение для вычисления половины отсчетов ДПФ исходной последовательности. Вспомним, что коэффициенты ДПФ образуют периодическую последовательность с периодом, равным количеству отсчетов преобразуемой последовательности. Поэтому $X_{10}[k]$ и $X_{11}[k]$ имеют период $N/2$, следовательно, вторая половина отсчетов может быть найдена с учетом этой периодичности при $k = \frac{N}{2}, N-1$:

$$\begin{aligned} X\left[k + \frac{N}{2}\right] &= X_{10}\left[k + \frac{N}{2}\right] + W_N^k W_N^{N/2} X_{11}\left[k + \frac{N}{2}\right] = \\ &= X_{10}[k] - W_N^k X_{11}[k]. \end{aligned}$$

Итак, ДПФ последовательности $x[n]$ может быть выражено через ДПФ четной и нечетной подпоследовательностей при всех значениях k .

Для примера рассмотрим 8-точечное ДПФ (так принято называть ДПФ от последовательности, содержащей $N = 8$ отсчетов). Вышеописанные операции можно представить структурной схемой, приведенной на рис. П.5.1 [3]. Видно, что любой отсчет последовательности 8-точечного ДПФ можно получить из двух отсчетов 4-точечных ДПФ при помощи операции, называемой «бабочкой» и изображаемой графом (рис. П.5.2). Очевидно, 4-точечные

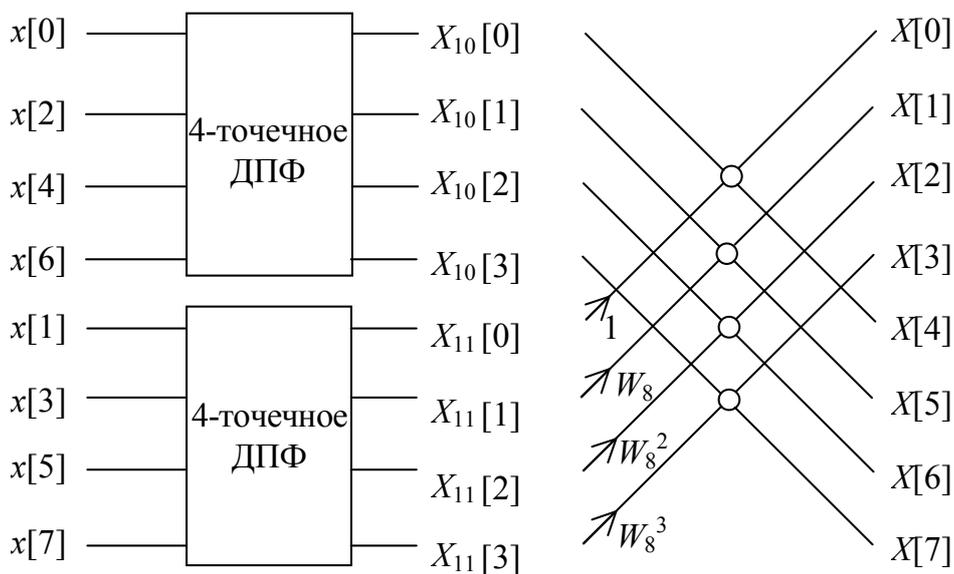


Рис. П.5.1

ДПФ можно таким же способом свести к 2-точечным. (Вообще, если $N/2$ четно, то $N/2$ -точечное ДПФ сводится к $N/4$ -точечным. При этом в последовательности x_{10} выделяются две подпоследовательности – с четными x_{100} и нечетными x_{101} номерами (аналогично в последовательности x_{11} выделяются две

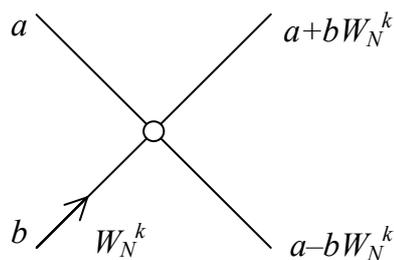


Рис. П.5.2

подпоследовательности – с четными x_{110} и нечетными x_{111} номерами и т.д.) Двухточечное ДПФ изображается «бабочкой» наиболее простого вида, где вес нижней входной ветви равен 1.

Граф-схема 8-точечного ДПФ представлена на рис. П.5.3. Нетрудно видеть, что правый «слой» схемы требует выполнения примерно $N/2$ умножений на комплексные числа («примерно», потому что умножение на $W^0 = 1$ на самом деле не выполняется) и N сложений/вычитаний. Это же справедливо и для всех остальных «слоев». Если $N = 2^r$, как предполагалось вначале, то в схеме

