

APPLIED
TIME
SERIES
ANALYSIS

VOLUME 1
BASIC TECHNIQUES

Robert K. Otnes

Electromagnetic Systems Laboratories, Sunnyvale, California

Loren Enochson

GenRad Inc.

Acoustics, Vibration and Analysis Division

A Wiley-Interscience Publication

John Wiley and Sons

New York · Chichester · Brisbane · Toronto
1978

Р. ОТНЕС
Л. ЭНОКСОН

ПРИКЛАДНОЙ
АНАЛИЗ
ВРЕМЕННЫХ
РЯДОВ

ОСНОВНЫЕ
МЕТОДЫ

Перевод с английского
В. И. ХОХЛОВА

под редакцией
И. Г. ЖУРБЕНКО

ИЗДАТЕЛЬСТВО «МИР»
МОСКВА 1982

к решению таких упражнений, читателю следует сначала просмотреть приложение А, с тем чтобы найти и переписать тестовую программу и программы, связанные с ней.

Изрядное количество материала, включенное в книгу для полноты изложения, можно без ущерба опустить при отборе материала для чтения лекций по этому предмету. Это относится, например, к большей части гл. 4 о расчете цифровых фильтров и большей части гл. 6, в которой обсуждаются детали быстрого преобразования Фурье. При первом знакомстве с книгой можно опустить и разделы гл. 9, посвященные многомерным процессам.

Нам хотелось бы поблагодарить профессоров Августина Х. Грея (Калифорнийский университет, Санта-Барбара), Лоуренса П. Мак-Нейми (Калифорнийский университет, Лос-Анджелес) и Руперта Стритса (Университет Калгари) за полезные замечания при чтении черновика рукописи; Стива Барбера, Лэрри Моултона и Эда Онстида за помощь в графическом оформлении книги; Луизу Бозин и Джуди Хелблинг за большую работу по подготовке рукописи; Беатрис Шубе, нашего терпеливого редактора; наконец, студентов, помогавших нам выявлять и устранять неточности в первоначальной редакции книги.

*Роберт К. Отнес
Лорен Эноксон*

Пало-Альто, Калифорния
Лос-Альто, Калифорния
Февраль 1978

ПРЕДВАРИТЕЛЬНЫЕ ПОНЯТИЯ

1.1. ВВЕДЕНИЕ

Обзор основных математических и статистических понятий, необходимых для чтения книги, приводится в двух первых главах. Содержание этой главы составляют математические методы, применяемые при анализе временных рядов. Среди них наиболее важный — преобразование Фурье и его модификации. Преобразование Фурье и его быстрая реализация на цифровых устройствах в последнее время обрели особую важность и служат основным инструментом, используемым в книге.

Понятия линейных систем, переходных функций и свертки, приведенные вместе, весьма важны для понимания дальнейшего материала.

В конце главы помещены упражнения. Для решения некоторых из них потребуются написать небольшие программы. Если читатель намерен решать задачи, ему лучше всего сразу ознакомиться с приложением А, в котором содержатся подпрограммы для ЭВМ, упрощающие написание программ. Кроме того, в упражнениях даны численные значения для тестов и соответствующие результаты. Прежде чем пытаться использовать программы, рекомендуется прогнать тесты и сверить ответы.

1.2. ОПЕРАЦИИ НАД КОМПЛЕКСНЫМИ ЧИСЛАМИ

Квадратный корень из -1 , как принято в технической литературе, будем обозначать буквой j (в отличие от обозначения i , принятого в математике): $j = \sqrt{-1}$, а комплексное число c — записывать в виде $c = a + jb$. Здесь a — действительная, а b — мнимая части c ; соответствующие обозначения: $\operatorname{Re} c = a$, $\operatorname{Im} c = b$. Число $c^* = \operatorname{Re} c - j \operatorname{Im} c$ называется комплексно-сопряженным с c .

Сумма, произведение и частное комплексных чисел вычисляются по формулам

$$\begin{aligned}(a + jb) + (g + jh) &= a + g + j(b + h), \\(a + jb)(g + jh) &= ag - bh + j(ah + bg), \\ \frac{a + jb}{g + jh} &= \frac{a + jb}{g + jh} \cdot \frac{g - jh}{g - jh} = \frac{ag + bh}{g^2 + h^2} + j \frac{bg - ah}{g^2 + h^2}\end{aligned}$$

(в последнем случае предполагается, что $g^2 + h^2 \neq 0$). Абсолютная величина, или модуль комплексного числа $c = a + jb$ обозначается $|c|$; имеем

$$|c| = \sqrt{cc^*} \text{ или } |a + jb| = \sqrt{a^2 + b^2},$$

ибо $cc^* = (a + jb)(a - jb) = a^2 + b^2$. Ясно, что

$$\left| \frac{a + jb}{g + jh} \right| = \frac{|a + jb|}{|g + jh|},$$

если $g^2 + h^2 \neq 0$.

Соотношение Эйлера. Следующая формула, известная под названием соотношения Эйлера:

$$e^{j\theta} = \exp(j\theta) = \cos \theta + j \sin \theta,$$

будет постоянно использоваться на протяжении всей книги. Ее доказательство можно найти во многих учебниках, поэтому оно здесь не приводится. Для $\theta = \pi$ соотношение можно переписать в виде $e^{j\pi} + 1 = 0$. Эта формула, несомненно, — одно из самых красивых выражений в математике; оно устанавливает связь пяти самых распространенных величин.

Для комплексного числа $(a + jb)$ можно ввести *полярные координаты* по правилу $r = |a + jb|$, $\operatorname{tg} \theta = b/a$; при этом $a + jb = r \exp(j\theta)$.

Примером функции комплексного переменного может служить $f(z) = z^2 + z + 1$, где $z = x + jy$ (x и y — действительные переменные). Ее можно переписать в виде

$$f(x, y) = (x + jy)^2 + (x + jy) + 1 = (x^2 - y^2 + x + 1) + j(2xy + y).$$

Пусть $f(z)$ — дробно-рациональная функция от z , т. е. отношение двух полиномов от z :

$$f(z) = \frac{N(z)}{D(z)}.$$

Если полиномы $N(z)$ и $D(z)$ не имеют общих корней, то *нулями* $f(z)$ называются корни уравнения $N(z) = 0$. Корни уравнения $D(z) = 0$ называются *полюсами* $f(z)$. Полюсы и нули дробно-рациональной функции определяют ее с точностью до постоянного множителя. Иными словами, если z_1^n, \dots, z_N^n — нули функции $f(z)$, а z_1^d, \dots, z_D^d — ее полюсы, то найдется такое число c , что

$$f(z) = c \prod_{i=1}^N (z - z_i^n) / \prod_{k=1}^D (z - z_k^d).$$

Хорошо известная основная теорема алгебры утверждает, что любой (отличный от постоянного) многочлен в поле комплексных чисел имеет нули.

1.3. ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Преобразование Фурье — главный инструмент, используемый в книге. Хотя бóльшая часть необходимого материала изложена в следующих главах, читателю, совершенно незнакомому с преобразованием Фурье, рекомендуем обратиться к книгам Колмогорова и Фомина (1976), Винера (1963) или Су (1967).

По-видимому, читателю знакома та форма преобразования Фурье, которая переводит временную функцию $x(t)$ в частотную функцию $X(f)$ посредством интегрального соотношения

$$X(f) = \int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft) dt.$$

Иногда его записывают в виде

$$X(f) = F[x(t)].$$

Исходная независимая переменная t для времени, измеряемая, как правило, в секундах, изменяется в интервале $(-\infty, \infty)$. Единицами измерения новой независимой переменной служат герцы, обозначаемые Гц. Интервалом ее изменения также будет $(-\infty, \infty)$. Иногда вместо f используется переменная ω ; соотношение между ними определяется равенством $\omega = 2\pi f$. Величина ω измеряется в *радианах* за единицу времени.

Размерности ω и f одинаковы и равны $1/t$. Как будет видно из дальнейшего, в некотором отношении функция $X(f)$ имеет свойства, похожие на «обращения» соответствующих свойств $x(t)$, и обратно

Для произвольной функции $x(t)$ не всегда существует функция $X(f)$. Можно привести целые классы функций, не имеющих преобразования Фурье. На практике это не столь уж важная проблема. Все функции, которые встретятся в этой книге, такое преобразование будут иметь.

Преобразование Фурье обратимо. Данное выше определение может быть обращено:

$$x(t) = \int_{-\infty}^{\infty} X(f) \exp(j2\pi ft) df.$$

Это соотношение иногда записывается в виде

$$x(t) = F^{-1}[X(f)].$$

Результат этого преобразования было бы лучше обозначать $\hat{x}(t)$, поскольку встречаются «патологические» случаи, в которых исходная функция $x(t)$ и результат $\hat{x}(t)$ обращения преобразо-

вания Фурье функции $x(t)$ незначительно отличаются. Пользуясь математической терминологией, можно сказать, что $x(t)$ и $\hat{x}(t)$ равны *почти всюду*. «Почти всюду» означает, что существует множество точек, в которых они не совпадают, но не найдется ни одного интервала, на котором они не совпадали бы¹⁾. Эта проблема для функций тех типов, которые будут рассматриваться ниже, большого значения не имеет.

Если вместо переменной f используется переменная ω , то приведенное выше уравнение примет вид

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) \exp(j\omega t) d\omega.$$

Дельта-функции. С преобразованием Фурье тесно связано понятие дельта-функции. Дельта-функцию $\delta(t)$ можно попытаться представить себе как функцию, график которой имеет бесконечную высоту, нулевую ширину и площадь, равную единице. Такое описание нельзя считать строгим в математическом смысле. С точки зрения использования дельта-функций полезными будут следующие два замечания: 1) они предназначаются для использования в интегралах и 2) должны аппроксимироваться последовательностью функций, предел которой обладает соответствующими свойствами.

Первое утверждение приводит к следующему математически более строгому определению дельта-функции:

$$x(t_0) = \int_{-\infty}^{\infty} x(t) \delta(t - t_0) dt.$$

Отметим, что левая часть этого равенства постоянна. Поэтому применение дельта-функции и интегрирование в правой части могут представляться в виде процесса, дающего значение $x(t)$ при определенном значении t_0 .

В качестве последовательности, о которой шла речь во втором утверждении, можно взять последовательность ($i = 1, 2, \dots$)

$$d_i(t) = \begin{cases} i, & \text{если } -\frac{1}{2i} \leq t < \frac{1}{2i}, \\ 0 & \text{в противном случае.} \end{cases}$$

График этой функции имеет высоту i , ширину $1/i$ и ограничивает площадь, равную 1.

¹⁾ Это неточное определение выражения «почти всюду». Тем, кого интересует этот вопрос, рекомендуются книги Колмогорова и Фомина (1976) и Дима и Маккина (1972).

Определим последовательность $x_i(t_0)$ с помощью интеграла

$$x_i(t_0) = \int_{-\infty}^{\infty} x(t) d_i(t-t_0) dt.$$

Очевидно, что

$$x_i(t_0) = i \int_{t_0-(2i)^{-1}}^{t_0+(2i)^{-1}} x(t) dt.$$

Для достаточно хорошей функции $x(t)$ и при достаточно большом i эту функцию можно аппроксимировать на малом интервале $(t_0-(2i)^{-1}, t_0+(2i)^{-1})$ ее значением в центре интервала. В этом случае предыдущее соотношение можно записать в виде

$$x_i(t_0) \cong ix(t_0) \int_{t_0-(2i)^{-1}}^{t_0+(2i)^{-1}} dx = ix(t_0) t \Big|_{t_0-(2i)^{-1}}^{t_0+(2i)^{-1}} = x(t_0).$$

Если приведенные выше соображения формализовать, то получится доказательство того факта, что последовательность $x_i(t_0)$ сходится к пределу $x(t_0)$.

Рассуждения такого типа применяются всякий раз, когда необходима формализация. Заметим, что функция $d_i(t)$, приведенная выше, определяется не единственным способом. Есть немало последовательностей, которые отвечают цели так же хорошо. Рассмотрим, например, последовательность $e_i(t)$, определенную следующим образом:

$$e_i(t) = \begin{cases} -|t|i^3 + i, & |t| \leq i^{-1}, \\ 0, & |t| > i^{-1}. \end{cases}$$

Графики этих функций представляют собой равнобедренные треугольники с высотой i и основанием $2/i$.

В дальнейшем такие последовательности и предельные переходы рассматриваться не будут. Вместо них всегда будет использоваться определение через интеграл.

Продемонстрируем использование дельта-функции в области частот. Предположим, что требуется получить обратное преобразование функции $\delta(f-f_0)$. Его можно записать и вычислить следующим образом:

$$\int_{-\infty}^{\infty} \delta(f-f_0) \exp(j2\pi ft) df = \exp(j2\pi f_0 t).$$

Таким образом, дельта-функция выделяет значение показательной функции при определенном значении f_0 . Полученное обратное преобразование представляет собой комплексную функцию времени. Если применить соотношение Эйлера, то результат

можно переписать в виде

$$\exp(j2\pi f_0 t) = \cos(2\pi f_0 t) + j \sin(2\pi f_0 t).$$

Непосредственно из этого равенства можно определить преобразования Фурье $\cos(2\pi f_0 t)$ и $\sin(2\pi f_0 t)$. Для этого сначала находят преобразование Фурье функции $\delta(f + f_0)$:

$$\begin{aligned} \int_{-\infty}^{\infty} \delta(f + f_0) \exp(j2\pi f t) df &= \exp(-j2\pi f_0 t) = \\ &= \cos(2\pi f_0 t) - j \sin(2\pi f_0 t). \end{aligned}$$

Теперь рассмотрим обратное преобразование Фурье суммы двух дельта-функций:

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\delta(f - f_0) + \delta(f + f_0)}{2} \exp(j2\pi f t) df &= \\ = \frac{1}{2} (\cos(2\pi f_0 t) + j \sin(2\pi f_0 t)) + \frac{1}{2} (\cos(2\pi f_0 t) - j \sin(2\pi f_0 t)) &= \\ = \cos(2\pi f_0 t). \end{aligned}$$

Значит, преобразование Фурье $\cos(2\pi f_0 t)$ есть сумма двух дельта-функций $\delta(f - f_0)$ и $\delta(f + f_0)$, каждая из которых берется с

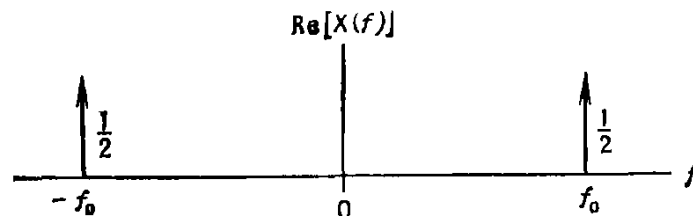


Рис. 1.1. Преобразование Фурье косинуса.

множителем $1/2$. Заметим, что функции — действительные, здесь j не фигурирует.

На рис. 1.1 изображено такое преобразование Фурье. Дельта-функции обычно обозначают стрелками, поскольку на графике бесконечную высоту изобразить невозможно. При этом их множители помещают обычно рядом со стрелками по одну сторону от них.

Этот график показывает, что \cos во временной области переходит в сумму двух дельта-функций в области частот, причем дельта-функции расположены во взятых с минусом и плюсом частотах, соответствующих частоте косинуса. Таким образом, вся информация о $\cos(2\pi f_0 t)$ после преобразования в область частот сконцентрирована в частотах $\pm f_0$ Гц.

Подобная процедура, примененная к $\sin(2\pi f_0 t)$, дает

$$\frac{\delta(f-f_0) - \delta(f+f_0)}{2j}.$$

Сравнение особенностей преобразований $\cos(2\pi f_0 t)$ и $\sin(2\pi f_0 t)$ предоставим читателю. Здесь отметим следующее: 1) преобразование $\cos(2\pi f_0 t)$ дает чисто вещественный результат, в то время как $\sin(2\pi f_0 t)$ — чисто мнимый; 2) оба преобразования имеют вид двух дельта-функций, расположенных в частотах f_0 Гц и $-f_0$ Гц; 3) преобразование косинуса — четная, а преобразование синуса — нечетная функции (определение см. ниже).

Действуя примерно в том же духе, рассмотрим обратное преобразование $\delta(f)$:

$$\int_{-\infty}^{\infty} \delta(f) \exp(j2\pi ft) dt = \exp(0) = 1.$$

Отсюда видно, что преобразование Фурье константы есть дельта-функция в 0 Гц. Тот же результат можно получить, если в преобразовании Фурье $\cos(2\pi f_0 t)$ положить значение f_0 равным нулю.

Симметрические и антисимметрические функции. Функцию $x(t)$ называют *симметрической* (или *четной*), если

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

Подобно этому, функцию $x(t)$ называют *антисимметрической* (или *нечетной*), если

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

Всякую действительную функцию можно представить в виде суммы двух функций, одна из которых симметрическая, а другая — антисимметрическая. Если, например, $x_s(t)$ и $x_a(t)$ — симметрическая и антисимметрическая составляющие функции $x(t)$, то определить их можно из соотношений

$$x(t) = x_s(t) + x_a(t), \quad x(-t) = x_s(t) - x_a(t).$$

Решая эти уравнения, получим

$$x_s(t) = \frac{x(t) + x(-t)}{2}, \quad x_a(t) = \frac{x(t) - x(-t)}{2}.$$

Например, если $x(t)$ — полином третьей степени от t , т.е. $x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$, то $x(-t) = a_0 - a_1 t + a_2 t^2 - a_3 t^3$, поэтому

$$x_s(t) = a_0 + a_2 t^2, \quad x_a(t) = a_1 t + a_3 t^3.$$

Если функция $x(t)$ записана в виде $x_s(t) + x_a(t)$ и показательная функция в преобразовании Фурье — в виде $\cos(2\pi ft)$ —

— $j \sin(2\pi ft)$), то определение преобразования можно переписать следующим образом:

$$\begin{aligned} X(f) &= \int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft) dt = \\ &= \int_{-\infty}^{\infty} [x_s(t) + x_a(t)] [\cos(2\pi ft) - j \sin(2\pi ft)] df = \\ &= \int_{-\infty}^{\infty} x_s(t) \cos(2\pi ft) dt + \int_{-\infty}^{\infty} x_a(t) \cos(2\pi ft) dt \\ &\quad - j \int_{-\infty}^{\infty} x_s(t) \sin(2\pi ft) dt - j \int_{-\infty}^{\infty} x_a(t) \sin(2\pi ft) dt. \end{aligned}$$

Необходимо отметить три дополнительных факта: 1) интеграл от нечетной функции по отрезку $[-A, A]$, где A — любое положительное число, равен 0; 2) произведение двух четных или двух нечетных функций есть четная функция; 3) произведение четной и нечетной функций есть нечетная функция.

Из предыдущего можно видеть, что произведения $x_a(t) \cos(2\pi ft)$ и $x_s(t) \sin(2\pi ft)$ — нечетные функции, и значит, интегралы от них равны 0. Поэтому

$$\begin{aligned} \int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft) dt &= \\ &= \int_{-\infty}^{\infty} x_s(t) \cos(2\pi ft) dt - j \int_{-\infty}^{\infty} x_a(t) \sin(2\pi ft) dt. \end{aligned}$$

Наконец, если $x(t)$ — четная функция, и следовательно, $x_a(t) = 0$ для всех t , то последний результат сведется к равенству

$$\int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft) dt = \int_{-\infty}^{\infty} x(t) \cos(2\pi ft) dt.$$

Причина, по которой мы на этом остановились, состоит в том, что преобразования четных функций всегда действительны и, кроме того, могут быть записаны как косинус-преобразования. Оба этих обстоятельства постоянно используются в примерах. В частности, преобразования симметрических функций изучать намного проще, так как их преобразования — действительные функции — легче изобразить графически.

Важнейшие соотношения, связанные с преобразованием Фурье. Те свойства преобразования Фурье, которые важны для нашей книги, сведены в табл. 1.1 и 1.2. Они составляют лишь малую

Таблица 1.1

Преобразование Фурье — функциональные соотношения

Функция	Преобразование Фурье
1. $x(t)$ $F^{-1}[X(f)]$ $\int_{-\infty}^{\infty} X(f) \exp(j2\pi ft) df$	$X(f)$ $F[x(t)]$ $\int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft) dt$
2. $y(t)$	$Y(f)$
3. $x(t)y(t)$ $\int_{-\infty}^{\infty} x(\tau)y(t-\tau) d\tau$	$\int_{-\infty}^{\infty} X(\eta)Y(f-\eta) d\eta$ $X(f)Y(f)$
4. $x(t-t_0)$ $x(t)e^{j2\pi ft_0}$	$X(f)e^{-j2\pi ft_0}$ $X(f-f_0)$
5. $\frac{dx(t)}{dt}$	$j2\pi fX(f)$
6. $\int_{-\infty}^t x(\tau) d\tau$	$\frac{1}{j\omega} X(f)$ в том случае, если $X(0)=0$
7. $\int_{-\infty}^{\infty} x^2(t) dt$	$\int_{-\infty}^{\infty} X(f) ^2 df$ (равенство Парсеваля)

долю всего класса теорем о преобразовании Фурье, имеющих в настоящее время в литературе, но их достаточно для решения тех задач, которые нам встретятся.

Для примера приведем вывод соотношения 4 из табл. 1.1. Пусть, как обычно, $X(f)$ есть преобразование Фурье $x(t)$ и требуется найти преобразование функции $x(t-t_0)$, т. е. функции $x(t)$, смещенной вправо на величину t_0 . По определению таким преобразованием служит

$$F[x(t-t_0)] = \int_{-\infty}^{\infty} x(t-t_0) \exp(-j2\pi ft) dt.$$

Если произвести замену переменной $\tau = t - t_0$, то $d\tau = dt$, и предыдущее выражение примет вид

$$F[x(t-t_0)] = \int_{-\infty}^{\infty} x(\tau) \exp[-j2\pi f(\tau + t_0)] d\tau.$$

Таблица 1.2

Преобразование Фурье — некоторые пары преобразований

$x(t)$	$X(f)$
1. $x(t) = \begin{cases} 1, & -P/2 \leq t < P/2 \\ 0 & \text{в противном случае} \end{cases}$	$\frac{\sin(\pi f P)}{\pi f}$
2. $\cos(2\pi f_0 t)$	$\frac{1}{2} [\delta(f - f_0) + \delta(f + f_0)]$
3. $\text{s.n}(2\pi f_0 t)$	$\frac{1}{2j} [\delta(f - f_0) - \delta(f + f_0)]$
4. $x(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$	$\frac{1}{2} \delta(f) + \frac{1}{j2\pi f}$
5. $x(t) = 1$ при всех t	$\delta(f)$
6. $x(t) = \delta(t)$	$X(f) = 1$ при всех f
7. $\exp(j2\pi f_0 t)$	$\delta(f - f_0)$
8. $\exp\left(-\frac{t^2}{2a^2}\right)$	$\sqrt{2\pi a^2} \exp\left(-\frac{(2\pi f a)^2}{2}\right)$
9. $\frac{2 \sin(2\pi B t)}{2\pi t}$	$X(f) = \begin{cases} 1, & -B \leq f < B \\ 0 & \text{в противном случае} \end{cases}$

Бесконечные пределы при этом не изменятся, поскольку они сдвинуты лишь на конечную величину t_0 .

Экспонента с показателем, содержащим t_0 , может быть вынесена за знак интегрирования:

$$\begin{aligned} F[x(t - t_0)] &= \exp(-j2\pi f t_0) \int_{-\infty}^{\infty} x(\tau) \exp(-j2\pi f \tau) d\tau = \\ &= \exp(-j2\pi f t_0) X(f), \end{aligned}$$

т. е., после того как эта экспонента вынесена, остается интеграл, который представляет собой просто определение преобразования $X(f)$.

Таким образом, запаздыванию во временной области соответствует умножение на комплексную экспоненту в области частот. Эта теорема о преобразовании Фурье — наиболее важный факт, который используется в работе с выборочными данными.

1.4. ПРЯМОУГОЛЬНЫЙ ИМПУЛЬС И ЕГО ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Преобразование первой из функций в табл. 1.2 также заслуживает особого рассмотрения. Как принято в этой книге, обозначим ее $u(t)$ и определим на временной прямой соотношением

$$u(t) = \begin{cases} 1, & -P/2 \leq t < P/2, \\ 0 & \text{в противном случае.} \end{cases}$$

Как показано на рис. 1.2,а, график этой функции имеет вид симметричного относительно оси $u(t)$ прямоугольника, имеющего ширину P .

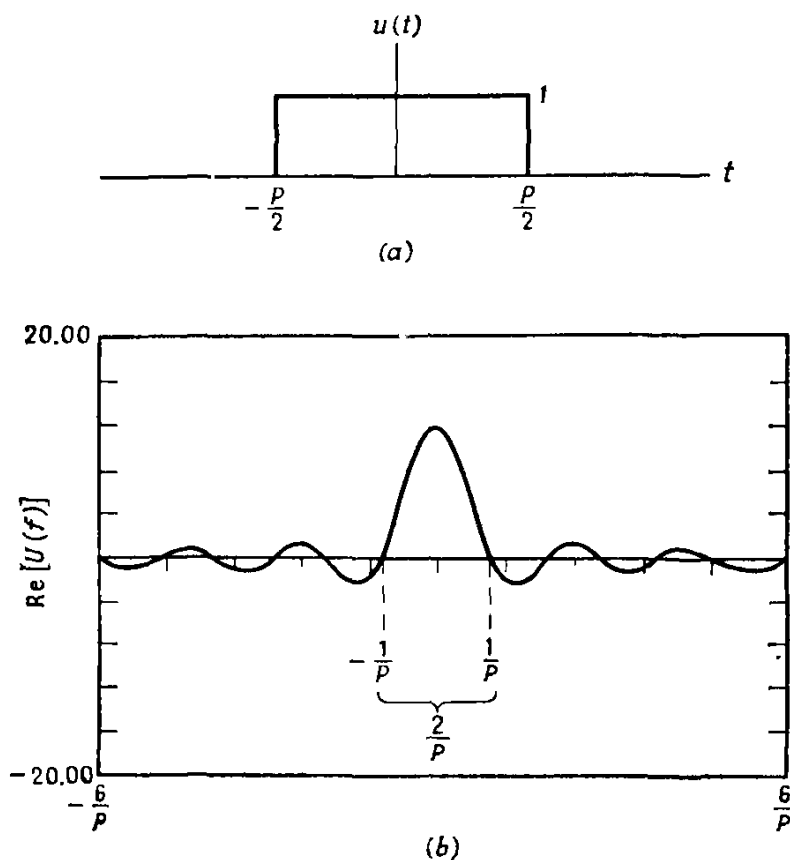


Рис. 1.2. а — прямоугольный импульс; б — его преобразование Фурье, $P = 12$.

Получить преобразование Фурье этой функции несложно:

$$\begin{aligned} U(f) &= \int_{-\infty}^{\infty} u(t) \exp(-j2\pi ft) dt = \int_{-P/2}^{P/2} \exp(-j2\pi ft) dt = \\ &= -\frac{1}{j2\pi f} \exp(-j2\pi ft) \Big|_{-P/2}^{P/2} = \frac{\sin(\pi f P)}{\pi f}. \end{aligned}$$

Оно оказывается четной функцией, представляющей собой произведение двух нечетных функций $\sin(\pi f P)$ и $1/(\pi f)$. Максимальное значение достигается при $f=0$. Его можно вычислить по правилу Лопиталья:

$$\lim_{f \rightarrow 0} \frac{\sin(\pi f P)}{\pi f} = \lim_{f \rightarrow 0} \frac{d}{df} (\sin(\pi f P)) \left(\frac{d}{df} (\pi f) \right)^{-1} = \lim_{f \rightarrow 0} \frac{\pi P \cos(\pi f P)}{\pi} = P.$$

Первые пересечения нулевого уровня происходят при $\pi f P = \pm \pi$, т. е. при $f = \pm P^{-1}$. Поэтому ширина центрального выступа равна $2P^{-1}$, и значит, ширина в области частот обратно пропорциональна ширине во временной области. Это свойство взаимной

сбратимости двух интервалов лучше всего видно для других трех преобразований, приведенных в табл. 1.2.

1. В пятой строке таблицы константа, равная единице на всей бесконечной временной области, преобразуется в дельта-функцию на частотной области. Последняя имеет бесконечно малую ширину.

2. Обратное, как показывает шестая строка таблицы, дельта-функция на временной области, имеющая бесконечно малую ширину, преобразуется в постоянную функцию в области частот, имеющую бесконечный носитель.

3. Как видно из восьмой строки, колоколообразная кривая во временной области переходит в колоколообразную кривую в частотной области. Если взять точки перегиба на каждой из этих кривых, то расстояние между ними для кривой во временной области будет равно $2a$, а расстояние для кривой в частотной области будет равно $(\pi a)^{-1}$, так что эти два расстояния обратно пропорциональны.

1.5. ЛИНЕЙНЫЕ СИСТЕМЫ

В этом параграфе рассматриваются функции $x(t)$ и $y(t)$, между которыми установлено *линейное соотношение*. В частности, здесь рассматриваются функциональные соотношения вида

$$D^{(1)} [y(t)] = D^{(2)} [x(t)],$$

где $D^{(1)}$ и $D^{(2)}$ — линейные дифференциальные операторы. Например, оператор $D^{(1)}$ может иметь вид

$$D^{(1)} = \left[a \frac{d^2}{dt^2} + b \frac{d}{dt} + c \right],$$

где a , b и c — постоянные. Предположим, что и для $x(t)$ имеется подобное выражение. Это означает, что $x(t)$ и $y(t)$ связаны *линейным дифференциальным уравнением с постоянными коэффициентами*.

Как будет показано ниже, существуют еще два способа записи такого соотношения — через *импульсную функцию отклика* и ее преобразование Фурье (или, иначе, передаточную функцию)¹⁾.

Проиллюстрируем содержание этих понятий с помощью одного примера. Допустим, что соотношение определено равенством

$$\frac{d^2 y(t)}{dt^2} + 2\zeta\omega_n \frac{dy(t)}{dt} + \omega_n^2 y(t) = x(t).$$

Здесь

$$D^{(1)} = \left[\frac{d^2}{dt^2} + 2\zeta\omega_n \frac{d}{dt} + \omega_n^2 \right], \quad D^{(2)} = 1.$$

¹⁾ Нередко импульсную функцию отклика называют импульсной характеристикой, а передаточную функцию — частотной характеристикой. — Прим. перев.

С помощью табл. 1.1 можно найти преобразование Фурье этого равенства и получить соотношение

$$-(2\pi f)^2 Y(f) + (j2\pi f) 2\zeta \omega_n Y(f) + \omega_n^2 Y(f) = X(f),$$

где $X(f)$ и $Y(f)$ суть преобразования Фурье соответственно $x(t)$ и $y(t)$. Полагая $\omega_n = 2\pi f_n$ и вынося общий член $Y(f)$ за скобки, после перегруппировки получим

$$H(f) = \frac{Y(f)}{X(f)} = \left[\omega_n^2 \left(1 - \left(\frac{f}{f_n} \right)^2 + \frac{j2\zeta f}{f_n} \right) \right]^{-1}.$$

Отношение $Y(f)/X(f)$, обозначаемое $H(f)$, определяет для каждой системы *передаточную функцию* между двумя функциями.

Об этом отношении можно говорить иначе — как о действии *фильтра* между $x(t)$ и $y(t)$: функция $y(t)$ получается *фильтрацией* функции $x(t)$, а $H(f)$ есть передаточная функция фильтра.

Коэффициент усиления фильтра равен абсолютному значению передаточной функции. В нашем примере для коэффициента усиления мы имеем следующее выражение:

$$|H(f)| = \left| \frac{Y(f)}{X(f)} \right| = \left[\omega_n^2 \left(\left[1 - \left(\frac{f}{f_n} \right)^2 \right]^2 + \left(\frac{2\zeta f}{f_n} \right)^2 \right)^{1/2} \right]^{-1}.$$

Фазовый угол (или просто *фаза*) передаточной функции определяется функцией $\varphi(f)$ по формуле

$$\operatorname{tg} \varphi(f) = \frac{\operatorname{Im} [H(f)]}{\operatorname{Re} [H(f)]}.$$

Так определенная фаза измеряется в радианах. Нередко эту величину переводят в градусы, умножая на $180/\pi$, и обозначают $\varphi_d(f)$.

В нашем примере

$$\varphi(f) = \operatorname{arc} \operatorname{tg} \left(- \frac{2\zeta f/f_n}{1 - (f/f_n)^2} \right).$$

На рис. 1.3 приведены графики коэффициента усиления и фазы этого фильтра для $f_n = 100$ и $\zeta = 0.1$. Функция коэффициента усиления достигает максимума примерно при 100 Гц. На самом деле максимальное значение достигается в точке f_p (ее часто называют *пиковой частотой*), определяемой по формуле $f_p = f_n \sqrt{1 - 2\zeta^2}$. Для этого примера $f_p = 100 \sqrt{1 - 0.02} \approx 98.99$ Гц. Если величина ζ близка к 0, что часто выполняется в физических системах, то f_p и f_n почти одинаковы.

Определение передаточной функции можно переписать и в другом виде:

$$|Y(f)| = |H(f)| |X(f)|.$$

Такая форма записи помогает выявить тот факт, что абсолютная величина преобразования Фурье $y(t)$ есть просто произведение абсолютной величины преобразования Фурье функции $x(t)$ и

модуля функции $H(f)$. Обратное преобразование Фурье равенства $Y(f) = H(f) X(f)$ дает *свертку* функции $x(t)$ и новой функции

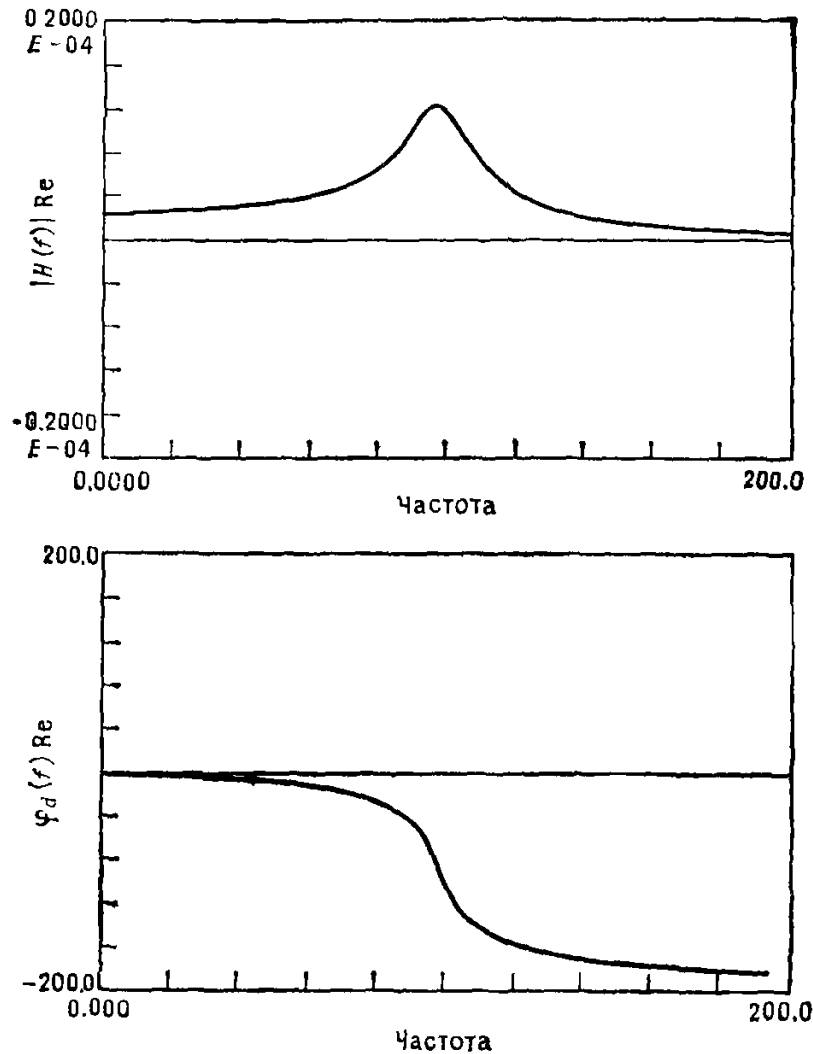


Рис. 1.3. Коэффициент усиления и фаза фильтра, определенного формулой (4.1) при $f_n = 100$ и $\zeta = 0.1$. Нижний индекс d у фазы означает градусы.

$h(t)$, представляющей собой обратное преобразование Фурье $H(f)$:

$$y(t) = \int_{-\infty}^{\infty} h(\tau) x(t - \tau) d\tau.$$

В нашем примере функция $h(t)$ находится из формулы

$$h(t) = \int_{-\infty}^{\infty} e^{j2\pi ft} \left[\omega_n^2 \left(1 - \left(\frac{f}{f_n} \right)^2 + j2\zeta \frac{f}{f_n} \right) \right]^{-1} df =$$

$$= \begin{cases} 0, & t < 0, \\ e^{-2\pi f_n \zeta t} \frac{\sin(2\pi f_n (1 - \zeta^2)^{1/2} t)}{2\pi f_n (1 - \zeta^2)^{1/2}}, & t \geq 0. \end{cases}$$

Эту функцию, представляющую собой обратное преобразование Фурье функции $H(f)$, часто называют *импульсной функцией отклика*. Если в качестве функции $x(t)$ взять дельта-функцию, то функция $y(t)$ будет равна $h(t)$.

Таким образом, с помощью интеграла свертки получаем

$$y(t) = (2\pi f_n \sqrt{1-\xi^2})^{-1} \int_0^t e^{-2\pi f_n \xi \tau} \sin 2\pi f_n \sqrt{1-\xi^2} \tau \cdot x(t-\tau) d\tau.$$

Конечные пределы интегрирования появились здесь в силу того, что как $h(t)$, так и $x(t)$ для отрицательных значений равны нулю.

Здесь лучше всего остановиться и подвести некоторые итоги. Если $x(t)$ и $y(t)$ связаны линейным дифференциальным уравнением, то существуют такие функции $h(t)$ и $H(f)$, что 1) функция $H(f)$ есть преобразование Фурье функции $h(t)$, 2) в частотной области $Y(f) = H(f) X(f)$, 3) во временной области выполняется соотношение

$$y(t) = \int_{-\infty}^{\infty} h(\tau) x(t-\tau) d\tau,$$

4) если $x(t) = \delta(t)$, то $y(t) = h(t)$, 5) исходному дифференциальному уравнению может быть сопоставлено эквивалентное ему интегральное уравнение.

1.6. СВЕРТКА И ФИЗИЧЕСКАЯ ОСУЩЕСТВИМОСТЬ

В предыдущем параграфе отмечалось, что для $y(t)$ и $x(t)$, связанных линейным соотношением, представляющим собой линейное дифференциальное уравнение с постоянными коэффициентами, существует такая функция $h(t)$, что

$$y(t) = \int_{-\infty}^{\infty} h(\tau) x(t-\tau) d\tau.$$

Интеграл такого вида имеет специальное название — *свертка*.

Поясним на примере, что происходит при этой операции. Допустим, что требуется вычислить $y(t)$ при $t = t_0$, т. е.

$$y(t_0) = \int_{-\infty}^{\infty} h(\tau) x(t_0 - \tau) d\tau.$$

Кроме того, предположим, что в качестве функций $h(t)$ и $x(t)$ выбраны следующие:

$$x(t) = \begin{cases} t & \text{при } t \geq 0, \\ 0 & \text{при } t < 0, \end{cases} \quad h(t) = \begin{cases} 2 & \text{при } t \geq 0, \\ 0 & \text{при } t < 0. \end{cases}$$

Их графики изображены на рис. 1.4, *a*; на рис. 1.4, *b* те же функции изображены в области изменения переменной τ . Как и следовало ожидать, $h(\tau)$ и $h(t)$ совпадают. Что же касается функции $x(t)$, то она отражается вдоль своей оси и смещается на t_0 единиц (подразумевается, что в этом примере величина t_0 положительна). Отдельное значение $y(t_0)$ получается перемножением двух получившихся функций и последующим интегрированием этого произведения.

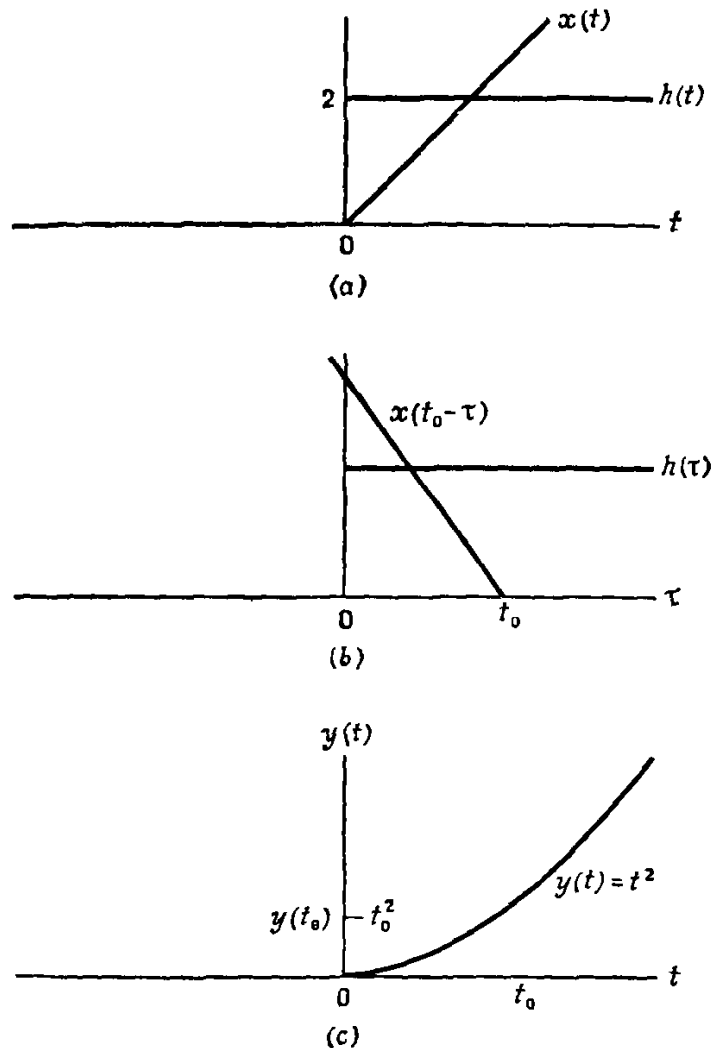


Рис. 1.4. Получение свертки для функций $h(t)$ и $x(t)$ в точке t_0 , рассматриваемых в примере. *a* — исходные функции; *b* — те же функции в области τ . Функция $x(t)$ обращается и сдвигается на t_0 единиц; *c* — окончательный результат — свертка функций $x(t)$ и $h(t)$ из примера.

Интегрирование проводится следующим образом:

$$y(t_0) = \int_0^{t_0} 2(t_0 - \tau) d\tau = 2t_0\tau - \tau^2 \Big|_0^{t_0} = 2t_0^2 - t_0^2 = t_0^2.$$

Следовательно, $y(t) = t^2$ для каждого t . Полученная для рассмотренных функций кривая изображена на рис. 1.4,с.

Важную роль играет теорема о коммутативности свертки, которая утверждает, что в определении свертки функции $x(t)$ и $h(t)$ можно поменять местами, т. е.

$$y(t) = \int_{-\infty}^{\infty} h(\tau) x(t - \tau) d\tau = \int_{-\infty}^{\infty} h(t - \tau) x(\tau) d\tau.$$

Для произведения функций во временной области справедливо двойственное правило, заключающееся в том, что произведение $y(t)$ двух функций $h(t)$ и $x(t)$ при преобразовании Фурье переходит в свертку функций $H(f)$ и $X(f)$ в частотной области, т. е. если $y(t) = h(t) x(t)$, то

$$Y(f) = \int_{-\infty}^{\infty} H(\eta) X(f - \eta) d\eta = \int_{-\infty}^{\infty} H(f - \eta) X(\eta) d\eta.$$

Как мы увидим, это соотношение имеет важное значение в изучении того, что подразумевается под конечным тестированием. Под *тестом* мы понимаем проведение эксперимента или испытания (например, структурного вибрационного испытания), при котором для последующего анализа некоторым методом записываются данные временного ряда.