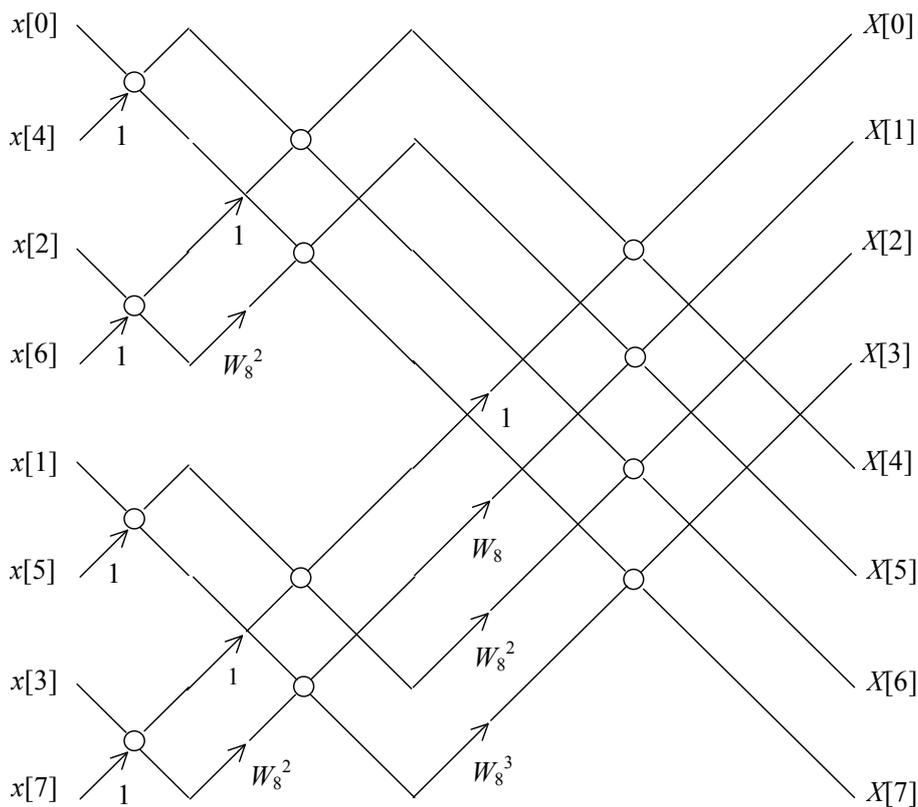


Рис. П 5.3

$r = \log_2 N$ «слоев», поэтому всего для реализации ДПФ по описанному алгоритму требуется $\frac{N}{2} \log_2 N$ (комплексных) умножений и $N \log_2 N$ сложений (вычитаний). Учитывая, что умножение – значительно более длительная и трудоемкая операция, чем сложение, обычно для простоты учитывается только количество умножений. Анализ выражения (П.5.1) показывает, что вычисление каждого отсчета ДПФ требует выполнения N умножений, а поскольку количество коэффициентов равно N , то вычислительные затраты на нахождение ДПФ по общей формуле (П.5.1) составляют N^2 умножений.

Таким образом, описанный в этом приложении алгоритм требует в $\frac{2N}{\log_2 N}$ раз меньше умножений, чем прямое вычисление ДПФ согласно (П.5.1), и по этой причине называется алгоритмом *быстрого преобразования Фурье*. Очевидно, выигрыш тем значительнее, чем больше N .

ФИЛЬТРЫ-ПРОТОТИПЫ И ЧАСТОТНЫЕ ПРЕОБРАЗОВАНИЯ

При изучении аналоговых фильтров-прототипов принято характеризовать фильтр квадратом его амплитудно-частотной характеристики. При этом рассматриваются только фильтры нижних частот, так как остальные фильтры могут быть получены из ФНЧ путем частотных преобразований. Наиболее часто применяются следующие виды фильтров-прототипов.

1. Фильтры Баттерворта

Амплитудно-частотная характеристика фильтра Баттерворта характеризуется *максимальной гладкостью* как в полосе пропускания, так и в полосе подавления. При заданном порядке фильтра требование максимальной гладкости эквивалентно требованию равенства нулю всех производных знаменателя квадрата АЧХ, кроме самой старшей производной; отсюда следует, что эта функция должна иметь вид

$$|H_B(j\Omega)|^2 = \frac{1}{1 + (\Omega/\Omega_c)^{2N}},$$

где Ω_c – частота среза (граничная частота), а N – порядок фильтра. Следовательно, передаточная функция фильтра Баттерворта удовлетворяет уравнению

$$H_B(p)H_B(-p) = \frac{1}{1 + (p/j\Omega_c)^{2N}},$$

причем полюсы правой части уравнения суть корни уравнения

$\frac{p}{j\Omega_c} = \sqrt[2N]{-1}$. Таких корней существует ровно $2N$, и расположены

они в p -плоскости равномерно на окружности радиуса Ω_c в соответствии с выражением

$$p_k = j\Omega_c \sqrt[2N]{-1}, \quad k = \overline{0, 2N-1}.$$

Для примера на рис. П.6.1 показано расположение полюсов квадрата модуля передаточной функции фильтра Баттерворта 3-го порядка. Для того чтобы фильтр был устойчивым, очевидно, полюсы, соответствующие передаточной функции $H_B(p)$, должны

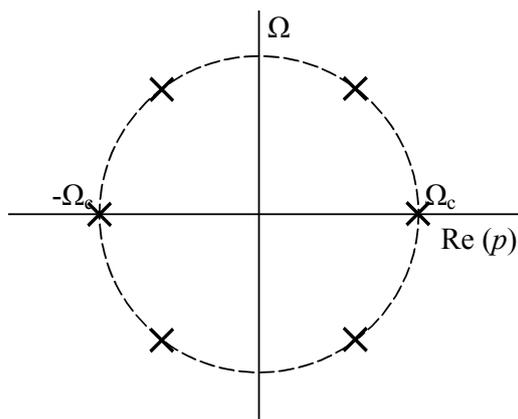


Рис. П.6.1

лежать в левой полуплоскости, а полюсы в правой полуплоскости будут соответствовать функции $H_B(-p)$. Вид АЧХ для фильтров Баттерворта различных порядков показан на рис. П.6.2. При любом N АЧХ остается монотонной, а ее значение на частоте Ω_c всегда равно $1/\sqrt{2}$. Заметим, что при трансформации фильтра Баттерворта в дискретный фильтр

методом билинейного преобразования окружность, на которой располагаются полюсы, отображается в *окружность* на z -плоскости, однако расположение полюсов дискретного фильтра по этой окружности уже не является равномерным.