Возвращаясь к дифференциальным уравнениям, приводившим к импульсным функциям отклика, заметим, что существуют условия, которые следует налагать на функцию $h(t)$, чтобы она соответствовала многим физическим ситуациям, а именно чтобы она стала *физически осуществимой*. Это достигается требованием невырожденности функции $h(t)$ лишь для времен, больших или равных 0, т. е.

$$h(t) = \begin{cases} 0 & \text{при } t < 0, \\ \text{любая-функция} & \text{при } t \geq 0. \end{cases}$$

Это условие эквивалентно требованию, чтобы функция $y(t)$ не давала отклика прежде, чем поступит $x(t)$.

1.7. ЭФФЕКТЫ КОНЕЧНОЙ ДЛИНЫ ВЫБОРКИ

Большинство рассмотренных до сих пор функций определялось на бесконечном временном интервале. Этим функциям соответствовали бы эксперименты бесконечной продолжительности. Но такие эксперименты проводить невозможно, и возникает естественный вопрос об эффектах, обусловленных конечной длиной записи.

Вероятно, наиболее полно ответил на него в своей работе Деннис Габор (1946), больше известный, пожалуй, своими исследованиями по голографии.

Конечная длина записи будет моделироваться нами при помощи прямоугольного импульса $u(t)$. Предположим, что функция $x(t)$ определена на интервале $-\infty < t < \infty$ и имеет преобразование Фурье $X(f)$. Если время наблюдения ограничено промежутком $-P/2 \leq t < P/2$, то сужение функции $x(t)$ на этот интервал, обозначаемое $x_p(t)$, можно записать в виде

$$x_p(t) = x(t) u(t).$$

Таким образом, после умножения $x(t)$ на прямоугольный импульс получилась функция $x_p(t)$, которая равна 0 вне интервала $-P/2 \leq t < P/2$. Это значит, что функция x_p отвечает выборке длины P .

Преобразование Фурье функции $u(t)$ известно, а для функции $x(t)$ оно задано по предположению. Поэтому можно определить функциональный вид $x_p(t)$. Поскольку функция $x_p(t)$ — произведение двух функций во временной области, ее преобразование Фурье, согласно третьей формуле в табл. 1.1, должно быть сверткой $X(f)$ и $U(f)$ в области частот. Обозначив символом $X_p(f)$ преобразование Фурье функции $x_p(t)$, получим

$$\begin{aligned} F[x_p(t)] &= X_p(f) = \int_{-\infty}^{\infty} X(\eta) U(f-\eta) d\eta = \\ &= \int_{-\infty}^{\infty} X(\eta) \frac{\sin(\pi(f-\eta)P)}{\pi(f-\eta)} d\eta = \int_{-\infty}^{\infty} X(f-\eta) \frac{\sin(\pi\eta P)}{\pi\eta} d\eta. \end{aligned}$$

Поэтому эффект перехода к конечной длине записи сводится, вообще говоря, к свертке преобразования исходной функции с бесконечной длиной записи и функции $(\sin x)/x$. Как станет видно из дальнейшего, это приводит к «размыванию».

Рассмотрим важный пример с косинусом, имеющим единичную амплитуду и частоту f_0 . Как было показано, преобразование Фурье этой функции на бесконечном интервале имеет вид

$$X(f) = 1/2 [\delta(f-f_0) + \delta(f+f_0)].$$

Поэтому для преобразования Фурье сужения косинуса на конечный интервал получим

$$\begin{aligned} \int_{-\infty}^{\infty} 1/2 [\delta(\eta-f_0) + \delta(\eta+f_0)] \frac{\sin(\pi(f-\eta)P)}{\pi(f-\eta)} d\eta = \\ = \frac{\sin(\pi(f-f_0)P)}{2\pi(f-f_0)} + \frac{\sin(\pi(f+f_0)P)}{2\pi(f+f_0)}. \end{aligned}$$

Эта функция изображена на рис. 1.5. На части *a* этого рисунка показаны исходные дельта-функции, а на части *b* — те же

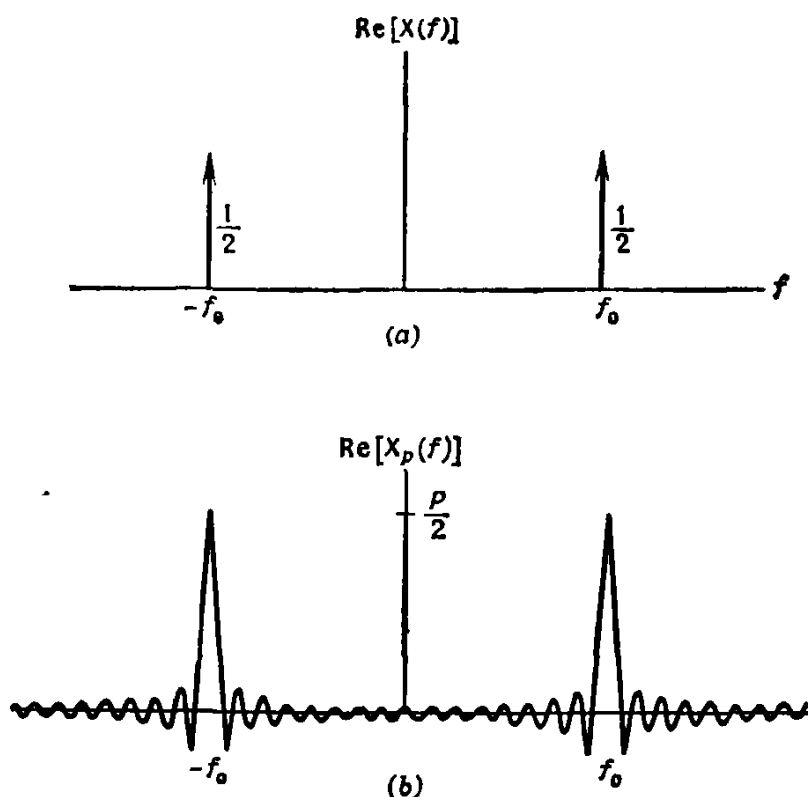


Рис. 1.5. Преобразование Фурье функции $\cos(2\pi f_0 t)$. *a* — без размывания, область неограниченна; *b* — с размыванием, область определения косинуса ограничена отрезком $-P/2, P/2$.

дельта-функции, но уже свернутые с функцией $(\sin x)/x$, появление которой обусловлено переходом к интервалу конечной длины. Следует обратить внимание на следующее:

1. Центры функций $(\sin x)/x$ располагаются в частотах $\pm f_0$, поэтому в этих точках располагаются максимумы¹⁾. Следовательно, большая часть информации концентрируется вблизи частот, $\pm f_0$ Гц.

2. С другой стороны, функция $X_P(f)$ нигде не обращается в нуль, за исключением самых крайних точек. Это и приводит к эффекту, называемому размыванием. Информация размывается по всей области частот, а не концентрируется точно в частотах $\pm f_0$ Гц.

3. Высота центрального выступа функции $(\sin x)/x$ равна $P/2$, а ширина $2/P$. Значит, если запись сделать длиннее, т. е. увеличить P , то высота главных выступов увеличится, а ширина —

¹⁾ Это не совсем точно, поскольку хвост каждой из функций $(\sin x)/x$ добавляется к центральному всплеску другой и, значит, вносит искажения. Для больших значений P это несущественно.

уменьшится. Если P устремить в бесконечность, то функции $(\sin x)/x$ перейдут в дельта-функции.

О синусоиде на бесконечном интервале $(-\infty, \infty)$ можно говорить как о чистом тоне, вся информация о котором сконцентрирована в двух дельта-функциях. Переход к конечному интервалу приводит из-за свертки с $(\sin x)/x$ к размыванию и потере чистоты тона. Таким образом, из-за использования во всяком эксперименте моделей с конечной длиной записи имеется только размытая частотная информация.

С конечной длиной записи связан и другой эффект. Он заключается в невозможности точного измерения частоты. Поскольку информация размыта, точно определить максимум главного всплеска практически невозможно и, значит, точно оценить частоту по конечной выборке синусоиды нельзя¹⁾. Эта неопределенность связана не только с длиной записи, но и с частотой синусоиды. Более высокие частоты определять легче, так как взаимное влияние двух функций $(\sin x)/x$ при их отдалении уменьшается.

1.8. ВЫБОРОЧНЫЕ ДАННЫЕ И ЧАСТОТА НАЙКВИСТА

Выборочными (квантованными) данными назовем данные, выбираемые через равные промежутки T , где T измеряется в секундах. Выборочная функция $x(i)$, как и ее непрерывный аналог, определяется формулой

$$x(i) = \int_{-\infty}^{\infty} x(\tau) \delta(iT - \tau) d\tau.$$

Таким образом, выборочная функция представляет собой последовательность дискретных величин, соответствующих моментам времени iT .

Преобразование Фурье функции $x(i)$, как и прежде, обозначим $X(f)$, однако его определение для выборочных данных станет другим:

$$X(f) = T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi fiT).$$

Это определение можно считать аппроксимацией предыдущего определения, хотя, пожалуй, лучше считать их определениями двух форм одного и того же преобразования (есть еще две формы, так что всего их четыре, см. Отнес и Эноксон (1972, с. 12)).

¹⁾ Если нет каких-то других сигналов и отсутствует шум, то для определения частоты f достаточно трех равноотстоящих точек. Точное значение частоты f равно $(2\pi T)^{-1} \arcsin \cos \{[x(t) + x(t - 2T)] / (2x(t - T))\}$. При этом предполагается, что $x(t - T) \neq 0$ и выборочные значения определены точно.

Одна из главных проблем, связанных с выборками, — явление *подмены частот*. Если с интервалом T производится выборка значений синусоиды с частотой, большей $1/(2T)$, то такую выборку можно перепутать с выборкой синусоиды меньшей частоты. Значит, наряду с истинной частотой в выборке появляется и другая (меньшая) — частота подмены. В качестве иллюстрации рассмотрим пример с синусоидой

$$x(t) = \sin(2\pi f_0 t), \quad f_0 = \frac{p+q}{2T},$$

где p — целое число, а q — дробная часть, меньшая единицы. Допустим, например, что $T = 0.005$ и $f_0 = 1025$ Гц. Тогда

$$(2T)^{-1} = 100, \quad p = 10, \quad q = 0.25.$$

Подставив выбранное значение f_0 и заменив t на iT , для $x(t)$ получим

$$\begin{aligned} \sin(2\pi f_0 t) &= \sin\left(2\pi \left(\frac{p+q}{2T}\right) iT\right) = \sin(\pi(p+q)i) = \\ &= \sin(\pi p i) \cos(\pi q i) + \cos(\pi p i) \sin(\pi q i) = \cos(\pi p i) \sin(\pi q i), \end{aligned}$$

поскольку $\sin(\pi p i) = 0$ для всех i .

Теперь в зависимости от четности p имеется две возможности.

1. Если p — четное, то $\cos(\pi p i) = 1$ для всех i , поэтому

$$x(i) = \sin(\pi q i).$$

2. Если p — нечетное, то

$$x(i) = (-1)^i \sin(\pi q i).$$

В первом случае оставшийся угол можно переписать в виде

$$\pi q i = \pi q i \frac{2T}{2T} = 2\pi \left(\frac{q}{2T}\right) iT.$$

В результате получается частота, равная $q/(2T)$. Обозначим ее f'_0 .

В нашем примере $p = 10$, и, следовательно, p — четное. Значит, полученная частота f'_0 составляет $(0.25)(100) = 25$ Гц. Таким образом, выборка синусоиды частоты 1025 Гц с интервалом $T = 0.005$ с, или, что то же, со скоростью 200 выб/с (выборки в секунду), не различается с синусоидой частоты 25 Гц!

Частоту $(2T)^{-1}$, играющую для цифровых данных особую роль, с этого момента будем обозначать специальным символом F (от англ. *folding*), т. е.

$$F = (2T)^{-1} = S/2,$$

где S — скорость выборки, измеряемая числом выборок в секунду. Ее часто называют *частотой Найквиста* в честь Найквиста, написавшего главную работу по этому вопросу (1924).

Возвращаясь ко второму случаю, в котором p — нечетное, получаем

$$\begin{aligned} x(i) &= (-1)^i \sin(\pi qi) = \cos(\pi i) \sin(\pi qi) = \\ &= \cos(\pi i) \sin(\pi qi) - \sin(\pi i) \cos(\pi qi) = \\ &= -\sin(\pi i - \pi qi) = -\sin(\pi i(1 - q)). \end{aligned}$$

Если умножить числитель и знаменатель на $2T$, то для оставшегося угла получим выражение

$$\pi i(1 - q) = \pi i(1 - q) \frac{2T}{2T} = 2\pi \left(\frac{1}{2T} - \frac{q}{2T} \right) iT = 2\pi (F - f_0) iT,$$

что опять указывает на возможность подмены частот.

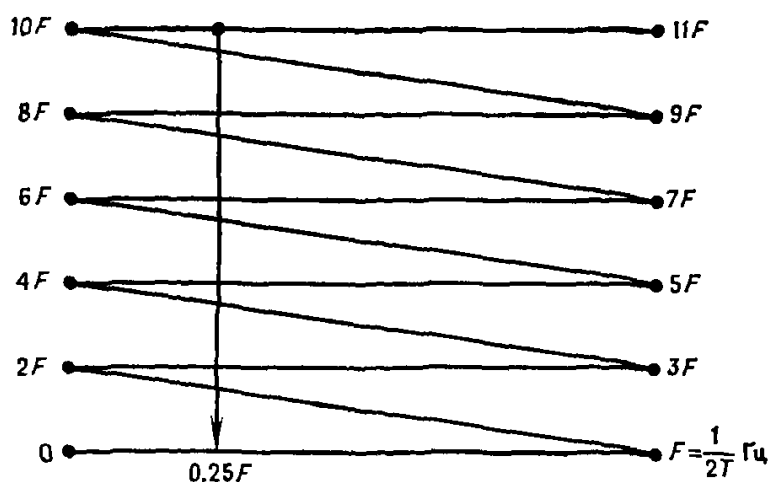


Рис. 1.6. Первые 11 складок диаграммы подмен Найквиста.

На рис. 1.6 эти результаты показаны на диаграмме подмен. Ось частот сложена гармошкой над интервалом $0 - F$ Гц кусками по F Гц. Чтобы показать, как работать с диаграммой, вернемся к нашему примеру. Частота 1025 Гц отвечает при скорости выборок 200 выб/с величине $10.25 F$. Если из этой точки опустить перпендикуляр, то получится $0.25 F$. То же значение получится, если взять $1.75 F$, $2.25 F$, $3.75 F$, $4.25 F$, $5.75 F$, $6.25 F$, $7.75 F$, $8.25 F$, $9.75 F$ и т. д.

Такую диаграмму очень часто полезно рисовать, чтобы по ней узнавать, где может происходить подмена частот для данного сигнала. Предположим, что в нашем примере появляется синусоида частоты 80 Гц, чего нельзя было ожидать, исходя из известной физической ситуации. В этом примере $F = 100$ Гц, и

сигнал в 80 Гц мог бы оказаться сигналом в 120 Гц¹⁾, получившим из-за подмены частот меньшее значение.

Теорема о выборке. Основной вопрос, который возникает при выборке временных функций, заключается в том, насколько хорошо выборочные данные представляют исходную функцию. Как утверждает *теорема о выборке*, такое представление будет очень хорошим, если выполняются определенные условия. Такая теорема приводится Шенноном (1963).

Эти условия, которые принимаются как гипотезы, состоят в следующем.

1. Функция $x(t)$ определена при $-\infty < t < \infty$.
2. Существует преобразование Фурье $X(f)$ функции $x(t)$.
3. Функция $X(f)$ имеет ограниченный носитель, т. е. $|X(f)| = 0$ при $|f| > F$ Гц.

Если эти условия выполняются, то справедлив следующий вывод.

Теорема²⁾. Пусть T' — такой интервал квантования, что $T' \leq \leq (2F)^{-1}$. Тогда функция $x(t)$ может быть восстановлена по последовательности $x(i)$ единственным образом, за исключением, быть может, изолированного множества точек.

Идея доказательства, которое мы не приводим, основана на следующем факте. Поскольку существует $X(f)$, множества точек $x(i)$, определенных соотношением

$$x(i) = \int_{-F}^F X(f) \exp(j2\pi fiT) df,$$

оказывается достаточно, чтобы построить по нему $X(f)$:

$$X(f) = T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi fiT).$$

Так как $X(f)$ можно восстановить по значениям $x(i)$, то функцию $x(t)$ можно восстановить по последовательности $x(i)$. Выражение для $x(i)$ можно рассматривать как свертку во временной области, соответствующую произведению $X(f)$ и прямоуголь-

¹⁾ Во многих приложениях переменный ток 110 В, 60 Гц преобразуется в постоянный ток посредством использования выпрямительных устройств. Обычно эти устройства используют полное волновое выпрямление, которое изменяет переменный ток в постоянный добавлением гармоник 120 Гц и выше, причем гармоника 120 Гц имеет наибольшую амплитуду. Если их должным образом не удалять, то они могут появиться в данных.

²⁾ Этот результат получен В. А. Котельниковым. — Прим. перев.

ного импульса в области частот:

$$\begin{aligned}
 x(t) &= \int_{-F}^F X(f) \exp(j2\pi ft) df = \\
 &= \int_{-F}^F T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi fiT) \exp(j2\pi ft) df = \\
 &= T \sum_{i=-\infty}^{\infty} x(i) \int_{-F}^F \exp[j2\pi f(t-iT)] df = \\
 &= T \sum_{i=-\infty}^{\infty} x(i) \frac{2 \sin[2\pi F(t-iT)]}{2\pi(t-iT)}.
 \end{aligned}$$

Заметим, что при $t = kT$

$$\frac{2 \sin[2\pi F(kT-iT)]}{2\pi(kT-iT)} = \begin{cases} 2F, & i = k, \\ 0, & i \neq k, \end{cases}$$

поэтому

$$x(kT) = Tx(k) 2F = x(k).$$

Таким образом, выражение, полученное для $x(t)$, полностью совпадает с выражением для $x(i)$ при $t = iT$. В промежуточных точках это выражение служит интерполяционной схемой, которая восстанавливает функцию $x(t)$ по величинам $x(i)$. Как и следовало ожидать, с наибольшими весами берутся те величины $x(i)$, индексы i которых соответствуют наиболее близким к t значениям.

1.9. ДИСКРЕТНОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ ДЛИНЫ N

Две формы преобразования Фурье уже рассматривались:

$$\begin{aligned}
 X(f) &= \int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft) dt, \\
 X(f) &= T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi fiT).
 \end{aligned}$$

Введем теперь новую форму преобразования Фурье, в которой как X , так и x дискретны, а пределы суммирования конечны. Оно определяется для $x(i)$ ($i = 0, 1, \dots, N-1$) следующим образом:

$$X(k) = T \sum_{i=0}^{N-1} x(i) \exp\left(-\frac{j2\pi ik}{N}\right).$$

Такая форма преобразования играет важнейшую роль в книге. Она не только позволяет вычислять преобразование Фурье и обратное к нему на цифровых ЭВМ. Как будет показано в гл. 6, при некоторых значениях N вычисления по этой формуле могут выполняться гораздо быстрее, чем это следует из ее общего вида.

Указанную форму преобразования можно отнести к преобразованиям второго типа (для дискретных $x(i)$ на бесконечном интервале). Для этого нужно только 1) положить $x(i) = 0$ для $i < 0$, $i > N-1$ и 2) определить величину $f_k = k/(NT)$. Если это сделано, то

$$\begin{aligned} X(f_k) &= T \sum_{i=0}^{N-1} x(i) \exp(-j2\pi f_k i T) = \\ &= T \sum_{i=0}^{N-1} x(i) \exp\left[-j2\pi \left(\frac{k}{NT}\right) i T\right] = \\ &= T \sum_{i=0}^{N-1} x(i) \exp\left(-\frac{j2\pi i k}{N}\right). \end{aligned}$$

Подобным образом определяется обратное преобразование Фурье.

$$x(i) = b \sum_{k=0}^{N-1} X(k) \exp\left(\frac{j2\pi i k}{N}\right),$$

где $b = (NT)^{-1}$ и $f_k = kb = k/(NT)$ (здесь $k = 0, \dots, N-1$). Следует обратить внимание на интерпретацию этих частот. Частота Найквиста равна $(2T)^{-1}$, поэтому соответствующее значение k определяется из равенства

$$\frac{k}{NT} = \frac{1}{2T},$$

или $k = N/2$.

Докажем обратимость введенных преобразований:

$$\begin{aligned} b \sum_{k=0}^{N-1} X(k) \exp\left(\frac{j2\pi i k}{N}\right) &= \\ &= b \sum_{k=0}^{N-1} \left[T \sum_{p=0}^{N-1} x(p) \exp\left(-\frac{j2\pi p k}{N}\right) \right] \exp\left(\frac{j2\pi i k}{N}\right) = \\ &= bT \sum_{p=0}^{N-1} x(p) \sum_{k=0}^{N-1} \exp\left(-j \frac{2\pi k(p-i)}{N}\right). \end{aligned}$$

Здесь $bT = N^{-1}$, а для внутренней суммы выполняется соотношение

$$\sum_{k=0}^{N-1} \exp\left(-j \frac{2\pi k(p-i)}{N}\right) = \begin{cases} N, & i = p, \\ 0, & i \neq p, \end{cases}$$

поскольку для $i = p$

$$\sum_{k=0}^{N-1} 1 = N,$$

а для $i \neq p$

$$\begin{aligned} \sum_{k=0}^{N-1} \exp\left(-j \frac{2\pi k (p-i)}{N}\right) &= \\ &= [1 - \exp(-j2\pi (p-i))] \left[1 - \exp\left(-\frac{j2\pi (p-i)}{N}\right)\right]^{-1} = 0. \end{aligned}$$

Таким образом, частота Найквиста соответствует середине последовательности $X(k)$.

При детальном рассмотрении этой формы преобразования Фурье обнаруживается ряд весьма интересных свойств. Первое из них — периодичность функций $x(i)$ и $X(k)$. Она понимается в том смысле, что для определения значений функций $x(i)$ и $X(k)$ вне множества $0, \dots, N-1$ достаточно брать значения их индексов по модулю N .

Например, допустим, что $i = pN + q$, где p и q — целые числа, причем $0 \leq q \leq N-1$. Подставим в выражение обратного преобразования Фурье

$$x(i) = b \sum_{k=0}^{N-1} X(k) \exp\left(\frac{j2\pi ik}{N}\right)$$

новое значение i :

$$\begin{aligned} x(pN + q) &= b \sum_{k=0}^{N-1} X(k) \exp\left(\frac{j2\pi k (pN + q)}{N}\right) = \\ &= b \sum_{k=0}^{N-1} X(k) \exp\left(\frac{j2\pi kq}{N}\right) \exp\left(\frac{j2\pi kpN}{N}\right). \end{aligned}$$

Крайний правый член в последней сумме равен единице, так как

$$\exp\left(\frac{j2\pi kpN}{N}\right) = \exp(j2\pi kp) = 1.$$

Следовательно, в этом случае

$$x(i) = x(q).$$

Аналогичное соотношение можно доказать и для прямого преобразования Фурье $x(i)$.

Рис. 1.7 иллюстрирует такую периодичность для $N = 16$.

Поскольку значениям k в промежутке $0, \dots, N/2$ соответствуют частоты, не превосходящие частоту Найквиста, закономерен вопрос о смысле, который имеют величины $X(k)$ при $k > N/2$.

Оказывается, этим величинам соответствуют отрицательные частоты. Покажем, как это получается. Предположим, что индекс k в определении величины $X(k)$ заменен на $-p$, т. е.

$$X(-p) = T \sum_{i=0}^{N-1} x(i) \exp\left(-\frac{j2\pi i(-p)}{N}\right).$$

После умножения экспоненциального члена на величину $\exp(-j2\pi i N/N)$, равную единице, получим

$$\begin{aligned} T \sum_{i=0}^{N-1} x(i) \exp\left(-\frac{j2\pi i(-p)}{N}\right) \exp\left(-\frac{j2\pi i N}{N}\right) &= \\ &= T \sum_{i=0}^{N-1} x(i) \exp\left(-\frac{j2\pi i(N-p)}{N}\right) = X(N-p). \end{aligned}$$

Таким образом, $X(-1) = X(N-1)$, $X(-2) = X(N-2)$ и т. д. Итак, второй половине преобразования отвечает преобразование для отрицательных значений k .

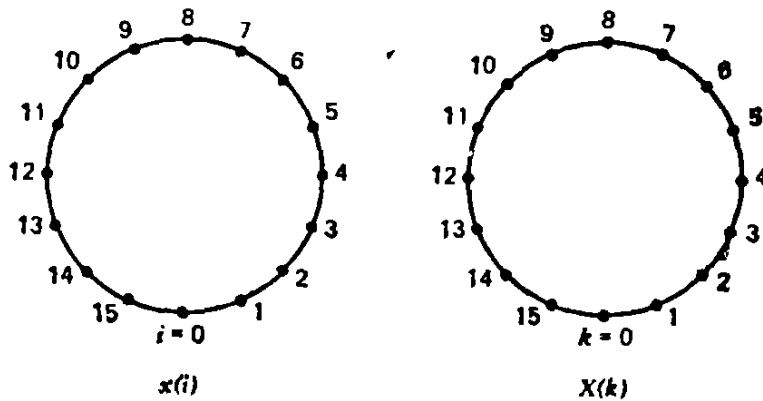


Рис. 1.7. Периодичность рядов $x(i)$ и $X(k)$ при $N = 16$. Частоте Найквиста соответствует $k = 8$.

Если $x(i)$ — действительные числа, что бывает довольно часто, то $X(-k) = X^*(k)$, как и следовало ожидать.

Второе важное свойство связано со сверткой. Предположим, что заданы величины $y(i)$ и для них преобразование Фурье определяется по формуле

$$Y(k) = T \sum_{i=0}^{N-1} y(i) \exp\left(-\frac{j2\pi ik}{N}\right).$$

Если частотная функция $Z(k)$ определяется соотношением

$$Z(k) = X(k) Y(k),$$

то вполне естественно ожидать, что ее обратное преобразование Фурье $z(i)$ будет сверткой $x(i)$ и $y(i)$.

В этом случае так и оказывается. Сложность заключается в том, что результирующая свертка тоже периодична. Приведем для нее формулы

$$z(i) = T \sum_{p=0}^{N-1} x(p) y(i-p) = T \sum_{p=0}^{N-1} x(i-p) y(p).$$

Индексы $(i-p)$ должны пониматься по модулю N , как показано на рис. 1.8. Оба множества значений $x(i)$ и $y(i)$ изображены как точки на окружностях. Значения $y(i)$ располагаются

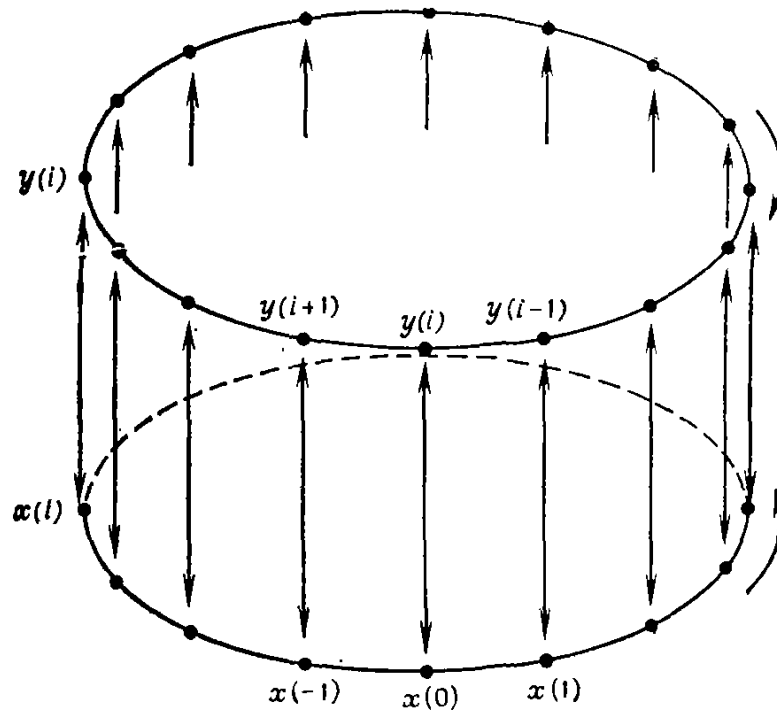


Рис. 1.8. Круговая свертка. Члены последовательности $y(i)$ располагаются в обратном по отношению к $x(i)$ порядке, причем напротив $y(i)$ располагается $x(0)$. Одно значение $z(i)$ получается суммированием всех попарных произведений противостоящих значений с последующим умножением на T .

на окружности в обратном порядке. Для того чтобы вычислить $z(i)$, все величины $y(i)$ так сдвигаются против часовой стрелки, чтобы напротив $x(0)$ расположилась величина $y(i)$. Все стоящие друг против друга величины перемножаются. Затем все произведения складываются и сумма умножается на T . Получается одно значение $z(i)$.

В качестве примера рассмотрим свертку функций

$$x(i) = \begin{cases} T^{-1}, & i = 7, \\ 0, & i \neq 7, \end{cases} \quad y(i) = 1 - \frac{i}{15},$$

где $i = 0, \dots, 15, N = 16$.

Они изображены на рис. 1.9 вместе со сверткой $z(i)$. Как видно из рисунка, значение $z(i)$ оказывается равным

$$T \sum_{p=0}^{N-1} x(p) y(i-p) = y((i-7) \pmod{16}),$$

поэтому

$$z(0) = y(9), \quad z(1) = y(10), \quad \dots, \quad z(7) = y(0), \quad \dots, \quad z(15) = y(8).$$

Таким образом, значения $z(i)$ равны значениям $y(i)$, сдвинутым на семь точек.

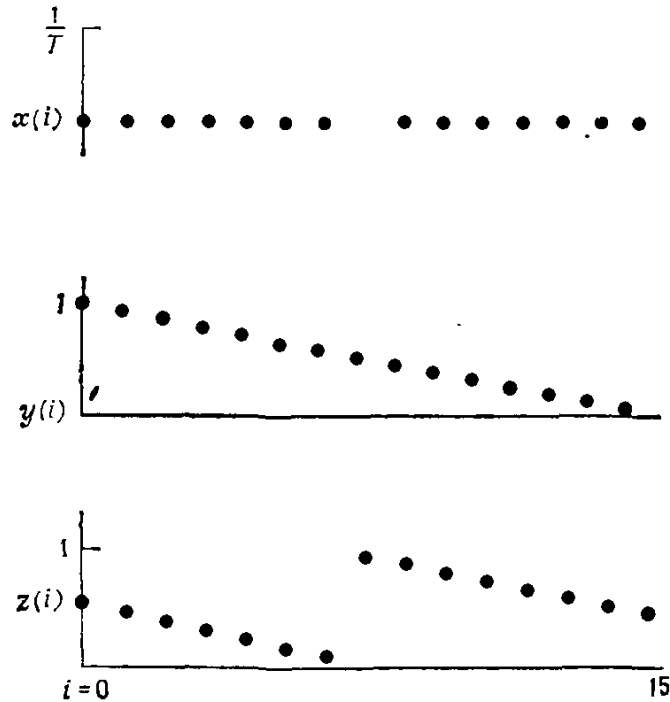


Рис. 1.9. Свертка функций $x(i)$ и $y(i)$, определенных в тексте. Здесь $N=16$.

С этой формой преобразования связана еще одна особенность. Использование свертки, полученной только что рассмотренным способом, при попытке построения моделей причинных систем может вызывать явление наложения, обусловленного периодичностью, и приводить к сильным искажениям данных. Такая опасность преодолевается добавлением нулей. Этот прием будет рассмотрен в гл. 6.

Вычисление дискретного преобразования Фурье. Обычная форма дискретного преобразования Фурье получается довольно просто. На рис. 1.10 приводится небольшая программа для вычисления преобразования 16 действительных величин.

Строки 4—14 отвечают за получение функции данных. Здесь задаются параметры f_0 , φ и вычисляются значения синусоиды

$$x(i) = \cos\left(\frac{2\pi i f_0}{N} + \varphi\right), \quad i = 0, \dots, 15.$$

```

C      МЕДЛЕННОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ
      DIMENSION XC(16),XS(16),X(16)
      FACT=2.*3.14159265
5      READ (5,10) ICASE,FO,PHI
10     FORMAT (I6,2F15.6)
      IF (ICASE.LE.0) STOP
      WRITE (6,15) ICASE,FO,PHI
25     FORMAT (23H1SLOW FOURIER TRANSFORM,/,
117HOTEST CASE NUMBER,I6,/,
26HOFO  =,F15.6,/,
36HOPHI =,F15.6,/,
436HO   K          REAL      IMAGINARY,/)
      DO 20 I=1,16
20     X(I)=COS(FACT*FLOAT(I-1)*FO+PHI)
      DO 30 KK=1,16
      K=KK-1
      SUMC=0.
      SUMS=0.
      DO 25 II=1,16
      I=II-1
      ANG=FACT*FLOAT(I*K)/16.
      SUMC=SUMC+X(II)*COS(ANG)
25     SUMS=SUMS-X(II)*SIN(ANG)
      XC(KK)=SUMC
      XS(KK)=SUMS
30     WRITE(6,10) K,SUMC,SUMS
      GO TO 5
      END

```

Рис. 1.10. FORTRAN-программа для получения синусоиды и вычисления ее преобразования Фурье.

```

SLOW FOURIER TRANSFORM
TEST CASE NUMBER      1
FO =                   0.250000
PHI =                  0.000000

```

K	REAL	IMAGINARY.
0	0.000002	0.000000
1	-0.000001	0.000001
2	-0.000001	-0.000001
3	-0.000000	-0.000001
4	8.000000	0.000001
5	-0.000001	-0.000001
6	0.000004	-0.000001
7	-0.000001	0.000001
8	-0.000001	0.000000
9	0.000001	0.000001
10	0.000001	0.000001
11	0.000000	0.000003
12	8.000000	-0.000001
13	0.000001	0.000001
14	0.000002	0.000001
15	-0.000001	-0.000001

Рис. 1.11. Распечатка результатов прогона теста при $FO=0.25$, $PHI=0$.

Параметр N фиксирован и равен 16. Это сделано в учебных целях и предохраняет от расхода неприемлемо большого машинного времени. Интервал T выбран равным единице, поэтому $F = 1/2$, а частоты f_k равны $k/16$ при $k = 0, \dots, 8$ и $1/2 - k/16$ при $k = 9, \dots, 15$.

В двойном цикле (строки 15—26) происходит вычисление преобразования Фурье и печать результатов. Запоминание результатов в ХС и ХS необязательно, однако оно упростит выполнение некоторых упражнений.

Эффективность этой программы возрастет, если синусы и косинусы будут вычисляться рекурсивно. Этот прием рассмотрен в гл. 4.

На рис. 1.11 приводятся результаты прогона тестовой программы со значениями $f_0 = 1/4$ и $\varphi = 0$. Обратите внимание, что, как и следовало ожидать, действительные части для $K = 4$ и $K = 12$ одинаковы.

1.10. z-ПРЕОБРАЗОВАНИЕ

• Определение z-преобразования может быть дано несколькими способами. По-видимому, проще всего сначала записать еще раз преобразование Фурье второго типа для функции $x(i)$:

$$X(f) = T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi f iT)$$

и заменить затем $\exp(j2\pi f T)$ на z :

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

Это преобразование можно использовать при решении многочисленных задач обработки цифровых сигналов.

Сразу отметим, что запись этого преобразования проще — вместо десяти символов для $\exp(j2\pi f T)$ нужен всего лишь один для z . Но дело не только в удобстве записи. Это преобразование стало появляться в усовершенствованных системах для решения задач, связанных с выборочными данными. Читателям, которые хотят глубже изучить этот вопрос, рекомендуем книгу Стейглица (1974)¹⁾.

Задачи анализа временных рядов, вообще говоря, могут решаться с помощью z-преобразования, и во многих случаях решения упрощаются. Однако в приложениях, затронутых в этой книге, обращение к z-преобразованию мало что меняет. Поэтому,

¹⁾ Изучение свойств соответствующих комплексных функций можно найти в книге Шабата (1976). — Прим. ред.

кроме редких случаев обращения к более удобной записи, это преобразование нигде появляться не будет.

Упражнения

1.1. Докажите, что

$$\left| \frac{a+jb}{g+jh} \right| = \frac{|a+jb|}{|g+jh|}.$$

1.2. Найдите преобразование Фурье следующих функций $x(t)$:

$$a) x(t) = \begin{cases} e^{-at} & \text{при } t \geq 0, \\ 0 & \text{при } t < 0. \end{cases}$$

Ответ: $(a + j2\pi f)^{-1}$.

$$b) x(t) = e^{-a|t|}.$$

Ответ: $2a/(a^2 + (2\pi f)^2)$.

с) (Треугольный импульс.)

$$x(t) = \begin{cases} \left(\frac{a}{P}\right)t + a & \text{при } -P < t < 0, \\ \left(-\frac{a}{P}\right)t + a & \text{при } 0 < t < P, \\ 0 & \text{при остальных значениях } t. \end{cases}$$

Ответ: $aP (\sin^2(\pi f P))/(\pi f P)^2$.

d) (Колоколообразный косинус.)

$$x(t) = \begin{cases} \frac{1}{2}(1 + \cos(\pi t/P)) & \text{при } -P < t < P, \\ 0 & \text{при остальных значениях } t. \end{cases}$$

Ответ:

$$\frac{\sin(2\pi f P)}{2\pi f} + \frac{P^2 2\pi f \sin(2\pi f P)}{\pi^2 - (2\pi f P)^2}.$$

1.3. Функции с разрывами часто имеют в преобразовании Фурье дельта-функции, как, например, функция 4 в табл. 1.2. Используя это преобразование и соотношение 4 в табл. 1.1, укажите другой способ вычисления преобразования Фурье прямоугольного импульса. Что происходит с дельта-функциями?

1.4. Дано дифференциальное уравнение

$$\frac{dy(t)}{dt} + ay(t) = x(t),$$

в котором $x(t) = 0$ при $t < 0$.

- Вычислите преобразование Фурье этого уравнения.
- Найдите передаточную функцию между двумя функциями.
- Определите коэффициент усиления передаточной функции.
- Определите фазу передаточной функции.
- Начертите графики коэффициента усиления и фазы как функций частот.
- Вычислите импульсную функцию отклика в предположении, что она существует.

1.5. Вычислите свертки функций $x(t)$ и $h(t)$. Во всех случаях $x(t)$, $h(t)$ и ответы равны нулю при t , меньших 0.

a) $x(t) = t$, $h(t) = t$ $[t^3/6]$.

b) $x(t) = e^t$, $h(t) = e^t$ $[te^t]$.

c) $x(t) = t + 1$, $h(t) = t - 1$ $[t^3/6 - t]$.

d) $x(t) = e^t$, $h(t) = \sin t$ $[1/2 (e^t - \sin t - \cos t)]$.

1.6. Вычислите интеграл

$$\int_{-\infty}^{\infty} \frac{\sin(\pi\eta P)}{\pi\eta} \frac{\sin(\pi(f-\eta)Q)}{\pi(f-\eta)} d\eta.$$

Ответ: $(\sin(\pi f P))/(\pi f)$ при $P < Q$; $(\sin(\pi f Q))/(\pi f)$ при $P > Q$.

Задачи для решения на ЭВМ.

1.7. Отперфорлируйте программу с рис. 1.10 и прогоните ее при следующих значениях:

Номер случая	f_0	φ
2	0.25	-1.5707963
3	0.5	0.0000000
4	0.3125	0.7853982

Объясните полученные результаты.

1.8. Напишите дополнительную программу, которая вычисляла бы и обратное преобразование Фурье после вычисления прямого преобразования. Прогоните все четыре случая.

1.9. Измените программу так, чтобы

$$x(i) = \begin{cases} 1 & \text{для } i = 0, 1, 2, 14, 15, \\ 0 & \text{для остальных } i, \end{cases}$$

и вычислите преобразование Фурье. Прокомментируйте результаты.

Глава 2

ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ И СТАТИСТИКИ

2.1. ВВЕДЕНИЕ

С приложениями анализа временных рядов тесно связаны некоторые статистические понятия. Цель этой главы состоит в том, чтобы определить такие крайне необходимые на практике ключевые параметры, как среднее, дисперсию и вероятностную плотность. Нередко приходится иметь дело с данными, испорченными нежелательным шумом, приводящим к определенным ошибкам при вычислениях различных параметров данных временного ряда. Однако мы ограничимся лишь теми понятиями, которые непосредственно относятся к анализу временных рядов. Вместе с тем эта книга — не учебник по математической статистике. Мы, как правило, только определяем термины и не доказываем теорем. Читателю, которому нужно более полное знакомство с основами статистики, советуем обратиться к книгам Большева и Смирнова (1965), Диксона и Мэсси (1969), содержащим полные и всеобъемлющие таблицы; теоретическим вопросам математической статистики посвящены книги Крамера (1975) и Фрэйзера (1958). В этих книгах даются более строгие математические определения статистических параметров. Например, корректное определение случайного процесса потребовало бы сложных математических построений, связанных с характером исследуемых величин. Но в наших приложениях, к счастью, те патологические функции, которые рассматриваются математиками, не встречаются. Поэтому можно сделать упрощающие допущения, которые не приводят к потере общности результатов.

1. Рассматриваемые функции ограничены и до преобразования в цифровую форму были непрерывными.

2. Случайный процесс, определяющий функцию, является эргодическим и стационарным.

Строгое математическое определение случайной величины требует в первую очередь ее измеримости в том смысле, в котором это понимается в теории меры. Это гарантируется допущением 1, поскольку ограниченные непрерывные функции образуют подмножество множества измеримых функций. Оставшуюся часть введения мы отведем обсуждению на интуитивном уровне допущения 2.

Вероятностная функция плотности (ВФП), среднее и дисперсия. Представим себе, что большое число одинаковых генераторов шума было включено в некоторый момент в прошлом и с тех пор безостановочно работает. С выходом всех этих генераторов связывают вероятностную функцию плотности $f(x, t)$, имеющую следующие характеристики. Вероятность того, что в определенный момент, скажем t_0 , выход q -го генератора сигналов $x_q(t_0)$ лежит в интервале между значениями a и b , определяется интегралом

$$P\{a \leq x_q(t_0) < b\} = \int_a^b f(x, t_0) dx.$$

Заметим, что интегрирование ведется по промежутку значений случайной величины. Математическое ожидание любой функции от x , обозначаемое $E[g(x)]$ (где $g(x)$ — та функция, ожидание которой ищется), определяется следующим образом:

$$E\{g[x(t_0)]\} = \int_{-\infty}^{\infty} g[x(t_0)] f(x, t_0) dx.$$

В частности, *истинные* или *множественные* среднее и дисперсия задаются формулами

$$\mu(t_0) = \int_{-\infty}^{\infty} x(t_0) f(x, t_0) dx \quad (2.1)$$

и

$$\sigma^2(t_0) = \int_{-\infty}^{\infty} [x(t_0) - \mu(t_0)]^2 f(x, t_0) dx.$$

Стандартным отклонением называют положительное значение квадратного корня из дисперсии.

Если случайный процесс является *стационарным*, то параметры $\mu(t_0)$ и $\sigma^2(t_0)$ не зависят от времени, т.е. для произвольных t_0 и t_1

$$\mu(t_0) = \mu(t_1) = \mu, \quad \sigma^2(t_0) = \sigma^2(t_1) = \sigma^2.$$

В дальнейшем параметр t_0 в среднем и в дисперсии нами указываться не будет, поскольку принимается допущение о стационарности.

Допущение об *эргодичности* позволяет заменить усреднение по ансамблю усреднением по времени. В примере с генераторами шума все генераторы были совершенно одинаковыми, поэтому знания лишь одной случайной функции отдельного генератора было бы достаточно, чтобы по выходу одного из них определить статистические параметры для всех. Так, выражение для сред-

него (2.1) можно заменить выражением

$$\mu_x = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P x(t) dt,$$

в котором временное усреднение ведется по одной траектории процесса. Подобное выражение для дисперсии имеет вид

$$\sigma_x^2 = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P [x(t) - \mu]^2 dt = E(x^2) - E^2(x). \quad (2.2)$$

Среднеквадратичное значение определяется формулой

$$\psi_x^2 = \mu_x^2 + \sigma_x^2.$$

Как сказано выше, свойство эргодичности позволяет заменять усреднения по множеству реализаций случайного процесса усреднениями по времени.

Центральным в этой книге является понятие *плотности спектра мощности* (ПСМ), которое формально будет определено ниже. А пока будем подразумевать под ПСМ функцию, представляющую собой разбиение или разложение общей среднеквадратичной мощности как функции частоты.

Заметим, что будет «правильно» связывать со случайным процессом функцию плотности спектра мощности. Дело в том, что статистические свойства и характеристики гауссовского процесса с нулевым средним полностью определяются (и притом наиболее кратко) спектром мощности этого процесса или, что эквивалентно, его автоковариационной функцией¹⁾. Это делает спектр мощности исключительно важным параметром, поскольку большинство процессов, встречающихся на практике, являются гауссовскими.

¹⁾ Здесь мы несколько выходим за рамки основного предмета этой книги. Статистические характеристики процесса известны, если известны все многомерные вероятностные функции плотности, определяющие совместные вероятности для процесса на всех временных промежутках. Например, автоковариационная функция для промежутка времени τ определяет корреляционные характеристики между амплитудами процесса в моменты t и $t + \tau$. Для гауссовских данных полностью определяют совместную (двумерную) вероятностную плотность величин амплитуд в моменты t и $t + \tau$ ковариация и две дисперсии. В общем случае это нужно распространить и на более высокие порядки.

При сборе данных в практических экспериментах или тестах можно вычислять только оценки спектра мощности (и другие параметры) и, пользуясь эргодичностью, давать оценки по временным средним, полученным по отдельным интервалам временной последовательности. Поэтому оценка ПСМ определяется по конечному временному промежутку. При замене отрезка конечным набором точек мы не получим, используя выборочные данные и дискретные формулы, существенного (со статистической точки зрения) расхождения непрерывной и дискретной версий. Такая замена производится, пока не нарушаются требования теоремы о выборках, позволяющей избегать существенной подмены частот.

2.2. ВЫБОРОЧНЫЕ ВЕЛИЧИНЫ И ОЦЕНКИ

Для выборочного среднего и выборочной дисперсии — оценок групповых параметров — выбирают отличающиеся от принятых в теории обозначения для среднего и дисперсии, поскольку длина записи реализаций случайной величины не может быть бесконечной и, следовательно, можно располагать только конечной выборкой. В частности, выборочное среднее для цифровых данных вычисляется по формуле

$$m = \bar{x} = \frac{1}{N} \sum_{i=0}^{N-1} x(i).$$

Несмещенную выборочную дисперсию s^2 получают так:

$$s^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} [x(i) - \bar{x}]^2.$$

Определение доверительных интервалов для выборочного среднего и выборочной дисперсии проводится с помощью критериев, точные формулы и таблицы для которых можно найти в книге Бендата и Пирсола (1971). Как правило, длинам записей, с которыми сталкиваются в анализе временных рядов, отвечает малая статистическая изменчивость этих двух параметров, и здесь трудностей не возникает.

Если делитель $N-1$ в выражении для s^2 заменить на N , то среднее значение s^2 станет равным $[(N-1)/N]\sigma^2$, т. е. оценка станет смещенной. Впрочем, для больших значений N эта ошибка мала. Обычно в анализе временных рядов значение N равно по крайней мере 100 и часто достигает величин порядка 10 000 и более. Поэтому вопрос о смещении, вызванном делением на N , как правило, несуществен.

Сделаем несколько замечаний о принятой в статистике системе обозначений. Групповые параметры по традиции обозначают греческими буквами, например μ и σ^2 для среднего и дисперсии. Оценки параметров обычно обозначаются соответствующими латинскими буквами, например m и s^2 . Несколько нарушая этот порядок, средние случайных величин часто помечают черточкой сверху, например, среднее величины x обозначают \bar{x} . В анализе временных рядов черточку обычно оставляют для временных средних, а обозначение математического ожидания $E(x)$ используют для более общих ситуаций. Наконец, представляется довольно разумным использование крышечки над параметрами как стандартного обозначения для оценки. Это, в частности, полезно при истощении запаса букв. Так, оценку спектра мощности обозначаем через $\hat{S}_x(k)$. Иногда мы будем вынуждены применять все перечисленные обо-

значения сразу и полагаем, что причина этого станет ясна читателю из контекста.

Теперь остановимся более подробно на некоторых важных свойствах статистических оценок. Мы рассмотрим смещение, дисперсию и состоятельность.

Смещение. Оценка $\hat{\varphi}$ параметра φ смещена, если среднее значение $\hat{\varphi}$ не совпадает с φ ;

$$\text{Смещение} = \Delta_{\hat{\varphi}} = E(\hat{\varphi}) - \varphi.$$

Выборочная дисперсия с делителем N служит примером смещенной оценки.

Дисперсия. Дисперсией оценки является величина

$$\sigma_{\hat{\varphi}}^2 = E[\hat{\varphi} - E(\hat{\varphi})]^2,$$

поэтому среднеквадратичную ошибку определяют следующим образом:

$$\Psi_{\hat{\varphi}}^2 = \sigma_{\hat{\varphi}}^2 + \Delta_{\hat{\varphi}}^2 = E[\hat{\varphi} - \varphi]^2.$$

Состоятельность. Оценка состоятельна, если ее дисперсия убывает с ростом объема выборки (длиной записи). Запишем это условие в виде равенства

$$\lim_{N \rightarrow \infty} \sigma_{\hat{\varphi}}^2 = 0.$$

Примером несостоятельной оценки может служить несглаженный выборочный спектр мощности. Если в качестве оценки спектра мощности вычисляется квадрат абсолютной величины преобразования Фурье (часто называемый периодограммой), то получается несостоятельная оценка. В этом случае дисперсия остается постоянной при возрастании длины записи.

Помимо того, примерами смещенных оценок служат обычные оценки спектральных функций. Возникающее в этом случае смещение интерпретируется как ошибка, размывающая полосу частот вследствие недостаточной спектральной разрешающей способности. Это явление подробно обсуждается в гл. 8.

Рассмотрим важный пример вычисления дисперсии для синусоиды. Предположив, что сигнал x определен формулой

$$x(t) = A \sin(2\pi f_c t + \varphi), \quad -\infty < t < \infty,$$

получим следующее выражение для дисперсии:

$$\begin{aligned}\sigma_x^2 &= \lim_{P \rightarrow \infty} \frac{1}{P} \int_{-P/2}^{P/2} A^2 \sin^2 (2\pi f_c t + \varphi) dt = \\ &= A^2 \lim_{P \rightarrow \infty} \frac{1}{P} \int_{-P/2}^{P/2} \sin^2 (2\pi f_c t + \varphi) dt = \\ &= A^2 \lim_{P \rightarrow \infty} \frac{1}{P} \left[\frac{P}{2} + \{\sin\text{-член} \leq 1\} \right] = \frac{A^2}{2}.\end{aligned}\quad (2.3)$$

Заметим, что для таких функций среднеквадратичная ошибка Ψ_x^2 и дисперсия σ_x^2 совпадают, поскольку среднее этих функций равно 0.

2.3. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Предположим, что случайная величина x имеет среднее μ и дисперсию σ^2 . Тогда говорят, что она распределена *нормально* или имеет *гауссовское* распределение, если ее вероятностная функция плотности $\varphi(x)$ определяется выражением

$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right]. \quad (2.4)$$

Вероятностная функция распределения $\Phi(x)$ такой гауссовской величины дается формулой

$$\Phi(x) = \int_{-\infty}^x \varphi(\xi) d\xi,$$

т. е.

$$\begin{aligned}\Phi(x_0) &= \mathbf{P} \{ -\infty < x \leq x_0 \}, & 1 - \Phi(x_0) &= \mathbf{P} \{ x_0 < x < \infty \}, \\ \Phi(x_1) - \Phi(x_0) &= \mathbf{P} \{ x_0 < x \leq x_1 \}\end{aligned}$$

(предполагается, что $x_0 < x_1$).

Для гауссовских величин справедлива важная теорема, которую мы приведем без доказательства.

Теорема. *Случайная величина, полученная линейным преобразованием из гауссовской величины, является гауссовской.*

В частности, величина

$$z = ax + b \quad (2.5)$$

— гауссовская, если x имеет гауссовское распределение со средним μ_x и дисперсией σ_x^2 , причем

$$\mu_z = a\mu_x + b \quad (2.6)$$

и

$$\sigma_x^2 = a^2 \sigma_x^2. \quad (2.7)$$

Применение преобразования (2.5) специального вида, задаваемого формулой

$$z = \frac{1}{\sigma_x} (x - \mu_x), \quad (2.8)$$

называют «стандартизацией». Из формул (2.6) и (2.7) следует, что $\mu_z = 0$, $\sigma_z^2 = 1$. С помощью этого преобразования можно составлять таблицы для тех гауссовских величин, которые фигурируют в приложении. Независимо от значений среднего и дисперсии исходных данных можно определять вероятности по таблице для величины с параметрами 0 и 1, а затем, обращая соотношение (2.8):

$$x = z\sigma_x + \mu_x,$$

переходить к значениям исходной величины. Приведем несколько часто используемых значений для функции распределения Φ при $\mu = 0$ и $\sigma = 1$:

$$\frac{1}{\sqrt{2\pi}} \int_{-1}^1 \exp\left(-\frac{x^2}{2}\right) dx = 0.682,$$

$$\frac{1}{\sqrt{2\pi}} \int_{-2}^2 \exp\left(-\frac{x^2}{2}\right) dx = 0.954,$$

$$\frac{1}{\sqrt{2\pi}} \int_{-3}^3 \exp\left(-\frac{x^2}{2}\right) dx = 0.997.$$

Мы полагаем, что читателю известно, как пользоваться таблицами распределений, и не останавливаемся на этом.

Нормальное распределение — понятие теоретическое; на практике обычно не встречается данных с действительно нормальным распределением. В основном это связано с областью значений случайной величины. Большинство данных берется из ограниченного интервала, в то время как нормально распределенные данные должны принимать значения на бесконечной прямой.

Предположим, что группа людей измеряет стержень длины 2 см, используя одну и ту же линейку, и что полученные измерения имеют нормальное распределение. Допустим также, что истинное среднее измерений равно 2 см, а дисперсия равна 0.25 см. Вероятность того, что результат измерения окажется меньше -0.5 см, равна

$$P\{x \leq -0.5\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{-0.5} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx.$$

После замены переменной $z = (x - \mu)/\sigma$ получим

$$P\{x \leq -0.5\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-(0.5+2)/0.25} \exp\left(-\frac{z^2}{2}\right) dz \approx 0.0000001.$$

Таким образом, есть конечная (хотя и малая) вероятность того, что результат измерения длины стержня окажется отрицательным. Разумеется, это абсурд. Тем не менее использование нормального распределения вполне оправданно, поскольку оно является приемлемым приближением к тем истинным распределениям, которые могут встречаться.

Предположение о нормальности (т. е. о том, что данные имеют гауссовское распределение) принимается по ряду причин. Главные из них следующие.

1. Нормальное распределение достаточно хорошо изучено. Это упрощает моделирование исследуемой ситуации и получение выводов о результатах статистических измерений.

2. Как показывает центральная предельная теорема (Крамер, 1975), распределение суммы случайных величин, вообще, с любым распределением стремится к нормальному, если суммируется «достаточное их число». Оказывается, что результаты фильтрации (см. гл. 3) представляют собой суммирование большого числа наблюдений и, следовательно, их распределение «стремится» к нормальному.

Перечислим несколько важных свойств нормального распределения:

1. Как плотность, так и распределение гауссовской случайной величины x полностью определяются средним μ_x и дисперсией σ_x^2 .

2. Моменты нормальной величины вычисляются по формулам

$$E[x^n] = \begin{cases} \mu, & n = 1, \\ \mu^2 + \sigma^2, & n = 2, \\ \mu(\mu^2 + 3\sigma^2), & n = 3, \\ \dots & \dots \end{cases}$$

3. Центральные моменты вычисляются по формулам

$$E[(x - \mu)^n] = \begin{cases} 0, & n = 1, 3, \dots, \\ 1 \cdot 3 \cdot \dots \cdot (n-1) \sigma^n, & n = 2, 4, \dots \end{cases}$$

Третий момент относительно среднего называют «асимметрией», а четвертый — «эксцессом». Асимметрия равна 0 для всякого симметричного распределения. Эксцесс, превосходящий гауссовский (равный $3\sigma^4$), указывает на значительное число данных с большими амплитудами. Распределения с таким эксцессом имеют более «толстые» хвосты функции плотности, чем нормальное. В § 2.10,

где приводятся примеры и применения вероятностных функций плотности, мы дадим различные примеры оценок вероятностной плотности.

Обратимся к последнему свойству случайных величин, связанному с ВФП, — *независимости*.

Две случайные величины x и y статистически независимы, если их совместная (двумерная) вероятностная функция плотности $p(x, y)$ распадается в произведение одномерных ВФП, т. е.

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

Другими словами, это свойство показывает, что со статистической точки зрения величины x и y не оказывают влияния друг на друга. Об этом пойдет речь в следующем параграфе, где будут рассмотрены корреляция и регрессия. А сейчас перейдем ко второму важному распределению.

Распределение хи-квадрат. Предположим, что x_1, x_2, \dots, x_n — независимые гауссовские величины с нулевым средним и единичной дисперсией. Определим величину χ_n^2 формулой

$$\chi_n^2 = \sum_{i=1}^n x_i^2.$$

Об этой величине говорят, что она имеет распределение хи-квадрат с n степенями свободы. Функция плотности χ_n^2 дается формулой

$$p(\chi_n^2) = \left[2^{n/2} \Gamma\left(\frac{n}{2}\right) \right]^{-1} (\chi_n^2)^{(n/2)-1} \exp[-\chi_n^2/2], \quad \chi_n^2 \geq 0,$$

где $\Gamma(n/2)$ есть гамма-функция.

Распределение хи-квадрат используется в анализе временных рядов главным образом при изучении поведения выборочной дисперсии и плотностей спектра мощности. Если ряд $x(i)$ имеет нулевое среднее и при вычислении выборочной дисперсии s^2 используется N независимых данных $x(i)$, то границы B_1 и B_2 , в которых с вероятностью $(1-\alpha)$ лежит истинное значение дисперсии, т. е.

$$P\{B_1 \leq \sigma^2 \leq B_2\} = 1 - \alpha,$$

определяются по формулам

$$B_1 = \frac{ns^2}{\chi_{n; 1-(\alpha/2)}^2}; \quad B_2 = \frac{ns^2}{\chi_{n; (\alpha/2)}^2}, \quad n = N - 1.$$

Отметим, что B_1 и B_2 суть функции от s^2 , α и χ_n^2 . Интервал (B_1, B_2) называют доверительным интервалом; о нем говорят также как о $[(1-\alpha) 100]\%$ -м доверительном интервале¹⁾.

¹⁾ Как принято в статистике, мы более свободно пользуемся термином «доверительный»; строго говоря, это уже не вероятность. Такое расширение смысла этого понятия можно найти у Фрэйзера (1958).

Рассмотрим пример (взятый из книги Блэкмена и Тьюки (1958)), в котором по N независимым наблюдениям нормально распределенной случайной величины ($N = 31$) требуется определить 90%-й доверительный интервал для выборочной дисперсии. Предположим, что выборочное среднее \bar{x} равно 58.61, а выборочная дисперсия s^2 равна 33.43. В этом случае для теоретического значения дисперсии σ^2 выполнено следующее равенство:

$$P \left\{ \frac{30s^2}{43.77} < \sigma^2 < \frac{30s^2}{18.49} \right\} = 0.90,$$

где $\chi_{30; 0.05}^2 = 18.49$ есть такое значение переменной χ_{30}^2 , что $P \{ \chi_{30}^2 < 18.49 \} = 0.05$, а $\chi_{30; 0.95}^2 = 43.77$ (соответственно $P \{ \chi_{30}^2 < 43.77 \} = 0.95$). После подстановки значения получаем следующие границы: $22.91 \leq \sigma^2 \leq 54.22$. Заметим, что значение истинной дисперсии попадает за пределы этого интервала примерно в десятой части случаев.

Важную роль в статистике играют и другие распределения вероятностей. Среди них можно, в частности, назвать 1) t -распределение Стьюдента, полезное в статистических выводах и при вычислении доверительных интервалов для истинных средних значений; 2) F -распределение, которое используется при проверке на равенство пар средних квадратов. Основные факты, касающиеся этих распределений, читатель может найти в книге Бендата и Пирсола (1971).

2.4. КОРРЕЛЯЦИЯ И РЕГРЕССИЯ

Статистические понятия корреляции и регрессии близко примыкают к понятиям когерентности и частотной функции отклика, широко используемым в анализе временных рядов (а также в технике). По этой причине мы остановимся на них подробнее.

Нам необходимо сначала определить понятие ковариации, которая представляет собой математическое ожидание произведения двух центрированных случайных величин, т. е.

$$\text{cov}(x, y) = \sigma_{xy} = E[(x - \mu_x)(y - \mu_y)].$$

Для ковариации двух независимых величин (т. е. таких, что $p(x, y) = p(x)p(y)$) выполняется важное равенство $\sigma_{xy} = 0$. Заметим, что обратное неверно. Однако в очень важном частном случае, когда две случайные величины имеют совместное (двумерное) гауссовское распределение, последнее представляется произведением двух одномерных гауссовских распределений. Поэтому для гауссовских данных справедлив следующий важный вывод: *из равенства нулю ковариации гауссовских величин вытекает их независимость*, т. е.

$$\sigma_{yx} = 0 \Rightarrow \Phi(x, y) = \Phi(x)\Phi(y).$$

Этот факт играет главную роль, вероятно, в 90% задач практического анализа временных рядов.

Корреляция — это нормированная ковариация, а именно

$$\rho_{xy} = \frac{\text{cov}(x, y)}{[\text{var}(x) \text{var}(y)]^{1/2}} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}.$$

Нетрудно показать, что коэффициент корреляции принимает значения в интервале, ограниченном ± 1 , т. е. $-1 \leq \rho_{xy} \leq 1$. Этот результат остается справедливым и для выборочной корреляции, если выборочная ковариация определяется формулой

$$s_{xy} = \frac{1}{N-1} \sum_{i=0}^{N-1} [x(i) - \bar{x}][y(i) - \bar{y}].$$

По аналогии с представлением о плотности спектра мощности как о разложении дисперсии по частотам под функцией *плотности кросс-спектра* (ПКС) подразумевают разложение по частотам ковариации.

С понятиями ковариации и корреляции тесно связан регрессионный анализ. Так называемую линию регрессии обычно определяют как линейное соотношение

$$y = \beta_0 + \beta_{xy}x$$

между двумя величинами. Если при сборе данных оценки b_0 и b_{xy} для величин β_0 и β_{xy} ищут с помощью критерия наименьших квадратов, то

$$b_{xy} = s_{xy}/s_x^2, \quad b_0 = \bar{y} - b_{xy}\bar{x}.$$

Полученная прямая

$$(y - \bar{y}) = b_{xy}(x - \bar{x})$$

называется линией регрессии¹⁾ y по x . Можно говорить также и о регрессии x по y . В этом случае

$$b_{yx} = s_{xy}/s_y^2.$$

Заметим, что

$$b_{xy}b_{yx} = \frac{s_{xy}^2}{s_x^2 s_y^2} = r_{xy}^2.$$

Для совместного гауссовского распределения линия регрессии допускает геометрическую интерпретацию. В этом важном случае коэффициенты регрессии суть тангенсы углов наклона главных осей эллипса постоянной вероятности.

¹⁾ Термин регрессия возник в некоторых давних исследованиях соотношения роста родителей и их детей. Было установлено, что рост «регрессирует» к среднему, т. е. высокие отцы и матери имеют более низких сыновей и дочерей, а низкие отцы и матери — более высоких.

Множественная регрессия. Двумерная регрессия без труда переносится на случай p величин (без потери общности можно считать среднее равным нулю):

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p.$$

Если предположить, что имеется N ($N > p$) наблюдений для каждой величины, то, вводя векторные обозначения

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \dots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{Np} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix},$$

решение задачи об оценке \mathbf{B} параметров β_i можно записать в виде (Андерсон (1963))

$$\mathbf{B} = (\mathbf{X}^* \mathbf{X})^{-1} \mathbf{X}^* \mathbf{Y},$$

где звездочка означает транспонирование.

В анализе частотных функций отклика важное значение приобретают некоторые другие множественные параметры, но мы пока отложим их изучение до гл. 9, где они понадобятся.

2.5. ФУНКЦИЯ ПЛОТНОСТИ СПЕКТРА МОЩНОСТИ

Как можно видеть из предыдущего, статистики второго порядка — дисперсия для одной случайной величины и ковариация для двух величин — имеют особое значение для величин с гауссовским распределением. Помимо этого в анализе временных рядов очень важная роль отведена разложению временных последовательностей по частотам. При оценивании таких важных параметров, как частота резонанса и декремент затухания, эта информация имеет решающее значение. В качестве другого применения можно назвать оценивание колебаний в таких экономических временных рядах, как индекс безработицы. Примерами использования плотности спектра мощности служат также обнаружение периодических колебаний температур в океане и определение цикла ритмических сокращений желудка коровы под воздействием определенных лекарств.

Среднеквадратичное значение на конечном отрезке определяется формулой

$$\psi_x^2 = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x^2(t) dt.$$

Размерность ψ_x^2 пропорциональна среднему квадрату энергии в единицу времени, который по определению есть мощность.

Плотность спектра мощности (ПСМ) $S_x(f)$ функции x — обобщение этого понятия. Ее можно понимать как функцию, интеграл от которой

$$\Psi_x^2(f_1, f_2) = 2 \int_{f_1}^{f_2} S_x(f) df, \quad 0 \leq f_1 \leq f_2,$$

равен среднеквадратичному значению в диапазоне частот от f_1 до f_2 . Таким образом,

$$\Psi_x^2 = \int_{-\infty}^{\infty} S_x(f) df.$$

Соответствующим выражением для S_x служит следующее:

$$S_x(f) = \lim_{P \rightarrow \infty} \frac{1}{P} \left| \int_{-P/2}^{P/2} x(t) e^{-i\omega t} dt \right|^2. \quad (2.9)$$

При непосредственном вычислении оценок ПСМ с помощью формулы (2.9) могут возникнуть известные осложнения. Если не проводить сглаживания по частотам, то полученные оценки окажутся статистически несостоятельными. Помимо того, как отмечалось выше, нужно учитывать и математические тонкости. Наряду с тем, что в прошлом не было известно эффективных вычислительных процедур, эти две причины делают невозможным построение приемлемой вычислительной процедуры на основе формулы (2.9). Однако с помощью быстрого преобразования Фурье вычислительные трудности были преодолены, правильное сглаживание по области частот привело к состоятельным оценкам и, наконец, использование записей конечной длины устранило математические проблемы.

Теорема, которую часто называют теоремой Винера — Хинчина, дает эквивалентное выражение

$$S_x(f) = \int_{-\infty}^{\infty} s_x(\tau) e^{-i\omega\tau} d\tau = 2 \int_0^{\infty} s_x(\tau) \cos(2\pi f\tau) d\tau,$$

где функция $s_x(\tau)$ (автоковариационная функция x) определяется соотношением

$$s_x(\tau) = \lim_{P \rightarrow \infty} \frac{1}{P} \int_{-P/2}^{P/2} x(t) x(t + \tau) dt, \quad -\infty < \tau < \infty,$$

причем предполагается, что процесс $x(t)$ имеет нулевое среднее.

Доказательство эквивалентности методов, основанных на этих двух формулах, не входит в нашу задачу. Его можно найти в нескольких монографиях (Бендат и Пирсол (1971), Винер (1963).

Существует и третий метод получения S_x , но его описание мы вынуждены отложить, поскольку он основан на понятии фильтрации, которое мы введем позже. А пока вернемся ко второму методу и поговорим о функции s_x . Ее можно рассматривать, учитывая принятое раньше предположение об эргодичности, как обобщение данного нами выше определения ковариации:

$$s_x(\tau) = E[x(t)x(t+\tau)] = \int_{-\infty}^{\infty} x(t)x(t+\tau)p(x,t,\tau)dt,$$

где $p(x,t,\tau)$ — соответствующая вероятностная функция плотности.

Допустим, что $n(t)$ — некоррелированный случайный шум с нулевым средним. Тогда для автокорреляционной функции справедливо равенство $s_n(\tau) = N\delta(\tau)$, и, следовательно, $S_n(f) = N$ для всех f . Таким образом, ПСМ некоррелированного случайного шума есть константа. Процессы такого типа называют *белым шумом*, поскольку мощность постоянна на любом промежутке частот, что очень напоминает белый свет, который в видимой части оптического спектра остается более или менее постоянным. Хотя для объяснения понятия ПСМ белый шум служит полезным математическим инструментом, следует заметить, что он физически нереализуем, поскольку имеет бесконечную дисперсию. Устройства, которые, как считается, создают белый шум, на самом деле производят шум, ПСМ которого остается постоянной лишь до определенной частоты, а затем на больших частотах убывает.

Предположим, что x — синусоидальный сигнал с частотой f_c , амплитудой A и фазой φ , т. е.

$$x(t) = A \sin(2\pi f_c t + \varphi).$$

Автоковариационная функция в этом случае вычисляется следующим образом:

$$\begin{aligned} s_x(\tau) &= \lim_{P \rightarrow \infty} \frac{1}{P} \int_{-P/2}^{P/2} A^2 \sin(2\pi f_c t + \varphi) \sin(2\pi f_c (t + \tau) + \varphi) dt = \\ &= \lim_{P \rightarrow \infty} \frac{1}{P} \left(\frac{A^2}{2} \right) \int_{-P/2}^{P/2} [\cos(2\pi f_c \tau) - \cos(2\pi f_c \tau + 4\pi f_c t + 2\varphi)] dt = \\ &= \lim_{P \rightarrow \infty} \frac{A^2}{2P} \left[t \cos(2\pi f_c \tau) + \frac{1}{4\pi f_c} \sin(2\pi f_c \tau + 4\pi f_c t + 2\varphi) \right]_{-P/2}^{P/2} = \\ &= \lim_{P \rightarrow \infty} \left\{ \frac{A^2}{2} \cos(2\pi f_c \tau) + \left(\frac{A^2}{8\pi f_c} \right) \left(\frac{1}{P} \right) \times \right. \\ &\quad \left. \times [\sin(2\pi f_c \tau + 2\pi f_c P + 2\varphi) - \sin(2\pi f_c \tau - 2\pi f_c P + 2\varphi)] \right\} = \\ &= \frac{A^2}{2} \cos(2\pi f_c \tau). \end{aligned}$$

Здесь мы учли, что выражение в квадратных скобках всегда меньше 2, поэтому второй член в предпоследнем выражении стремится к нулю, поскольку величина P , стоящая в знаменателе, стремится к бесконечности.

Отметим, что в окончательном результате фазовый угол исчез. Так как фазовый угол был произвольным, такой же результат получился бы и при замене синуса косинусом. Таким образом, независимо от начальной фазы автокорреляция для синусоидального сигнала всегда есть косинус с нулевой фазой. Вообще, фазовая информация в процессе получения автокорреляции теряется.

2.6. КАК ВЫЧИСЛЯТЬ СРЕДНЕЕ И ДИСПЕРСИЮ

Среднее и дисперсия вычисляются в принципе несложно. Однако, как иногда это случается при вычислениях на ЭВМ, возникают неожиданные и нетривиальные проблемы. Например, для вычисления выборочного среднего обычно берут формулу

$$\bar{x}_k = \left(\sum_{i=0}^{k-1} x(i) \right) + x(k), \quad k = 0, 1, \dots, N-1, \quad \bar{x} = \frac{1}{N} \bar{x}_{N-1},$$

т. е. накапливают значения данных из последовательности. Если сумма k величин будет существенно больше $x(k)$, то ошибка округления может оказаться того же порядка, что и значение $x(k)$. В этом случае применение вычислений с плавающей точкой или оперирование блоками с плавающей точкой из чисел с фиксированной точкой (что часто реализуется на мини-ЭВМ) может стать проблематичным. Другой метод, дающий более точные результаты, состоит в вычислении частных сумм следующим образом. Предположим, что число данных равно степени 2, $N = 2^m$. Тогда вычисление происходит в m стадий:

$$\begin{aligned} \bar{x}_1(i) &= 1/2 [x(2i) + x(2i+1)], \quad i = 0, \dots, \frac{N}{2} - 1, \\ \bar{x}_2(i) &= 1/2 [\bar{x}_1(2i) + \bar{x}_1(2i+1)], \quad i = 0, \dots, \frac{N}{4} - 1, \dots \\ &\dots, \bar{x} = \bar{x}_m(0) = 1/2 [\bar{x}_{m-1}(0) + \bar{x}_{m-1}(1)], \end{aligned}$$

т. е. каждый раз мы последовательно складываем парами числа исходных данных. В качестве пояснения рассмотрим пример, в котором $N = 8$ и $m = 3$. Массив данных состоит из значений $x(0)$, $x(1)$, \dots , $x(7)$. Вычисления проводятся в три этапа.

1. Вычисление значений

$$\begin{aligned} \bar{x}_1(0) &= 1/2 [x(0) + x(1)], & \bar{x}_1(1) &= 1/2 [x(2) + x(3)], \\ \bar{x}_1(2) &= 1/2 [x(4) + x(5)], & \bar{x}_1(3) &= 1/2 [x(6) + x(7)]. \end{aligned}$$

2. Вычисление с использованием предыдущих значений

$$\bar{x}_2(0) = 1/2 [\bar{x}_1(0) + \bar{x}_1(1)], \quad \bar{x}_2(1) = 1/2 [\bar{x}_1(2) + \bar{x}_1(3)].$$

3. Определение выборочного среднего

$$\bar{x} = \bar{x}_2(0) = 1/2 [\bar{x}_2(0) + \bar{x}_2(1)].$$

Этот метод позволяет избежать ошибки округления, если данные (как функция индекса i) случайны. Когда ведутся вычисления с плавающей точкой или блоками с плавающей точкой из чисел с фиксированной точкой, можно после каждой итерации заново нормировать массив, и значения данных никогда не поглощаются шумом, вызванным округлением. Эти замечания можно отнести ко всякому суммированию последовательности величин.

Естественно, сделанные замечания относятся также к вычислению суммы квадратов и, вообще, к вычислениям перекрестных произведений. Для вычисления дисперсии применяются несколько иные методы. Наименее чувствительна к ошибке округления формула

$$s_x^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} [x(i) - \bar{x}]^2.$$

Вычитание \bar{x} из каждого значения данных устраняет «динамическую составляющую» результирующей величины, что делает менее вероятным существенное влияние ошибок округления на общее суммирование. С другой формулой,

$$s_x^2 = \frac{1}{N-1} \left[\left(\sum_{i=0}^{N-1} x^2(i) \right) - N\bar{x}^2 \right], \quad (2.10)$$

связана проблема, заключающаяся в том, что в ней как $\sum x^2(i)$, так и $N\bar{x}^2$ может оказаться большим числом, а s_x^2 — маленьким. Это делает весьма вероятным положение, при котором ошибка округления разности в правой части (2.10) будет того же порядка, что и значение s_x^2 .

Сущность проблемы состоит в том, что такую величину можно сдвинуть вправо (за младшие значащие разряды) регистра-накопителя. Такое явление известно под названием *антипереполнения*.

В следующих главах будут обсуждаться и многие другие проблемы, обусловленные использованием ЭВМ при расчетах. Почти во всех случаях эти проблемы будут связаны с антипереполнением и накоплением ошибки округления.

2.7. ВЕРОЯТНОСТНЫЕ ГИСТОГРАММЫ

Из данных можно получать также выборочные вероятностные функции плотности или гистограммы. В отличие от \bar{x} и s^2 они строятся неоднозначно, поскольку зависят от некоторых определяющих эти функции параметров. Вычисление гистограммы происходит следующим образом: интервал изменения величины x , скажем $a < x < b$, делят на k подынтервалов равной длины (их называют «интервалами группировки») так, чтобы полная область изменения x распалась на $k+2$ интервалов. После этого берут все данные и фиксируют число попаданий на каждый интервал. Изображая графически числа попаданий на каждый интервал, получают гистограмму.

Представим это более формально. Обозначим символом $\{N_j\}$ множество целых чисел, полученное при подсчете попаданий $\{x(i)\}$ в j -й интервал. Пусть $c = (b - a)/k$ и $d_j = a + jc$. Тогда $\{N_j\}$ определяются следующим образом:

j	N_j
0	[число таких x , что $x < a$],
...	...
j	[число таких x , что $d_{j-1} \leq x < d_j$],
...	...
k	[число таких x , что $d_{k-1} \leq x < b$],
$(k+1)$	[число таких x , что $x \geq b$].

Сказанное поясняет рис. 2.1. Довольно часто $\{N_j\}$ называют карманами. Один из методов указанной сортировки на цифровых

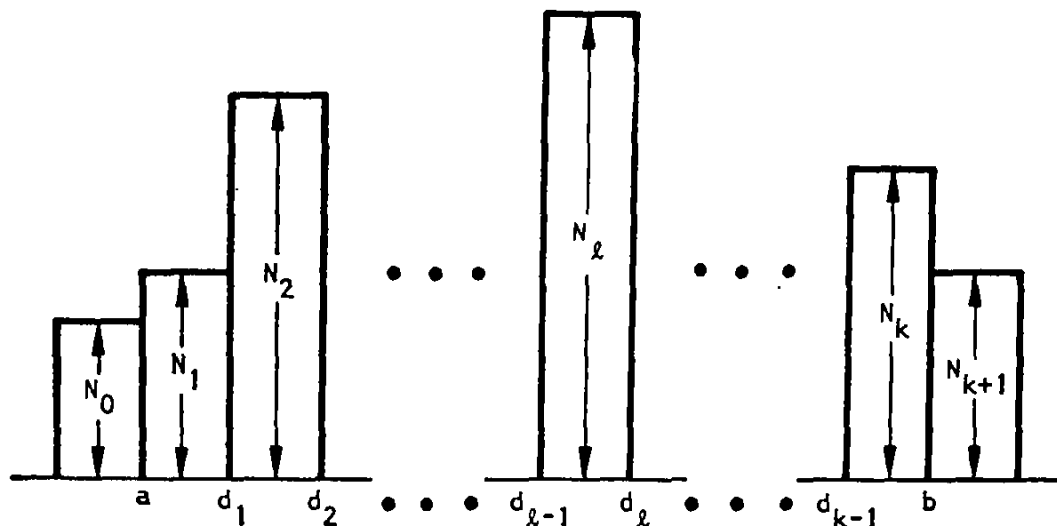


Рис. 2.1. Построение гистограммы.

ЭВМ состоит в переборе всех $x(i)$, $i = 1, \dots, N$, по очереди со следующими проверками:

- 1) если $x(i) < a$, то прибавить 1 к N_0 ;
- 2) если $x(i) \geq b$, то прибавить 1 к N_{k+1} ;
- 3) если предыдущие условия не выполняются и, следовательно, $a \leq x(i) < b$, то вычислить

$$j = \left[\frac{x(i) - a}{c} \right] + 1. \quad (2.11)$$

Здесь скобки $[]$ обозначают целую часть (т. е. в выражении $j = [a]$ число j представляет собой наибольшее целое, не превосходящее a). Получив по формуле (2.11) значение j , остается лишь прибавить 1 к N_j . Такой прием нетрудно осуществить на большинстве ЭВМ.

На полученных результатах можно построить последовательности трех видов. Первая из них — гистограмма, т. е. последовательность $\{N_j\}$, взятая без всяких изменений. Вторая последовательность $\{P_j\}$, где P_j — выборочная вероятность того, что $d_{j-1} \leq x < d_j$, определяется формулой

$$P_j = \frac{N_j}{N}, \quad j = 0, 1, \dots, k+1.$$

Третья последовательность называется выборочной вероятностной функцией плотности (ВФП), члены которой $\{p_j\}$, $j = 0, 1, \dots, k+1$, вычисляются по формуле

$$p_j = \frac{N_j k}{N(b-a)}.$$

Эти числа можно интерпретировать как значения производной функции распределения в серединах соответствующих интервалов.

Приведенной выше процедуре должен предшествовать выбор конкретных значений a , b и k . Естественно, возникает вопрос: каким должен быть критерий выбора этих трех параметров? Правильного однозначного ответа на него нет. Выбор определяется главным образом предположением о распределении, которому подчиняются данные, и способом сбора данных. Если данные получаются с помощью системы, подобной гипотетической системе сбора данных, из гл. 3, то на них накладываются два ограничения: данные имеют ограниченный промежуток значений и внутри этого промежутка располагаются лишь на конечном числе уровней. Для цифрового преобразователя, имеющего 128 уровней, очевидно, бессмысленно брать $k > 128$, поскольку некоторые из уровней окажутся пустыми. Нетрудно представить также ситуацию, в которой при распределении отсчетов (или их преобразованных эквивалентов) аналого-цифрового преобразования по интервалам выборочной функции может исказиться истинная картина.

Предположим, что цифровой преобразователь имеет 16 уровней и никакого преобразования в технические единицы не произво-

дится. Далее, предположим, что все уровни 0, ..., 15 совершенно равноправны, так что вероятность попадания на каждый из них a p_i o_i одинакова и равна $1/16$. Если теперь выбрать k равным 12, a и b равными 0 и 15 соответственно, то получится следующее распределение отсчетов по уровням:

Карман	Интервал	Содержимое уровней
1	0–1.25	0, 1
2	1.25–2.50	2
3	2.50–3.75	3
4	3.75–5.00	4
5	5.00–6.25	5, 6
6	6.25–7.50	7
7	7.50–8.75	8
8	8.75–10.00	9
9	10.00–11.25	10, 11
10	11.25–12.50	12
11	12.50–13.75	13
12	13.75–15.00	14

Ожидаемое число попаданий в карманы 1, 5 и 9 в два раза превосходит остальные, поэтому в гистограмме появятся значительные смещения в соответствующих интервалах, что приведет к неверному результату.

Рассмотренный сейчас пример показывает, с какой осторожностью нужно приступать к вычислениям функций плотностей, и делает очевидным необходимость тщательного изучения критериев, позволяющих избегать подобных ловушек. Один критерий для выбора параметров возникает при попытке дать ответ в другой, но связанной с рассмотренной задачей, а именно в задаче о том, как определять, подчиняются ли данные нормальному (гауссовскому) распределению.

Критерий согласия хи-квадрат для проверки на гауссовость. Согласно формуле (2.4), вероятностная функция плотности нормального распределения, обозначенная φ , определяется формулой

$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Вероятностная функция распределения есть интеграл от функции плотности:

$$\begin{aligned} P\{x \leq X\} &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(X-\mu)/\sigma} \exp\left(-\frac{t^2}{2}\right) dt = \Phi\left(\frac{X-\mu}{\sigma}\right). \end{aligned}$$

Вероятность того, что значение случайной величины лежит в интервале $[\alpha, \beta)$, дает формула

$$P\{\alpha \leq x < \beta\} = \Phi\left(\frac{\beta - \mu}{\sigma}\right) - \Phi\left(\frac{\alpha - \mu}{\sigma}\right).$$

Нормальное распределение принимается как гипотеза во многих исследованиях и возникает естественным образом во многих теоретических расчетах. Поэтому нередко желательно знать, действительно ли по отношению к собранным данным справедливо предположение о гауссовости. Одна из процедур, предназначенных для проверки такой гипотезы, известна под названием *критерия согласия хи-квадрат*. Его основу составляет использование в качестве меры расхождения наблюдаемой и теоретической функций плотности статистики хи-квадрат. При этом гипотеза об эквивалентности проверяется с помощью исследования выборочного распределения хи-квадрат. Частоты, которые следует ожидать для попаданий в j -й интервал группировки данных, имеющих гауссовское распределение, называют ожидаемыми частотами и обозначают F_j . Расхождение между наблюдаемыми и ожидаемыми частотами равно $(N_j - F_j)$. Чтобы получить полное расхождение, следует учесть каждый интервал, так как

$$\sum_{j=0}^{k+1} N_j = \sum_{j=0}^{k+1} F_j = N.$$

Сумма всех расхождений должна равняться 0. Заметим, что F_j , вообще говоря, не целые числа. Они вычисляются следующим образом:

$$\begin{aligned} F_0 &= N\Phi\left(\frac{a-m}{s}\right), \\ &\dots \\ F_j &= N\left\{\Phi\left(\frac{a+jc-m}{s}\right) - \Phi\left(\frac{a+c(j-1)-m}{s}\right)\right\}, \\ &\dots \\ F_{k+1} &= N\left(1 - \Phi\left(\frac{b-m}{s}\right)\right). \end{aligned}$$

Образуем величину

$$X^2 = \sum_{j=0}^{k+1} \frac{(N_j - F_j)^2}{F_j}. \quad (2.12)$$

Если выполняются определенные условия, то эта величина имеет распределение χ^2 и может быть сопоставлена с теоретическим распределением, которое обозначается $\chi_{n; \alpha}^2$.

Распределение χ^2 , которое мы ввели в этой главе раньше, рассмотрено во многих книгах, например у Бендата и Пирсола (1971). Оно зависит от некоторого числа n квадратов независимых случайных величин (называемого числом степеней свободы (ст. св.)). Значение n равно разности числа $k + 2$, если карманы используются все до единого, и числа разных независимых линейных ограничений, наложенных на наблюдения. Одно такое ограничение присутствует всегда, поскольку известные частоты для первых $(k + 1)$ интервалов группировки определяют последнюю оставшуюся (напомним, что сумма всех частот равна N). Два дополнительных ограничения обусловлены подгонкой теоретической нормальной функции плотности к частотной гистограмме наблюдаемых данных. Это связано с тем, что при вычислении $\{F_j\}$ используются не истинные, а выборочные значения среднего и дисперсии. В результате следует вычитать еще двойку из числа степеней свободы. Итак, если используются все $\{N_j\}$, то

$$n = (k + 2) - 3 = k - 1.$$

На самом деле число n может быть и меньше, так как карманы, для которых $F < 2$, должны объединяться с другими. Детали такой операции мы обсудим позднее.

После того как число ст. св. n получено, проверка гипотезы проводится следующим образом. Допустим, что величина x имеет нормальное распределение. Группируя выборочные наблюдения по $(k + 2)$ интервалам группировки и определяя значения частот F_j для каждого интервала, основываясь на выборочных значениях среднего и дисперсии, вычисляют величину (2.12). Всякое отклонение выборочной функции плотности от нормальной приводит к возрастанию значения X^2 . Гипотеза о том, что данные имеют нормальное распределение, принимается, если

$$X^2 \leq \chi_{n; \alpha}^2.$$

В этом случае говорят, что гипотеза принимается с уровнем значимости α . Если же значение X^2 превосходит $\chi_{n; \alpha}^2$, то гипотеза отвергается с тем же уровнем значимости α . Широко используются общепринятые уровни значимости 5, 10, 20% (соответствующие доверительным уровням 95, 90 и 80%). Особые значения доверительного уровня выбираются исследователями реже. Авторы, как правило, если не оговорено противное, предпочитают другим значениям α традиционное значение, равное 5%.

Когда для проверки на гауссовость выбирается критерий согласия хи-квадрат, можно воспользоваться выражением для числа интервалов группировки при данном значении N , предложенном Кендаллом и Стъртом (1973). Они, предполагая, что данные не-

коррелированы и $\alpha = 0.05$, получили следующую формулу:

$$\text{Число интервалов группировки} = 1.87 (N - 1)^{2/5}.$$

Значения этой функции приводятся в табл. 2.1. Выше отмечалось, что смещения в гистограмме могут быть значительными, если

Таблица 2.1.

Наименьшие оптимальные числа
интервалов группировки k
для выборки объема N при $\alpha = 0.05$

N	k	N	k
200	16	20 000	94
400	20	40 000	129
600	24	70 000	162
800	27	100 000	187
1 000	30	200 000	247
1 500	35	400 000	326
2 000	39	700 000	407
4 000	57	1 000 000	470
7 000	65	1 140 000	500
10 000	74		

число интервалов группировки станет сравнимым с числом уровней отсчетов цифрового преобразователя.