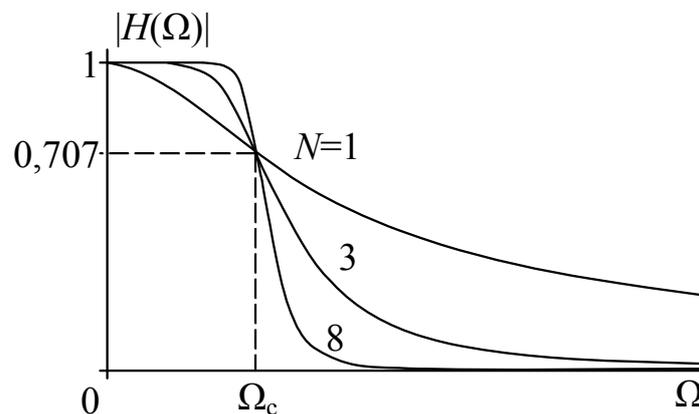


Рис. П.6.2

2. Фильтры Чебышёва

Аппроксимация желаемой АЧХ, основанная на полиномах Чебышёва, обеспечивает более быстрый спад в переходной полосе частот за счет колебаний АЧХ в полосе пропускания (фильтры Чебышёва 1-го рода) или в полосе подавления (фильтры Чебышёва 2-го рода).

Полином Чебышёва порядка n удовлетворяет выражению

$$T_n(x) = \cos[n \arccos(x)]$$

при $x \leq 1$ и выражению

$$T_n(x) = \operatorname{ch}[n \operatorname{Arch}(x)]$$

при $x \geq 1$. Кроме того, для полиномов Чебышёва справедливы рекуррентные соотношения

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Полином Чебышёва при $x \leq 1$ осциллирует между значениями -1 и 1 , а при $x > 1$ монотонно возрастает. Поэтому функция

$$|H_T(\Omega)|^2 = \frac{1}{1 + \varepsilon^2 T_N^2(\Omega/\Omega_n)}$$

при $\Omega < \Omega_n$ (в полосе пропускания для фильтра 1-го рода) осциллирует между значениями 1 и $1/(1 + \varepsilon^2)$, а за пределами этой полосы монотонно убывает.

Передаточная функция фильтра Чебышёва удовлетворяет уравнению

$$H_T(p)H_T(-p) = \frac{1}{1 + \varepsilon^2 T_N^2(p/j\Omega_c)},$$

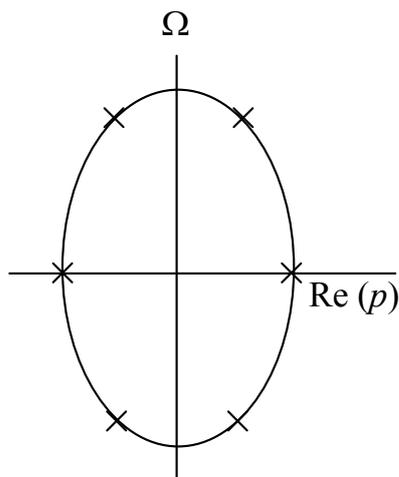


Рис. П.6.3

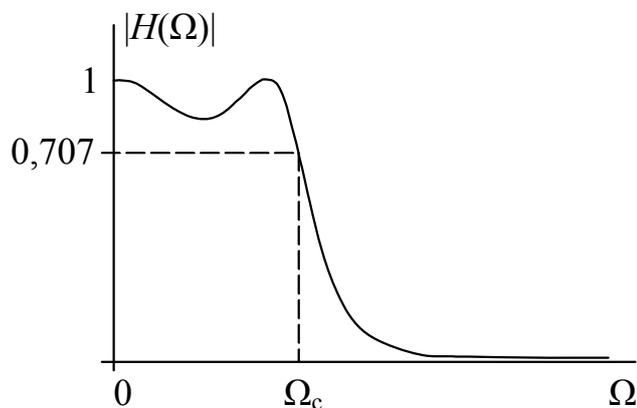


Рис. П.6.4

причем полюсы правой части уравнения лежат в p -плоскости на эллипсе, заключенном между окружностями радиусов $a = \frac{\alpha^{1/N} - \alpha^{-1/N}}{2}$ и $b = \frac{\alpha^{1/N} + \alpha^{-1/N}}{2}$, где $\alpha = \frac{1}{\varepsilon} + \sqrt{1 + \frac{1}{\varepsilon^2}}$.

Для примера на рис. П.6.3 показано расположение полюсов квадрата модуля передаточной функции фильтра Чебышёва 3-го

порядка. Полюсы, соответствующие передаточной функции $H_T(p)$ лежат в левой полуплоскости, а полюсы в правой полуплоскости соответствуют функции $H_T(-p)$. Вид АЧХ для фильтра Чебышёва 3-го порядка показан на рис. П.6.4.