Статистики, применяя критерий хи-квадрат, обычно пользуются эмпирическим правилом, согласно которому каждому интервалу должна соответствовать частота, не меньшая 2. Такое ограничение позволяет выбрать разумные значения a и b . Наименьшего числа попаданий следует ожидать для крайних карманов. Поэтому значение a определяется из уравнения

$$2 = N \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(a-m)/s} e^{-t^2/2} dt \right\},$$

которое можно разрешить относительно a . После этого значение b находят по простой формуле

$$b = 2m - a,$$

а параметр k — по формуле

$$k = [\text{число интервалов группировки}] - 2.$$

После того как эти три параметра определены, можно вычислять выборочную функцию плотности и ожидаемые частоты для нормального распределения.

На следующем шаге с помощью сравнения решается вопрос о том, принять или отвергнуть гипотезу о гауссовском распре-

делении данных. Как упоминалось выше, обычно эту процедуру сравнения X^2 и χ_n^2 ; α начинают с вычисления параметра α' , определенного равенством

$$\alpha' = P \{ X^2 > \chi_n^2; \alpha' \}$$

Этот параметр зависит только от X^2 и n . После вычисления его сравнивают с заранее выбранным значением α и проводят проверку на гауссовость, опираясь на следующее правило: гипотеза при $\alpha' \leq \alpha$ принимается, при $\alpha' > \alpha$ отвергается. Укажем один метод вычисления величины α' : для нечетных n применяем формулу

$$\alpha' = 2\Phi(X) - 1 - \sqrt{\frac{2}{\pi}} e^{-X^2/2} \left[\sum_{r=1}^{(n-1)/2} \frac{X^{2r-1}}{1 \cdot 3 \cdot 5 \dots (2r-1)} \right]$$

и для четных n — формулу

$$\alpha' = 1 - e^{-X^2/2} \left[1 + \sum_{r=1}^{(n-2)/2} \frac{X^{2r}}{2 \cdot 4 \cdot 6 \dots (2r)} \right].$$

Хотя существуют более эффективные и точные алгоритмы, как правило, для задач анализа временных рядов вполне достаточно предложенного.

2.8. ВЕРОЯТНОСТНЫЕ ФУНКЦИИ ПЛОТНОСТИ ПИКОВ

С определением пиков, максимумов и минимумов связаны известные терминологические трудности. При изучении экстремальных значений данного ряда представляют интерес по крайней мере три различных вопроса:

распределение наибольшего (или наименьшего) значения в ряде длины p ,

распределение наибольшего значения между двумя пересечениями нулевого уровня,

распределение значений пиков.

Рассмотрим задачу первого типа. Предположим, что имеется N независимых наблюдений $\{x_i\}$, имеющих функцию плотности $f(x)$ и функцию распределения $F(x)$. Тогда функция плотности наибольшего значения имеет вид

$$f(x_{\max}, N) = Nf(x) [F(x)]^{N-1}.$$

Функция распределения значения x_{\max} определяется формулой

$$F(x_{\max}, N) = [F(x)]^N,$$

а математическое ожидание x_{\max} — формулой

$$E[x_{\max}] = \int_{-\infty}^{\infty} x N f(x) F^{N-1}(x) dx. \quad (2.13)$$

Допустим, что x_i ($i = 1, \dots, N$) — независимые величины, равномерно распределенные на $[-1/2, 1/2]$. Тогда

$$f(x_i) = \begin{cases} 1 & \text{при } -1/2 \leq x_i \leq 1/2, \\ 0 & \text{в остальных случаях;} \end{cases}$$

$$F(x_i) = \begin{cases} 0 & \text{при } x < -1/2, \\ x + 1/2 & \text{при } -1/2 \leq x \leq 1/2, \\ 1 & \text{при } x > 1/2; \end{cases}$$

$$F(x_{\max}, N) = \begin{cases} 0 & \text{при } x_{\max} < -1/2, \\ (x + 1/2)^N & \text{при } -1/2 \leq x_{\max} \leq 1/2, \\ 1 & \text{при } x_{\max} > 1/2; \end{cases}$$

$$E[x_{\max}] = \int_{-1/2}^{1/2} xN(x + 1/2)^{N-1} dx = \frac{N-1}{2(N+1)}.$$

Хотя вычисление $E[x_{\max}]$ по формуле (2.13) может в теории вызвать затруднения, нахождение x_{\max} по выборке совсем просто. На практике, даже если и не интересоваться x_{\max} и x_{\min} самими по себе, все же разумно искать их с помощью какой-нибудь программы, поскольку это нередко оказывается полезным при проверке данных с целью удаления неправдоподобных значений.

Задача анализа пиков второго типа встречается при изучении множества наибольших значений между пересечениями нулевого уровня. Найти такие значения нетрудно. Это можно сделать с помощью следующего типичного алгоритма. Как всегда, предполагается, что ряд $\{x_i\}$ имеет N точек.

1. Вычислить таблицу значений $\{I_k\}$, $k = 1, \dots, K$, в которой I_k определяются парами неравенств

$$x_{I_k} \leq 0 \quad \text{и} \quad 0 < x_{(I_k+1)}$$

или

$$0 < x_{I_k} \quad \text{и} \quad x_{(I_k+1)} \leq 0.$$

2. Определить минимум значений x_i , $i = (I_k + 1), \dots, I_{k+1}$, если $0 < x_{I_k}$, и максимум в другом случае.

3. Далее, если $x_{I_k} < 0$, то найти максимум x_i , $i = (I_k + 1), \dots, I_{k+1}$; в противном случае найти минимум этих же значений.

Полученные значения максимумов и минимумов можно использовать двояко. Во-первых, применить приемы предыдущего параграфа и построить выборочные функции плотности отдельно для максимумов и минимумов. Во-вторых, можно сравнить значения абсолютных значений минимумов с максимумами. Для данных многих типов статистические свойства пересечений нулевого уровня хорошо известны, поэтому максимумы (или минимумы) между

соседними пересечениями нулевого уровня, как правило, поддаются анализу. Несколько более сложные проблемы возникают в третьем случае, указанном в начале этого параграфа, а именно при изучении пиков.

Под пиком здесь подразумевается наибольшее значение между двумя относительными минимумами. Например, на рис. 2.2 есть

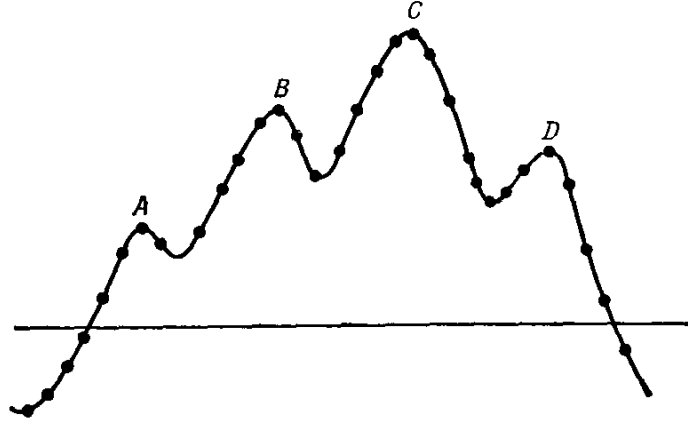


Рис. 2.2. Различные типы пиковых значений.

только один пик второго типа (точка C), а пиков третьего типа — целых три (A, B и D). Приведем процедуру для выделения таких значений.

1. Для максимумов — найти все такие x_p , что

$$x_p - x_{p-1} > 0, \quad x_p - x_{p+1} > 0.$$

2. Аналогично для минимумов — найти все такие x_q , что

$$x_q - x_{q-1} < 0, \quad x_q - x_{q+1} < 0.$$

Эти множества максимумов и минимумов также можно обрабатывать с помощью приемов, изложенных в предыдущем параграфе.

2.9. МНОГОМЕРНЫЕ ФУНКЦИИ ПЛОТНОСТИ

Для двух рядов данных $\{x(i)\}$ и $\{y(i)\}$ можно вычислить их совместную выборочную ВФП, или, иначе, выборочную гистограмму. Такое вычисление производится при помощи деления обоих интервалов изменения данных на k_1 и k_2 подынтервалов соответственно, как это показано на рис. 2.3. Вообще, в двумерном случае потребуется больший по сравнению с одномерным случаем объем памяти. Например, если для самого ряда $\{x(i)\}$ потребуется $k_1 + 2$ ячеек, а для $\{y(i)\}$ потребуется $k_2 + 2$ ячеек, то необходимое число карманов для их совместного распределения может оказаться равным $(k_1 + 2)(k_2 + 2)$. Это количество может быть чрезмерно большим, особенно если обрабатывается r таких рядов. В этом случае из них можно выбрать $r(r-1)/2$

различных пар рядов. Предполагая, что $k_1 = k_2 = \dots = k_r = k$, получим такое выражение для общего необходимого объема памяти ЭВМ:

$$\frac{r(r-1)}{2} (k+2)^2.$$

При $r = 50$ и $k = 100$ это число равно 12 744 900, в то время как для одномерных гистограмм с теми же параметрами нужно всего лишь 5100 ячеек. (См. рис. 2.3.)

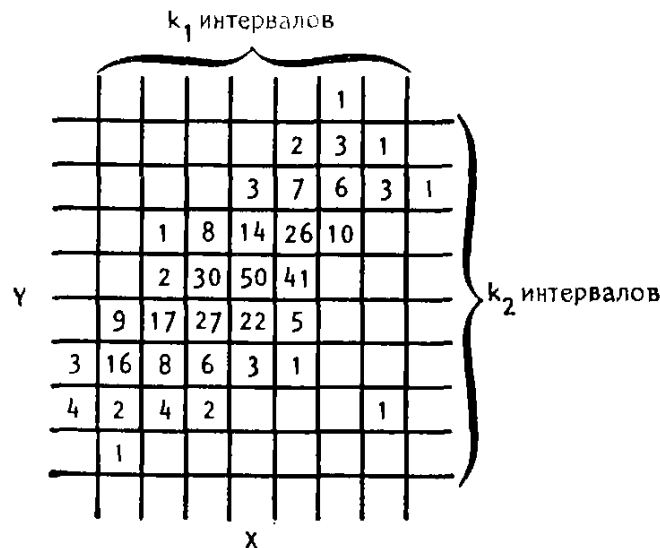


Рис. 2.3. Выборочная гистограмма для совместной функции плотности.

Добиться наглядного представления результатов на дисплее довольно трудно. Машинные распечатки выборочных совместных гистограмм неудобны из-за обилия цифр, от которых рябит в глазах. Изобразить результаты графически тоже трудно. С некоторым успехом можно пользоваться контурными графиками, но их построение не так уж просто. Несомненно, создание трехмерных дисплеев в будущем должно повысить значение совместных гистограмм. Однако пока полностью удовлетворительных методов представления результатов на дисплее нет, и это сильно усложняет построение выборочных гистограмм, мешает интерпретировать данные.

2.10. ПРИМЕРЫ И ПРИМЕНЕНИЯ ВЕРОЯТНОСТНЫХ ФУНКЦИЙ ПЛОТНОСТИ

На рис. 2.4 представлена выборочная гистограмма, построенная по 1001 значению равномерного белого шума, полученного на цифровой ЭВМ. Теоретическая функция плотности определяется формулой

$$f(x) = \begin{cases} 1, & \text{если } -1/2 \leq x < 1/2, \\ 0, & \text{в противном случае.} \end{cases}$$

В этом случае теоретические значения среднего и дисперсии равны

$$\begin{aligned}\mu &= \int_{-\infty}^{\infty} x f(x) dx = 0, \\ \sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \\ &= \int_{-1/2}^{1/2} x^2 dx = \frac{x^3}{3} \Big|_{-1/2}^{1/2} = \frac{1}{24} + \frac{1}{24} = \frac{1}{12} = 0.08333333.\end{aligned}$$

С этими значениями отлично согласуются величины, полученные по выборке и представленные в нижнем правом углу рисунка.

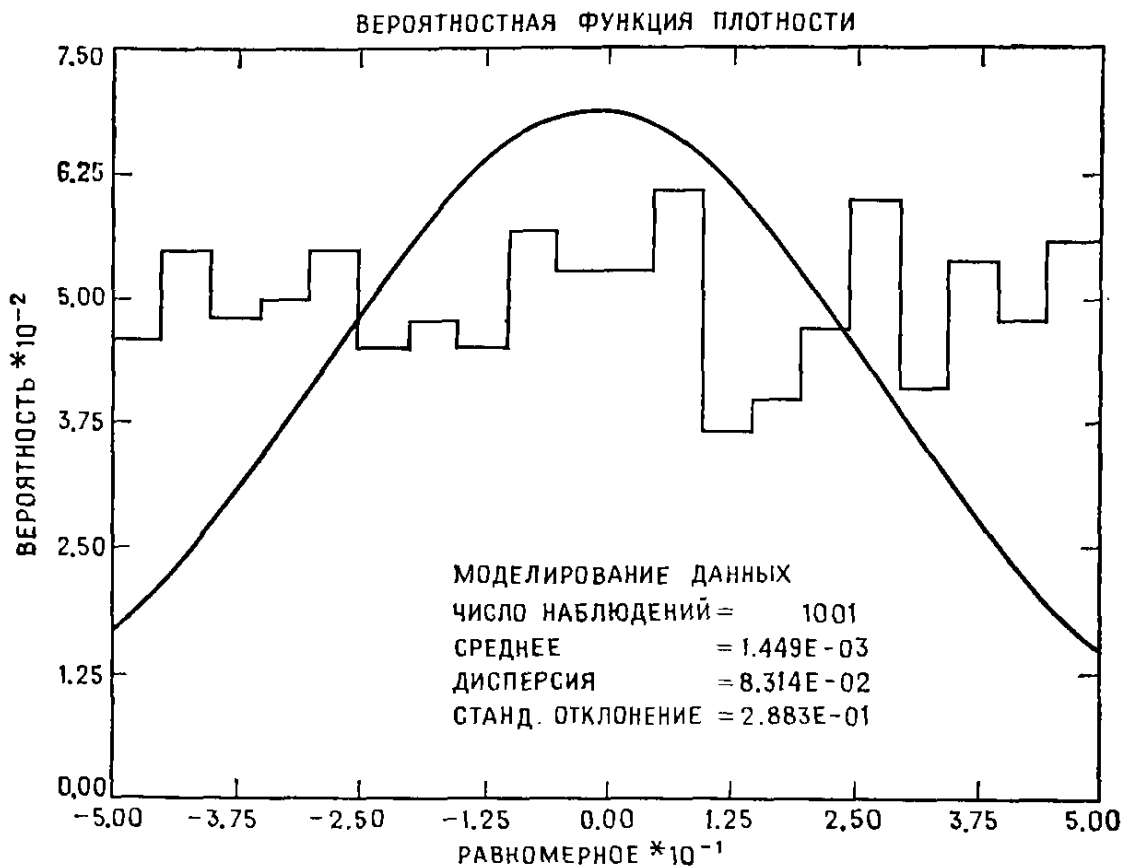


Рис. 2.4. Выборочная функция плотности последовательности значений равномерно распределенного белого шума.

Значение P_j выборочной гистограммы в j -м интервале есть отношение числа наблюдений, попавших в j -й карман, ко всему числу наблюдений:

$$P_j = \frac{N_j}{N} = (\text{выборочная вероятность того, что } [d_{j-1} \leq x < d_j]).$$

Поэтому математическое ожидание P_j равно

$$E[P_j] = \int_{-\infty}^{\infty} P_j f(x) dx = \int_{d_{j-1}}^{d_j} f(x) dx.$$

В случае рассматриваемого равномерного шума

$$E[P_j] = \int_{d_{j-1}}^{d_j} dx = d_j - d_{j-1}.$$

В частности, в нашем случае эта разность равна примерно 0.05.

Отметим, что на выборочную гистограмму наложена кривая плотности гауссовского распределения с теми же значениями выборочного среднего и дисперсии. Очевидно, что они совершенно не согласуются, и критерий χ^2 здесь не нужен.

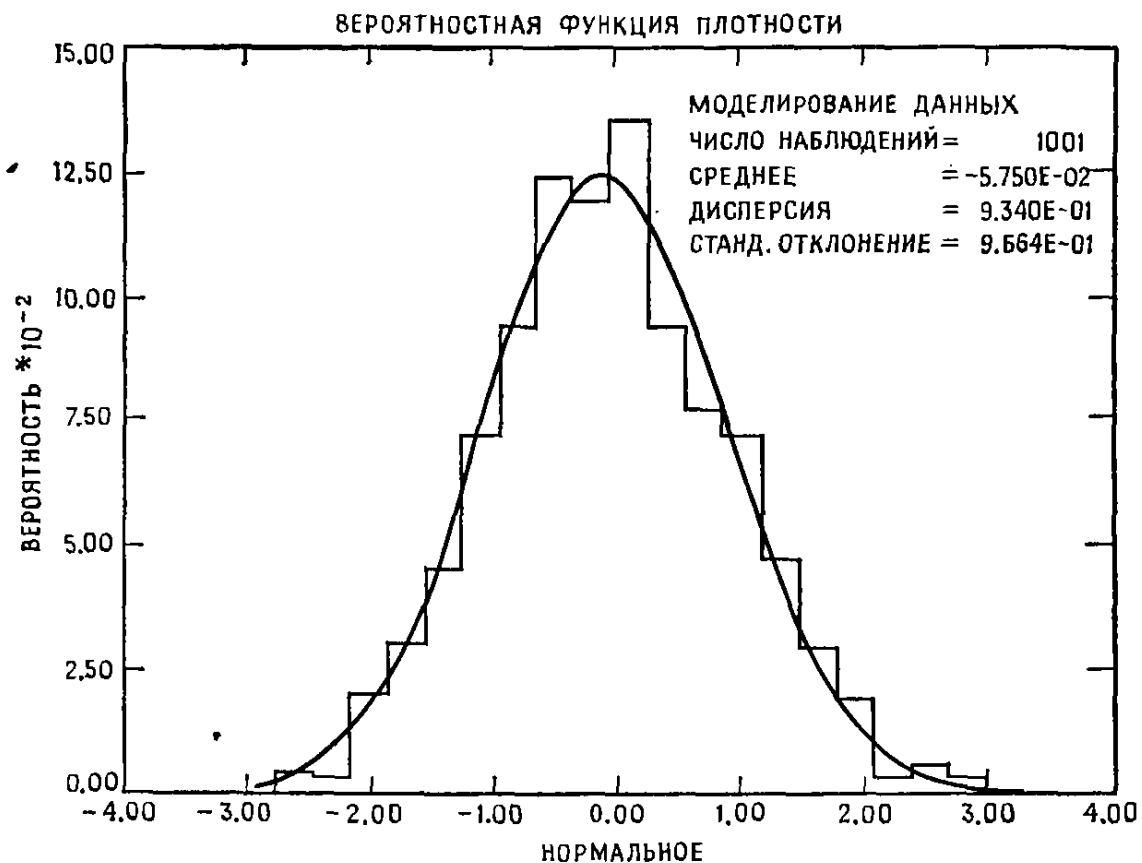


Рис. 2.5. Выборочная функция плотности последовательности значений гауссовского белого шума.

На рис. 2.5 показано, как такой критерий используется в случае гауссовского шума, полученного с помощью цифрового генератора псевдослучайных чисел. Математическое ожидание и дисперсия равны 0 и 1 соответственно. Выборочная дисперсия, рав-

ная 0.934, хотя и немного меньше теоретической, но лежит в разумных пределах ожидаемой величины. Данные согласуются с гипотезой о гауссовости на 95%-м доверительном уровне (параметры критерия: выборочная величина $X^2 = 18.39$, $P\{\chi^2, 1000, 0.05\} = 0.69825$).

Теперь рассмотрим следующий случай, в котором ряд данных образован значениями синусоиды, полученными по формуле

$$x(i) = \sin(0.15i), \quad i = 0, \dots, 1000.$$

Случайная синусоида с постоянными коэффициентами имеет такие статистические параметры:

$$\begin{aligned} \mu &= 0, \quad \sigma^2 = 1/2, \\ f(x) &= \begin{cases} \frac{1}{\pi \sqrt{1-x^2}}, & -1 \leq x \leq 1, \\ 0, & |x| > 1. \end{cases} \end{aligned} \quad (2.14)$$

Заметим, что на части интервала $(-1, 1)$ выборочная функция плотности больше единицы. Однако интеграл от вероятностной функции плотности по всей или по какой-нибудь части интервала будет меньше или равен единице.

На рис. 2.6 изображена выборочная гистограмма. Выборочные значения среднего и дисперсии, а также гистограмма находятся в хорошем соответствии с (2.14).

Большой интерес представляет следующий пример, в котором гауссовский шум пропускается через узкополосный фильтр. На достаточно коротких отрезках данных выход такого фильтра довольно похож на простую синусоиду. Однако при вычислении выборочной функции плотности мы не увидим сходства с функцией плотности синусоиды. Это показывает рис. 2.7, на котором изображена выборочная функция плотности, построенная по 60 точкам, полученным на выходе полоснопропускающего фильтра, у которого расстояние между точками половинной мощности составляет 1% общей допустимой полосы частот (т. е. 1% частоты Найквиста), а центр частот расположен в точке, соответствующей 15% частоты Найквиста. Как видно из рисунка, эта выборочная функция плотности не похожа на функции плотностей, изображенные на рис. 2.6 и 2.5.

На рис. 2.8 приведена выборочная функция плотности для таких же данных, но взятых на более длинном промежутке. В этом случае выборочная функция плотности приближает гауссовскую функцию плотности, пожалуй, хорошо. Здесь критерий χ^2 для проверки гауссовости неприменим, поскольку не выполняется предположение о независимости значений данных.

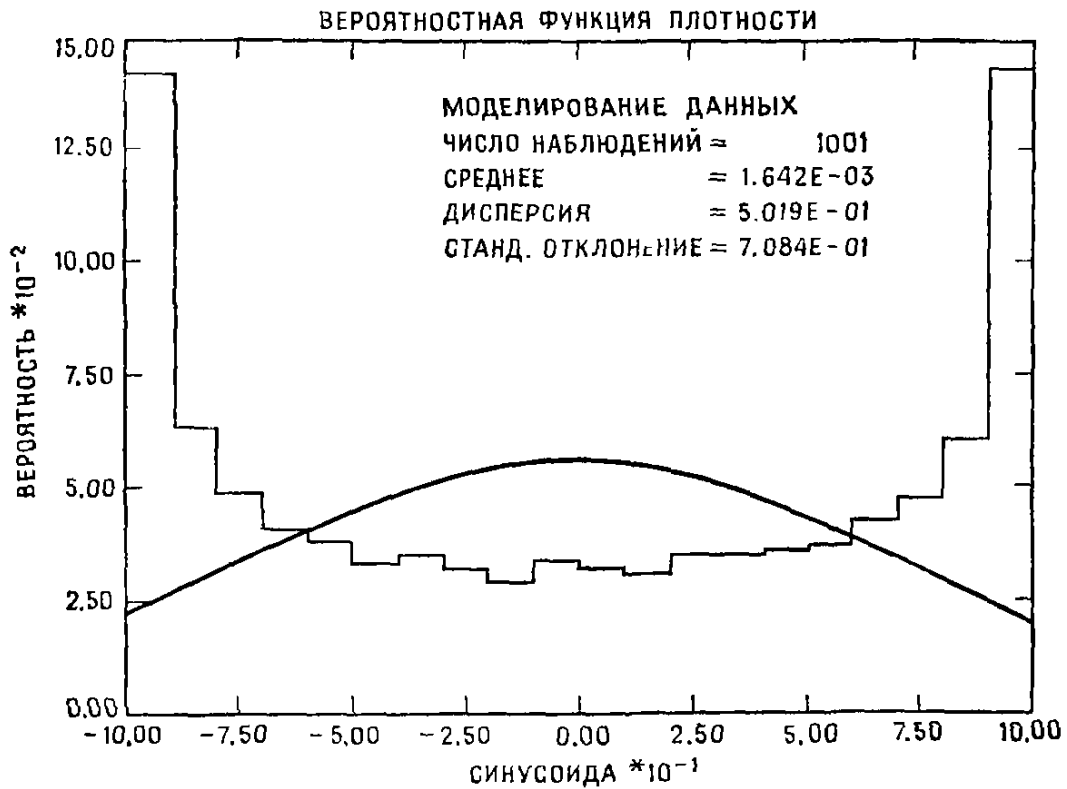


Рис. 2.6. Выборочная функция плотности синусоиды.

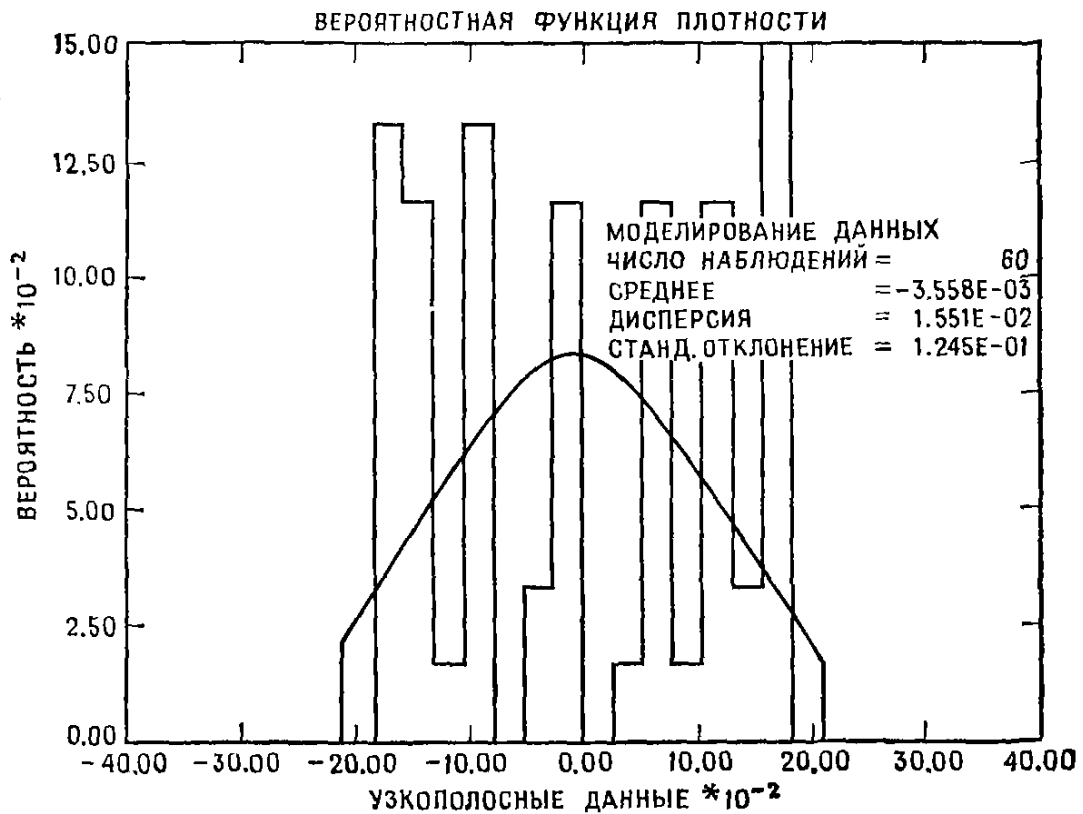


Рис. 2.7. Выборочная функция плотности для небольшого участка процесса, полученного узкополосной фильтрацией.

Этого вполне можно было бы избежать, если N (которое в этом случае равно 10 001) выбирать таким образом, чтобы учитывать возможность зависимых наблюдений. В таком случае N сократилось бы до 1% первоначального значения, т. е. до 100 наблюдений. В результате произошло бы соответствующее уменьшение значения выборочного χ^2 разницы выборочной функции плотности и плотности гауссовского распределения, построенной по выборочному среднему и дисперсии.

На самом деле в условиях применимости центральной предельной теоремы гауссовский шум или, вообще, шум с любым рас-



Рис. 2.8. Выборочная функция плотности для большого участка процесса, полученного узкополосной фильтрацией. Сюда включены данные, полученные для рис. 2.7.

пределением после линейной фильтрации сходится к гауссовскому шуму.

Применение критерия χ^2 в таком случае может оказаться не совсем простым делом, поскольку в выборе точно определить число независимых оценок трудно.

Уточнение данных с помощью вероятностных гистограмм. Один из способов уточнения данных состоит в том, чтобы изобразить все данные на графике и просмотреть их. Для больших объемов данных это утомительно, требует больших затрат времени и средств. В другом способе уточнения данных, более экономичном, используются выборочные функции плотности. Вообще говоря,

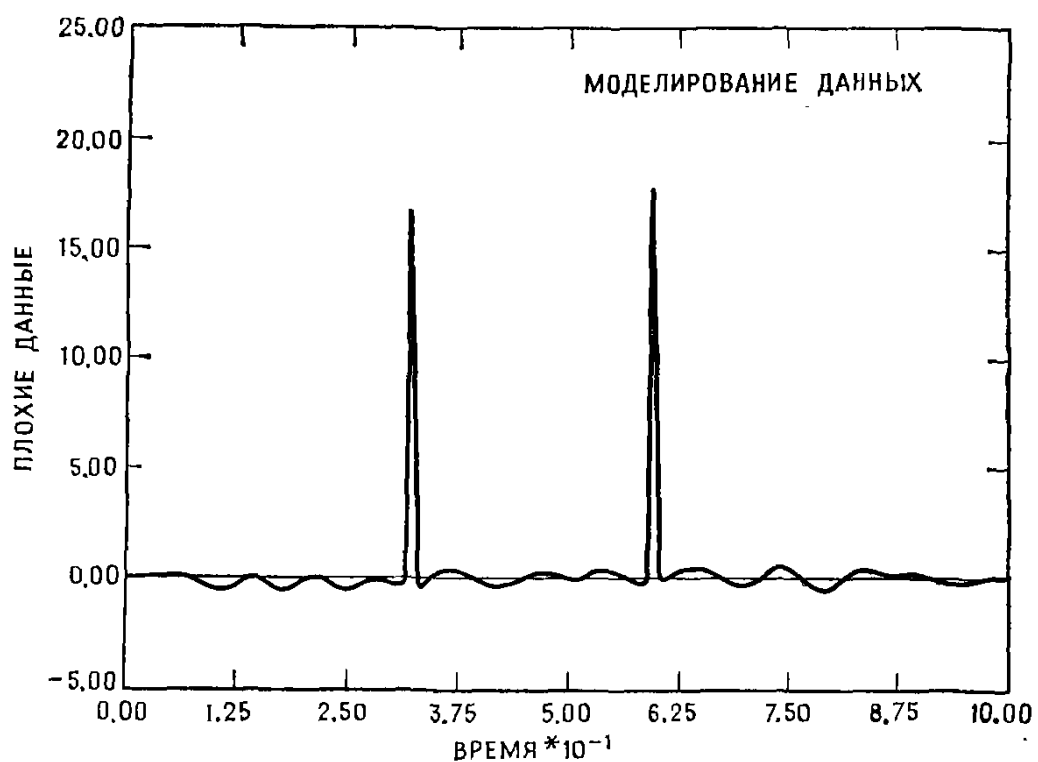


Рис. 2.9. Временная последовательность с двумя неправдоподобными значениями.

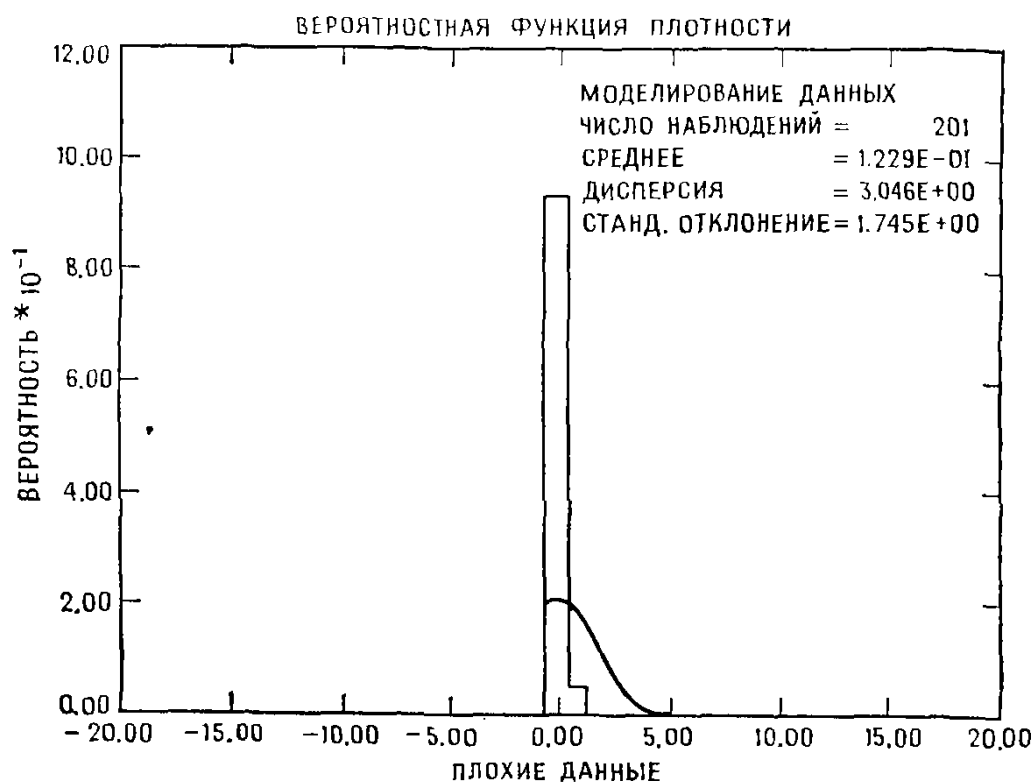


Рис. 2.10. Выборочная функция плотности данных, изображенных на рис. 2.9. Обратите внимание, как данные группируются вместе.

выборочные функции плотности несут бóльшую информацию о природе данных и выявляют проблемы разного характера. Для больших количеств данных вывод выборочной функции плотности по затратам может сравниться с полной распечаткой данных или с получением их графика.

С помощью гистограмм быстро выявляются неправдоподобные значения. На рис. 2.9 показаны данные, полученные таким способом, чтобы появлялись неправдоподобные значения. Соответствующая выборочная функция плотности представлена на рис. 2.10. Как видно, большинство данных тесно сгруппировалось вместе, а две неправдоподобные точки расположились далеко справа.

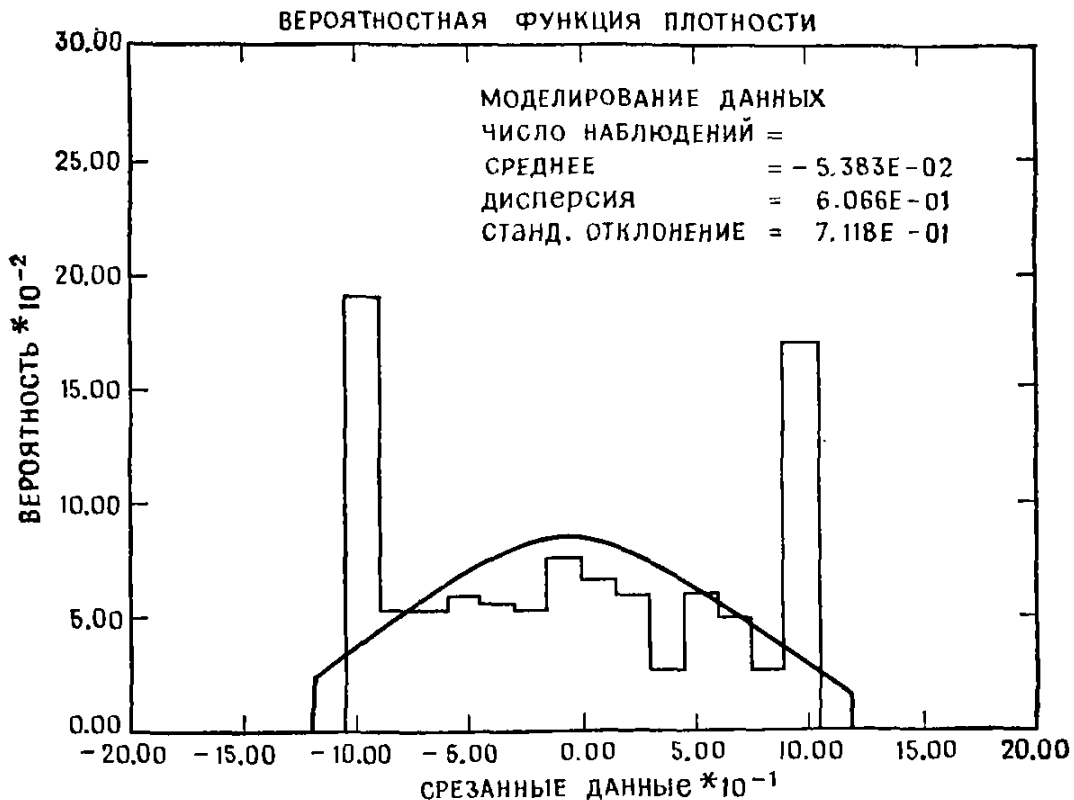


Рис. 2.11. Выборочная функция плотности срезанных данных.

Все промежуточные карманы пусты. Такая картина типична для данных с неправдоподобными значениями, поскольку наличие неправдоподобных данных загоняет все настоящие данные в один или два кармана.

Выборочная функция плотности, порождающая проблему другого характера, представлена на рис. 2.11. Здесь крайние карманы содержат значительное число наблюдений, в то время как ни один из оставшихся не содержит наблюдений. Это указывает на «срезание». Срезание может происходить по нескольким причинам. Чаще всего оно происходит из-за ухода аналогового усилителя в точки, где нарушается линейность, или из-за такой

установки аналого-цифрового преобразователя, при которой не покрывается весь возможный интервал изменения данных в некотором канале.

Чтобы произвести процедуру уточнения данных, рекомендуется делать следующее. Предполагается, что для получения данных используется программа, в которой калибровка задается самим пользователем.

1. Задайте калибровку и при ней вычислите выборочную функцию плотности и основные статистические параметры для каждого файла полученных на выходе данных.

2. Прежде чем начинать дальнейшую обработку, просмотрите полученную выборочную функцию плотности. При этом нужно учесть следующее:

а) Необходимо следить за временем (или какой-то другой независимой переменной) с тем, чтобы установить, совпадает ли обработанный промежуток с истинным и согласуется ли число точек для выборочной функции плотности с предполагаемым.

б) Следует проверить статистики для каждой зависимой переменной. Необходимо установить, насколько отвечают физической реальности выборочная функция плотности с точки зрения тех проблем, которые только что были затронуты, и тех, которые появятся сами, а также среднее, дисперсия, максимум и минимум.

Предложенная процедура не гарантирует, что данные станут хорошими, но она экономно и быстро выявляет возникающие проблемы.

Упражнения

2.1. Докажите, что $E(s^2) \neq \sigma^2$, если $s^2 = N^{-1} \sum_{i=0}^{N-1} (x_i - \bar{x})^2$.

2.2. Докажите, что дисперсия синусоиды $r(t)$ с амплитудой A и произвольной фазой равна $A^2/2$.

2.3. Докажите соотношение (2.2), т. е.

$$\sigma_x^2 = E(x^2) - E^2(x)$$

Ответ:

$$\begin{aligned} \sigma^2 &= E(x - \mu)^2 = E(x^2 - 2x\mu + \mu^2) = \\ &= E(x^2) - 2E(x)\mu + \mu^2 = \\ &= E(x^2) - 2\mu^2 + \mu^2 = E(x^2) - \mu^2. \end{aligned}$$

2.4. Докажите, что корреляция ограничена следующими пределами?

$$-1 \leq \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} \leq 1$$

Ответ:

$$\begin{aligned} \text{cov}(x, y) &= E(x - \mu_x)(y - \mu_y) = E(xy - x\mu_y - y\mu_x + \mu_x\mu_y) = \\ &= E(xy) - \mu_x\mu_y - \mu_y\mu_x + \mu_x\mu_y = \\ &= E(xy) - \mu_x\mu_y, \end{aligned}$$

$$\begin{aligned} \text{cov}^2(x, y) &= E^2(xy) - 2E(xy)\mu_x\mu_y + \mu_x^2\mu_y^2 \geq \\ &\geq E(x^2)E(y^2) - E(x^2)\mu_y^2 - E(y^2)\mu_x^2 + \mu_x^2\sigma_y^2. \end{aligned}$$

2.5. Вычислите ковариационную функцию синусоиды.

2.6. Приведите формулу для sip-члена в формуле (2.3).

Задачи для решения на ЭВМ.

2.7. Смоделируйте равномерный шум и повторите рис. 2.5.

2.8. Смоделируйте гауссовский шум и повторите рис. 2.6.

2.9. Смоделируйте синусоиду и повторите рис. 2.7.

Глава 3

СБОР И ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ДАННЫХ

3.1. ВВЕДЕНИЕ

В этой главе мы остановимся на некоторых основных задачах, возникающих в начальной стадии анализа временных рядов. Данные, которые будут дальше рассматриваться, распадаются по типу на два разных класса — *непрерывные* (или *аналоговые*) данные и данные, имеющие по существу *цифровой* характер.

Для работы более удобны, пожалуй, данные второго типа. К ним можно отнести сведения об уровне цен на бирже, статистику занятости, годовые уровни осадков и многие другие. Даже если поиск и регистрация таких данных могут оказаться утомительным занятием, то со статистической точки зрения или с точки зрения приложений никаких сложностей не возникает.

Исследованию же непрерывных данных сопутствует ряд проблем. Среди них измерение данных, передача и (или) их запись, преобразование из аналоговой формы в цифровую.

Вопросам измерений посвящена книга Мэгреба и Бломквиста (1971). Задача здесь сводится, по сути дела, к конструированию или выбору *датчика* — устройства, преобразующего измеряемую физическую величину в электрический потенциал. Как правило, хотя и не всегда, полученный в результате потенциал пропорционален измеряемой физической величине. К таким приборам относятся микрофоны, гидрофоны, акселерометры и устройства, измеряющие давление.

Обычно датчик выбирается так, чтобы его передаточная функция была линейной в пределах предусмотренного промежутка. Идеальное линейное устройство создать невозможно. Интервал линейности ограничен, и большинству датчиков в большей или меньшей степени присуще наличие *гистерезиса*. Последнее означает, что при изменении элемента входа от начальной точки до конечной с возвращением в начальную точку две соответствующие траектории выхода, которые должны быть одинаковыми, в точности не совпадут. Такой график называется *гистерезисной кривой* инструмента. Для ряда хорошо сконструированных аппаратов получающаяся ошибка пренебрежимо мала.

Даже и совершенно линейный датчик может иметь передаточную функцию, заглушающую вход на некоторых промежутках частот. Например, передаточные функции для некоторых микро-

фонов выглядят почти как полоснопропускающие фильтры. Иначе говоря, они пропускают широкую полосу информации, но заглушают частоты, близкие к нулю или превосходящие определенную точку отсечки. Внутри полосы пропускания передаточные характеристики могут изменяться крайне незначительно. На этом мы еще остановимся, поскольку иногда такое глушение желательно компенсировать.

3.2. СБОР ДАННЫХ

На рис. 3.1 изображена типичная система для сбора данных в цифровой форме.

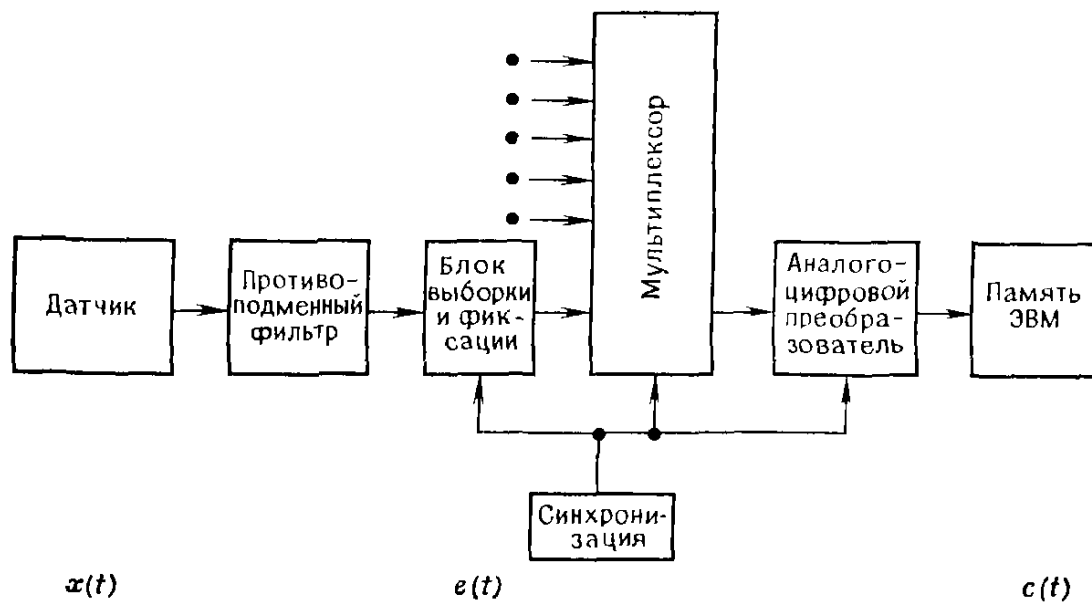


Рис. 3.1. Система для обработки цифровых данных. Все операции управляются синхронизирующей схемой.