3. Эллиптические фильтры (фильтры Золотарева–Кауэра)

Эллиптические фильтры обеспечивают еще более быстрый спад АЧХ в переходной полосе за счет осцилляций как в полосе пропускания, так и в полосе подавления и основаны на аппроксимации квадрата АЧХ функциями вида [1]

$$|H_{\varepsilon}(\Omega)|^2 = \frac{1}{1 + \varepsilon^2 U_N^2(\Omega/\Omega_n)},$$

где $U_N(\cdot)$ – эллиптическая функция Якоби. Более подробное рассмотрение этой аппроксимации, а также аппроксимации Бесселя содержится, например, в [7].

При построении аналоговых фильтров пользуются частотными преобразованиями [14], позволяющими получить фильтр заданного вида из имеющегося фильтра нижних частот.

Пусть известна передаточная функция $H(p)$ аналогового фильтра нижних частот с частотой среза Ω_c . Введем новую частотную переменную $\bar{\Omega} = \Omega/\lambda$, где λ – масштабный коэффициент. Тогда подстановка в комплексную частотную характеристику фильтра $H(j\Omega)$ вместо Ω новой переменной $\lambda\bar{\Omega}$ даст новую функцию частоты $H(j\lambda\bar{\Omega}) = H'(j\bar{\Omega})$, которая является комплексной частотной характеристикой фильтра нижних частот, но уже с частотой среза $\bar{\Omega}_c = \Omega_c/\lambda$. В самом деле, на частоте среза АЧХ имеет значение $1/\sqrt{2}$, т. е. $|H(j\Omega_c)| = |H(j\lambda\bar{\Omega}_c)| = |H'(j\bar{\Omega}_c)| = 1/\sqrt{2}$, следовательно, коэффициент λ определяется отношением старой и новой частот среза $\lambda = \Omega_c/\bar{\Omega}_c$. Преобразование передаточной функции происходит путем замены переменных $p = \lambda\bar{p}$.

Преобразование ФНЧ в фильтр верхних частот сопровождается инверсией (обращением) оси частот, поэтому частотное преобразование для этого случая определяется выражением $p = \lambda/\bar{p}$, где коэффициент λ определяется так же, как в первом случае, отношением старой и новой частот среза $\lambda = \Omega_c/\bar{\Omega}_c$.

Преобразование ФНЧ в полосовой фильтр выполняется путем подстановки

$$p = \frac{1}{B} \left(\frac{-}{p} + \frac{\omega_0^2}{p} \right),$$

где B и ω_0^2 – константы, определяемые частотой среза исходного ФНЧ и заданными граничными частотами $\bar{\Omega}_{п1}$ и $\bar{\Omega}_{п2}$ полосы пропускания полосового фильтра согласно выражениям [14]

$$B = \frac{\bar{\Omega}_{п2} - \bar{\Omega}_{п1}}{\Omega_c}, \quad \omega_0^2 = \bar{\Omega}_{п1} \bar{\Omega}_{п2}.$$

Для преобразования ФНЧ в режекторный (полосно-заграждающий) фильтр выполняется подстановка

$$p = \frac{B \bar{p}}{\bar{p}^2 + \omega_0^2}$$

с константами B и ω_0^2 , определяемыми аналогично случаю полосового фильтра.

ДИСКРЕТНОЕ ПРЕОБРАЗОВАНИЕ ГИЛЬБЕРТА

В теории аналоговых сигналов и цепей широко используется метод комплексных амплитуд, суть которого состоит в представлении гармонического колебания $A \cos(\Omega_0 t + \varphi)$ комплексным колебанием

$Ae^{j(\Omega_0 t + \varphi)} = A \cos(\Omega_0 t + \varphi) + jA \sin(\Omega_0 t + \varphi)$, которое можно изобразить вектором длины A на комплексной плоскости, повернутым при $t = 0$ на угол φ относительно ее вещественной оси и вращающимся против часовой стрелки с угловой скоростью Ω_0 . Поскольку гармоническое колебание можно представить в виде

$$A \cos(\Omega_0 t + \varphi) = A \frac{e^{j(\Omega_0 t + \varphi)} + e^{-j(\Omega_0 t + \varphi)}}{2},$$

то его спектральная плотность представляет собой пару δ -функций, смещенных влево и вправо по оси частот на Ω_0 , и домноженных на $A\pi$ (проверьте!).

Комплексному колебанию $Ae^{j(\Omega_0 t + \varphi)}$ соответствует спектральная плотность в виде одной δ -функции $2A\pi\delta(\Omega - \Omega_0)$, смещенной в область положительных частот. Таким образом, при замене гармонического колебания $A \cos(\Omega_0 t + \varphi)$ комплексным колебанием $Ae^{j(\Omega_0 t + \varphi)}$ происходит подавление его спектральной плотности в области отрицательных частот и удвоение – в области положительных частот.