Выход датчика поступает сначала в противоположенный фильтр. Если эффективная скорость отсчетов для этого выхода примерно равна S , то, чтобы предотвратить явление подмены частот для высокочастотной информации, фильтр должен гасить все частоты свыше $S/2 = F$ Гц. Другими словами, передаточная функция идеального противоположенного фильтра в этом случае имела бы вид

$$H(f) = \begin{cases} 1, & -F < f < F, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Однако построить фильтр, имеющий в точности такую передаточную функцию, невозможно. Здесь возникают два рода осложнений. Во-первых, передаточная функция приведенного вида физически неосуществима. Чтобы это устранить, следует ввести задержку в виде ненулевого фазового угла. Во-вторых, идеального блока, устраняющего все частоты вне промежутка

$|f| > F$, не существует. Ребра или углы графика должны быть сглажены, и насколько бы незначительной ни была информация с частотами, превосходящими $|F|$, полностью избавиться от нее невозможно.

Другая проблема заключается в том, что при сравнении двух временных функций (что бывает часто необходимо) два фильтра, которые при этом используются, должны быть по возможности одинаковыми. В противном случае различие амплитудных или фазовых характеристик может послужить причиной ошибок, которые возникнут при вычислениях передаточной функции, корреляционной функции и относительной задержки между двумя входами.

Типичную форму низкочастотного фильтра, применяемого для устранения подмены частот, имеет фильтр Баттеруорта. Квадрат модуля передаточной функции этого фильтра имеет вид

$$|H(f)|^2 = \left[1 + \left(\frac{2\pi f}{2\pi B} \right)^{2M} \right]^{-1}.$$

Изобразив это соотношение графически, нетрудно установить, что $|H(0)|^2 = 1$ для $f = 0$ и $|H(B)|^2 = 1/2$ для $f = B$, величина $|H(f)|^2$ остается примерно постоянной в промежутке $0 < f < B$ вплоть до малой окрестности B , а при $f > B$ быстро убывает с ростом f . Нетрудно видеть также, что при $f \gg B$

$$|H(f)|^2 \approx \left(\frac{B}{f} \right)^{2M}.$$

Например, для $f = 8B$ и $M = 6$

$$|H(8B)|^2 \approx (1/8)^{12} = 1.46 \cdot 10^{-11} \approx -108.4 \text{ дБ}.$$

Фазовая характеристика фильтра Баттеруорта приближенно линейна вдоль полосы частот. Это означает, что фаза соответствует простой задержке.

Более детально фильтрация будет рассмотрена в следующих двух главах. Та терминология, которая связана главным образом с цифровыми фильтрами, в большинстве случаев применима и к аналоговым.

За противоположенным фильтром на рис. 3.1 следует устройство выборки и фиксации. Хорошим теоретическим описанием такого устройства служит формула взятия замера

$$x(i) = \int_{-\infty}^{\infty} x(t) \delta(iT - t) dt. \quad (3.1)$$

Осуществить операцию, идеально схожую с дельта-функцией, невозможно. Поэтому дельта-функцию заменяют «щелевой» функцией с конечным носителем. Если щель имеет прямоугольную

форму, то формула (3.1) преобразуется к виду

$$x(i) = \frac{1}{a} \int_{iT - a/2}^{iT + a/2} x(t) dt,$$

где параметр a — ширина щели. Вычисление по этой формуле есть, по сути дела, определение среднего значения функции $x(t)$ в промежутке от $iT - a/2$ до $iT + a/2$. Естественно, что ширина a должна быть меньше T , и чем она будет меньше, тем лучше.

Значение $x(i)$ обычно фиксируется блоком выборки и фиксации в виде разности потенциалов заряженного конденсатора и хранится до тех пор, пока мультиплексор не выберет его для пересылки в аналого-цифровой преобразователь. Мультиплексор

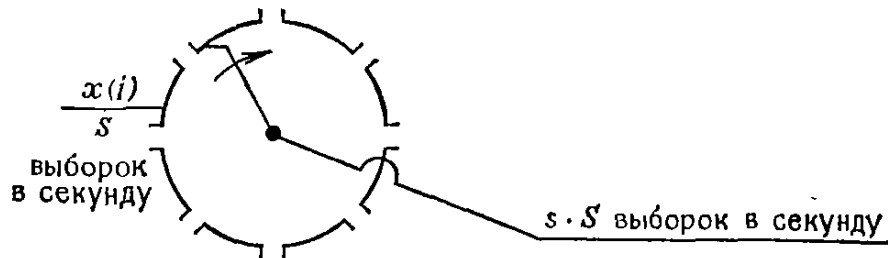


Рис. 3.2. Идеализированный мультиплексор, изображенный здесь в виде поворотного переключателя, имеет восемь сегментов — по одному на каждую обрабатываемую функцию.

можно представлять себе как поворотный переключатель, изображенный на рис. 3.2, со скоростью отсчетов S для каждой функции. Для мультиплексора с s сегментами на выходе вырабатывается sS значений в секунду. Раньше многие мультиплексоры представляли собой электромеханические переключатели. Современные мультиплексоры целиком изготавливаются на базе электроники.

Иногда в целях экономии различные для каждой функции устройства выборки и фиксации, расположенные перед мультиплексором, заменяют одним блоком выборки и фиксации, располагая его за мультиплексором. В этом случае отсчеты функций производятся не одновременно, как это происходит в случае распределения устройств выборки и фиксации по одному на каждый вход, а последовательно. Это приводит к небольшому относительному смещению времен взятия отсчетов для разных функций; чем плотнее отсчеты подаются на мультиплексор, тем меньше величина такого смещения.

Кадр данных, содержащий по одному отсчету каждой функции, получается во время одного полного поворота мультиплексора. С каждым кадром связан момент времени, в который этот кадр преобразуется в цифровую форму. Позднее, во время обработки, обычно генерируется «время» и добавляется в каждый кадр.

Поступающие с мультиплексора отсчеты в аналоговой форме подаются на цифровой преобразователь для перевода их в цифровой формат. Прежде чем обсуждать этот формат, имеет смысл остановиться на методах цифрового представления данных.

3.3. ЦИФРОВОЕ ПРЕДСТАВЛЕНИЕ ИНФОРМАЦИИ

Термином *бит*, предложенным Дж. В. Тьюки, называют одно-разрядное *двоичное целое число*. Таким образом, бит — это величина, принимающая только два значения — 0 и 1. В виде набора битов могут быть представлены числа и параметры. Имеются два основных способа представления — двоичное представление и представление с помощью *двоично-десятичного кодирования* (ДДК).

В двоичном представлении n битов b_0, b_1, \dots, b_{n-1} интерпретируют¹⁾ как число

$$(-1)^{b_0} \sum_{p=1}^{n-1} b_p 2^{n-1-p}.$$

Приведем несколько типичных примеров для $n = 10$.

Двоичная запись	Цифровая запись	
	$b=0$	$b=9$
0000000000	0	0
0000000001	1	0.001953125
1000000001	-1	0.001953125
0111111111	511	0.998046875
1111111111	-511	-0.998046875

Это определение можно модифицировать, введя понятие *двоичной точки*. Двоичной точке отвечает, по сути дела, умножение числа на 2^{-b} , где b — положительное (или отрицательное) целое число. Например, для $b=9$ те же двоичные комбинации можно интерпретировать совершенно другим способом (см. табл. выше).

В ДДК-представлении для десятичных цифр используются группы по 4 бита:

0000	0	0101	5
0001	1	0110	6
0010	2	0111	7
0011	3	1000	8
0100	4	1001	9

Например, если на передней панели цифрового вольтметра указано значение $+9.75$, то в это же время с обратной стороны

¹⁾ Существуют и другие способы определения.

посредством электрических соединений выход реализуется ДДК-представлением 0 1001 0111 0101.

Двоично-десятичное кодирование не столь эффективно, как двоичное представление. Чтобы получать числа в промежутке ± 999 , при ДДК потребуется тринадцать битов, в то время как для числового промежутка ± 1023 в двоичном представлении достаточно лишь одиннадцати. Недостаток двоичного представления связан с необходимостью перехода на некотором этапе к более привычной десятичной записи.

Память ЭВМ содержит большие массивы битов, записанных различными способами,— на ферритовых сердечниках, магнитных дисках, лентах или на магнитном барабане. Эти биты группируются в комплекты стандартной длины, называемые *словами*. Как правило, слово содержит от 4 до 60 битов. Длина слова колеблется в зависимости от типа ЭВМ и составляет 4—16 битов для микро-ЭВМ, 12—32 для мини-ЭВМ и 32—60 битов для стандартных ЭВМ.

При использовании памяти на магнитной ленте принято оперировать терминами *символ* или *байт*. Ленту с семью дорожками можно представлять себе как прямоугольный массив битов, в котором поперек ленты расположены семерки битов, а вдоль — сотни битов в одном дюйме. Каждая семерка битов подразделяется на один *символ* и его *контрольную сумму*. Последняя является младшим значащим битом суммы остальных шести¹⁾.

Если ЭВМ имеет 36-битовые слова, то для их запоминания требуется по шесть смежных символов на магнитной ленте.

Для ленты с девятью дорожками по ширине ленты расположено девять битов; восемь из них образуют один байт, а девятый, как и в случае с семью дорожками, предназначен для контрольной суммы. Поэтому для ЭВМ с 32-битовыми словами одно слово запоминается на ленте в виде четырех последовательных байтов.

Непосредственно в ЭВМ слова допускают две интерпретации в зависимости от того, используется ли представление чисел с *фиксированной точкой* или же с *плавающей точкой*. Представление чисел с фиксированной точкой только что описано — последовательность битов понимается как число, которое следует умножить на масштабный множитель; здесь для представления используется все слово целиком, а масштабный множитель выбирается заранее (в машине его нет).

Представление с плавающей точкой, наоборот, уже в самом себе содержит двоичную точку. Остановимся на двух стандартных форматах слов для представления информации с плавающей точкой.

¹⁾ Существуют два разных способа вычисления контрольной суммы, известные как проверки на *четность* или *нечетность*.

надцатеричном масштабировании намного шире, чем в случае 36-битовых слов с двоичным масштабированием.

Вместе с тем 32-битовый формат оказывается менее точным, чем 36-битовый. Помимо основного различия на три бита в мантиссе эти форматы отличаются также *нормализованной* формой записи чисел. Числа с плавающей точкой называют нормализованными, если наибольший значащий бит мантиссы находится в крайней левой позиции. В случае 36-битового слова бит m_1 за некоторыми исключениями¹⁾ равен 1. Следовательно, в этой форме записи все 27 битов мантиссы всегда будут значащими. В 32-битовом формате с умножением на степень шестнадцати такое положение возможно не всегда. Первые 15 целых чисел в двоичной и шестнадцатеричной системах имеют вид

	Двоичная запись	Шестнадцатеричная запись	Двоичная запись	Шестнадцатеричная запись
			8	1000
1	0001	1	9	1001
2	0010	2	10	1010
3	0011	3	11	1011
4	0100	4	12	1100
5	0101	5	13	1101
6	0110	6	14	1110
7	0111	7	15	1111
				8
				9
				A
				B
				C
				D
				E
				F

Все эти числа имеют один и тот же масштабный множитель, равный 16. Если теперь представить, что слева от этих чисел стоит двоичная точка, то легко понять, что невозможно эти числа сдвинуть левее. Такое явление приводит к уменьшению среднего числа битов, в которых может запоминаться информация.

В некоторых 32-битовых машинах подобная вещь происходит и при округлении. Именно после окончания умножения с плавающей точкой округление производится также при помощи арифметической процедуры со степенью шестнадцати.

Все это приводит к тому, что 32-битовые машины не обеспечивают такой точности, как 36-битовые. Если перейти к десятичной записи, то определить число значащих десятичных цифр для 36-битового слова с плавающей точкой можно из равенства $\log_{10} 2^{27} = 8.13$.

Статистическое изучение рассмотренных представлений показывает, что при шестнадцатеричном масштабировании и округлении теряется в среднем около трех из 24 битов мантиссы 32-битового слова. Поэтому в среднем десятичные числа в этом

¹⁾ Таким исключением, например, служит число 0.

случае имеют $\log_{10} 2^{21} = 6.32$ значащих цифр. Таким образом, разница в точности для 32- и 36-битовых форматов составляет примерно две десятичные цифры.

Часто оказывается, что шести десятичных цифр недостаточно для некоторых расчетов, проводимых при обработке временных рядов. В этих случаях необходимо переходить к вычислениям с *двойной точностью*.

3.4. АНАЛОГО-ЦИФРОВОЕ ПРЕОБРАЗОВАНИЕ

Чтобы избежать в дальнейшем путаницы, при анализе процесса преобразования в цифровую форму ограничимся только одной входной функцией, хотя пользоваться будем несколькими ее представлениями. Перечислим их: $x(t)$ — исходная физическая величина; $x_e(t)$ — выход датчика в вольтах, соответствующий $x(t)$; $C(t)$ — выход $x_e(t)$ датчика, переведенный в *непрерывные* отсчеты; $c(i)$ — выход $x_e(t)$ датчика, переведенный в округленные отсчеты; $e(i)$ — погрешность, равная $C(iT) - c(i)$.

Фиктивная функция $C(t)$ введена здесь только для удобства. Это должно сделать более понятными построения графиков. Как функция $C(t)$, так и функция $c(i)$ измеряются в одних единицах — *отсчетах*. Функция $C(t)$ есть просто результат линейного преобразования функции $x_e(t)$ вида

$$C(t) = Ax_e(t) + B.$$

Например, если номинальный диапазон изменения величины $x_e(t)$ лежит в пределах от -5 до $+5$ В и ему соответствует интервал изменения функции $C(t)$ от 0 до 1023, то $A = 102.3$, $B = 511.5$.

На рис. 3.3, *a* изображена функция $C(t)$ и соответствующие значения $c(i)$, фактически получающиеся на выходе цифрового преобразователя. Две эти функции связаны соотношением

$$c(i) = [C(iT) + 0.5^-].$$

Здесь квадратные скобки означают, что берется значение наибольшего целого числа, не превосходящего значения выражения внутри скобок. Добавление к функции $C(iT)$ величины 0.5^- вызвано необходимостью *округления*, поскольку без него происходило бы постоянное *срезание* истинных значений $C(iT)$.

Погрешность $e(i)$ определяется как разность $c(i)$ и $C(t)$:

$$e(i) = C(iT) - c(i).$$

Если цифровой преобразователь работает правильно, то значения $e(i)$ будут ограничены интервалом $(-1/2, 1/2)$. На рис. 3.3, *b* показана последовательность погрешностей для функций $C(t)$ и $c(i)$, изображенных на рис. 3.3, *a*. Последнюю формулу можно

переписать в виде

$$c(i) = C(iT) + e(i).$$

С этой точки зрения функция $c(i)$ есть сумма функции $C(iT)$ и погрешности, или, иначе, *шума*.

Теперь перед нами стоит задача охарактеризовать функцию $e(i)$. Это сделать не так просто, поскольку в одних случаях функция $e(i)$ ведет себя как случайная величина, а в других она может оказаться сильно коррелированной последовательностью или четной периодической функцией.

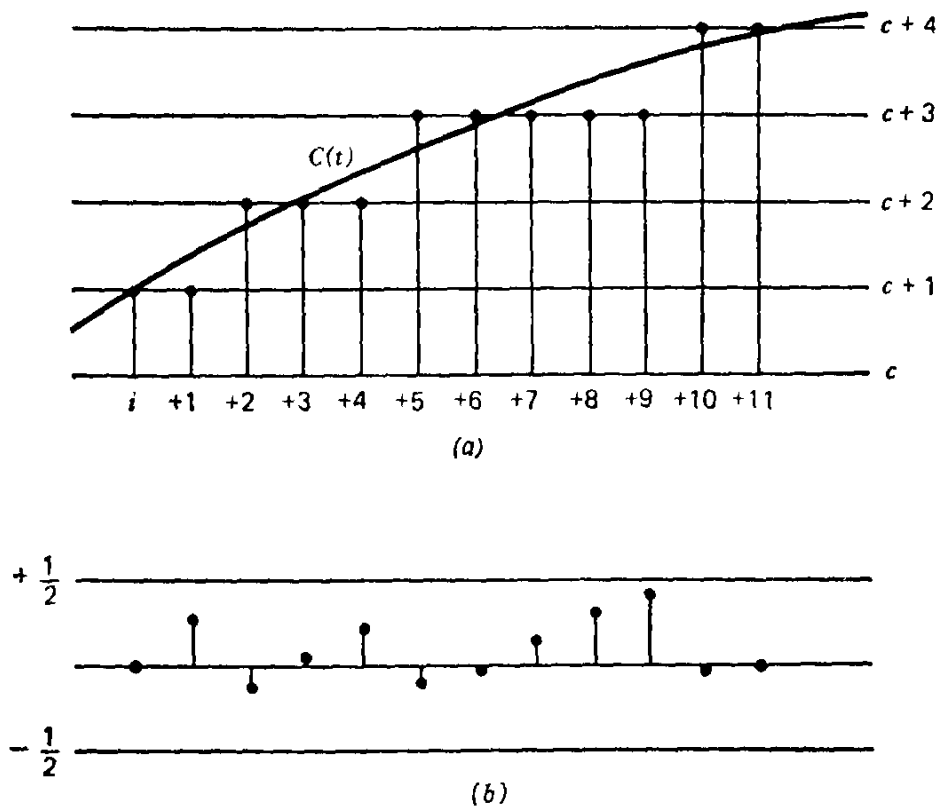


Рис. 3.3. Функция $C(t)$ пересекает четыре уровня отсчета. a — функция $C(t)$ и соответствующая ей функция $c(i)$, полученная на выходе преобразователя в цифровую форму; b — погрешность $e(i)$.

Рассмотрим пример, в котором функция $C(t)$ есть синусоида с частотой, равной частоте Найквиста:

$$C(t) = 511.25 \left[\cos\left(\frac{\pi t}{T}\right) + 1 \right] + 0.25.$$

В этом случае для $C(iT)$ (обратите внимание на замену независимой переменной) получим равенства

$$\begin{aligned} C(iT) &= 511.25 [\cos(\pi i) + 1] + 0.25 = \\ &= \begin{cases} 1022.5 + 0.25 = 1022.75 & \text{для четных } i, \\ 0.25 & \text{для нечетных } i. \end{cases} \end{aligned}$$

Следовательно,

$$c(i) = \begin{cases} 1023, & \text{если } i \text{ четное,} \\ 0, & \text{если } i \text{ нечетное,} \end{cases} \quad e(i) = \begin{cases} -0.25, & \text{если } i \text{ четное,} \\ 0.25, & \text{если } i \text{ нечетное.} \end{cases}$$

Таким образом, погрешность также есть синусоида с частотой, равной частоте Найквиста, и значит, она периодична.

Другая крайняя возможность — полностью случайная функция $e(i)$. Такая функция встречается во многих ситуациях. Этот случай представляет большой интерес, и мы остановимся на нем подробнее.

Сделаем следующие предположения:

1) Функция $e(i)$ случайна, и значит, $E[e(i)e(i+p)] = 0$ для $p \neq 0$.

2) Значения $e(i)$ имеют равномерное распределение.

Второе предположение означает, что все значения между $-1/2$ и $1/2$ равновероятны. Поэтому вероятностная функция плотности $p(e)$ имеет вид

$$p(e) = \begin{cases} 1, & -1/2 \leq e < 1/2, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Вычислим μ_e и σ_e^2 , т. е. среднее и дисперсию величины $e(i)$:

$$\mu_e = E[e] = \int_{-\infty}^{\infty} ep(e) de = 0,$$

$$\sigma_e^2 = E[(e - \mu_e)]^2 = \int_{-\infty}^{\infty} (e - \mu_e)^2 p(e) de = \int_{-1/2}^{1/2} e^2 de = \frac{e^3}{3} \Big|_{-1/2}^{1/2} = 1/12.$$

Размерностями среднего, равного 0, и дисперсии, равной $1/12$, будут линейные и квадратные единицы соответственно. Для того чтобы перевести дисперсию в физические единицы, нужно определить соответствующий коэффициент. Если один отсчет соответствует p физическим единицам, то дисперсия станет равной $p^2/12$. Например, если промежуток отсчетов $0-1023$ соответствует интервалу $0-5$ В, то каждому отсчету соответствует $5/1023 = 4.8876 \cdot 10^{-3}$ В/отсчет, а дисперсия будет равна $1.9907 \cdot 10^{-6}$ В². Как правило, полный диапазон цифрового преобразователя является степенью 2. Если всему диапазону цифрового преобразователя соответствует E физических единиц, то

$$\sigma_x^2 = \left(\frac{E}{2^n}\right)^2 \frac{1}{12} = \frac{E^2 2^{-2n}}{12},$$

где n — число цифр в слове цифрового преобразователя. Отметим, что эта ошибка экспоненциально убывает с ростом n . С каждым дополнительным битом в слове цифрового преобразователя эта

ошибка убывает в четыре раза (или на 6 децибел). Таким образом, теоретически шум можно сделать произвольно малым, увеличивая n . Этот результат детально рассмотрен Оливером, Пирсом и Шенноном (1948).

На практике стоимость цифрового преобразователя зависит среди прочего от числа битов. Крайние значения этого числа суть 6 и 18, но в большинстве работ по временным рядам применяются цифровые преобразователи с 10 и 12 битами. Шестибитовые преобразователи используются в устройствах, работающих с частотами порядка мегагерц, например в устройствах обработки радарных сигналов, где 18-битовые преобразователи оказались бы, пожалуй, слишком медленными. Восемнадцать битов соответствуют точности примерно в $5^{1/2}$ десятичных цифр. Устройства с таким уровнем очень трудно изготавливать и эксплуатировать. Довольно часто то же говорят и о 15-битовых цифровых преобразователях.

Таблица 3.1.

Диапазон отсчетов и число десятичных цифр
для нескольких двоичных цифровых преобразователей

Число битов	Диапазон отсчетов	Число десятичных цифр
6	63	1.8
8	255	2.4
10	1 023	3.0
12	4 095	3.6
15	32 767	4.5
18	262 143	5.4

В табл. 3.1 приводятся диапазоны отсчетов и числа десятичных цифр для шести цифровых преобразователей, относящихся к рассмотренному типу со степенью двух.

Довольно просто можно вычислить плотность спектра мощности шума (погрешности). Поскольку предполагается, что шум некоррелирован, он будет иметь плоский спектр. В частности, произведение плотности спектра мощности на ширину интервала от 0 до частоты Найквиста должно быть равно дисперсии. Это следует из формулы

$$\sigma_x^2 = 2 \int_0^F S_x(f) df = 2S_x(f) F,$$

если предположить, что функция $S_x(f)$ постоянна. Значит,

$$S_x(f) = \frac{\sigma_x^2}{2F} = T\sigma_x^2.$$

Для диапазона цифрового преобразователя, соответствующего E физическим единицам, плотность спектра мощности шума, преобразованного в цифровую форму, имеет вид

$$S_e(f) = \frac{E^2 2^{-2nT}}{12}.$$

Интересно сравнить этот случай с теоретической плотностью спектра мощности синусоиды. Можно показать, что синусоида максимальной амплитуды имеет пик в точке $A^2 P/4$. В данном случае $A = E/2$, поэтому амплитуда пика станет равной $E^2 P/16$. Отношение пика ПСМ синусоиды к среднему значению ПСМ шума равно

$$\left(\frac{E^2 P}{16}\right) / \left(\frac{E^2 2^{-2nT}}{12}\right) = \frac{3}{4} \frac{P 2^{2n}}{T}.$$

Поскольку $P = NT$, получаем величину $3/4 N 2^{2n}$.

После подстановки сюда часто встречающихся значений $N = 1024$ и $n = 10$ и перехода к децибелам получается примерно 90 дБ. Но это приближенное значение. Вычислительный шум, обусловленный округлениями, и другие факторы приводят к снижению этого значения до 60 дБ, что является типичным результатом, получаемым цифровыми спектроанализаторами с приведенными выше параметрами. Это один из методов определения *динамического диапазона* анализаторов плотности спектра мощности.

3.5. ДРУГИЕ ПОГРЕШНОСТИ

Перечислим наиболее важные проблемы, связанные с применением систем сбора данных. Среди них назовем в первую очередь проблемы посторонних электрических шумов, вызванных внешними условиями, а также шумов, обусловленных применением усилителей; проблемы, связанные с искажениями, имеющими причиной слишком высокие уровни сигналов, и чрезмерный уровень шума, возникающий вследствие слишком низкого уровня сигналов. Отметим также наложение высокочастотной информации в тех случаях, когда или не применяются противоподменные фильтры, или применяются низкокачественные фильтры. Кроме того, в устройствах выборки возникают щелевые ошибки и дрожание. Наконец, укажем на нелинейность и пропадание знаков в цифровом преобразователе, шум в цифровом преобразователе и ошибки операций.

Чтобы избежать этих проблем или уменьшить их значение, следует обратить внимание на три важных вопроса.

1. Прежде чем заказывать оборудование, все нужно взвесить с точки зрения конечного результата: какова природа решаемой задачи? Каковы параметры, которые нужно вычислять? Насколько точными должны быть их значения, чтобы получились исчерпы-

вающие ответы? Только после решения этих вопросов можно уточнить состав оборудования. На практике не всегда можно ответить на все эти вопросы, но экспериментатор должен задавать их себе, иначе он обнаружит, что характер решения определяется не поиском решения, а оборудованием.

2. Система должна быть как можно более автоматизированной. Как правило, система сбора данных управляется ЭВМ. Последняя должна комплектоваться и программироваться так, чтобы быть подчиненной эксперименту. Например, ЭВМ может выдавать такие ключевые параметры, как скорость выборки и калибровочные множители. Если их не регистрировать независимо, то станет весьма вероятным появление ошибки или при задании системе параметров, или при записи их значений.

3. Систему необходимо сделать самоконтролирующей и периодически проводить контрольные проверки. В решающих экспериментах они должны проводиться как до, так и после прогона теста. Сравнительно просто к системе можно присоединить цифро-аналоговый преобразователь (ЦАП). Имея один или несколько таких преобразователей, можно прогнать такой круговой тест: получить с помощью ЭВМ системы некоторые цифровые данные и преобразовать их в аналоговую форму посредством цифро-аналоговых преобразователей; затем выход преобразователей ввести в систему как данные, выбрать их и преобразовать в цифровую форму; после этого сравнить начальную функцию и конечный результат. Такого типа круговые проверки могут выявить нелинейности, некоторые виды исчезновения знаков, нарушения в наборе коэффициентов цифрового преобразования и некоторые другие особенности. Такую процедуру настоятельно рекомендуется проводить всегда, когда для этого есть возможность.

Щелевая ошибка. Упомянутая выше *щелевая ошибка* возникает вследствие того, что взятие замеров непрерывной функции производится не с помощью дельта-функции, а при помощи функции с конечным носителем, подобной той, которая встречалась в § 3.2. С этого момента будем предполагать, что в результате операции с прямоугольной щелью, примененной к $x(t)$, получается функция $y(t)$:

$$y(t) = \frac{1}{a} \int_{t-a/2}^{t+a/2} x(\tau) d\tau.$$

Эта функция есть свертка $x(t)$ с прямоугольным импульсом

$$u^*(t) = \begin{cases} a^{-1}, & -\frac{a}{2} \leq t < \frac{a}{2}, \\ 0 & \text{при остальных значениях } t. \end{cases}$$

Звездочка здесь поставлена для того, чтобы подчеркнуть отличие этой функции от стандартного прямоугольного импульса, приведенного в гл. 1. В частотной области преобразование Фурье $Y(f)$ равно произведению $X(f)$ и $U^*(f)$:

$$Y(f) = X(f) U^*(f),$$

где функция $U^*(f)$ имеет вид

$$U^*(f) = \frac{\sin(\omega a/2)}{\omega a/2}.$$

Предположим, что $a = T$, т. е. ширина щели в точности совпадает с интервалом выборки, используемым при цифровом преобразовании. Получающаяся в результате функция изображена на рис. 3.4.

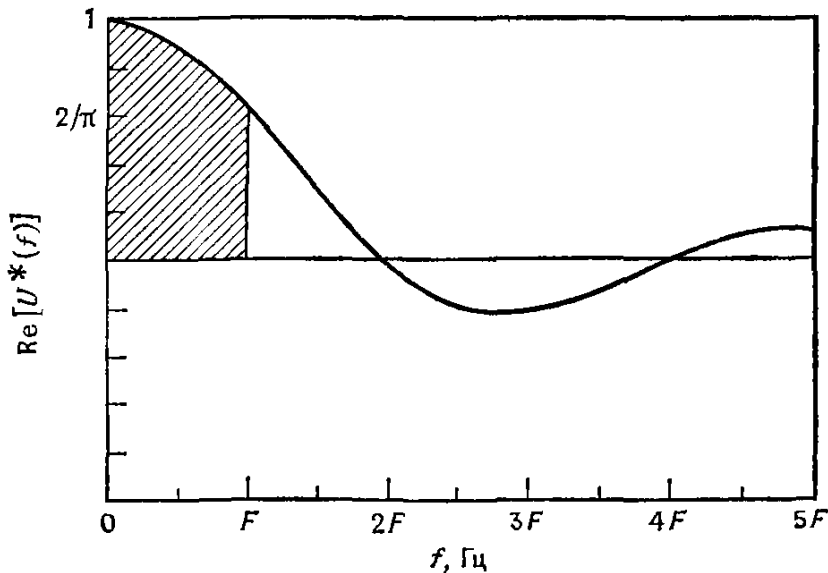


Рис. 3.4. Щелевой прямоугольный импульс при $a = T$.

Как видно из рисунка, на основном интервале $0 - F$ Гц происходит некоторое затухание. Значение $U^*(f)$ уменьшается от единицы в нуле до $2/\pi$ в частоте Найквиста. Поскольку $U^*(f)$ есть по существу передаточная функция щелевой операции, понятно, что в этом случае происходит дальнейшее затухание информации более высоких частот.

Такое положение может смягчаться уменьшением ширины щели a . Следует также помнить, что прямоугольная щелевая функция есть только идеализация настоящей функции. Углы такой функции, вероятно, закруглены, и ее вершина приподнимается.

Дрожание — это процесс, в силу которого выборки происходят не в точно заданные моменты, как это должно быть. Иначе говоря, выборки происходят не в моменты $t = iT$, а в моменты $iT + J(t)$, где $J(t)$ — случайный процесс. В качестве примера можно было

бы рассмотреть случайную функцию с распределением

$$p[J(t)] = \begin{cases} 1/T, & -T/2 \leq J(t) < T/2, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Другими словами, предполагается, что истинное время выборки распределено равномерно в интервале длины T с центром в точке iT .

Анализ дрожания представляет собой довольно сложную задачу. Здесь обычно ссылаются на работу Балакришнана (1962), который характеризовал дрожание следующим образом. Определим $y(i)$ как $x(i)$ с дрожанием:

$$y(i) = x[i + J(i)].$$

Балакришнан показал, что односторонняя плотность спектра мощности $y(i)$, обозначенная $G_y(f)$, равна

$$G_y(f) = |C(f)|^2 G_x(f) + \frac{R_x(0) - a^2}{F},$$

где $G_x(f)$ — односторонняя ПСМ функции $x(i)$ ($= 2S_x(f)$), $C(f)$ — характеристическая функция распределения величины $J(i)$, $R_x(0)$ — значение в нуле корреляционной функции $x(i)$,

$$a^2 = \int_0^F |C(f)|^2 G_x(f) df.$$

При этом предполагается, что для $n \neq m$

$$E[J(n)J(m)] = 0,$$

т. е. дрожание не коррелирует с самим собой. Член a^2 предназначен для того, чтобы полная мощность оставалась той же самой. Полная мощность $[R_y(0)]$, получаемая интегрированием $G_y(f)$, окажется равной $R_x(0)$, и, стало быть, дрожание на нее не повлияет.

Функция $C(f)$, определенная как характеристическая функция вероятностной функции распределения, имеет вид

$$C(f) = \int_{-\infty}^{\infty} p[J] \exp(j2\pi fJ) dJ.$$

Балакришнан рассмотрел три вида распределений $J(i)$:

1. Дрожание с гауссовским распределением, имеющим дисперсию σ^2 . В этом случае

$$|C(f)|^2 = \exp(4\pi^2 f^2 \sigma^2).$$

2. Дрожание с равномерным распределением

$$p[J(t)] = \begin{cases} (\gamma T)^{-1}, & \text{если } |J(t)| < \gamma T/2, \\ 0, & \text{если } |J(t)| \geq \gamma T/2. \end{cases}$$

В этом случае

$$|C(f)|^2 = \left(\frac{\sin(\pi\gamma f T)}{\pi\gamma f T} \right)^2.$$

3. Дрожание с распределением

$$p[J(i)] = \begin{cases} P & \text{при } J(i) = -\gamma T, \\ 1-P & \text{при } J(i) = \gamma T. \end{cases}$$

В этом случае

$$C(f) = (1-P)^2 + P^2 + 2P(1-P) \cos(4\pi\gamma f T).$$

Наиболее важны два первых случая. Для каждого из них характерно следующее. Во-первых, поведение $|C(f)|^2$ при возрастании f приводит к затуханию $G_x(f)$. Значит, ошибка увеличивается с ростом значений частот. Во-вторых, к процессу добавляется ровный белый шум, которому отвечает слагаемое $(R_x(0) - a^2)/F$.

Балакришнан отмечал, что сделать какой-то вывод о затухании с ростом частот в третьем случае невозможно. Однако в большинстве практических ситуаций хорошими моделями оказываются случаи 1 и 2.

3.6. ПЕРЕХОД К ФИЗИЧЕСКИМ ЕДИНИЦАМ

В этом параграфе мы остановимся на *шаговой* и *синусоидальной калибровках* — широко распространенных процедурах перевода в технические единицы.

В обоих случаях предлагается перед записью данных исключить из системы датчик. Это связано с тем, что датчик, как правило, не может обеспечить эталонной точности.

При шаговой калибровке используется набор уровней напряжений, называемых шагами. Например, для датчика, напряжения которого изменяются в диапазоне от 0 до 5 В, такими уровнями могут служить 0.0, 2.5 и 5.0 В¹⁾. Напряжение каждого уровня подается в течение установленного промежутка времени, равного, например, одной секунде, и преобразуется в цифровую форму. Для удобства последующих вычислений результаты следует располагать так, чтобы их нельзя было спутать. Если используется магнитный носитель, то этого можно достичь, например, записы-

¹⁾ В этой процедуре датчик нередко заменяют набором калибровочных сопротивлений, по одному на каждый шаг.

вая результаты на магнитную ленту отдельными файлами. Кроме того, можно начинать и прекращать запись результатов для каждого шага при помощи *прерываний* работы устройства, подающего соответствующие напряжения.

Сначала в ЭВМ вычисляются средние значения результатов, полученных для каждого уровня. Получающиеся по одному для каждого уровня средние числа обозначим a_1, \dots, a_N (предполагается, что всего используется N шагов).

После этого в ЭВМ вводятся величины в физических единицах для каждого уровня, по одной для каждого числа a_k . Эти величины обозначим p_1, \dots, p_N .

Допустим теперь, что в систему вводятся данные для обработки, которые поступают с цифрового преобразователя в отсчетах. Они составляют последовательность $c(i)$, рассмотренную ранее. С учетом всей полученной выше информации эта последовательность преобразуется в последовательность калиброванных данных $x(i)$.

Для этой цели обычно используется линейная интерполяция, сводящаяся к следующим действиям. Сначала определяется такое K , что $a_K \leq c(i) < a_{K+1}$, а затем по формуле

$$x(i) = p_K + (p_{K+1} - p_K) \frac{c(i) - a_K}{a_{K+1} - a_K}$$

получают соответствующее $c(i)$ калиброванное значение. При этом для крайних точек вычисления производят по таким же формулам, полагая $K=1$, если $c(i) < a_1$, и $K=N-1$, если $c(i) > a_N$. В последних случаях интерполяционная формула будет выступать в качестве экстраполяции.

При синусоидальной калибровке в систему вводится синусоидальный сигнал с известной амплитудой и на большом числе периодов производится запись и преобразование в цифровую форму. При этом определяется среднее значение

$$\bar{d} = \frac{1}{N} \sum_{i=1}^N d(i)$$

последовательности $d(i)$ чисел, полученных при преобразовании синусоиды в цифровую форму. Помимо этого, вычисляется выборочная дисперсия по приводившейся ранее формуле

$$s_d^2 = \frac{1}{N-1} \sum_{i=1}^N (d(i) - \bar{d})^2.$$

По этим величинам определяют три точки:

$$a_3 = \bar{d} + \sqrt{2}s_d, \quad a_2 = \bar{d}, \quad a_1 = \bar{d} - \sqrt{2}s_d,$$

которые отвечают положительному пику, нулю и отрицательному пику синусоиды соответственно (все эти величины измерены в отсчетах). Затем вводятся соответствующие физические единицы и применяются приведенные выше формулы для интерполяции и экстраполяции.

Если бы не было шума и использовалось целое число периодов синусоиды, то эта процедура была бы точной. Однако избежать шума невозможно. Нельзя также начать и остановиться в одной и той же точке периода синусоиды. Поэтому для того, чтобы не получить ошибок или по крайней мере свести их к минимуму, берется большое число периодов.

3.7. РЕДАКТИРОВАНИЕ ДАННЫХ С НЕПРАВДОПОДОБНЫМИ ЗНАЧЕНИЯМИ

Большинство систем сбора информации вносит в данные ложные значения. Это может происходить по многим причинам — из-за потери сигналов в линиях связи, сбоев в цифровом преобразователе, сбоев в устройстве задания форматов. Неправдоподобные значения, возникающие в результате этих сбоев, могут вызвать значительные трудности при последующем анализе. Одно неправдоподобное значение, равное максимально представимой в цифровом преобразователе величине, может повысить предельный уровень шума и, значит, привести к значительному изменению плотности спектра мощности. Близко расположенная пара таких значений может дать несколько ложных частот в ПСМ.

По этим причинам лучше всего с помощью схемы предварительного преобразования предельных данных организовать поиск и удаление неправдоподобных значений. К сожалению, трудно определить точно, какие данные неправдоподобны. Общих процедур для автоматического удаления таких данных нет.

Поэтому разумно запастись программами редактирования или, если это необходимо, средствами ручной коррекции. Все машинные программы, предназначенные для этих целей, действуют следующим образом. Сначала для поиска неправдоподобных значений просматриваются все данные. Если такие значения обнаружены, то они выводятся на печать в виде таблицы или графика. После просмотра этих распечаток исследователь вводит очередную программу, чтобы удалить плохие значения, заменить их или вовсе не использовать. Последнее лучше всего, если, правда, это можно сделать.

Для устранения неправдоподобных значений предложено несколько схем автоматического редактирования, но ни одна из них не является полностью удовлетворительной. Приведем две такие схемы, чтобы пояснить, какого рода операции в этом случае производятся.

В первой из этих схем, представленной на рис. 3.5, используются два цифровых RC-фильтра для сглаживания. Эти фильтры подробно рассмотрены в гл. 4. Они производят низкочастотную фильтрацию исходной функции и вырабатывают сглаженные оценки. При рассмотрении обеих схем неявно предполагается, что нужные данные имеют «плавный» характер, а неправдоподобные значения — «резкий». Определим две последовательности $[\bar{x}(i)]^2$ и $\overline{x^2(i)}$ как последовательности сглаженных данных с последующим

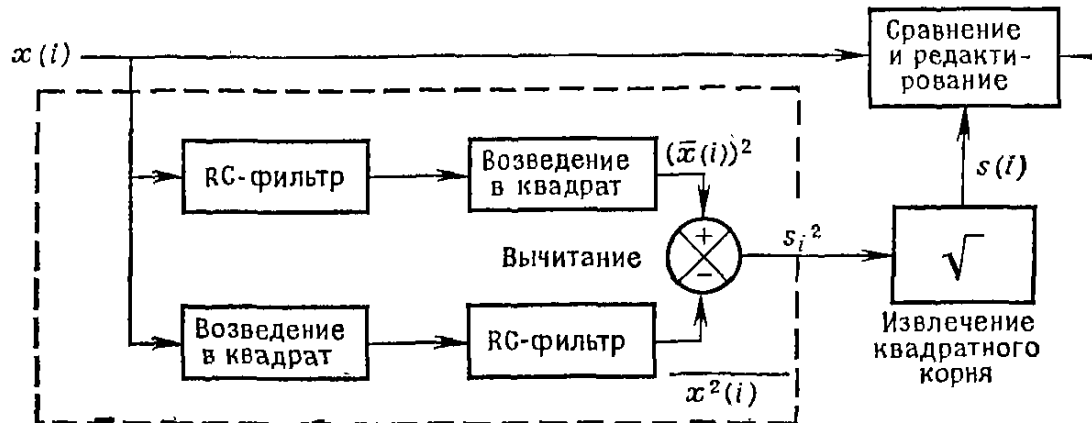


Рис. 3.5. Схема редактирования данных с неправдоподобными значениями.

возведением их в квадраты и сглаженных данных, предварительно возведенных в квадраты, соответственно.

Задача той части схемы, которая на рисунке ограничена пунктирной линией, состоит в выработке постоянно обновляемого значения выборочной дисперсии $s^2(i)$. Это значение определяется по формуле

$$s^2(i) = \overline{x^2(i)} - [\bar{x}(i)]^2.$$

Извлекая из него квадратный корень, получают стандартное отклонение.

На следующем шаге проверяют очередное значение ряда $x(i+1)$. Это значение считается хорошим, если

$$\bar{x}(i) - ks(i) < x(i+1) < \bar{x}(i) + ks(i). \quad (3.2)$$

Подходящий для данных параметр k выбирается самим исследователем. Как правило, он колеблется в пределах от 3 до 9, но для начала лучше выбирать его равным 6.

Плохие значения $x(i+1)$ могут заменяться значениями $\hat{x}(i+1)$, вычисленными по формуле

$$\hat{x}(i+1) = 2x(i) - x(i-1),$$

т. е. полученными, по сути дела, линейной экстраполяцией.

Эта процедура требует некоторых дополнительных уточнений. Для числа последовательных экстраполяций следует с самого начала установить предел, чтобы избежать постоянной экстраполяции. Дело в том, что последовательная экстраполяция для ряда плохих точек может достаточно далеко увести полосу (3.2), и в результате в нее не попадут те хорошие данные, которые в конце концов появятся.

Вторую схему редактирования данных с неправдоподобными значениями, частично принадлежащую Дж. В. Тьюки, называют процедурой «Тьюки 53X». Основная идея этой процедуры состоит в том, чтобы получать оценку гладкого куска кривой, которую затем вычитают из данных; после удаления гладкой составляющей или тренда выделить неправдоподобные значения гораздо легче.

В этой процедуре используется тот факт, что *медиана* есть *робастная* оценка среднего. Если все данные упорядочить в порядке возрастания, то медианой будет просто то значение, которое в этом ряде данных займет место посередине. Например, оценку среднего роста 101 студента можно получить следующим образом: отперфорировать рост каждого студента на 80-колоночных перфокартах, ввести эти карты в ЭВМ и определить среднее и медиану; среднее вычисляется обычным способом, а медиана — при помощи упорядочения результатов в порядке возрастания и выбора значения, оказавшегося посередине. Если при перфорации не было серьезных ошибок, то эти значения скорее всего окажутся довольно близкими. Посмотрим теперь, что произойдет, если один из студентов отперфорировывает на карте свой рост равным 100 футам. В этом случае среднее возрастет примерно на 1 фут, в то время как медиана останется практически неизменной.

При вычислении оценки гладкой составляющей в процедуре Тьюки 53X это свойство робастности медианы используется дважды. Сама процедура выглядит следующим образом:

1. По последовательности $x(i)$ строится новая последовательность $x'(i)$. Сначала определяется медиана данных $x(1), \dots, x(5)$. Это значение становится в новой последовательности элементом $x'(3)$. Теперь из пяти первых данных убирается $x(1)$ и добавляется $x(6)$. Медиана этого нового множества есть $x'(4)$. Так продолжается до тех пор, пока все данные не будут исчерпаны. По сравнению с последовательностью $x(i)$ новая последовательность будет короче на четыре элемента. Медиана всегда выбирается из группы пяти последовательных значений.

2. Почти таким же способом по последовательности $x'(i)$ строится последовательность $x''(i)$. Единственное различие в том, что теперь медиана выбирается из *троек* последовательных значений.

3. На последнем этапе строится последовательность $x'''(i)$ по правилу

$$x'''(i) = 1/4 x''(i-1) + 1/2 x''(i) + 1/4 x''(i+1),$$

отвечающему сглаживающему фильтру Хэннинга.

После этого, как и раньше, для элементов последовательности $x(i) - x'''(i)$ проверяется условие $|x(i) - x'''(i)| > k$, где k — заранее выбранное число. Если оно выполняется, то соответствующую величину заменяют интерполированным значением.

3.8. УДАЛЕНИЕ ТРЕНДА

Иногда из некоторых временных рядов нужно удалить линейный или медленно меняющийся тренд. Такого рода тренды наблюдаются в рядах, например, при суммировании одной или нескольких компонент, приводящем к ошибкам двух типов. Во-первых, при неправильной калибровке нулевой точки в каждый момент отбора данных будет возникать небольшая ошибка. После суммирования эта постоянная величина даст прямую. Такой линейный тренд может привести к большим ошибкам при определении плотности спектра мощности и в связанных с этим вычислениях.

Ошибка второго типа возникает из-за возрастания в процессе суммирования мощности, соответствующей низкочастотному шуму. Как правило, такой шум в данных всегда есть. При суммировании он обретает форму случайного, но медленно меняющегося тренда. Насколько быстро меняется такой тренд, до некоторой степени зависит от интервала квантования.

Наилучшим способом удаления тренда служит применение высокочастотных фильтров, которые будут предметом обсуждения гл. 4. Полиномиальный тренд можно удалять методом наименьших квадратов. Это делается следующим образом. Как обычно, предположим, что $x(i)$ ($i = 0, \dots, N-1$) — выборка функции на постоянном интервале T и что требуется подобрать многочлен вида

$$\hat{x}(i) = \sum_{k=0}^K (iT)^k c_k,$$

в котором $\hat{x}(i)$ — оцениваемые значения многочлена, присутствующего в ряде $x(i)$ в виде тренда. Для этой цели применяют стандартный способ, при котором определяется функция от коэффициентов многочлена

$$E(c) = \sum_{i=0}^{N-1} \left(x(i) - \sum_{k=0}^K (iT)^k c_k \right)^2, \quad (3.3)$$

называемая погрешностью. Ее можно минимизировать, применив стандартный прием, а именно приравняв все частные производные

по c_k нулю и решив относительно них систему уравнений

$$\frac{\partial E}{\partial c_j} = \sum_{i=0}^{N-1} 2 \left[x(i) - \sum_{k=0}^K (iT)^k c_k \right] [-(iT)^j] = 0, \quad j=0, \dots, K.$$

Эту систему можно переписать в виде

$$\sum_{k=0}^K c_k \sum_{i=0}^{N-1} (iT)^{k+j} = \sum_{i=0}^{N-1} (iT)^j x(i), \quad j=0, \dots, K.$$

Решить ее относительно c при больших значениях K трудно. К счастью, довольно редко приходится решать эту систему при K , большем 3 или 4. Если требуется удаление многочленов только низких порядков, то решение соответствующей системы методом обратной¹⁾ матрицы можно свести к непосредственному вычислению коэффициентов с использованием памяти ЭВМ. Например, таким решением в случае $K=0$ будет коэффициент $c_0 = N^{-1} \sum_{i=0}^{N-1} x(i)$.

В случае $K=1$ решение дают формулы

$$c_0 = \frac{2(2N-1) \sum_{i=0}^{N-1} x(i) - 6 \sum_{i=0}^{N-1} ix(i)}{N(N+1)},$$

$$c_1 = \frac{12 \sum_{i=0}^{N-1} ix(i) - ((N-1)/2) \sum_{i=0}^{N-1} x(i)}{TN(N^2-1)}.$$

Вычисления значительно упрощаются, если N — нечетное число. Разумеется, и при четном N последнюю точку можно удалить. Оказывается, что при нечетном числе точек и интервале $(-N/2, N/2)$ изменения независимой переменной i многие члены обращаются при вычислениях в нуль, в частности все те, которые имеют противоположные значения. Это сокращает число необходимых вычислений и с арифметической точки зрения улучшает их²⁾.

Далее знак Σ будет обозначать суммирование от $-N/2 + 1/2$ до $N/2 - 1/2$, т. е.

$$\Sigma = \sum_{i=-N/2+1/2}^{N/2-1/2}.$$

¹⁾ Замечание. Почти всегда вместо вычисления обратной матрицы лучше применять другие методы решения систем.

²⁾ Обращаем внимание читателя на то, что N здесь уже другое. С этого момента мы сменили обычную индексацию от 0 до $N-1$. Теперь используется $N+1$ точек.

Для первого случая ($K = 1$)

$$d_0 = \frac{\sum x(t)}{N}, \quad d_1 = \frac{\sum tx(t)}{\sum t^2},$$

для $K = 2$

$$d_0 = \frac{\sum t^2 \sum t^2 x(t)}{(\sum t^2)^2 - N \sum t^4}, \quad d_1 = \frac{\sum tx(t)}{\sum t^2},$$

$$d_2 = \frac{\sum t^2 \sum x(t) - N \sum t^2 x(t)}{[(\sum t^2)^2 - N \sum t^4]},$$

для $K = 3$

$$d_0 = \frac{\sum x(t) \sum t^4 - \sum t^2 \sum t^2 x(t)}{N \sum t^4 - (\sum t^2)^2}, \quad d_1 = \frac{\sum t^4 \sum t^3 x(t) - \sum t^6 \sum tx(t)}{[(\sum t^4)^2 - \sum t^2 \sum t^6]},$$

$$d_2 = \frac{\sum t^2 \sum x(t) - N \sum t^2 x(t)}{[(\sum t^2)^2 - N \sum t^4]}, \quad d_3 = \frac{\sum t^4 \sum tx(t) - \sum t^2 \sum t^3 x(t)}{[(\sum t^4)^2 - \sum t^2 \sum t^6]}$$

и для $K = 4$ (сначала вычисляется d_4 , затем d_3 и т.д.)

$$d_0 = (\sum x(t) - d_2 \sum t^2 - d_4 \sum t^4) / N,$$

$$d_1 = \left\{ \sum tx(t) [\sum t^2 \sum t^6 - (\sum t^4)^2] - [\sum t^4 \sum t^2 \sum t^3 x(t) - \sum t^4 \sum tx(t)] \right\} / \left\{ \sum t^2 [\sum t^2 \sum t^6 - (\sum t^4)^2] \right\},$$

$$d_2 = [\sum t^2 \sum x(t) - N \sum t^2 x(t) - d_4 (\sum t^2 \sum t^4 - N \sum t^6)] / [(\sum t^2)^2 - N \sum t^4],$$

$$d_3 = [\sum t^2 \sum t^3 x(t) - \sum t^4 \sum tx(t)] / [\sum t^2 \sum t^6 - (\sum t^4)^2],$$

$$d_4 = \left\{ [N \sum t^4 x(t) - \sum t^4 \sum x(t)] [(\sum t^2)^2 - N \sum t^4] + \right.$$

$$\left. + [\sum t^2 \sum x(t) - N \sum t^2 x(t)] (\sum t^4 \sum t^2 - N \sum t^6) \right\} /$$

$$\left\{ (\sum t^2 \sum t^4 - N \sum t^6)^2 - [(\sum t^4)^2 - N \sum t^8] [(\sum t^2)^2 - N \sum t^4] \right\},$$

где $(N + 1)$ — число равноотстоящих друг от друга выборок $x(t)$ (N может быть четным),

$$\sum t^2 = \frac{N(N^2 - 1)}{12}, \quad \sum t^4 = \frac{N(N^2 - 1)(2N - 7)}{240},$$

$$\sum t^6 = \frac{N(N^2 - 1)(3N^4 - 18N^2 + 31)}{1344},$$

$$\sum t^8 = \frac{N(N^2 - 1)(5N^6 - 55N^4 + 239N^2 - 381)}{11520}, \quad t = \frac{N}{2}, \quad \cdot 2$$

Нужный полином имеет вид

$$\hat{x}(t) = \sum_{k=0}^K t^k d_k, \quad t = -\frac{N}{2}, \dots, \frac{N}{2}.$$

Отметим, что все вычисления проводятся с целыми числами; время T нигде не появляется. Работа в целочисленном формате также увеличивает точность

Вычисление коэффициентов происходит в два этапа

1 Вычисление моментов $\Sigma t^k x(t)$, где $k=0, 1, \dots, K$

2 Вычисление коэффициентов d_0, \dots, d_K по приведенным выше формулам

Для этих вычислений рекомендуется использовать двойную точность. В 32-битовых ЭВМ это следует делать обязательно

Хотя вычисление и использование коэффициентов d_0, \dots, d_K удобно, сами по себе они играют незначительную роль. Многие исследователи предпочитают работать с коэффициентами c_0, \dots, c_K . Одной из причин этого служит тот факт, что d_0 не является средним значением, за исключением случаев $K=0$ или 1

Переход от коэффициентов d к коэффициентам c производится следующим образом. Сначала определяется величина M . $M = -NT/2$. После этого t заменяют величиной $[(M + iT)/T]$, где снова $t=0, \dots, N$. Теперь формула для $\hat{x}(t)$ принимает вид

$$\hat{x}(t) = \sum_{k=0}^K \left(\frac{M + iT}{T} \right)^k d_k, \quad t = 0, \dots, N.$$

Можно расписать ее для случая $K=4$:

$$d_0 + \left(\frac{M + iT}{T} \right) d_1 + \left(\frac{M + iT}{T} \right)^2 d_2 + \left(\frac{M + iT}{T} \right)^3 d_3 + \left(\frac{M + iT}{T} \right)^4 d_4.$$

Раскрывая скобки и собирая члены при одинаковых степенях $(iT)^k$, получаем

$$\begin{aligned} \hat{x}(t) = & \left(d_0 + \frac{Md_1}{T} + \frac{M^2d_2}{T^2} + \frac{M^3d_3}{T^3} + \frac{M^4d_4}{T^4} \right) + \\ & + iT \left(\frac{d_1}{T} + \frac{2Md_2}{T^2} + \frac{3M^2d_3}{T^3} + \frac{4M^3d_4}{T^4} \right) + \\ & + (iT)^2 \left(\frac{d_2}{T^2} + \frac{3Md_3}{T^3} + \frac{6M^2d_4}{T^4} \right) + \\ & + (iT)^3 \left(\frac{d_3}{T^3} + \frac{4Md_4}{T^4} + (iT^4) \frac{d_4}{T^4} \right). \end{aligned}$$

Приравнивая полученные коэффициенты коэффициентам c_k , имеем

$$c_0 = d_0 + \frac{Md_1}{T} + \frac{M^2d_2}{T^2} + \frac{M^3d_3}{T^3} + \frac{M^4d_4}{T^4},$$

$$c_1 = \frac{d_1}{T} + \frac{2Md_2}{T^2} + \frac{3M^2d_3}{T^3} + \frac{4M^3d_4}{T^4}, \quad c_2 = \frac{d_2}{T^2} + \frac{3Md_3}{T^3} + \frac{6M^2d_4}{T^4},$$

$$c_3 = \frac{d_3}{T^3} + \frac{4Md_4}{T^4}, \quad c_4 = \frac{d_4}{T^4}.$$

По этим же формулам вычисляются коэффициенты и при меньших K . В этих случаях соответствующие коэффициенты d_k полагают равными 0.

Подгонка и удаление среднего. Подгонка и удаление среднего — самый частный случай полиномиальной фильтрации и удаления полиномиального тренда. Обычно эти операции объединяют с вычислениями корреляционной функции и плотностей спектра мощности. Само вычисление выглядит до смешного просто:

$$c_0 = m = \frac{1}{N} \sum_{i=0}^{N-1} x(i).$$

Довольно легко понять, что использования в некоторых машинах вычислений с обычной точностью для этой цели может оказаться недостаточным. Вспоминается один пример из приложений, в котором на 32-битовой машине с одинарной точностью в режиме плавающей точки суммировалось около 14 000 величин со средним, равным приблизительно 100. Сработала программа антипереполнения, и с некоторого момента вычисления суммы все новые данные были просто потеряны.

Поскольку по сравнению с вычислением ПСМ с одинарной точностью стоимость вычисления среднего с двойной точностью мала, для среднего рекомендуется программно задавать двойную точность. Исключение составляют только случаи с малыми значениями N равными, например 1024. Для них вполне достаточно и одинарной точности.

Упражнения

3.1. Смоделируйте квантизацию синусоиды 10-битовым цифровым преобразователем при помощи программы

```

DIMENSION E(1000)
:
:
C   ВВОД PHI И FACT
:
:
DO 10 I=1, 1000
ANG=FLOAT(I-1)*FACT+PHI
CT=511.5*(SIN(ANG)+1.)
CI=FLOAT(IFIX(CT+0.49999))
10  E(I)=CT-CI

```

- а) Опишите характеристики величины СТ.
 б) То же самое сделайте для СИ.
 в) Выполните программу, дополнив ее вычислениями среднего и дисперсии для величины Е, а также вычислением коэффициентов корреляции для соседних величин Е при следующих значениях FACT и PHI:

FACT	PHI
0.0015708	0.0000
3.0753802	0.5218
1.5826000	0.7538

Объясните полученные ответы.

3.2. Напишите программу, моделирующую дрожание при преобразовании синусоиды в цифровой формат. Рекомендуется следующее:

- а) Получить последовательность $x(i) = \sin(2\pi f_0 i T)$, выбирая $T = 1$ и $f_0 = 0.5/8$. В этом случае f_0 составит $1/8F$.
 б) Смоделировать дрожание, выбирая новое значение iT равным

$$(iT)_{\text{новое}} = (iT)_{\text{исходное}} + q [u(i) - 0.5],$$

где $u(i)$ — случайный шум с равномерным распределением, полученным с помощью процедуры TDRAND, а q , $0 \leq q \leq 1$, — масштабный множитель. Случаю $q = 0$ соответствует отсутствие дрожания.

в) Получить последовательность длины 64 и вычислить преобразование Фурье при помощи процедуры FFTRAN. Не забудьте перед этим действительным частям присвоить значения $x(i)$, а мнимым 0.

г) Просчитать модель при $q = 0, 0.5$ и 1. Прокомментировать результаты.

3.3. При синусоидальной калибровке результаты получаются точными, если используется целое число циклов (ошибки преобразования в цифровую форму при этом не учитываются). Вычислите верхние границы для максимальной ошибки \bar{d} и \bar{s}_d^2 как функции N , T и частоты калибровочной синусоиды. Как ведут себя ошибки с ростом N ?

3.4. В этой задаче проверяется арифметика ЭВМ на антипереполнение. Напишите тестовую программу, которая должна выполнять следующую последовательность действий:

- а) Вычислить $x(i) = 99.5 + u(i)$, где $u(i)$ — случайное число, полученное с помощью процедуры TDRAND. Чему равно математическое ожидание отдельной реализации этой последовательности?
 б) Вычислить для $k = 1, 2, \dots, 12$

$$\bar{x}_k = \frac{1}{N_k} \sum_{i=1}^{N_k} x(i),$$

где $N_k = [10^{(k+1)/2}]$.

- в) Чему равны математическое ожидание и дисперсия \bar{x}_k как функции N_k ?
 г) Приведите график \bar{x}_k как функции N_k и объясните результат.
 е) Если позволяет машинное время, еще раз прогоните программу, но суммирование производите с двойной точностью.
 3.5. Докажите (3.3).

Глава 4

КОНСТРУКЦИИ ЦИФРОВЫХ ФИЛЬТРОВ

4.1. ОСНОВНЫЕ ПОНЯТИЯ

В этой главе рассмотрены конструкции цифровых фильтров и ряд характерных приемов, которыми реализуются фильтры наиболее общих типов.

Прежде чем перейти к обсуждению фильтров более высоких порядков (полоснопропускающих, заграждающих, фильтров высоких и низких частот), подробно рассмотрим главные элементы их конструкций — импульсные фильтры с бесконечными областями отклика (ИБО) первого и второго порядка.

Войти в круг вопросов цифровой фильтрации поможет введенный в гл. 1 фильтр с непрерывной областью, определенный дифференциальным уравнением

$$\ddot{y}(t) + 2\zeta\omega_n\dot{y}(t) + \omega_n^2 y(t) = x(t), \quad (4.1)$$

где $x(t)$ — функция на входе, а $y(t)$ — выход фильтра. Как правило, такой фильтр предназначается для *полосной* фильтрации, причем центр полосы располагается около ω_n рад или $f_n = \omega_n/(2\pi)$ Гц.

Предположим, что заданы нулевые начальные условия. Тогда применение преобразования Фурье приводит к уравнению

$$-\omega^2 Y(f) + 2\zeta\omega_n j\omega Y(f) + \omega_n^2 Y(f) = X(f),$$

где $\omega = 2\pi f$. Значит, передаточная функция равна

$$\frac{Y(f)}{X(f)} = H(f) = \frac{1}{-\omega^2 + 2\zeta\omega_n j\omega + \omega_n^2}.$$

Передаточная функция представляет собой, вообще говоря, комплексную функцию от частоты. Как правило, эта независимая переменная принимает действительные значения. Исключение составляют случаи нулей и полюсов фильтра. Наиболее общим видом передаточной функции служит отношение

$$H(f) = \frac{N(f)}{D(f)}$$

взаимно простых многочленов $N(f)$ и $D(f)$ от f . Нулями и полюсами будут решения уравнений $N(f) = 0$ и $D(f) = 0$ соответственно.

Как подчеркивалось в гл. 1, нули и полюсы определяют дробно-рациональную функцию типа $H(f)$ с точностью до постоянного множителя.

В нашем примере нулей нет, а полюсами будут решения уравнения

$$\omega_n^2 \left[1 - \left(\frac{f}{f_n} \right)^2 + \frac{j2\zeta f}{f_n} \right] = 0.$$

Перепишав его в виде

$$f^2 - j2\zeta f_n f - f_n^2 = 0,$$

получим решения

$$\begin{aligned} f_1, f_2 &= \frac{j2\zeta f_n \pm \sqrt{(-j2\zeta f_n)^2 + 4f_n^2}}{2} = \\ &= f_n [j\zeta \pm \sqrt{1 - \zeta^2}] = \pm f_n \sqrt{1 - \zeta^2} + j\zeta f_n. \end{aligned}$$

Поэтому передаточную функцию в рассматриваемом примере можно переписать в виде

$$H(f) = - \frac{1}{(2\pi)^2 (f - f_1) (f - f_2)}.$$

В общем случае такой способ записи приводит к представлению передаточной функции вида

$$H(f) = c \frac{\prod_{l=1}^L (f - f_{zl})}{\prod_{k=1}^K (f - f_{pk})},$$

в котором f_{zl} — нули, а f_{pk} — полюсы.

Напомним также, что обратное преобразование Фурье $h(t)$ функции $H(f)$ связывает $y(t)$ и $x(t)$ посредством интегрального уравнения

$$y(t) = \int_{-\infty}^{\infty} h(\tau) x(t - \tau) d\tau, \quad (4.2)$$

в котором интеграл есть свертка функций $h(t)$ и $x(t)$.

Заслуживает внимания сопоставление соотношений (4.1) и (4.2). Первое представляет собой *дифференциальное* уравнение, связывающее $x(t)$ и $y(t)$, а последнее — определяющее ту же связь *интегральное* уравнение. Как будет видно из дальнейшего, подобная связь возникает и для цифровых фильтров. В частности, дифференциальным уравнениям будут соответствовать *разностные* уравнения, а свертке (4.2) — соответствующие бесконечные суммы.

Необходимо обратить внимание также на следующие важные обстоятельства. Во-первых, главной характеристикой типов фильтров является коэффициент усиления, т. е. функция $|H(f)|$. Во-вторых, передаточную функцию с точностью до постоянного множителя определяют нули и полюсы. Кроме того, с помощью полюсов устанавливается *устойчивость* фильтра, означающая, грубо говоря, что отклик на импульсную функцию ограничен сверху убывающей экспонентой.

4.2. ФИЛЬТРЫ ПЕРВОГО ПОРЯДКА

Наиболее общий вид цифрового фильтра первого порядка определяется разностным уравнением

$$y(i) = \sum_{k=0}^K b_k x(i-k) - a_1 y(i-1). \quad (4.3)$$

Текущий выход фильтра определяется линейной комбинацией K предшествовавших значений входа, текущим значением входа и одним предыдущим значением выхода. Этот фильтр и называют фильтром первого порядка потому, что для получения нового значения выхода из линейной комбинации значений входов вычитают только одно предыдущее значение выхода (умноженное на постоянный множитель). Вообще, порядок фильтра определяет число входящих в разностное уравнение значений предшествовавших выходов.

В табл. 4.1 приводятся коэффициенты для нескольких важных случаев. Как видно из этой таблицы, фильтр низких частот задается разностным уравнением

$$y(i) = (1 - \alpha) x(i) + \alpha y(i-1).$$

В качестве основных средств анализа фильтров принято использовать передаточную функцию, ее коэффициент усиления и фазу. Вычислить передаточную функцию можно несколькими способами. Приступая к работе с цифровыми фильтрами, имеет смысл для начала провести это вычисление кружным путем. Запишем определение преобразования Фурье

$$Y(f) = T \sum_{i=-\infty}^{\infty} y(i) \exp(-j2\pi f iT),$$

а затем заменим $y(i)$ правой частью предыдущего уравнения:

$$Y(f) = T \sum_{i=-\infty}^{\infty} [(1 - \alpha) x(i) + \alpha y(i-1)] \exp(-j2\pi f iT).$$

Таблица 4.1

Коэффициенты фильтров пяти типов, полученных из общего уравнения (4.3) для фильтра первого порядка

	b_0	b_1	a_1
Интегратор	T	0	-1
Дифференциатор ^{a)}	$\frac{1}{T}$	$-\frac{1}{T}$	0
Низкочастотный фильтр	$(1-\alpha)$	0	$-\alpha$
Фильтр высоких частот	$(1-\alpha)$	0	α
Фильтр Конелла ^{b)} (высоких частот)	$\left(1-\frac{\alpha}{2}\right)$	$-\left(1-\frac{\alpha}{2}\right)$	$-(1-\alpha)$
		$0 < \alpha < 1$	

^{a)} Пожалуй, не совсем честно называть дифференциатор фильтром первого порядка, поскольку он не имеет члена a_1

^{b)} Фильтр Конелла был изобретен Ноутессом (C. V. Notess).

Правую часть полученного уравнения разобьем на две суммы:

$$Y(f) = (1-\alpha) T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi f iT) + \\ + \alpha T \sum_{i=-\infty}^{\infty} y(i-1) \exp(-j2\pi f iT).$$

Первая из них представляет собой преобразование Фурье последовательности $x(i)$:

$$(1-\alpha) \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi f iT) = (1-\alpha) X(f).$$

Вторая сумма — почти преобразование Фурье. Чтобы показать это, сделаем замену индекса суммирования $l = i - 1$ и перепишем ее в виде

$$\alpha T \sum_{i=-\infty}^{\infty} y(i-1) \exp(-j2\pi f iT) = \\ = \alpha T \sum_{l=-\infty}^{\infty} y(l) \exp(-j2\pi f (l+1) T) = \\ = \alpha \exp(-j2\pi f T) T \sum_{l=-\infty}^{\infty} y(l) \exp(-j2\pi f l T) = \\ = \alpha \exp(-j2\pi f T) Y(f).$$

После этой замены можно вынести за знак суммирования экспоненту. В результате получится произведение этой экспоненты на преобразование Фурье последовательности $y(i)$; разумеется, совершенно безразлично, используется ли в определении преобразования $Y(f)$ переменная i или l .

Собирая теперь полученные результаты, приходим к соотношению

$$Y(f) = (1 - \alpha) X(f) + \alpha \exp(-j2\pi fT) Y(f),$$

из которого определяем отношение $Y(f)/X(f)$:

$$H(f) = \frac{Y(f)}{X(f)} = \frac{1 - \alpha}{1 - \alpha \exp(-j2\pi fT)}.$$

Проведенные вычисления можно значительно упростить, если воспользоваться теоремой из гл. 1, в которой утверждалось следующее.

Теорема. Если $X(f)$ — преобразование Фурье последовательности $x(i)$, то преобразованием Фурье сдвинутой последовательности $x(i - k)$ будет служить $X(f) \exp(-j2\pi f k T)$.

Ее можно непосредственно применять к разностному уравнению. Например, полученное только что выражение для $H(f)$ можно получить прямо из разностного уравнения, производя замены

$$x(i - k) \rightarrow X(f) \exp(-j2\pi f k T), \quad y(i - l) \rightarrow Y(f) \exp(-j2\pi f l T).$$

После этого остается найти только отношение $Y(f)/X(f)$. Это можно сделать перестановкой всех членов, содержащих $Y(f)$, в одну часть, а содержащих $X(f)$ — в другую. Затем надо поделить обе части на произведение $X(f)$ и коэффициента, полученного при $Y(f)$.

Для того чтобы найти коэффициент усиления и фазу фильтра низких частот, нужно разделить в знаменателе действительную и мнимую части:

$$H(f) = \frac{1 - \alpha}{1 - \alpha \exp(-j2\pi fT)} = \frac{1 - \alpha}{(1 - \alpha \cos(2\pi fT)) + j\alpha \sin(2\pi fT)}.$$

Получится следующее выражение для коэффициента усиления.

$$\begin{aligned} |H(f)| &= \sqrt{\frac{(1 - \alpha)^2}{(1 - \alpha \cos(2\pi fT))^2 + (\alpha \sin(2\pi fT))^2}} = \\ &= \frac{1 - \alpha}{\sqrt{1 - 2\alpha \cos(2\pi fT) + \alpha^2}}. \end{aligned}$$

Аналогично находится фаза

$$\varphi(f) = \operatorname{arctg} \left(\frac{\alpha \sin(2\pi fT)}{1 - \alpha \cos(2\pi fT)} \right).$$

Для некоторых значений α коэффициент усиления этого фильтра графически изображен на рис. 4.1. Заметим, что $|H(0)| = 1$ для всех значений α . Коэффициент $\alpha_0 = (1 - \alpha)$ выбран таким лишь затем, чтобы сделать коэффициент усиления равным единице в 0 Гц, т. е. чтобы нормировать его.

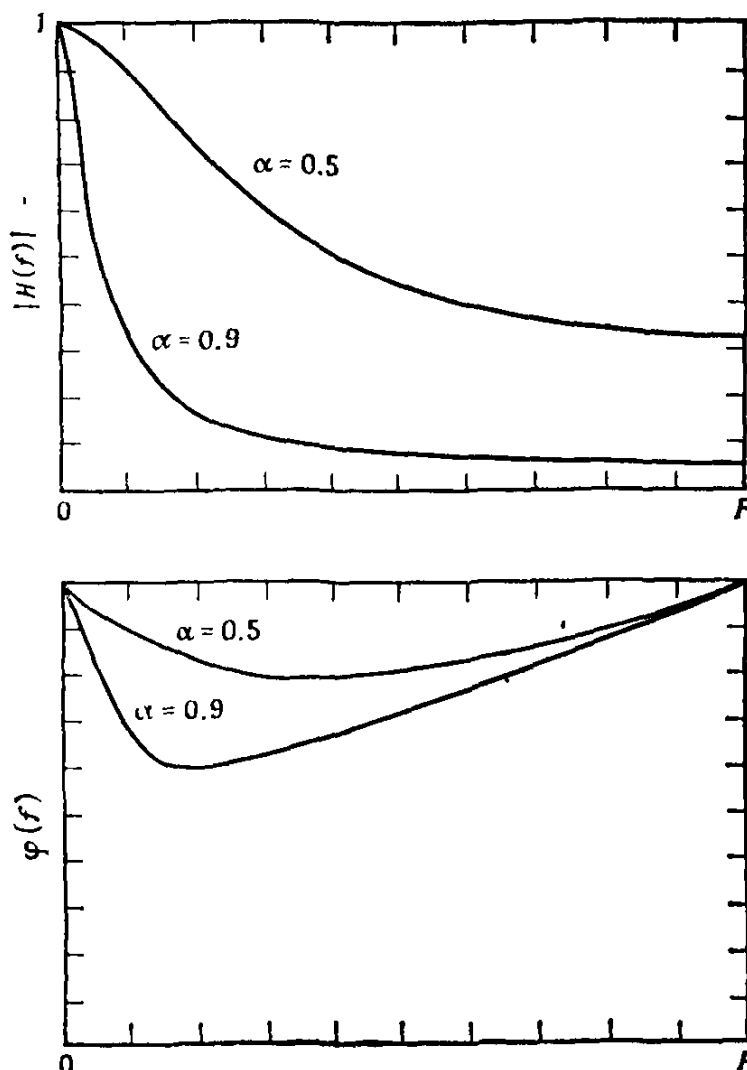


Рис. 4.1. Графики коэффициента усиления и фазы для низкочастотного фильтра первого порядка.

Импульсную функцию отклика $h(i)$ этого фильтра можно получить, применяя обратное преобразование Фурье к функции $H(f)$:

$$h(i) = \int_{-F}^F H(f) \exp(j2\pi f iT) df = \int_{-F}^F \frac{(1 - \alpha) \exp(j2\pi f iT)}{[1 - \alpha \exp(-j2\pi f T)]} df.$$

Найти функцию $h(i)$ можно и не проводя этого сложного интегрирования: следует только получить $h(i)$ как отклик для последовательности $1/T, 0, 0, \dots$. По определению $h(i) = 0$ при

$i < 0$. Поэтому

$$h(0) = \frac{1-\alpha}{T} + 0,$$

$$h(1) = (1-\alpha)0 + \alpha \frac{1-\alpha}{T} = \alpha \frac{1-\alpha}{T},$$

$$h(2) = (1-\alpha)0 + \alpha \left(\alpha \frac{1-\alpha}{T} \right) = \alpha^2 \frac{1-\alpha}{T},$$

...

$$h(i) = \alpha^i \frac{1-\alpha}{T}.$$

Теперь становится понятной причина того, что значение α выбрано в интервале $(0, 1)$. Для таких значений α члены последовательности $h(i)$ *убывают* в геометрической прогрессии. Для значений α , больших единицы, члены этой последовательности *возрастают* в геометрической прогрессии. Например, в случае $\alpha = 0.5$ получается последовательность $0.5/T, 0.25/T, 0.125/T, \dots, 2^{-i}/(2T)$, а в случае $\alpha = 2$ — последовательность $-1/T, -2/T, -4/T, \dots, -2^i/T$.

При $\alpha = 2$ значения последовательности растут настолько быстро, что вскоре исчерпывают емкость любого арифметического устройства с фиксированной точкой и приводят к *переполнению*. Отметим в заключение, что это и есть пример неустойчивости цифрового фильтра.

4.3. ФИЛЬТРЫ ВТОРОГО ПОРЯДКА

Начнем с фильтра второго порядка

$$y(i) = b_0 x(i) - a_1 y(i-1) - a_2 y(i-2).$$

Это не самый общий случай; в последнем могут присутствовать нули, т. е. в уравнении для фильтра появились бы слагаемые $b_1 x(i-1)$ и $b_2 x(i-2)$.

Необходимо запомнить следующие важные выводы, касающиеся фильтров второго порядка. Эти результаты будут доказаны чуть ниже.

1. В наибольшей степени определяют характер фильтра члены a_1 и a_2 . Для того чтобы фильтр был устойчивым, они должны попадать в относительно небольшую треугольную область на плоскости пар возможных значений коэффициентов (a_1, a_2) .

2. Этими же членами, т. е. a_1 и a_2 , определяются и полюсы фильтра. При этом возникают два случая: 1) полюсы действительны, 2) полюсы комплексны. Последний случай наиболее интересен.

Первым шагом в доказательстве этих утверждений будет вычисление передаточной функции для исходного уравнения. Используя теорему о сдвиге, получаем соотношение

$$Y(f) = b_0 X(f) - a_1 Y(f) \exp(-j2\pi fT) - a_2 Y(f) \exp(-j4\pi fT),$$

из которого определяется передаточная функция

$$H(f) = \frac{Y(f)}{X(f)} = \frac{b_0}{1 + a_1 \exp(-j2\pi fT) + a_2 \exp(-j4\pi fT)}.$$

Переходя к обозначению $z = \exp(j2\pi fT)$, перепишем знаменатель в виде

$$1 + a_1 z^{-1} + a_2 z^{-2}.$$

Полюсами фильтра будут корни уравнения

$$z^2 + a_1 z + a_2 = 0,$$

полученного приравниванием нулю знаменателя и умножением обеих частей уравнения на z^2 . Этими корнями будут значения z_1 и z_2 , определенные формулой

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

Возникают два случая.

Первый случай. Если $a_1^2 - 4a_2 \geq 0$, то корни будут разными действительными числами, за исключением случая $a_1^2 - 4a_2 = 0$, в котором эти действительные корни совпадают.

Второй случай. Если $a_1^2 - 4a_2 < 0$, то корнями будут два комплексно-сопряженных числа.

Границу этих двух случаев определяет уравнение для выражения под радикалом

$$a_1^2 - 4a_2 = 0.$$

Это уравнение параболы, изображенной на рис. 4.2.

Действительные полюсы. В первом случае исходное уравнение для фильтра можно переписать в виде

$$y(i) = b_0 x(i) + (c_1 + c_2) y(i-1) - c_1 c_2 y(i-2),$$

где

$$a_1 = -(c_1 + c_2), \quad a_2 = c_1 c_2.$$

Как будет видно чуть позже, c_1 и c_2 суть корни уравнения $z^2 + a_1 z + a_2 = (z - c_1)(z - c_2)$. Естественно, выбрать произвольно их нельзя. Имея значения a_1 и a_2 , числа c_1 и c_2 можно вычислить, записав c_1 в виде a_2/c_2 и подставив его в выражение для a_1 :

$$a_1 = -\left(\frac{a_2}{c_2} + c_2\right);$$

следовательно, $c_2^2 + c_2 a_1 + a_2 = 0$ и

$$c_2 = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2}}{2}$$

и, значит,

$$c_1 = \frac{2a_2}{-a_1 \pm \sqrt{a_1^2 - 4a_2}}.$$

Заметим, что оба коэффициента c_1 и c_2 будут действительными, если $a_1^2 - 4a_2 > 0$. Разумеется, поэтому и определялись

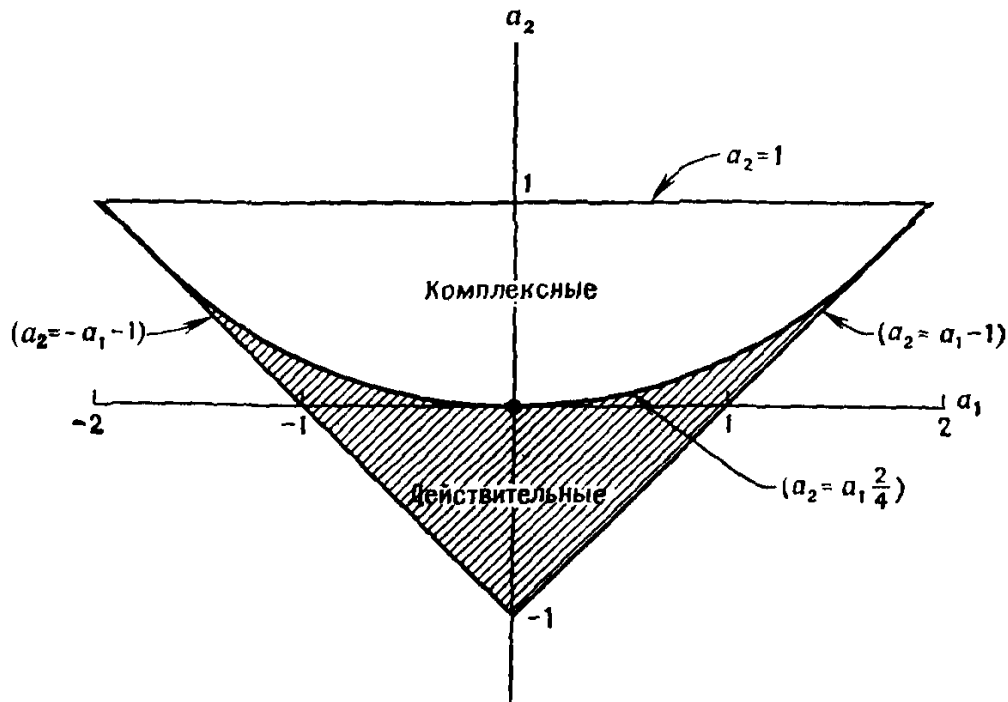


Рис 4.2. Диаграмма устойчивости для фильтра второго порядка. Чтобы фильтр был устойчивым, необходимо, чтобы пары коэффициентов (a_1, a_2) попадали внутрь указанной области, имеющей форму треугольника.

коэффициенты c_1 и c_2 ; они соответствуют случаю с действительными полюсами.

В этом случае передаточная функция фильтра имеет вид

$$H(f) = \frac{b_0}{1 - (c_1 + c_2) \exp(-j2\pi fT) + c_1 c_2 \exp(-j4\pi fT)}.$$

Знаменатель можно представить в виде произведения, поэтому

$$H(f) = \left(\frac{\sqrt{b_0}}{1 - c_1 \exp(-j2\pi fT)} \right) \left(\frac{\sqrt{b_0}}{1 - c_2 \exp(-j2\pi fT)} \right).$$

Теперь ясно, что в этом случае передаточная функция может быть записана в виде произведения

$$H(f) = H_1(f) H_2(f)$$

передаточных функций двух одновременно применяемых фильтров первого порядка

$$H_1(f) = \frac{\sqrt{b_0}}{1 - c_1 \exp(-j2\pi fT)} \quad \text{и} \quad H_2(f) = \frac{\sqrt{b_0}}{1 - c_2 \exp(-j2\pi fT)}.$$

Иными словами, если $u(i)$ — промежуточная функция при вычислении $y(i)$, т. е.

$$u(i) = \sqrt{b_0}x(i) + c_1u(i-1), \quad y(i) = \sqrt{b_0}u(i) + c_2y(i-1),$$

то

$$U(f) = H_1(f)X(f), \quad Y(f) = H_2(f)U(f),$$

где $U(f)$ — преобразование Фурье функции $u(i)$. Следовательно,

$$Y(f) = H_2(f)[H_1(f)X(f)].$$

Заметим, что $H_1(f)$ и $H_2(f)$ можно поменять ролями; конечный результат при этом остается тем же.

Как следует из рассмотрения фильтров первого порядка, абсолютные значения обоих коэффициентов c_1 и c_2 должны быть ограничены единицей. В противном случае по крайней мере один из составляющих фильтров будет неустойчивым. Если абсолютное значение c_2 меньше единицы, т. е. $-1 < c_2 < 1$, то

$$-1 < \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2}}{2} < 1.$$

Умножая эту цепочку неравенств на 2 и прибавляя ко всем частям неравенств a_1 (знаки неравенств при этих операциях на противоположные не изменятся), получим цепочку

$$a_1 - 2 < \pm \sqrt{a_1^2 - 4a_2} < a_1 + 2.$$

Возводя в квадраты обе части второго неравенства, получим неравенство

$$a_1^2 - 4a_2 < a_1^2 + 4a_1 + 4,$$

которое после упрощения принимает вид

$$a_2 > -a_1 - 1.$$

Производя подобные действия с первым неравенством, т. е. возводя в квадрат обе его части (при этом получится неравенство $a_1^2 - 4a_1 + 4 > a_1^2 - 4a_2$) и упрощая результат, придем к неравенству

$$a_2 > a_1 - 1.$$

Эти результаты показывают, что значения a_1 , a_2 , удовлетворяющие неравенствам, должны быть расположены выше каждой

из прямых

$$a_2 = a_1 - 1, \quad a_2 = -a_1 - 1$$

на плоскости (a_1, a_2) . Как явствует из рис. 4.2, отрезки этих прямых служат нижними сторонами треугольника устойчивости.

Комплексные полюсы. Гораздо более интересен случай с комплексными полюсами. Определим фильтр уравнением

$$y(i) = b_0 x(i) + [2 \exp(-\alpha) \cos \beta] y(i-1) + [-\exp(-2\alpha)] y(i-2),$$

т. е. положим

$$a_1 = -2 \exp(-\alpha) \cos \beta, \quad a_2 = \exp(-2\alpha).$$

Для данных a_1 и a_2 параметры α и β вычисляются по формулам

$$\alpha = -\frac{1}{2} \ln(a_2),$$

$$\beta = \arccos \left(- \left[\frac{a_1}{2 \exp(-\alpha)} \right] \right) = \arccos \left(- \left[\frac{a_1}{2 \sqrt{a_2}} \right] \right).$$

Из этих формул получаются два ограничения на a_1 и a_2 : из уравнения для α — ограничение $0 < a_2$, а из уравнения для β — ограничение

$$\left| - \frac{a_1}{2 \sqrt{a_2}} \right| < 1.$$

Если хотя бы одно из этих условий не выполняется, то корни будут действительными. Возводя в квадрат обе части последнего неравенства, получим условие $a_1^2 < 4a_2$, которому соответствует область над параболой $a_1^2 = 4a_2$.

Квадратное уравнение для полюсов

$$z^2 + a_1 z + a_2 = 0$$

в рассматриваемом случае имеет вид

$$z^2 - [2 \exp(-\alpha) \cos \beta] z + \exp(-2\alpha) = 0.$$

Поэтому корни вычисляются по формуле

$$z_{1,2} = \frac{2 \exp(-\alpha) \cos \beta \pm \sqrt{4 \exp(-2\alpha) \cos^2 \beta - 4 \exp(-2\alpha)}}{2}.$$

Вынося из-под знака квадратного корня $4 \exp(-2\alpha)$ и упрощая полученное выражение, имеем

$$z_{1,2} = \exp(-\alpha) \left[\cos \beta \pm \sqrt{\cos^2 \beta - 1} \right] = \exp(-\alpha) [\cos \beta \pm j \sin \beta] =$$

$$= \exp(-\alpha) \exp(\pm j\beta) = \exp(-\alpha \pm j\beta).$$

Наконец, подстановкой $z = \exp(j2\pi fT)$ получаем

$$j2\pi fT = -\alpha \pm j\beta, \quad f_{1,2} = \frac{1}{2\pi T} [\pm \beta + j\alpha].$$

Если фильтр второго порядка имеет действительные полюсы, то они всегда будут расположены или на прямой $f=0$, или на прямой $f=F$. В последнем случае один из полюсов попадет и на прямую $f=-F$. В остальных случаях полюсы заполняют часть комплексной области над интервалом $(-F, F)$.

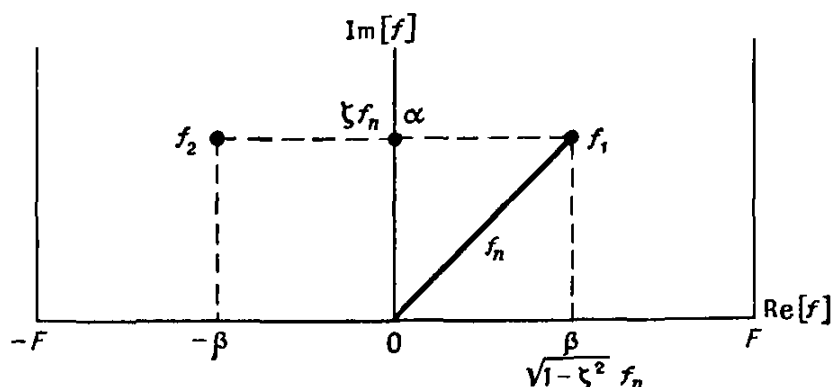


Рис. 4.3. Расположение полюсов фильтра второго порядка. Предполагается, что величина α положительна.