Известно, что спектральная плотность вещественного сигнала обладает свойством сопряженной симметрии, иначе говоря, ее вещественная часть четна, а мнимая – нечетна. Подавляя все спектральные составляющие с отрицательными частотами и удваивая – с положительными, получим комплексное колебание, называемое *аналитическим сигналом* [11]. Метод аналитического сигнала представляет собой обобщение метода комплексных амплитуд на сигналы произвольного вида. Особенно он удобен для описания узкополосных колебаний, так как в этом случае аналитическому сигналу соответствует вектор на комплексной плоскости, медленно меняющаяся длина которого соответствует огибающей узкополосного сигнала, а угол между вектором и вещественной осью равен его начальной фазе. Термин «аналитический сигнал» неприменим к дискретным сигналам, так как на дискретной временной оси не

определена операция дифференцирования (а следовательно, и понятие аналитической функции), однако это не препятствует использованию комплексного представления дискретных сигналов, вполне аналогичных аналитическому сигналу.

Поставим в соответствие вещественному дискретному сигналу $x[n]$ сигнал $x^+[n]$, спектральная плотность которого равна нулю при отрицательных частотах и удвоенной спектральной плотности сигнала $x[n]$ – при положительных. Очевидно, $x^+[n]$ является комплексной последовательностью $x^+[n] = x[n] + j\hat{x}[n]$, причем спектральные плотности вещественной и мнимой частей должны, очевидно, удовлетворять соотношениям

$$\hat{X}(e^{j\omega}) = \begin{cases} -jX(e^{j\omega}), & 0 \leq \omega < \pi, \\ jX(e^{j\omega}), & -\pi \leq \omega < 0, \end{cases}$$

поэтому мнимая часть $\hat{x}[n]$ может быть получена из вещественной $x[n]$ при помощи ЛИС-цепи с комплексной частотной характеристикой

$$H(e^{j\omega}) = \begin{cases} -j, & 0 \leq \omega < \pi, \\ j, & -\pi \leq \omega < 0 \end{cases}$$

и импульсной характеристикой

$$h[n] = \frac{1}{2\pi} \int_{-\pi}^0 je^{j\omega n} d\omega - \frac{1}{2\pi} \int_0^{\pi} je^{j\omega n} d\omega = \begin{cases} \frac{2 \sin^2(\pi n/2)}{\pi n}, & n \neq 0, \\ 0, & n = 0, \end{cases}$$

изображенной на рис. П. 7.1. Такая цепь называется дискретным преобразователем Гильберта и является идеальным фазовращателем, так как увеличивает на $\pi/2$ фазу всех спектральных составляющих сигнала с отрицательными частотами и уменьшает на $\pi/2$ фазу всех остальных.

Заметим, что, в отличие от «аналогового» преобразователя Гильберта с импульсной характеристикой $h(t) = \frac{1}{\pi t}$, содержащей разрыв второго рода в нуле, дискретный преобразователь Гильберта имеет ограниченную импульсную характеристику при всех значениях аргумента. Тем не менее дискретный преобразователь

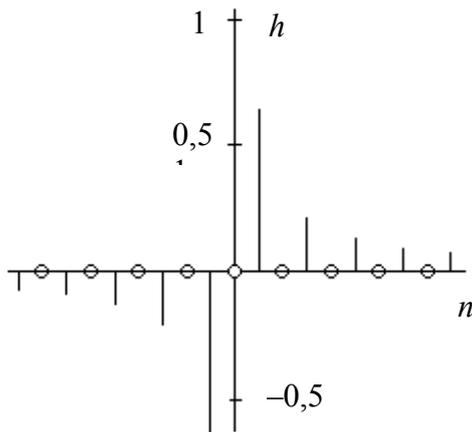


Рис. П.7.1

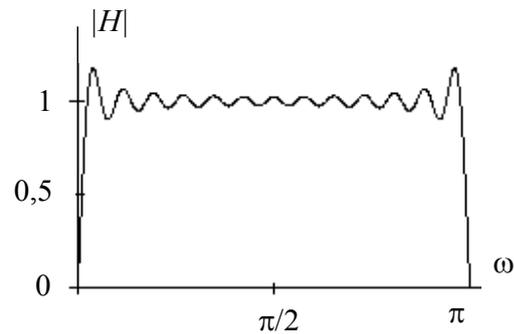


Рис. П.7.2

Гильберта также нельзя реализовать *точно* по причине некаузальности и бесконечной протяженности. Даже если бы его можно было реализовать, он был бы неустойчивым в силу слишком медленного убывания импульсной характеристики (сравните с идеальным ФНЧ в примере 2.10). Приближенная реализация дискретного преобразователя Гильберта представляет собой каузальный трансверсальный фильтр, полученный усечением импульсной характеристики и сдвигом ее вправо по временной оси. Следствием усечения являются гиббсовские осцилляции АЧХ, для уменьшения которых можно воспользоваться каким-либо окном.

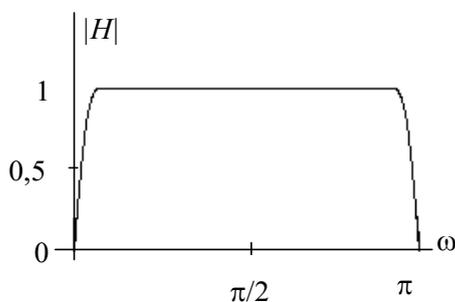


Рис. П.7.3

На рис. П.7.2 показана амплитудно-частотная характеристика «усеченного» дискретного преобразователя Гильберта с гиббсовскими осцилляциями вследствие усечения (длина импульсной характеристики равна 51 отсчету), а на рис. П.7.3 – амплитудно-частотная характеристика, соответствующая этой же импульсной характеристике, умноженной на окно Хэмминга. Стоит

отметить, что на фазочастотную характеристику дискретного преобразователя Гильберта не влияют ни усечение импульсной характеристики, ни применение окна (см. п. 3.8).

СПИСОК ИСПОЛЬЗУЕМЫХ СОКРАЩЕНИЙ

АКП	автокорреляционная последовательность
АКФ	автокорреляционная функция
АЛУ	арифметико-логическое устройство
АМ	амплитудная модуляция
АР	авторегрессионная (модель)
АРСС	(модель) авторегрессии – скользящего среднего
АРУ	автоматическая регулировка усиления
АЦП	аналого-цифровой преобразователь
АЧХ	амплитудно-частотная характеристика
БИС	большая интегральная схема
БИХ	бесконечная импульсная характеристика
БПО	блок предварительной обработки
БПФ	быстрое преобразование Фурье
ВРК	временное разделение каналов
ГДШ	генератор дискретного шума
ГОС	генератор опорных сигналов
ГПП	генератор периодической последовательности
ДПФ	дискретное преобразование Фурье
ЗУ	запоминающее устройство
КИХ	конечная импульсная характеристика
КЛП	коэффициенты линейного предсказания
КРК	кодированное разделение каналов
КЧХ	комплексная частотная характеристика
ЛИС	линейная инвариантная к сдвигу (цепь)
ЛЧМ	линейная частотная модуляция
МЛБ	матричный логический блок
НПД	нуль-полосная диаграмма
ОБП	одна боковая полоса
ОВОВ	«ограниченный вход – ограниченный выход»
ОДПФ	обратное дискретное преобразование Фурье
ОЗУ	оперативное запоминающее устройство
ПАВ	поверхностные акустические волны
ПВМ	программируемая вентиляционная матрица
ПДП	прямой доступ к памяти
ПЗУ	постоянное запоминающее устройство
ПКМБ	программируемый коммутируемый матричный блок
ПЛИС	программируемая логическая интегральная схема

ПЛИМ	программируемая логическая матрица
ПМЛ	программируемая матричная логика
СПМ	спектральная плотность мощности
СС	(модель) скользящего среднего
УВХ	устройство выборки-хранения
ФАР	фазированная антенная решетка
ФНЧ	фильтр нижних частот
ФПП	фильтр погрешности предсказания
ФЧХ	фазочастотная характеристика
ЦАП	цифроаналоговый преобразователь
ЦОС	цифровая обработка сигналов
ЦПЭ	центральный процессорный элемент
ЦСА	цифровой спектральный анализ
ЦСП	цифровой сигнальный процессор
ЦФ	цифровой фильтр
ЦФОС	цифровой фильтр основной селекции
ЧРК	частотное разделение каналов
ША	шина адресов
ШД	шина данных

ЛИТЕРАТУРА

1. *Голд Б., Рейдер Ч.* Цифровая обработка сигналов. – М.: Сов. радио, 1973. – 368 с.
2. *Оппенгейм А.В., Шафер Р.В.* Цифровая обработка сигналов: Пер. с англ. – М.: Связь, 1979. – 416 с.
3. *Рабинер Л., Гоулд Б.* Теория и применение цифровой обработки сигналов: Пер. с англ. – М.: Мир, 1978. – 848 с.
4. *Гольденберг Л.М., Матюшкин Б.Д., Поляк М.Н.* Цифровая обработка сигналов. – М.: Радио и связь, 1990. – 256 с.
5. *Гольденберг Л.М., Матюшкин Б.Д., Поляк М.Н.* Цифровая обработка сигналов: Справочник. – М.: Радио и связь, 1985. – 312 с.
6. *Куприянов М.С., Матюшкин Б.Д.* Цифровая обработка сигналов: процессоры, алгоритмы, средства проектирования. – СПб.: Политехника, 1999. – 592 с.
7. *Куприянов М.С. и др.* Техническое обеспечение цифровой обработки сигналов: Справочник. – СПб.: Политехника, 2000. – 752 с.
8. *Солонина А.И., Улахович Д.А., Яковлев Л.А.* Цифровые процессоры обработки сигналов фирмы Motorola. – СПб.: БХВ-Петербург, 2000. – 512 с.
9. *Цикин Ю.А.* Дискретно-аналоговая обработка сигналов. – М.: Радио и связь, 1982. – 160 с.
10. *Колмогоров А.Н., Фомин С.В.* Элементы теории функций и функционального анализа. – М.: Наука, 1972. – 496 с.
11. *Френкс Л.* Теория сигналов. – М.: Связь, 1970. – 248 с.
12. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров. – М.: Мир, 1970. – 720 с.
13. *Баскаков С.И.* Радиотехнические цепи и сигналы. – М.: Высш. шк., 1988. – 448 с.
14. *Антонью А.* Цифровые фильтры: анализ и проектирование. – М.: Радио и связь, 1983. – 320 с.
15. *Левин Б.Р.* Теоретические основы статистической радиотехники. – М.: Радио и связь, 1989. – 656 с.
16. *Даджион Д., Мерсеро Р.* Цифровая обработка многомерных сигналов / Пер. с англ. – М.: Мир, 1988. – 488 с.
17. *Вострецов А.Г.* Правило обнаружения сигнала на фоне импульсных помех и шумов неизвестной интенсивности для автоматизированных систем контроля // Радиотехника и электроника. – 1997. – Т. 42. – № 6. – С. 706 – 711.
18. *Вострецов А.Г., Васюков В.Н.* Эффект «джиттера» в устройствах дискретной обработки сигналов // Докл. СО АН ВШ. – 2000. – № 1. – С. 62 – 70.
19. *Сиберт У.М.* Цепи, сигналы, системы: В 2-х ч. Ч.2: Пер. с англ. – М.: Мир, 1988. – 360 с.