Расположение полюсов показано на рис. 4.3. Если перейти к собственной частоте и коэффициенту затухания, то

$$\alpha = 2\pi\zeta f_n, \quad \beta = 2\pi\sqrt{1-\zeta^2} f_n.$$

Теперь остается проверить выражение для знаменателя передаточной функции

$$D(f) = \{1 - \exp[-j2\pi T(f - f_1)]\} \{1 - \exp[-j2\pi T(f - f_2)]\}.$$

Заметим, что как только f примет значение f_1 или f_2 , один из сомножителей обратится в 0, поскольку соответствующая экспонента станет равной единице. Таким образом,

$$D(f_1) = D(f_2) = 0,$$

что и следовало ожидать.

Более того, если подставить выражения для f_1 и f_2 в формулу для $D(f)$, то получатся исходные коэффициенты. Действительно,

$$\begin{aligned} D(f) &= \left\{1 - \exp\left[-j2\pi T\left(f - \frac{-\beta + j\alpha}{2\pi T}\right)\right]\right\} \times \\ &\quad \times \left\{1 - \exp\left[-j2\pi T\left(f - \frac{+\beta + j\alpha}{2\pi T}\right)\right]\right\} = \\ &= [1 - \exp(-j2\pi f T) \exp(-j\beta - \alpha)] \times \\ &\quad \times [1 - \exp(-j2\pi f T) \exp(j\beta - \alpha)]. \end{aligned}$$

Раскрывая скобки, получим

$$\begin{aligned} D(f) &= 1 - \exp(-j2\pi fT) [\exp(-j\beta - \alpha) + \exp(j\beta - \alpha)] + \\ &\quad + \exp(-j4\pi fT) \exp(-j\beta - \alpha + j\beta - \alpha) = \\ &= 1 - \exp(-j2\pi fT) [2 \exp(-\alpha) \cos \beta] + \\ &\quad + [\exp(-2\alpha)] \exp(-j4\pi fT) = \\ &= 1 - 2 \exp(-\alpha) \cos \beta z^{-1} + \exp(-2\alpha) z^{-2}, \end{aligned}$$

т. е. получился тот знаменатель передаточной функции, который и должен быть.

Импульсная функция отклика. Для того чтобы получить последнюю границу треугольника устойчивости, изображенного на рис. 4.2, нужно найти единичный импульсный отклик фильтра второго порядка с комплексными полюсами (второй случай).

Предоставим читателю¹⁾ проверку того, что при входе

$$x(i) = \begin{cases} 1/T, & i=0, \\ 0 & \text{при остальных } i, \end{cases}$$

и при условии, что $y(i) = 0$ для $i < 0$, выходом для $x(i)$ фильтра

$$y(i) = b_0 x(i) + 2 \exp(-\alpha) \cos \beta y(i-1) - \exp(-2\alpha) y(i-2)$$

будет значение ($i \geq 0$)

$$y(i) = \frac{b_0}{T} \exp(-i\alpha) \left[\frac{\cos \beta \cos(i\beta - \beta) + (1 - 2 \cos^2 \beta) \cos i\beta}{1 - \cos^2 \beta} \right].$$

Хотя эта запись и допускает дальнейшее упрощение, мы все же остановимся на ней, поскольку так немного проще вычислять несколько первых значений $y(i)$.

Обратим внимание на то, что части формулы, содержащие α и β , разделились. Это поможет выяснить роль, которую играет каждая из этих величин. Заметим также, что α входит в экспоненту, а β — в тригонометрическое выражение.

Тригонометрический сомножитель периодичен и ограничен сверху величиной $2/(1 - \cos^2 \beta)$. Поэтому с точки зрения устойчивости важна экспонента.

Экспоненциальный член определяет три случая: устойчивость при $\alpha > 0$, маргинальную устойчивость при $\alpha = 0$ и неустойчивость при $\alpha < 0$.

При $\alpha > 0$ величина $\exp(-i\alpha)$ экспоненциально *убывает*, при $\alpha = 0$ экспоненциальный член исчезает, при $\alpha < 0$ величина $\exp(-i\alpha)$ экспоненциально *возрастает*.

Граница устойчивости $\alpha = 0$ дает третью сторону треугольника устойчивости, поскольку после подстановки в уравнение

¹⁾ См. упр. 4.6.

$a_2 = \exp(-2\alpha)$ значения $\alpha = 0$ получается равенство $a_2 = 1$, определяющее верхнюю сторону треугольника с рис. 4.2.

Фильтр второго порядка как осциллятор. Фильтр второго порядка с комплексными полюсами при $\alpha = 0$ может использоваться в качестве генератора синуса или косинуса. Соответствующее уравнение имеет вид

$$y(i) = 2(\cos \beta) y(i-1) - y(i-2), \quad i = 0, 1, \dots,$$

где начальные условия определяются таблицей

	Для синуса	Для косинуса
$y(-1)$	$-\sin \beta$	$\cos \beta$
$y(-2)$	$-\sin 2\beta$	$\cos 2\beta$

Как синус, так и косинус будет иметь частоту $f = \frac{\beta}{2\pi T}$ Гц.

Рассмотрим в качестве примера случай с синусом:

$$y(0) = 2 \cos \beta (-\sin \beta) - \sin 2\beta = 0,$$

$$y(1) = 2 \cos \beta (0) - (-\sin \beta) = \sin \beta,$$

$$y(2) = 2 \cos \beta \sin \beta - (0) = \sin 2\beta,$$

$$y(3) = 2 \cos \beta \sin 2\beta - \sin \beta = 4 \cos^2 \beta \sin \beta - \sin \beta = \\ = 4(1 - \sin^2 \beta) \sin \beta - \sin \beta = 3 \sin \beta - 4 \sin^3 \beta = \sin 3\beta.$$

Для общего случая доказательство получается по индукции. Предположим, что

$$y(i-1) = \sin(i-1)\beta, \quad y(i-2) = \sin(i-2)\beta.$$

Тогда

$$y(i) = 2 \cos \beta \sin(i-1)\beta - \sin(i-2)\beta.$$

Если равенство

$$\sin(i-2)\beta = \sin(i-1)\beta \cos \beta - \cos(i-1)\beta \sin \beta$$

подставить в выражение для $y(i)$, то последнее примет вид

$$y(i) = 2 \cos \beta \sin(i-1)\beta - \sin(i-1)\beta \cos \beta + \cos(i-1)\beta \sin \beta = \\ = \cos \beta \sin(i-1)\beta + \cos(i-1)\beta \sin \beta = \sin i\beta,$$

что и завершает доказательство по индукции... Доказательство в случае косинуса проводится аналогично.

Точный фильтр низких частот второго порядка. Фильтр с передаточной функцией

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\sin(\omega T/2)}{A}\right)^4},$$

где $A = \sin(\pi BT)$, называют фильтром низких частот Баттеруорта (Баттеруорт (1930)). Аналоговая форма этого фильтра рассмат-

ривалась в гл. 3. Там указывалось, что величины

$$|H(0)|^2 = \frac{1}{1+0} = 1,$$

$$|H(B)|^2 = \frac{1}{1 + \left(\frac{\sin(\pi BT)}{\sin(\pi fT)}\right)^4} = \frac{1}{2}$$

помогают определить характеристики фильтра. Величина $|H(f)|^2$ при таких f , что $\pi BT < \pi fT < \pi$, стремится к нулю. Величина $(\sin(\pi fT)/\sin(\pi BT))^4$ больше 1, поэтому $|H(f)|^2$ убывает с ростом πfT и достигает минимума, когда $\pi fT = \pi/2$.

Коэффициенты a_1 и a_2 этого фильтра можно определить несколькими способами. Будем их вычислять «в лоб». Сразу заметим, что это довольно сложно.

Передаточную функцию можно переписать в виде

$$\frac{1}{1 + \left(\frac{\sin(\omega T/2)}{A}\right)^4} = \frac{A^4}{A^4 + \sin^4(\omega T/2)} =$$

$$= \frac{A^4}{A^4 + \left(\frac{1 - \cos(\omega T)}{2}\right)^2} = \frac{4A^4}{4A^4 + (1 - 2\cos(\omega T) + \cos^2(\omega T))} =$$

$$= \frac{4A^4}{\cos^2(\omega T) - 2\cos(\omega T) + (1 + 4A^4)}.$$

Без такого перехода от половинного угла прямое применение метода было бы невозможным. Чтобы сделать следующий шаг, заметим, что знаменатель преобразования Фурье разностного уравнения любого двухполюсного фильтра можно записать в виде $(s + pz^{-1} + qz^{-2})$. Величины s , p и q определяют искомые коэффициенты. Найдем их. Квадрат модуля знаменателя передаточной функции должен быть равным

$$(s + pz^{-1} + qz^{-2})(s + pz + qz^2) =$$

$$= (s^2 + p^2 + q^2) + (pq + sp)(z + z^{-1}) + sq(z^2 + z^{-2}).$$

Поскольку $\cos(\omega T) = (z + z^{-1})/2$, $\cos^2(\omega T) = 1/4(z^2 + 2 + z^{-2})$, знаменатель можно переписать в виде

$$\cos^2(\omega T) = 2\cos(\omega T) + (1 + 4A^4) =$$

$$= 1/4(z^2 + 2 + z^{-2}) - 2\left(\frac{z + z^{-1}}{2}\right) + (1 + 4A^4) =$$

$$= 1/4\{(z^2 + z^{-2}) - 4(z + z^{-1}) + [4(1 + 4A^4) + 2]\}.$$

Приравнивая коэффициенты при одинаковых членах и отбрасы-

вая множитель $1/4$, получим три уравнения:

$$\begin{aligned}sq &= 1, \\pq + sp &= -4, \\s^2 + p^2 + q^2 &= (6 + 16A^4) = C,\end{aligned}$$

из которых нетрудно получить решения

$$a_1 = qp, \quad a_2 = q^2,$$

где

$$\begin{aligned}p &= \sqrt{\frac{C+2 - \sqrt{(C+2)^2 - 64}}{2}}, \\q &= \left(-\frac{4}{p} + \sqrt{\left(\frac{4}{p}\right)^2 - 4}\right) / 2\end{aligned}$$

и $C = 6 + 16A^4$.

На рис. 4.4 приводится специальная программа для вычисления весов a_1 , a_2 и b_0 этого фильтра. Обозначения в тексте сопоставляются переменным в программе согласно следующей таблице:

<i>В тексте</i>	<i>B</i>	<i>T</i>	<i>a₁</i>	<i>a₂</i>	<i>b₀</i>	<i>A</i>	<i>C+2</i>	<i>p</i>	<i>q</i>
<i>В программе</i>	BW	T	A1	A2	B0	A	C	P	Q

На такой простой результат нами потрачено довольно много усилий. Можно, пожалуй, таким же способом решать задачу

```

1      SUBROUTINE LP2 (BW,T,A1,A2,B0)
2      DOUBLE PRECISION A,C,P,PP,Q,QQ
3      A = SIN (3.1415927*T*BW)
4      C = 8.0D0 + 16.0D0*(A**4)
5      PP = DSQRT (C*C - 64.0D0)
6      P = DSQRT ((C - PP)/2.0D0)
7      QQ = (4.0D0/P)**2 - 4.0D0
8      Q = -2.0D0/P + DSQRT(QQ)/2.0D0
9      A1 = P*Q
10     A2 = Q*Q
11     B0 = 1.0 + A1 + A2
12     RETURN
13     END

```

Рис. 4.4. Программа для получения весов низкочастотного фильтра второго порядка.

для фильтров Баттеруорта и более высоких порядков, но это едва ли необходимо. Как мы вскоре увидим, для реализации фильтров более высоких порядков есть значительно более простой способ

4.4 ФИЛЬТРЫ БОЛЕЕ ВЫСОКИХ ПОРЯДКОВ

Под фильтром более высокого порядка подразумевается фильтр, выражающийся через три или более предыдущих члена. Напри-

мер, фильтр

$$y(i) = b_0 x(i) - a_1 y(i-1) - a_2 y(i-2) - a_3 y(i-3)$$

является фильтром третьего порядка.

В этом параграфе будут затронуты четыре вида реализаций фильтра: прямая, последовательная (или каскадная), параллельная и цепная. В тех приложениях, которые нами рассматриваются, наиболее важную роль играет последовательная форма реализации.

Фильтр M -го порядка прямого вида имеет разностное уравнение

$$y(i) = \sum_{k=0}^M b_k x(i-k) - \sum_{m=1}^M a_m y(i-m).$$

Если к этому уравнению, как и раньше, применить преобразование Фурье, то получится уравнение

$$Y(f) = X(f) \sum_{k=0}^M b_k \exp(-j2\pi f k T) - Y(f) \sum_{m=1}^M a_m \exp(-j2\pi f m T).$$

Передаточная функция имеет вид

$$H(f) = \frac{Y(f)}{X(f)} = \frac{\sum_{k=0}^M b_k \exp(-j2\pi f k T)}{1 + \sum_{m=1}^M a_m \exp(-j2\pi f m T)}.$$

Числитель и знаменатель суть многочлены M -го порядка от $z = \exp(j2\pi f T)$. В частности,

$$N(f) = \sum_{k=0}^M b_k z^{-k}, \quad D(f) = 1 + \sum_{m=1}^M a_m z^{-m}.$$

В этой главе было показано, что уравнение $D(f) = 0$ имеет корни f_1, \dots, f_M . Поэтому многочлен $D(f)$ можно переписать в виде

$$D(f) = a_0 \prod_{i=1}^M \{1 - \exp[-j2\pi T(f - f_k)]\},$$

где a_0 — постоянный множитель.

Корни распадаются на два типа.

Чисто мнимые корни. В этом случае корни f_k имеют вид $j\alpha_k/(2\pi T)$ или $j\alpha_k/(2\pi T) + F$, где $\alpha_k > 0$. Соответствующие сомножители в знаменателе запишутся в виде

$$\left\{ 1 - \exp \left[-j2\pi T \left(f - \frac{j\alpha_k}{2\pi T} \right) \right] \right\} = [1 - \exp(-\alpha_k) z^{-1}].$$

Если корень имеет вид $jc_k + F$, то результат $[1 + \exp(-\alpha_k) z^{-1}]$ будет соответствовать фильтру высоких частот.

Комплексные корни. В этом случае корни встречаются парами. Если $f_k = (j\alpha_k + \beta_k)/(2\pi T)$, где $\beta_k > 0$, то найдется такое k' , что $f_{k'} = (j\alpha_k - \beta_k)/(2\pi T)$. Два соответствующих сомножителя в знаменателе принимают вид

$$\left\{ 1 - \exp\left[-j2\pi T\left(f - \frac{j\alpha_k + \beta_k}{2\pi T}\right)\right] \right\} \left\{ 1 - \exp\left[-j2\pi T\left(f - \frac{j\alpha_k - \beta_k}{2\pi T}\right)\right] \right\} = \\ = 1 - 2\exp(-\alpha_k) \cos \beta_k z^{-1} + \exp(-2\alpha_k) z^{-2}.$$

Учитывая полученные результаты, можно утверждать, что найдутся такие числа P и Q (Q — четное), что $M = P + Q$ и

$$D(f) = \prod_{p=1}^P (1 - a_{1p} z^{-1}) \prod_{q=1}^{Q/2} (1 - a_{1q} z^{-1} - a_{2q} z^{-2}),$$

где a_{1p} , a_{1q} и a_{2q} — действительные числа. Это означает, что многочлен $D(f)$ всегда можно разложить в произведение сомножителей первой и второй степеней с действительными коэффициентами.

До сих пор мы пользовались только фактами из алгебры, вытекающими из некоторых элементарных теорем. Посмотрим теперь, как эти факты преломляются в теории фильтров. Допустим для простоты (без большой потери общности), что в числителе передаточной функции стоит только одна константа, т. е.

$$H(f) = \frac{b_0}{D(f)}.$$

Определим функции $H_r(f)$ формулой

$$H_r(f) = \begin{cases} \frac{b_0^{1/R}}{1 - a_{1p} z^{-1}}, \\ \frac{b_0^{1/R}}{1 - a_{1q} z^{-1} - a_{2q} z^{-2}}, \end{cases}$$

где $R = P + Q/2$, $r = q = 1, \dots, Q/2$, $p = r - Q/2, \dots, R$. Теперь передаточная функция запишется в виде

$$H(f) = \prod_{r=1}^R H_r(f).$$

Нетрудно видеть, что функции $H_r(f)$ суть передаточные функции, соответствующие фильтрам первого и второго порядка. Таким образом, нами доказан следующий результат.

Теорема. *Всякий фильтр более высокого порядка можно реализовать последовательностью фильтров первого и второго порядка.*

Уравнением для r -го фильтра служит

$$u^{(r)}(i) = b'_0 u^{(r-1)}(i) - a_1 u^{(r)}(i-1) - a_2 u^{(r)}(i-2),$$

где

$$b'_0 = b_0^{1/R}, \quad u^{(0)}(i) = x(i), \quad u^{(R)}(i) = y(i).$$

Другими словами, входом r -го фильтра будет последовательность $u^{(r-1)}(i)$, а выходом $u^{(r)}(i)$. Заметим, что некоторые веса a_{2r} фильтров могут быть равными нулю.

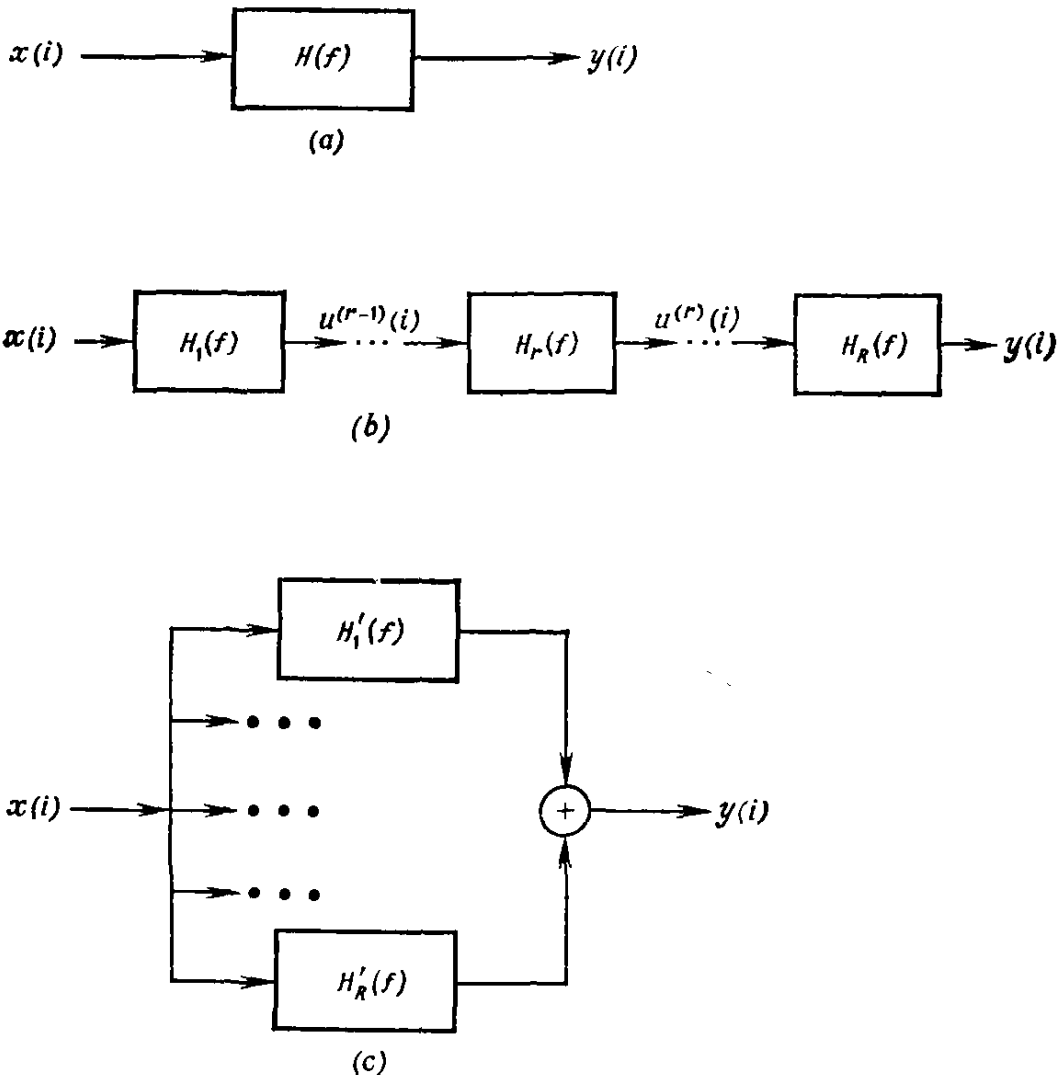


Рис. 4.5. Три способа реализации одного и того же фильтра. a — прямой; b — последовательный; c — параллельный.

Прямая и последовательная формы фильтра показаны соответственно на рис. 4.5, a и b . Последовательный фильтр часто называют *каскадным* фильтром.

В последовательном соединении всякие два фильтра коммутативны, и по крайней мере на ЭВМ с очень большим числом битов никаких различий при их перестановках не возникает. Число таких перестановок равно $R!$. Если же число битов

уменьшается, то, как будет показано в следующей главе, из-за шума, порождаемого самими фильтрами, порядок их расположения уже начинает приобретать значение, поскольку некоторые фильтры могут генерировать более сильный шум, чем другие, а некоторые наоборот, определенным образом ослаблять шум, создаваемый другими фильтрами.

Далее, оказывается, что прямой фильтр в определенных случаях может быть заменен фильтром параллельного типа, схема которого изображена на рис. 4.5, с. Отдельный блок такого фильтра определяется передаточной функцией

$$H_r(f) = \frac{N_r(f)}{D_r(f)},$$

где $D_r(f)$ — тот же многочлен, что и в последовательной схеме.

Общая передаточная функция этого фильтра имеет вид

$$H(f) = \sum_{r=1}^R H_r(f) = \frac{N_1(f)}{D_1(f)} + \dots + \frac{N_R(f)}{D_R(f)}.$$

Эту функцию можно записать в виде дробно-рациональной функции, если сложить все члены:

$$H(f) = \frac{\sum_{r=1}^R N_r(f) \prod_{k=1}^R D_k(f)/D_r(f)}{\prod_{r=1}^R D_r(f)}.$$

В этом случае числитель может оказаться равным b_0 . Сама проблема реализации параллельного фильтра и состоит в поиске таких функций $N_r(f)$ в числителях, чтобы это имело место, т. е.

$$b_0 = \sum_{r=1}^R N_r(f) \prod_{k=1}^R \frac{D_k(f)}{D_r(f)}.$$

Используя частные производные, такое решение можно найти методами комплексного анализа, если все полюсы (т. е. корни общей функции $D(f)$) будут попарно различны.

Из рассмотренных определений понятно, что можно строить массу промежуточных типов фильтров, частично являющихся последовательными, частично прямыми и частично параллельными. Пожалуй, нет никакого смысла останавливаться на этом подробнее.

В *цепном* фильтре все операции располагаются, грубо говоря, в виде цепи. Фрагмент такого фильтра изображен на рис. 4.6. Цепным фильтрам посвящены работы Грэя и Мэркела (1973), Митры и Шервуда (1973). Разумеется, форма таких фильтров

определяется неоднозначно. Существует немало способов для представления в виде цепного фильтра даже простого фильтра второго порядка, а с учетом фильтров более высоких порядков число способов, вообще, становится безмерным.

При таком большом типовом разнообразии фильтров (включающем четыре основных типа) возникает естественный вопрос о наилучшем выборе.

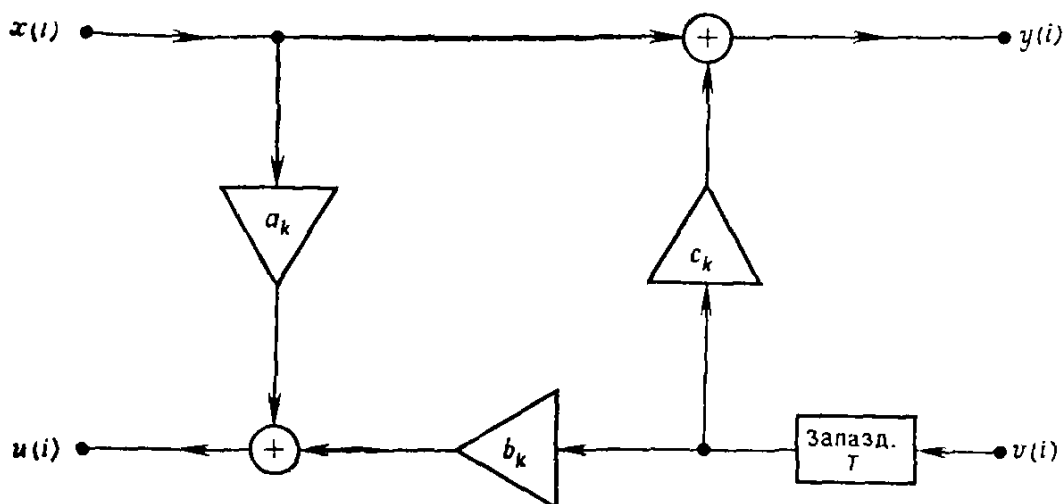


Рис. 4.6. Один участок цепного фильтра.

Ответ на этот вопрос зависит от многих факторов, в частности от характера обрабатываемой информации и оборудования, используемого при обработке. Однако если ограничиться сферой приложений, освещенных в этой книге, т. е. анализом тестовых данных на ЭВМ широкого назначения, производящих операции в режиме плавающей точки, то основу выбора можно свести к следующему.

1. Наилучшей формой фильтра во всех отношениях является последовательная.

2. С вычислительной точки зрения наихудшей является прямая форма фильтра, однако с точки зрения стоимости вычислений, производимых при фильтрации, она наилучшая.

3. Те фильтры, которые встречаются в этой книге, всегда можно представить в виде последовательного фильтра, имеющего не более одного входящего в него фильтра первого порядка.

4. При использовании цепной реализации физические фильтры для ЭВМ широкого назначения, пожалуй, непригодны.

5. Параллельная реализация не имеет никаких преимуществ, и поэтому не рекомендуется.

Короче говоря, если нет достаточных оснований для выбора других возможностей, то при работе на ЭВМ широкого назначения самым предпочтительным будет выбор последовательной формы фильтра.

Если же используется арифметика с фиксированной точкой, то положение коренным образом меняется.

1. Прямая форма редко оказывается пригодной, если полоса частот узкая.

2. Наименьшую ошибку дает параллельная форма (см., например, работу Джексона (1970)).

3. При последовательной форме большое значение для уменьшения общего шума, порождаемого фильтром, имеет порядок расположения составляющих его фильтров. Основные результаты по этому вопросу также можно найти у Джексона (1970).

4. Некоторые типы цепных реализаций дают, по-видимому, наилучшую возможность для получения наименьшего уровня собственного шума фильтра.

Устойчивость. Если ограничиться только последовательной формой реализации фильтра, то вопрос о расчете устойчивости будет довольно простым. Чтобы вникнуть в суть проблемы, вернемся к упомянутому в начале этой главы фильтру третьего порядка. Три коэффициента этого фильтра определяют трехмерное пространство. Множество коэффициентов устойчивых фильтров третьего порядка будет заключено в некотором тетраэдре этого пространства. Уже в этом случае границы области устойчивости определить не так просто. Еще более значительно возрастают трудности при попытке определить границы таких областей для прямых форм фильтров четвертого и пятого порядка. Основные результаты, решающие эти проблемы, были получены в работах Джури и Блэнхарда (1961) и Джури (1961).

Для последовательной формы такого же фильтра проблема оказывается совсем простой. Каждый составляющий его фильтр можно рассматривать отдельно от других. Весь фильтр целиком будет устойчивым, если каждый из составляющих его фильтров тоже будет устойчивым.

Если отнести сказанное к полюсам, то это будет означать, что все они должны лежать в той части верхней плоскости частот, которая изображена на рис. 4.3. *Замечание:* во многих технических книгах вместо f пользуются величиной $j2\pi f$. В этом случае график поворачивается на 90° против часовой стрелки и устойчивые полюсы переходят на левую полуплоскость. Кроме того, многие, если не почти все авторы, обращаясь в книгах к дискретной частоте, пользуются тем фактом, что устойчивые полюсы располагаются в единичном круге на z -плоскости.

Наш выбор f -плоскости в противовес $j2\pi f$ - и z -плоскостям продиктован главной темой книги — изучением плотности спектра мощности. Такой выбор позволяет легче интерпретировать как характеристики фильтров, так и соответствующие спектральные плотности.

4.5. ИДЕАЛЬНЫЕ БАЗИСНЫЕ ФИЛЬТРЫ

По традиции под *базисными фильтрами* понимают фильтры, которые широко применяются при обработке данных и сигналов. К ним относят главным образом фильтры низких и высоких

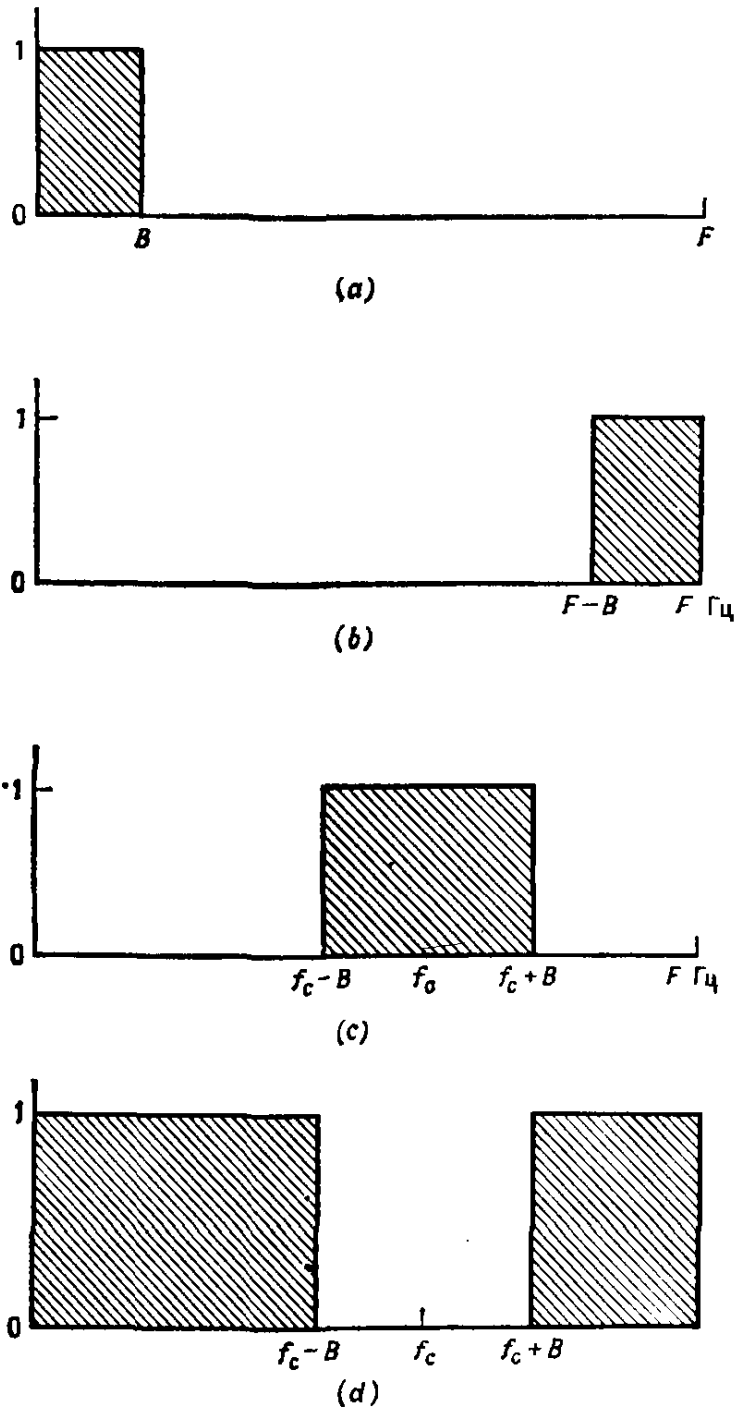


Рис. 4.7. Квадраты абсолютных значений передаточных функций. *a* — идеального низкочастотного фильтра; *b* — идеального фильтра высоких частот; *c* — идеального полоснопропускающего фильтра и *d* — идеального заграждающего фильтра. Штриховка показывает, какая часть данных пропускается фильтром,

частот, полоснопропускающие и заграждающие фильтры. Последние иногда называют также *вырезающими* фильтрами.

В идеале эти фильтры должны были бы иметь следующие характеристики.

1. Идеальный фильтр низких частот должен пропускать всю частотную информацию в диапазоне $(0, B)$ Гц и подавлять всю информацию в диапазоне (B, F) Гц.

2. Идеальный фильтр высоких частот должен иметь обратные характеристики, т. е. пропускать частоты в диапазоне $(F - B, F)$ Гц и поглощать их в диапазоне $(0, F - B)$ Гц.

3. Полоснопропускающий фильтр должен пропускать частоты только в диапазоне от $(f_c - B)$ до $(f_c + B)$ Гц.

4. Идеальный заграждающий фильтр должен поглощать всю частотную информацию в диапазоне от $(f_c - B)$ до $(f_c + B)$ Гц, оставляя всю информацию вне этого диапазона без изменения.

Полосы пропускания фильтров низких и высоких частот имеют ширину B Гц, для полоснопропускающего и заграждающего фильтров эта ширина равна $2B$ и $(F - 2B)$ Гц соответственно. На рис. 4.7 изображены модули передаточных функций идеальных фильтров. На практике добиться таких точных откликов фильтров невозможно по двум причинам. Первая состоит в том, что получить идеальную форму фильтра можно только в пределе при стремящемся к бесконечности N . Поэтому фактически используются лишь аппроксимации идеальных фильтров. Вторая причина — ограниченное число данных, используемое на практике. Это равносильно тому, что истинная передаточная функция есть в действительности свертка теоретической передаточной функции с функцией $(\sin x)/x$, появление которой обусловлено переходом к конечной длине записи. И все-таки можно так реализовать фильтры и применять их, что для практических потребностей они будут столь же хороши, как и идеальные фильтры.

Реализация фильтров, аппроксимирующих идеальные фильтры. Для точной реализации перечисленных выше фильтров потребовалось бы бесконечное число членов. Поэтому идеальные формы приближают с помощью многочленов, как это и делается в электрических или электронных фильтрах. В случае электрического фильтра эти многочлены суть просто многочлены от f . В случае цифровых устройств эти многочлены являются тригонометрическими функциями f .

Например, показанный на рис. 4.7, а идеальный фильтр низких частот можно аппроксимировать для электрических устройств фильтром Баттеруорта

$$|H(f)|^2 = \frac{1}{1 + (f/B)^{2M}},$$

а для цифровых устройств — фильтром Баттеруорта

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\sin(\pi f T)}{\sin(\pi B T)} \right)^{2M}}$$

или

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\operatorname{tg}(\pi f T)}{\operatorname{tg}(\pi B T)} \right)^{2M}}.$$

Во всех трех случаях

$$|H(f)|^2 = \begin{cases} 1 & \text{при } f=0, \\ 1/2 & \text{при } f=B, \\ <1/2 & \text{при } f > B, \end{cases}$$

Степень точности приближения увеличивается с ростом M , который сопровождается, естественно, ростом стоимости и самого фильтра, и процесса фильтрации.

Причина, по которой эти функции выбираются в качестве аппроксимации передаточной функции идеального фильтра низ-

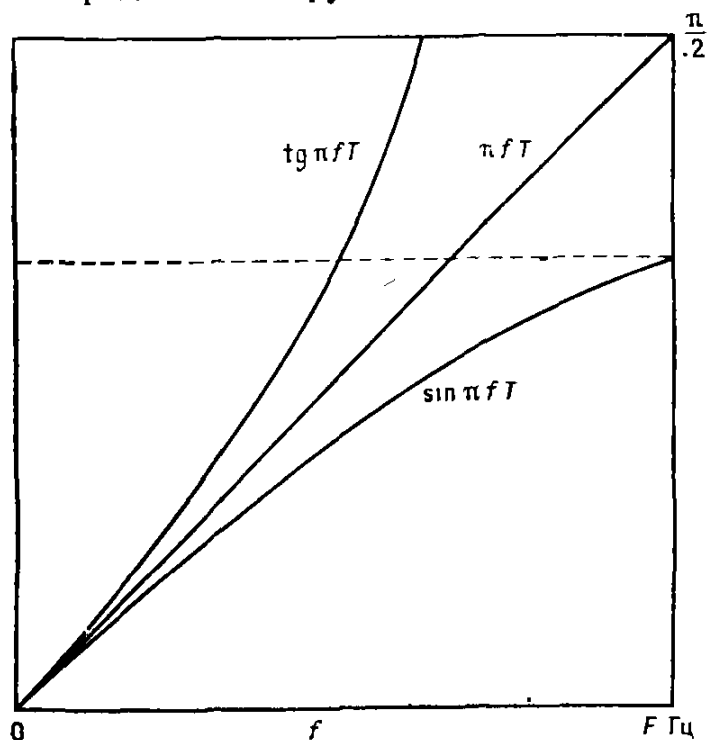


Рис. 4.8. Сравнение поведения функций $\pi f T$, $\operatorname{tg}(\pi f T)$ и $\sin(\pi f T)$.

ких частот, состоит в том, что функции $\sin(\pi f T)$ и $\operatorname{tg}(\pi f T)$ в интервале $(0, F)$ изменения частоты f ведут себя почти так же, как $\pi f T$ (см. рис. 4.8).

Теперь возникает задача определения весов a_m и b_k для прямой реализации фильтра (или для других эквивалентных реализаций) с передаточной функцией, аппроксимирующей передаточные функции идеальных фильтров.

4.6. СИГУСНЫЙ ФИЛЬТР НИЗКИХ ЧАСТОТ БАТТЕРУОРТА

Синусный фильтр низких частот — не только самый простой из фильтров высоких порядков, но во многих приложениях и один из наиболее полезных. Поэтому остановимся на нем подробно. Попутно с изучением процедуры получения коэффициентов фильтров этого типа рассмотрим подпрограмму LPSB из приложения А, предназначенную для этой цели.

Квадрат модуля передаточной функции синусного фильтра Баттеруорта определен формулой

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\sin(\pi fT)}{\sin(\pi BT)} \right)^{2M}}.$$

Фильтр не имеет нулей, число его полюсов равно M .

В конце § 4.3 рассматривалась реализация этого фильтра при $M=2$. При этом отмечалось, что использованный там для поиска коэффициентов метод нельзя применить при M , большем 2.

В этом параграфе будет применен метод, основными составляющими которого являются следующие действия:

1. Определяются полюсы фильтра в виде решений уравнения

$$D(f) = 1 + (\sin(\pi fT)/\sin(\pi BT))^{2M} = 0.$$

2. Из $2M$ корней этого уравнения M попадают в верхнюю, а другие M — в нижнюю полуплоскости. Все корни, попавшие в нижнюю полуплоскость, отвечают неустойчивым фильтрам; они возникли при умножении передаточной функции $H(f)$ на сопряженную ей функцию $H^*(f)$. Так или иначе используются полюсы только из верхней полуплоскости, поскольку они определяют устойчивый фильтр.

3. На этом наиболее сложном этапе определяются решения α_m и β_m уравнения

$$\sin\left(\frac{\beta_m + j\alpha_m}{2}\right) = a_m + jb_m,$$

которые будут найдены ниже. Решение этого уравнения требует привлечения комплексного арксинуса.

4. После нахождения полюсов (α_m, β_m) , $m=1, \dots, M$, их подставляют в квадратные трехчлены вида

$$1 - 2 \exp(-\alpha_m) \cos \beta_m z^{-1} + \exp(-2\alpha_m) z^{-2},$$

чтобы определить коэффициенты

$$a_{1m} = 2 \exp(-\alpha_m) \cos \beta_m, \quad a_{2m} = -\exp(-2\alpha_m),$$

где $m=1, \dots, M/2$ и M — четное. В случае нечетного M вычисления для $m=1, \dots, (M-1)/2$ производятся так же, а для $m=(M+1)/2$ формула для многочлена меняется.

Детали этого метода будут рассматриваться параллельно с программой для вычисления весов фильтра.

Сначала запишем уравнение $D(f) = 0$ в виде

$$s^{2M} + 1 = 0,$$

где $s = \sin(\pi fT) / \sin(BT)$.

Это уравнение имеет $2M$ решений. Из них M (попавших в нижнюю полуплоскость) не используются вовсе, а из оставшихся непосредственно используется только $M/2$ решений.

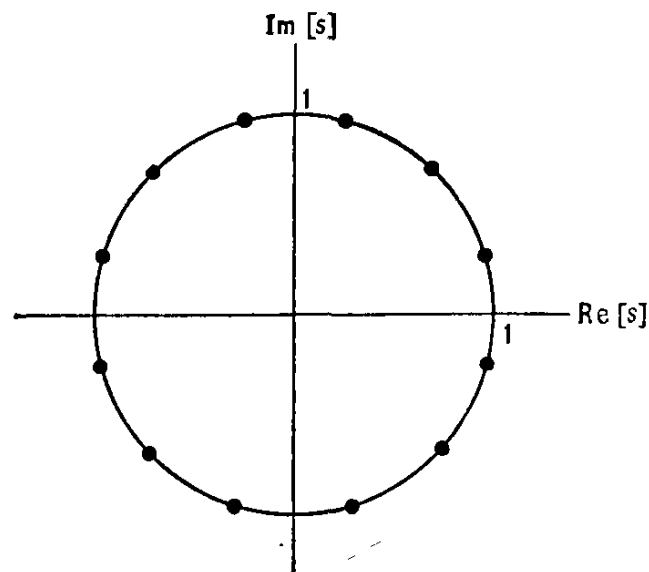


Рис. 4.9. Корни уравнения $s^{2M} + 1 = 0$ при $M = 6$. Обратите внимание на равные расстояния между корнями; ни одного действительного корня нет.

Решения этого уравнения всегда мнимые или комплексные. На рис. 4.9 показано расположение корней при $M = 6$. Как видно из рисунка, все корни располагаются на единичной окружности. Эти числа можно получить с помощью программы

```
SECTOR=3.14159265/FLOAT(M)
WEDGE=SECTOR/2.
DO 5 I=1,M1
FN=I-1
ANG=FN*SECTOR+WEDGE
:
```

Параметр $M1$ выбирается равным $M/2$ в случае четного M и равным $(M-1)/2$ в случае нечетного M . В следующей программе

последовательно вычисляются углы в радианах:

	1
WEDGE	1
SECTOR+WEDGE	2
2.*SECTOR+WEDGE	3
⋮	
FLOAT(M/2-1)*SECTOR+WEDGE	M1

Например, в случае $M=6$ эти значения, переведенные в градусы, образуют последовательность углов 15° , 45° и 75° .