20. *Введение* в цифровую фильтрацию / Под ред. Р. Богнера, А. Константинодиса. – М.: Мир, 1976. – 216 с.
21. *Каппелини В., Константинодис А.Дж., Эмилиани П.* Цифровые фильтры и их применение. – М.: Энергоатомиздат, 1983.
22. *Марпл-мл.С.Л.* Цифровой спектральный анализ и его приложения: Пер. с англ. – М.: Мир, 1990. – 584 с.
23. *Бендат Дж., Пирсол А.* Прикладной анализ случайных данных: Пер. с англ. – М.: Мир, 1989. – 540 с.
24. *ADSP-2100 Family User Manual.* <http://www.analog.com>
25. *TMS320 DSP Development Support. Reference Guide.* Texas Instruments Incorporated. Literature Number: SPRU011F May 1998. <http://www.ti.com>
26. *Smith S.V.* The Scientist and Engineer's Guide to Digital Signal Processing. – San Diego: California Tech. Publ., California. <http://www.dspguide.com>
27. *Стешенко В.Б.* ПЛИС фирмы ALTERA: проектирование устройств обработки сигналов. – М.: ДОДЭКА, 2000. – 128 с.
28. *Применение* цифровой обработки сигналов: Пер. с англ. / Под ред. Э.Оппенгейма. – М.: Мир, 1980. – 550 с.
29. Прием и обработка сигналов в авиационных радиоустройствах / Под ред. В.В. Криницина. – М.: Транспорт, 1992. – 328 с.
30. *Рабинер Л.Р., Шафер Р.В.* Цифровая обработка речевых сигналов: Пер. с англ. – М.: Радио и связь, 1981. – 496 с.
31. *Оппенгейм А., Шафер Р., Стокхэм Т.* Нелинейная фильтрация сигналов, представленных в виде произведения и свертки // Тр. ин-та инженеров по электронике и радиотехнике. – 1968. – Т. 56. – № 8. – С. 5 – 46.
32. *Ратынский М.В.* Основы сотовой связи. – М.: Радио и связь, 2000. – 248 с.
33. *Маркел Дж.Д., Грэй А.Х.* Линейное предсказание речи: Пер. с англ. – М.: Связь, 1980. – 308 с.
34. *Прокис Дж.* Цифровая связь. – М.: Радио и связь, 2000. – 798 с.

СОДЕРЖАНИЕ

Предисловие	5
1. ВВЕДЕНИЕ	8
2. МАТЕМАТИЧЕСКИЕ ОСНОВЫ ЦОС	11
2.1. Элементарные последовательности	11
2.2. Пространство последовательностей	13
2.3. Базисы пространства последовательностей	19
2.4. Последовательности и дискретные цепи	22
2.5. Дискретная свертка и ее свойства	26
2.6. Представление последовательностей полиномами и формальными рядами	29
2.7. z -Преобразование	32
2.8. Единичная окружность	38
2.9. Преобразование Фурье	40
2.10. Дискретное преобразование Фурье	46
Упражнения	54
Вопросы для самоконтроля	55
3. СТРУКТУРНЫЕ СХЕМЫ И РАЗНОСТНЫЕ УРАВНЕНИЯ ЛИС-ЦЕПЕЙ	58
3.1. Структура цепи с конечной импульсной характеристикой	58
3.2. Разностное уравнение КИХ-цепи	60
3.3. Разностное уравнение и структурная схема рекурсивной цепи	62
3.4. Разностное уравнение и структура ЛИС-цепи конечного порядка	67
3.5. Строение импульсной характеристики ЛИС-цепи конечного порядка	69
3.6. Комплексная частотная характеристика ЛИС-цепи конечного порядка	76
3.7. Все пропускающие и минимально-фазовые цепи	79
3.8. ЛИС-цепи с линейной фазочастотной характеристикой	81
3.9. Рекурсивная вычислимость	86
3.10. Описание цепей направленными графами и матричными разностными уравнениями	89
3.11. Формы реализации ЛИС-цепей	95
Упражнения	99
Вопросы для самоконтроля	102

4. СЛУЧАЙНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ И ЛИС-ЦЕПИ	103
4.1. Стационарные случайные последовательности	103
4.2. Воздействие стационарных случайных последовательностей на ЛИС-цепи	107
Упражнения	112
Вопросы для самоконтроля	112
5. МНОГОМЕРНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ И ЛИС-ЦЕПИ	113
5.1. Многомерные последовательности	113
5.2. Многомерные ЛИС-цепи	118
5.3. Многомерные случайные последовательности и ЛИС-цепи	125
Упражнения	127
Вопросы для самоконтроля	128
6. СИНТЕЗ ЦИФРОВЫХ ФИЛЬТРОВ	129
6.1. Взаимосвязь аналоговых, импульсных и дискретных сигналов (спектральные и энергетические соотношения)	129
6.2. Задача синтеза цифрового фильтра как задача аппроксимации	136
6.3. Синтез КИХ-фильтров	138
6.3.1. Метод взвешивания (метод функций окна)	138
6.3.2. Метод частотной выборки	143
6.3.3. Метод быстрой свертки	146
6.4. Синтез БИХ-фильтров методом аналого-цифровой трансформации	150
6.4.1. Метод инвариантности импульсной характеристики	151
6.4.2. Метод билинейного преобразования	156
6.4.3. Частотные преобразования цифровых фильтров	160
Упражнения	164
Вопросы для самоконтроля	164
7. РЕАЛИЗАЦИЯ ЦИФРОВЫХ ФИЛЬТРОВ	166
7.1. Различия между дискретными и цифровыми цепями	166
7.2. Дрожание отсчетов (джиттер)	168
7.3. Представление чисел в цифровых устройствах	176
7.4. Эффекты квантования	178
7.4.1. Квантование сигналов	178
7.4.2. Квантование коэффициентов цифрового фильтра	184
7.4.3. Предельные циклы	187
7.5. Аппаратурная и программная реализация цифровых фильтров	190
Упражнения	192
Вопросы для самоконтроля	193
8. ЦИФРОВОЙ СПЕКТРАЛЬНЫЙ АНАЛИЗ И ЕГО ПРИМЕНЕНИЕ	194
8.1. Краткий обзор методов ЦСА	195
8.2. Периодограммный метод ЦСА	196
8.3. Коррелограммный метод ЦСА	199
8.4. Параметрические модели случайных процессов	202
8.5. Спектральное оценивание по методу авторегрессии	206
Упражнения	209
Вопросы для самоконтроля	209

9. СИГНАЛЬНЫЕ ПРОЦЕССОРЫ И ИХ ПРИМЕНЕНИЕ	211
9.1. Архитектурные особенности сигнальных процессоров.....	212
9.2. Средства разработки приложений для сигнальных процессоров	218
9.3. Альтернатива сигнальным процессорам – программируемые логические интегральные схемы.....	220
Вопросы для самоконтроля.....	222
10. ЦИФРОВАЯ ОБРАБОТКА ВРЕМЕННЫХ И ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ СИГНАЛОВ В СИСТЕМАХ ПОДВИЖНОЙ РАДИОСВЯЗИ	223
10.1. Цифровая согласованная фильтрация.....	223
10.2. Интерполяция и децимация дискретных сигналов.....	226
10.3. Перенос спектра дискретного сигнала.....	229
10.4. Формирование сигнала с одной боковой полосой.....	230
10.5. Цифровая демодуляция сигналов.....	231
10.6. Цифровая обработка пространственно-временных сигналов.....	232
10.7. Цифровая обработка речевых сигналов.....	238
10.7.1. Модели речевых сигналов	238
10.7.2. Кратковременное преобразование Фурье	241
10.7.3. Принцип действия вокодера	243
10.7.4. Гомоморфная обработка речевых сигналов.....	245
10.8. Цифровая обработка сигналов в системах сотовой связи.....	247
Вопросы для самоконтроля.....	255
ЗАКЛЮЧЕНИЕ	256
Приложение 1. Основные сведения из теории функций комплексного переменного	257
Приложение 2. Свойства z -преобразования	260
Приложение 3. Прямое и обратное ДПФ.....	269
Приложение 4. Структура КИХ-цепи на основе интерполяционной формулы Лагранжа	270
Приложение 5. Быстрое преобразование Фурье	272
Приложение 6. Фильтры-прототипы и частотные преобразования	277
Приложение 7. Дискретное преобразование Гильберта.....	282
СПИСОК ИСПОЛЬЗУЕМЫХ СОКРАЩЕНИЙ	285
ЛИТЕРАТУРА	287

УЧЕБНОЕ ИЗДАНИЕ

Василий Николаевич Васюков

**ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ
И СИГНАЛЬНЫЕ ПРОЦЕССОРЫ
В СИСТЕМАХ ПОДВИЖНОЙ РАДИОСВЯЗИ**

Учебник

Редактор *Т.П. Петроченко*
Технический редактор *Н.В. Белова*
Художник-дизайнер *А. В. Волошина*
Компьютерная верстка *В. Ф. Ноздрева*

Лицензия ИД № 04303 от 20.03.01
Подписано в печать 16.05.06
Формат 60×90 1/16. Бумага офсетная
Уч.-изд. л. 21,9. Печ. л. 18,25
Тираж 3000 экз. (II-й з-д 501–750 экз.). Заказ № 659

Издательство Новосибирского государственного
технического университета
630092, г. Новосибирск, пр. К. Маркса, 20
Тел. (383) 346-31-87
E-mail: office@publish.nstu.ru

Отпечатано в типографии
Новосибирского государственного технического университета
630092, г. Новосибирск, пр. К. Маркса, 20



Васюков Василий Николаевич – доктор технических наук, профессор. Имеет около 100 опубликованных работ, в том числе два учебника, несколько учебных пособий.

На преподавательской работе с 1983 года. Специализировался в области теории сигналов и цепей (дисциплины «Радиотехнические цепи и сигналы», «Цифровая обработка сигналов и сигнальные процессоры в системах подвижной радиосвязи», «Теория электрической

связи» и др.)

Область научных интересов – статистические методы обработки и анализа дискретных сигналов и изображений.