Обозначим полученные корни $s_1, \dots, s_{M/2}$. На следующем этапе необходимо найти такие α_m и β_m , что

$$\sin\left(\frac{\beta_m + j\alpha_m}{2}\right) = \sin(\pi BT) s_m.$$

Вычисление комплексного арксинуса. Предположим, что рассматривается уравнение, подобное только что приведенному:

$$\sin(u + jv) = a + jb,$$

где a и b известны, а u и v требуется определить. Прежде всего применим формулу для синуса суммы углов:

$$\sin(u + jv) = \sin u \cos jv + \cos u \sin jv.$$

Поскольку $\sin j\theta = j \operatorname{sh} \theta$ и $\cos j\theta = \operatorname{ch} \theta$, после очевидных преобразований получим

$$\left(1 - \frac{a^2}{\operatorname{ch}^2 v}\right) \operatorname{sh}^2 v = b^2$$

или

$$(\operatorname{ch}^2 v - a^2) \operatorname{sh}^2 v = b^2 \operatorname{ch}^2 v.$$

Применяя теперь тождество $\operatorname{ch}^2 v - \operatorname{sh}^2 v = 1$, приходим к квадратному уравнению

$$\operatorname{sh}^4 v + (1 - a^2 - b^2) \operatorname{sh}^2 v - b^2 = 0.$$

Вводя обозначение C для коэффициента $c = 1 - a^2 - b^2$, перепишем уравнение в виде

$$\operatorname{sh}^4 v + C \operatorname{sh}^2 v - b^2 = 0.$$

Можно показать, что для реализации фильтра Баттеруорта коэффициент C зависит только от частоты, соответствующей точке отсечки фильтра. Полученное уравнение решается относительно $\operatorname{sh}^2 v$:

$$\operatorname{sh}^2 v = \frac{-C \pm \sqrt{C^2 + 4b^2}}{2}.$$

Для удобства, а также для последующего перехода к программе обозначим один корень буквой D , т. е.

$$D = \frac{-C + \sqrt{C^2 + 4b^2}}{2}$$

(необходим только вариант с плюсом). Поскольку

$$\operatorname{sh}^2 v = \left(\frac{e^v - e^{-v}}{2} \right)^2 = D,$$

извлекая квадратный корень из обеих частей, получим новое уравнение

$$e^v - e^{-v} = 2\sqrt{D}.$$

Если умножить его на e^v , то получится уравнение

$$e^{2v} - 2\sqrt{D} e^v - 1 = 0,$$

решив которое относительно e^v , получим корни

$$e^v = \frac{2\sqrt{D} \pm \sqrt{4D+4}}{2} = \sqrt{D} \pm \sqrt{D+1}.$$

Определим параметр E равенством

$$E = e^v = \sqrt{D} + \sqrt{D+1}.$$

Очевидно, что для действительных значений v нужно выбирать знак плюс. В этом случае v можно найти из уравнения

$$v = \ln E = \ln [\sqrt{D} + \sqrt{D+1}].$$

Поскольку $\cos u \operatorname{sh} v = b$, имеем соотношение

$$u = \arccos \left(\frac{b}{\operatorname{sh} v} \right).$$

Вспоминая, что $\operatorname{sh}^2 v = D$, получим решение

$$u = \arccos \left(\frac{b}{\sqrt{D}} \right).$$

Фортран-программа для фильтра низких частот. Рассмотренный выше метод для комплексного арксинуса можно применить в случае уравнения

$$\sin \left(\frac{\beta_m + j\alpha_m}{2} \right) = \sin (\pi BT) s_m,$$

которое решается относительно α_m и β_m .

Величина b_m для каждого корня s_m вычисляется по формуле

$$b_m = \sin (\pi BT) \operatorname{Im} [s_m] = \sin (\pi BT) \sin \theta_m,$$

где угол θ_m есть переменный параметр ANG, определенный выше (величину a_m вычислять не нужно). В программе величина $\sin(\pi BT)$ заменяется параметром FACT, поэтому

$$B = \text{FACT} \cdot \text{SIN}(\text{ANG})$$

где параметр FACT вычисляется заранее по формуле

$$\text{FACT} = \text{SIN}(3.14159265 \cdot \text{BW} \cdot \text{T})$$

(здесь BW — частота отсечки B, а T — выборочный интервал в секундах).

Полученное выше решение, выраженное через α и β , примет вид

$$\frac{\beta + j\alpha}{2} = u + jv.$$

Очень важно, что в знаменателе стоит двойка. Остается определить на каждом шаге коэффициенты

$$\begin{aligned} a_1 &= -2 \exp(-\alpha) \cos \beta = -2 \exp(-2v) \cos 2u = \\ &= -2 \exp(-2 \ln E) \cos \left[2 \arccos \left(\frac{B}{\sqrt{D}} \right) \right] = -2E^{-2} \left(\frac{2B^2}{D} - 1 \right), \\ a_2 &= \exp(-2\alpha) = \exp(-2v) = E^{-4}. \end{aligned}$$

Соберем все необходимые для вычислений формулы:

$$\begin{aligned} A &= \text{FACT} \cdot \cos(\text{ANG}), \\ B &= \text{FACT} \cdot \sin(\text{ANG}), \\ C &= 1 - A^2 - B^2 = 1 - (\text{FACT})^2, \\ D &= \frac{-C + \sqrt{C^2 + 4B^2}}{2}, \\ E &= \sqrt{D} + \sqrt{D+1}, \\ a_1 &= -2E^{-2} \left(\frac{2B^2}{D} - 1 \right), \\ a_2 &= E^{-4}. \end{aligned}$$

Из них видно, что для реализации фильтра Баттеруорта параметр A не нужен. Все эти соотношения, записанные в виде программы на Фортране, представлены на рис. 4.10. По сравнению с той длинной процедурой, которая потребовалась для получения необходимых для вычисления формул, сама программа оказывается гораздо короче.

Сравните эту программу с аналогичной подпрограммой LPSB, приведенной в приложении А. Они совершенно одинаковы, за

исключением того, что в подпрограмме LPSB задана двойная точность.

На рис. А.11 в приложении А приведен тест для случая $M=5$. Скорость выборки равна 200 выб/с, поэтому $T=0.005$, а $F=100$. Частота отсечки выбирается равной 5 Гц. Она состав-

```

M1 = M/2
FACT = SIN(3.14159265*BW*T)
DØ 5 = I-1 M1
FN = I-1
B = FACT*SIN(FN*SECTØR+WEDGE)
C = 1. - FACT*FACT
D = 0.5*(-C+SQRT(C*C+4*B*B))
E = SQRT(D+1.)+SQRT(D)
A1(I) = -2.*(E**-2)*(2.*B*B/D-1.)
5 A2(I) = E**-4

```

Рис. 4.10. Фортран-программа для получения весов фильтра.

ляет 5% частоты F . Оказывается, что исчисление в процентах удобнее непосредственного задания величины B . Например, величины B и S , приведенные ниже в таблице, приводят к одним и тем же весам фильтров:

B	S	F	%
0.5	20	10	5
1	40	20	5
5	200	100	5
100	4000	2000	5
500	20000	10000	5

Причина этого кроется в том, что B здесь выбирается всегда равным 5% от F , т.е. их соотношение остается все время постоянным.

При $M=5$ в фильтр входят три составляющие: две из них с двойными комплексными полюсами, а оставшаяся — с одним.

Постоянный множитель b_0 для всех трех составляющих остается одинаковым. Посмотрим, как это получается. Запишем формулу для передаточной функции рассматриваемого фильтра в виде

$$H(f) = \prod_{p=1}^3 \frac{b_0}{1 + a_{1p} \exp(-j2\pi fT) + a_{2p} \exp(-j4\pi fT)}$$

Полоса пропускания полагается равной диапазону от 0 до f_c Гц. Естественно предположить, что коэффициент усиления в этой полосе равен единице. В частности, должно иметь место равенство $H(0) = 1$. Так как $\exp(0) = 1$, из формулы для передаточной

функции получим соотношение

$$1 = b_0^3 / \prod_{p=1}^3 (1 + a_{1p} + a_{2p}),$$

из которого следует, что для выбранных в тесте параметров имеет место соотношение

$$\begin{aligned} b_0 &= \sqrt[3]{\prod_{p=1}^3 (1 + a_{1p} + a_{2p})} = \\ &= \sqrt[3]{(1 + a_{11} + a_{21})(1 + a_{12} + a_{22})(1 + a_{13})} = 0.041955. \end{aligned}$$

График коэффициента усиления передаточной функции фильтра в децибелах приведен на рис. А.11, б. Выделим следующие значения:

f	$10 \log_{10} H(f) ^2$	$ H(f) ^2$
0	0.0000	1.0
5	-3.0103	0.5
100	-110.5357	$1/1.1313 \cdot 10^{11}$

Точку половинной мощности иногда называют „3-х децибеловой“ точкой, поскольку в ней коэффициент усиления уменьшается по сравнению со значением в полосе пропускания примерно на 3 дБ.

4.7. СИЛУСНЫЙ ФИЛЬТР ВЫСОКИХ ЧАСТОТ БАТТЕРУОРТА

Фильтр высоких частот можно получить почти таким же способом, каким получают фильтр низких частот. В этом случае необходимо реализовать передаточную функцию

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\sin(\pi fT + \pi/2)}{\sin(\pi BT)} \right)^{2M}}.$$

Критические значения для этой функции сведены в таблицу

f	$H(f)^2$
0	$< \frac{1}{2}$
$F - B$	$\frac{1}{2}$
F	1

Таким образом, в диапазоне частот от 0 до $(F - B)$ Гц вся или почти вся информация гасится, а в диапазоне частот от $(F - B)$ до F Гц она проходит сравнительно беспрепятственно.

Кроме дополнительного угла $\pi/2$, эта передаточная функция ничем иным от передаточной функции фильтра низких частот

не отличается. Поэтому для получения фильтра высоких частот потребуются только небольшие изменения в предыдущем алгоритме. В частности, изменяется основная формула для углов:

$$\frac{\beta + j\alpha + \pi}{2} = u + jv,$$

т. е. $\beta = 2u + \pi$, $\alpha = 2v$. При этом угол α остался прежним. Веса a_{2p} получаются точно так же, поскольку их вычисление требует знания только величин α_p . Для вычисления же весов a_{1p} формулы нужно изменить:

$$\begin{aligned} a_1 &= -2 \exp(-\alpha) \cos \beta = -2 \exp(-\alpha) \cos \left(2u + \frac{\pi}{2}\right) = \\ &= -2 \exp(-\alpha) (\cos 2u \cos \pi - \sin 2u \sin \pi) = 2 \exp(-\alpha) \cos 2u. \end{aligned}$$

Единственное отличие от случая фильтра низких частот сводится к перемене знака у весов a_{1p} на противоположный. Другими словами, зная веса для фильтра низких частот с точкой отсечки B Гц, можно определить веса для фильтра высоких частот с точкой отсечки $(F - B)$ Гц, просто меняя знаки у коэффициентов a_{1p} . Заметим, что ширина полосы пропускания в обоих случаях одна и та же.

Оказывается, что постоянный множитель b_0 также остается прежним. Предлагаем читателю самостоятельно выяснить причину этого явления.

4.8. ПОЛОСНОПРОПУСКАЮЩИЕ ФИЛЬТРЫ

Полоснопропускающий фильтр предназначен для пропускания информации в полосе частот ширины B с центром в точке f_c .

Такой фильтр можно реализовать несколькими способами. Наверное, самый простой из них — тот, который изображен на рис. 4.11. Из него видно, что полосная фильтрация проводится при помощи двух фильтров низких частот, и последовательность $x(i)$ обрабатывается в двух ветвях схемы одновременно. В верхней происходит умножение $x(i)$ на $\sin(2\pi f_c iT)$, низкочастотная фильтрация и повторное умножение на $\sin(2\pi f_c iT)$. В нижней ветви происходит то же самое, но умножение производится на $\cos(2\pi f_c iT)$. В итоге на выходе получается функция

$$\begin{aligned} y(i) &= \sin(2\pi f_c iT) \{h(i) * [x(i) \sin(2\pi f_c iT)]\} + \\ &+ \cos(2\pi f_c iT) \{h(i) * [x(i) \cos(2\pi f_c iT)]\}, \end{aligned}$$

где $h(i)$ — единичная импульсная функция отклика, одинаковая для обоих фильтров. Если $h(i)$ есть дельта-функция, т. е.

$$h(i) = \begin{cases} 1/T, & i=0, \\ 0 & \text{в остальных случаях,} \end{cases}$$

то полное действие фильтра состоит в получении значений

$$\begin{aligned} y(i) &= \sin(2\pi f_c iT) [x(i) \sin(2\pi f_c iT)] + \\ &\quad + \cos(2\pi f_c iT) [x(i) \cos(2\pi f_c iT)] = \\ &= x(i) (\sin^2(2\pi f_c iT) + \cos^2(2\pi f_c iT)) = x(i). \end{aligned}$$

Таким образом, в диапазонах пропускания обоих фильтров действие полного фильтра сводится к действию одного из них.

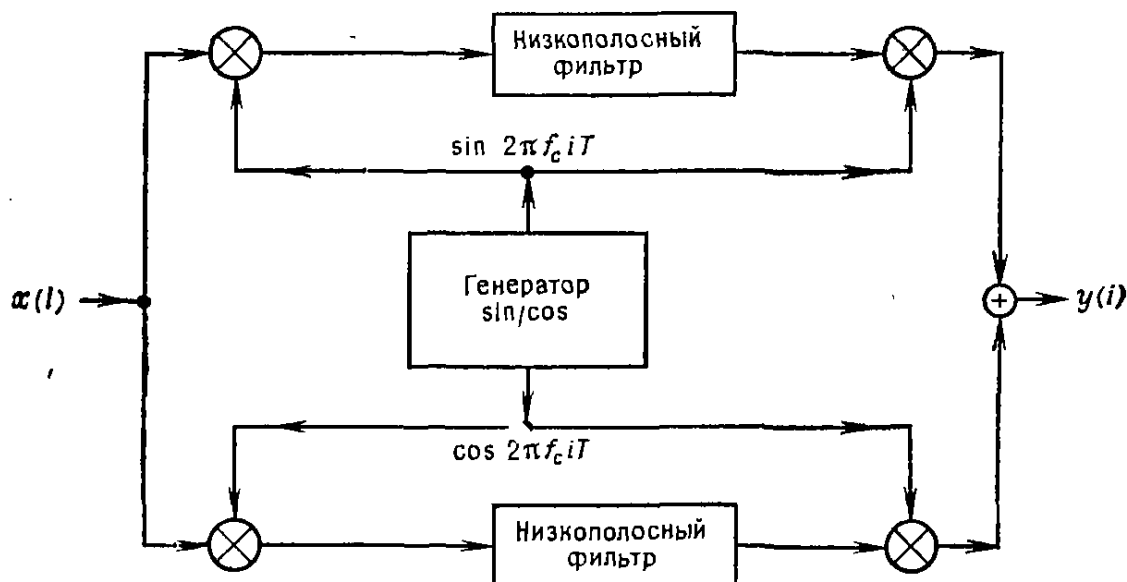


Рис. 4.11. Полоснопропускающий фильтр, реализованный посредством двух низкочастотных фильтров и генератора \sin/\cos .

В истинном смысле происходящего помогает разобраться переход в определении к комплексной записи:

$$y(i) = \exp(-j2\pi f_c iT) \{h(i) * [x(i) \exp(j2\pi f_c iT)]\}.$$

Внутренняя экспонента умножается на $x(i)$ во временной области. Как было показано в гл. 1, этому соответствует сдвиг преобразования $X(f)$ в области частот. В частности, преобразованием Фурье функции $x(i) \exp(j2\pi f_c iT)$ служит

$$\begin{aligned} T \sum_{i=-\infty}^{\infty} x(i) \exp(j2\pi f_c iT) \exp(-j2\pi f iT) &= \\ &= T \sum_{i=-\infty}^{\infty} x(i) \exp[-j2\pi (f - f_c) iT] = X(f - f_c). \end{aligned}$$

Таким образом, информация, располагавшаяся в частоте f_c или около нее, смещается в 0 Гц. Эту информацию два фильтра низких частот пропускают, а выходящую за пределы нужного диапазона частот удаляют. Наконец, умножение на $\exp(-j2\pi f_c iT)$ возвращает полученную информацию в первоначальное положение.

Чтобы полоса пропускания получилась равной B Гц, точки половинной мощности обоих низкочастотных фильтров следует положить равными $B/2$ Гц.

Синус и косинус можно получать несколькими способами. Для ЭВМ, имеющих слова больших размеров, может оказаться вполне достаточной рекурсивная процедура, рассмотренная в § 4.3. Но более надежной процедурой служит другая. В ней сначала получают последовательность углов $\lambda(i) = 2\pi f_c i T$, последовательно прибавляя $2\pi f_c T$ к предыдущему углу: $\lambda(i) = \lambda(i-1) + 2\pi f_c T$; каждый раз, когда угол $\lambda(i)$ превосходит 2π , из него вычитается 2π (т. е. значения углов берутся по модулю 2π); соответствующий углу $\lambda(i)$ индекс в таблице с 2048 точками для одного периода синусоиды определяется по формуле $\lambda(i) 2048 / (2\pi)$. На практике, как правило, ограничиваются только первой четвертью периода синусоиды (т. е. 513 точками в нашем примере), а по величине $\lambda(i) 2048 / (2\pi)$ определяют ту четверть синусоиды, в которую угол $\lambda(i)$ попадает фактически.

Лучше всего, пожалуй, предложенный алгоритм подходит в тех случаях, когда коэффициенты заданы аппаратно, как это бывает, например, в микропроцессоре, и для определения весов фильтра программных возможностей нет. Поэтому, чтобы получить полоснопропускающие фильтры, нередко оказывается достаточным этого алгоритма и нескольких наборов весов для фильтров низких частот.

Что же касается программной реализации фильтра такого типа, то ее можно получить почти так же как и фильтр низких частот. Передаточная функция в этом случае выглядит иначе:

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\cos(2\pi f T) - C}{S} \right)^M},$$

где

$$C = \cos(2\pi f_c T) \cos(\pi B T), \quad S = \sin(2\pi f_c T) \sin(\pi B T).$$

Читатель без труда проверит, что

$$\left| H\left(f_c - \frac{B}{2}\right) \right|^2 = \left| H\left(f_c + \frac{B}{2}\right) \right|^2 = \frac{1}{2},$$

и что максимум передаточной функции достигается в частоте f_m , определяемой равенством

$$\cos(2\pi f_m T) = \cos(2\pi f_c T) \cos(\pi B T).$$

Для малых значений B это равенство эквивалентно соотношению $f_m \approx f_c$. Поэтому полоса пропускания простирается от $(f_c - B/2)$ до $(f_c + B/2)$ Гц.

Отметим также, что в передаточной функции вместо $\sin(\pi f T)$ появился $\cos(2\pi f T)$ и, что еще более важно, появился множи-

тель 2 для угла. Кроме того, в показателе степени вместо $2M$ теперь используется просто M , которое, как оказывается, для фильтров низких частот и для вырезающих фильтров должно быть четным.

Предложенный в предыдущем параграфе алгоритм можно приспособить с некоторыми изменениями и для получения полоснопропускающих фильтров. Сначала знаменатель $|H(f)|^2$ полагается равным нулю и находятся полюсы. Для удобства заменим косинус синусом с дополнительным углом:

$$1 + \left(\frac{\sin(2\pi fT + \pi/2) - C}{S} \right)^M = 0.$$

Это уравнение можно переписать в виде $s^M + 1 = 0$, где

$$s = \frac{\sin(2\pi fT + \pi/2) - C}{S}.$$

У этого уравнения есть M решений. Хотя из них используется только половина, т. е. $M/2$ решений, в самой реализации фильтра будут использованы пары полюсов, определенные каждым из $M/2$ решений.

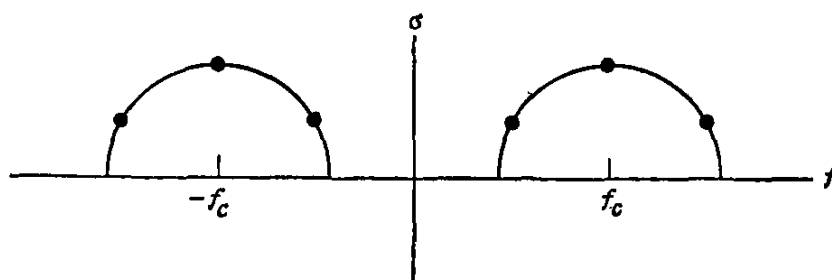


Рис. 4.12. Расположение полюсов шестиполюсного полоснопропускающего фильтра.

Как это происходит, показано на рис. 4.12, где изображены полюсы для шестиполюсного фильтра ($M = 6$). Уравнение $s^6 + 1 = 0$ имеет 6 решений, но из них используется только три; появление других трех, отраженных относительно оси σ , обусловлено членами C и S .

Итак, получаем основное уравнение

$$\sin \left[\left(\beta_m + \frac{\pi}{2} \right) + j\alpha_m \right] = Ss_m + C = a_m + jb_m.$$

Оно имеет точно такой же вид, как и в случае фильтра низких частот. Отличие состоит лишь в том, что в этом уравнении фигурирует угол $\pi/2$ и что a_m и b_m стали теперь другими. В программе на Фортране эти изменения учитываются следующим образом.

1. Вычисляются C и S . При этом в программе их лучше назвать CC и SS , поскольку C уже в программе использовалось:

```
ANG2=2.*3.14159265*FC*T
CC=COS(ANG2)*COS(FACT)
SS=SIN(ANG2)*SIN(FACT)
```

Переменная FC соответствует f_c . Переменная $FACT$ остается без изменений.

2. Угол $SECTOR$ умножается на 2. Теперь он вычисляется с помощью оператора

```
SECTOR=2.*3.14159265/FLOAT(M)
```

3. Вычисление переменных A , B и C изменяется следующим образом:

```
ANG=FN*SECTOR+WEDGE
A=SS*COS(ANG)+CC
B=SS*SIN(ANG)
C=1.-A*A-B*B
```

4. Вычисление переменных G и H (из программы $LPSB$) нужно изменить следующим образом:

```
G=2.*SQRT(1.-B*B/D)/E
IF(A.LT.0.)G=-G
H=-1./(E**2)
```

На рис. 4.13 показана программа $LPSB$ с внесенными уточнениями. Для нее выбрано более подходящее в этом случае название $BNPS$ ¹⁾.

Изменение 4, приведенное только что, нуждается в некоторых пояснениях. Вспомним, что

$$a_{1m} = -2 \exp(-\alpha_m) \cos \beta_m \quad a_{2m} = \exp(-2\alpha_m).$$

Вместо формулы

$$\frac{\beta_m + j\alpha_m}{2} = u_m + jv_m$$

в этом случае используется формула

$$\beta_m + \frac{\pi}{2} + j\alpha_m = u_m + jv_m.$$

¹⁾ От английского *bandpass*. — Прим перев.

```

SUBROUTINE BNPS (MM,T,BW,FC,A1,A2,BZERO)
DIMENSION A1(1),A2(1),B0(1),B1(1),B2(1)
1,FREQ(1),ABZ(1),PHS(1)
DDOUBLE PRECISION A,B,C,D,E,G,H,GN,FACT,WEDGE,SECTOR,ANG
1,CC,SS,ANG2
DATA B0 /1./,B1 /0./, B2 /0./
FACT=2.*3.14159265*T*BW
ANG2=2.*3.14159265*T*FC
CC=DCOS(ANG2)*DCOS(FACT)
CCC=CC
FREQ(1)=ATAN2(SQRT(1.-CCC*CCC),CCC)/6.2831853
SS=DSIN(ANG2)*DSIN(FACT)
M=MM
M1=M/2
A=M
HTRAN=0.
SECTOR=2.00*3.1415926500/A
WEDGE=SECTOR/2.00
DO 5 I=1,M1
FN=I-1
ANG=FN*SECTOR+WEDGE
A=SS*DCOS(ANG)+CC
B=SS*DSIN(ANG)
C=1.00-(A**2+B**2)
D=0.5*(-C+DSQRT(C*C+4.00*B**2))
E=DSQRT(D+1.00)+DSQRT(D)
G=2.00*DSQRT(1.00-B*B/D)/E
IF (A.LT.0.00) G=-G
H=-1.00/(E**2)
A1(I)=-G
A2(I)=-H
CALL TTRAN (A1(I),A2(I),B0,B1,B2,2,1,1.,FREQ,ABZ,PHS)
HTRAN=HTRAN+ABZ(1)
CONTINUE
BZERO=10.**(-HTRAN/(20.*FLOAT(M1)))
RETURN
END

```

Рис. 4.13. Программа вычисления весов синусного полоснопропускающего фильтра. FC -- центральная частота фильтра, величина BW равна половине ширины полосы пропускания. Обратите внимание на то, что вычисления в программе имеют двойную точность.

Поэтому, опуская индекс m , получаем

$$\beta = u - \frac{\pi}{2}, \quad \alpha = v.$$

Следовательно, формулы для коэффициентов изменятся следующим образом:

$$\begin{aligned}
 a_1 &= -2 \exp(-\alpha) \cos \beta = -2 \exp(-v) \cos \left(u - \frac{\pi}{2} \right) = \\
 &= 2 \exp(-\ln E) \sin u = 2E^{-1} \sin \left[\arccos \left(\frac{B}{\sqrt{D}} \right) \right] = \\
 &= \pm 2E^{-1} \sqrt{1 - \cos^2 \left[\arccos \left(\frac{B}{\sqrt{D}} \right) \right]} = \pm 2E^{-1} \sqrt{1 - \frac{B^2}{D}}, \\
 a_2 &= \exp(-2\alpha) = E^{-2}.
 \end{aligned}$$

В программе BNPS есть и другие небольшие отличия. Или в самой частоте F_c , или около нее, а именно в точке $f_m = \arccos(CS)/(2\pi T)$ передаточная функция достигает максимума. Для передаточной функции каждой составной двухполюсной части фильтра в программе BNPS с помощью подпрограммы TTRAN отдельно вычисляется значение в децибелах при f_m ; затем эти значения суммируются для определения общего постоянного множителя b_0 .

4.9. ЗАГРАЖДАЮЩИЕ ФИЛЬТРЫ

В этом параграфе будет рассмотрен заграждающий фильтр особого типа. В силу этого результаты носят частный характер и не всегда точно описывают свойства других фильтров заграждающего типа.

Фильтр заграждающего типа — близкий родственник полосно-пропускающему фильтру. Как позже выяснится, полюсы обоих фильтров совпадают. Но между всеми предшествующими фильтрами и заграждающим фильтром имеется существенное различие: помимо полюсов заграждающий фильтр имеет нули.

Получить передаточную функцию заграждающего фильтра нетрудно. Для этого следует из единицы вычесть выражение для передаточной функции полоснопропускающего фильтра:

$$|H(f)|^2 = 1 - \frac{1}{1 + \left(\frac{\cos(2\pi fT) - C}{S}\right)^M}.$$

Если это вычитание провести, то получится передаточная функция

$$|H(f)|^2 = \left(\frac{\cos(2\pi fT) - C}{S}\right)^M / \left(1 + \left(\frac{\cos(2\pi fT) - C}{S}\right)^M\right).$$

Необходимо обратить внимание на следующее.

1. Эта передаточная функция имеет такой же знаменатель, что и передаточная функция полоснопропускающего фильтра. Следовательно, полюсы этих фильтров одинаковы.

2. Передаточная функция удовлетворяет условию $|H(f)|^2 \geq 0$. В противном случае ее рассматривать было бы неправомерно.

3. Точки половинной мощности для полосы заграждения точно совпадают с точками половинной мощности полоснопропускающего фильтра.

Передаточная функция полоснопропускающего фильтра и соответствующая ей передаточная функция заграждающего фильтра показаны на рис. 4.14.

Как уже отмечалось, заграждающий фильтр имеет нули. До сих пор все рассматривавшиеся фильтры допускали каскад-

ную реализацию, с частям которой отвечали уравнения вида

$$y(i) = b_0x(i) - a_1y(i-1) - a_2y(i-2).$$

Заграждающему фильтру отвечает разностное уравнение вида

$$y(i) = b_0x(i) + b_1x(i-1) + b_2x(i-2) - a_1y(i-1) - a_2y(i-2),$$

поэтому задача осложняется точным определением величин b_0 , b_1 и b_2 для каждой части фильтра. Впрочем, сделать это нетрудно.

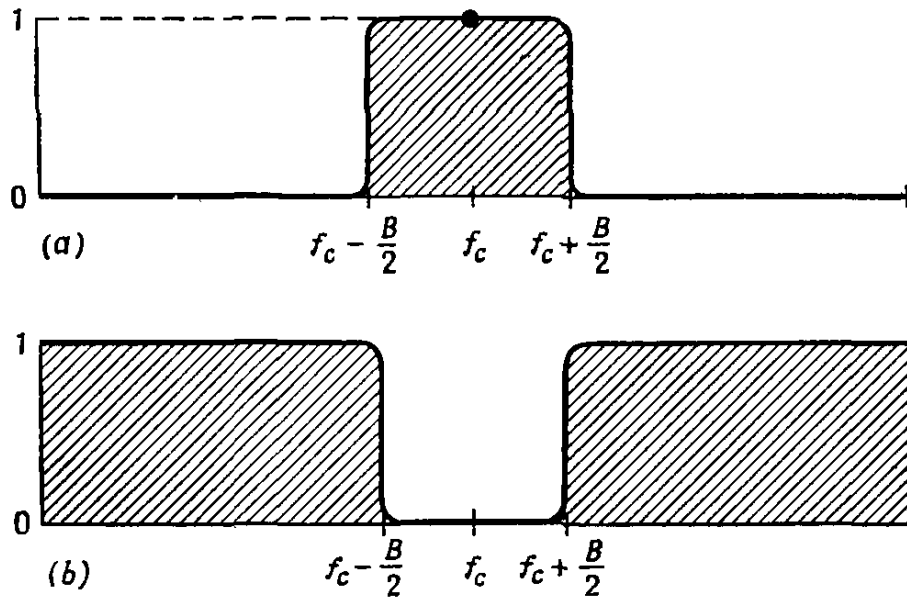


Рис. 4.14. Полоснопропускающий фильтр (а) и полученный из него полоснозаграждающий фильтр (б).

Так же, как и полюсы, нули получают, приравнивая определяющую фильтр передаточную функцию нулю и находя значения, удовлетворяющие полученному уравнению. В данном случае уравнение имеет вид

$$\left(\frac{\cos(2\pi fT) - C}{S} \right)^M = 0.$$

Найти его решения совсем просто. Все корни этого уравнения равны, и для всех нулей решение будет одним и тем же, а именно

$$\cos(2\pi f_m T) = C = \cos(2\pi f_c T) \cos(\pi BT),$$

т. е. $2\pi f_m T = \pm \arccos C$. Знаки плюс и минус появились из-за того, что $\cos(-x) = \cos x$. Поэтому решения для f_m могут отличаться знаками, хотя решения исходного уравнения и одинаковы.

На следующем шаге определяются значения b_0 , b_1 и b_2 . Для одного нуля (на каждую часть каскадной реализации приходится по два нуля) имеем равенство

$$1 - bz_m^{-1} = 0,$$

поэтому

$$b = z_m = \exp(j2\pi T f_m) = \exp(\pm j \arccos C).$$

Поскольку для одной части каскадной реализации частота f_m может иметь два знака, общие коэффициенты получаются из соотношений

$$\begin{aligned} & (1 - bz^{-1})(1 - b^*z^{-1}) = \\ & = [1 - \exp(j \arccos C) z^{-1}][1 - \exp(-j \arccos C) z^{-1}] = \\ & = 1 - [\exp(j \arccos C) + \exp(-j \arccos C)] z^{-1} + \\ & \quad + [\exp(j \arccos C) (-j \arccos C)] z^{-2} = \\ & = 1 - 2 \cos[\arccos C] z^{-1} + \exp(0) z^{-2} = \\ & = 1 - 2Cz^{-1} + z^{-2}. \end{aligned}$$

Из полученного выражения с точностью до постоянного множителя получаются величины b_0 , b_1 и b_2 :

$$b_0 = 1, \quad b_1 = -2C, \quad b_2 = 1.$$

Этот результат справедлив для каждой части фильтра второго порядка. Естественно, эти величины при нормировке претерпевают изменения. При нормировке сначала вычисляется значение передаточной функции фильтра с полученными выше коэффициентами в точке $f=0$ или $1/(2T)$. Предположим, что это значение равно $|H'(0)|^2$, где штрих указывает на то, что фильтр ненормирован. Затем требуют, чтобы для нормированного фильтра имело место равенство $|H(0)|^2 = 1$. Этого можно добиться, полагая

$$|H(f)|^2 = \left[\frac{1}{|H'(0)|^2} \right] |H'(f)|^2,$$

т. е. умножая передаточную функцию ненормированного фильтра на постоянную $1/|H'(0)|^2$, которая обеспечивает выполнение равенства $|H(0)|^2 = 1$. В итоге коэффициентами для каждой каскадной части будут служить

$$b_0 = 1 \frac{1}{|H'(0)|^{2/M}}, \quad b_1 = \frac{-2C}{|H'(0)|^{2/M}}, \quad b_2 = \frac{1}{|H'(0)|^{2/M}}.$$

Значение $|H'(0)|^2$ нетрудно вычислить, применив подпрограмму TTRAN, включенную в приложение А. При этом следует величины b_0 , b_1 и b_2 положить равными 1, $-2C$ и 1. Кроме того, для получения заграждающего фильтра не требуется написания особой программы: чтобы получить рекурсивные части фильтра¹⁾, можно воспользоваться программой получения полоснопропу-

¹⁾ Рекурсивная часть уравнения фильтрации — входящая в него линейная комбинация значений выходов, иерекурсивная часть — линейная комбинация входов. — Прим. перев.

скающих фильтров, а для определения весов нерекурсивных членов нужно непосредственно обратиться к приведенным выше формулам.

4.10. ТАНГЕНСНЫЕ ФИЛЬТРЫ

Заменяв в синусных фильтрах синус на тангенс, можно получить совершенно новое семейство фильтров. Например, при такой замене передаточная функция синусного фильтра низких частот Баттеруорта будет иметь вид

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\operatorname{tg}(\pi fT)}{\operatorname{tg}(\pi BT)} \right)^{2M}}.$$

Вычисление значений этой функции в точках 0, B и F дает следующие значения:

$$|H(0)|^2 = 1, \quad |H(B)|^2 = \frac{1}{2}, \quad |H(F)|^2 = 0.$$

Если эту передаточную функцию изобразить в виде графика, то внешне он будет очень напоминать график для синусного фильтра Баттеруорта с теми же значениями B и M . В деталях графики будут различаться. Если перейти к графикам для функций в децибелах, то можно увидеть, что тангенсный фильтр в полосе заграждения имеет лучшие характеристики: передаточная функция окажется в ней значительно ниже передаточной функции синусного фильтра.

Такое улучшение характеристики сопряжено с ростом стоимости. Дело в том, что все тангенсные фильтры имеют нули. Поэтому вычисления становятся более дорогими. Подробнее на этом мы остановимся ниже.

Тангенсные фильтры можно реализовать несколькими способами. Рассмотрим здесь два из них: метод определения нулей и полюсов и метод двустороннего z -преобразования. Сначала обратимся к определению полюсов; оказывается, что нули найти очень просто.

В случае фильтра низких частот Баттеруорта для отыскания полюсов необходимо решить уравнение

$$1 + \left(\frac{\operatorname{tg}(\pi fT)}{\operatorname{tg}(\pi BT)} \right)^{2M} = 0.$$

Обозначим для него символами $s_1, \dots, s_{M/2}$ корни из единицы на верхней полуплоскости, а искомые решения обозначим $\beta_m + j\alpha_m$. Таким образом, нужно решить уравнение

$$\operatorname{tg} \left(\frac{\beta_m + j\alpha_m}{2} \right) = \operatorname{tg}(\pi BT) s_m.$$

Это можно сделать, непосредственно применив полученные ранее результаты для комплексного арксинуса. При этом понадобятся незначительные изменения. Поскольку $\operatorname{tg} x = \sin x / \cos x$, предыдущее выражение можно записать в виде

$$\frac{\sin ((\beta_m + j\alpha_m)/2)}{\cos ((\beta_m + j\alpha_m)/2)} = \operatorname{tg} (\pi BT) s_m.$$

От косинуса можно избавиться, возведя обе части в квадрат и заменив после этого $\cos^2 x$ на $1 - \sin^2 x$. Теперь остается только решить уравнение относительно синуса.

Значительно более простой метод двустороннего z -преобразования предложен Кайзером (1963). В этом методе используется тот факт, что

$$\operatorname{tg} (\pi fT) = \frac{1}{j} \frac{1 - z^{-1}}{1 + z^{-1}}.$$

Свое название двустороннее z -преобразование получило от конформного отображения

$$y = \frac{1}{j} \frac{1 - 1/x}{1 + 1/x} = \frac{1}{j} \frac{x - 1}{x + 1}.$$

Его называют *дробно-линейным*¹⁾ *отображением* или *преобразованием*. Поскольку вместо $z^{-1/2}$, как это было в случае с синусом, в выражении для тангенса стоит z^{-1} , получить формулы для коэффициентов гораздо проще. Действительно, полагая

$$\operatorname{tg} (\pi f_m T) = \frac{1}{j} \frac{1 - z_m^{-1}}{1 + z_m^{-1}} = \operatorname{tg} (\pi BT) s_m = a_m + jb_m,$$

т. е.

$$\frac{1}{j} \frac{1 - z_m^{-1}}{1 + z_m^{-1}} = a_m + jb_m,$$

получаем для z_m^{-1} выражение

$$z_m^{-1} = \frac{(1 + b_m) - ja_m}{(1 - b_m) + ja_m}.$$

Поскольку должно выполняться соотношение

$$(1 - z_m z^{-1})(1 - z_m^* z^{-1}) = 1 + a_{1m} z^{-1} + a_{2m} z^{-2}$$

¹⁾ В англоязычной литературе это преобразование называют билинейным. — Прим. перев.

для коэффициентов a_{1m} и a_{2m} получим формулы

$$\begin{aligned} a_{1m} &= -z_m - z_m^* \frac{(1-b_m) + ja_m}{(1+b_m) - ja_m} \frac{(1-b_m) - ja_m}{(1+b_m) + ja_m} = \\ &= -2 \frac{(1-b_m^2) - a_m^2}{(1+b_m)^2 + a_m^2}, \\ a_{2m} &= z_m z_m^* = \frac{(1-b_m)^2 + a_m^2}{(1+b_m)^2 + a_m^2}. \end{aligned}$$

Запишем эти действия в виде алгоритма.

1. Задать число полюсов фильтра M и частоту отсечки B .
2. Вычислить числа a_m и b_m :

```

ANG=3.14159265*B*T
FACT=SIN(ANG)/COS(ANG)
M1=M-M/2
SECTOR=3.14159265/FLOAT(M)
WEDGE=SECTOR/2.
DO 5 I=1, M1
FN=I-1
ANG=FN*SECTOR+WEDGE
AM=FACT*SIN(ANG)
BM=FACT*COS(ANG)

```

3. Вычислить знаменатель:

```

AMS=AM*AM
DEN=(1.+BM)**2+AMS

```

4. Вычислить a_{1m} :

```

A1(I)=-2.*((1.-BM*BM)-AMS)/DEN

```

5. Вычислить a_{2m} :

```

A2(I)=((1.-BM)**2+AMS)/DEN

```

6. CONTINUE.

Целиком эта программа представлена на рис. 4.15 в виде подпрограммы LPTB. Последовательность команд обращения к этой подпрограмме такая же, как для LPSB. Осталось обсудить еще два пункта программы: отыскание нулей и нормирующего множителя.

Для того чтобы получить нули, числитель и знаменатель исходной передаточной функции

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\operatorname{tg}(\pi f T)}{\operatorname{tg}(\pi B T)} \right)^{2M}}$$

следует умножить на $\cos^{2M}(\pi fT)$:

$$|H(f)|^2 = \frac{\cos^{2M}(\pi fT)}{\cos^{2M}(\pi fT) + \frac{\sin^{2M}(\pi fT)}{\operatorname{tg}^{2M}(\pi fT)}}.$$

Поэтому нулями будут решения уравнения

$$\cos^{2M}(\pi fT) = 0,$$

которые равны $f = \pm F$.

```

SUBROUTINE LPTB (MM,T,B,A1,A2,BZERO)
DIMENSION A1(1),A2(1)
DOUBLE PRECISION ANG,FACT,SECTOR,WEDGE,FN,AM,BM,AMS,DEN
M=MM
ANG=3.14159265*B*T
FACT=DSIN(ANG)/DCOS(ANG)
M1=M-M/2
F=1.
FN=M
SECTOR=3.14159265D0/FN
WEDGE=SECTOR/2.00
DO 5 I=1,M1
FN=I-1
ANG=FN*SECTOR+WEDGE
AM=FACT*DSIN(ANG)
BM=FACT*DCOS(ANG)
AMS=AM*AM
DEN=(1.00+BM)**2+AMS
A1(I)=-2.00*((1.00-BM*BM)-AMS)/DEN
A2(I)=((1.00-BM)**2+AMS)/DEN
F=F*(1.+A1(I)+A2(I))/4.
5 CONTINUE
BZERO=F**(1./FLOAT(M1))
RETURN
END

```

Рис. 4.15. Подпрограмма LPTB для получения низкочастотных тангенсных фильтров Баттеруорта. Величина M ($=MM$) должна быть четной.

Так же, как и в случае синусного заграждающего фильтра, для одиночной составной части фильтра

$$b = z_m = \exp(j2\pi T f_m) = \exp(\pm j2\pi T F) = \exp(\pm j\pi) = -1.$$

Для остальных частей каскадной реализации имеем

$$(1 - bz^{-1})(1 - b^*z^{-1}) = (1 - z^{-1})(1 + z^{-1}) = 1 + 2z^{-1} + z^{-2},$$

т. е. b_0 , b_1 и b_2 пропорциональны 1, 2 и 1 соответственно.

Нормировка проводится тем же способом, что и раньше.

Тангенсный фильтр высоких частот. Передаточная функция тангенсного фильтра высоких частот имеет вид

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\operatorname{ctg}(\pi fT)}{\operatorname{tg}(\pi BT)}\right)^{2M}}.$$

Получим ее значения в точках 0, $(F - B)$ и F :

$$\begin{aligned} |H(0)|^2 &= 0, \\ |H(F - B)|^2 &= \frac{1}{1 + \left(\frac{\operatorname{tg}(\pi BT)}{\operatorname{tg}(\pi fT)}\right)^{2M}} = \frac{1}{2}, \\ |H(F)|^2 &= 1. \end{aligned}$$

Интересно сравнить упрощенные записи передаточных функций фильтров низких и высоких частот. В случае низких частот передаточная функция имеет вид

$$\frac{(\sin(\pi BT) \cos(\pi fT))^{2M}}{(\sin(\pi BT) \cos(\pi fT))^{2M} + (\cos(\pi BT) \sin(\pi fT))^{2M}},$$

а в случае высоких

$$\frac{(-\cos(\pi BT) \sin(\pi fT))^{2M}}{(\sin(\pi BT) \cos(\pi fT))^{2M} + (-\cos(\pi BT) \sin(\pi fT))^{2M}}.$$

С точностью до знака полюсы этих фильтров совпадают. Из этого следует важный вывод: *за исключением разницы в знаках, коэффициенты рекурсивных фильтров низких и высоких частот одинаковы*. Теперь рассмотрим передаточные функции при $B = F/2$. При этом значении точки отсечки передаточные функции примут значения

$$\frac{(\sin(\pi/4) \cos(\pi fT))^{2M}}{(\sin(\pi/4) \cos(\pi fT))^{2M} + (\cos(\pi/4) \sin(\pi fT))^{2M}} = \frac{(\cos(\pi fT))^{2M}}{(\cos(\pi fT))^{2M} + \sin(\pi fT)^{2M}}$$

для фильтра низких частот и

$$\begin{aligned} &\frac{(-\cos(\pi/4) \sin(\pi fT))^{2M}}{(\sin(\pi/4) \cos(\pi fT))^{2M} + (-\cos(\pi/4) \sin(\pi fT))^{2M}} = \\ &= \frac{(\sin(\pi fT))^{2M}}{(\cos(\pi fT))^{2M} + (-\sin(\pi fT))^{2M}} \end{aligned}$$

для фильтра высоких частот. В этом частном случае полюсы будут решениями уравнения

$$\cos(\pi fT) = s_m \sin(\pi fT),$$

где s_m — корень из единицы. Важная особенность, связанная с этим случаем, состоит в том, что $a_{1m} = 0$ для всех m . Это непосредственно следует из того, что $a_{1m} > 0$ при $B < F/2$ и $a_{1m} < 0$ при $B > F/2$. Поэтому коэффициенты a_{1m} должны быть равными нулю при $B = F/2$.

Правило выбора знаков коэффициентов тангенсных фильтров низких и высоких частот можно сформулировать следующим образом. Предположим сначала, что a_{1m} и a_{2m} ($m = 1, \dots, M/2$) — коэффициенты для фильтра низких частот в случае $B < F/2$.

Тогда коэффициенты a_{2m} для фильтра высоких частот остаются теми же, а коэффициенты a_{1m} будут определяться таблицей

Случай низких частот		Случай высоких частот	
$B < F/2$	$B > F/2$	$(F-B) > \frac{F}{2}$	$(F-B) < \frac{F}{2}$
a_{1m}	$-a_{1m}$	$-a_{1m}$	a_{1m}

В обоих случаях веса ненормированных нерекурсивных фильтров определяются по таблице

Случай низких частот		Случай высоких частот	
b_{0m}	1		1
b_{1m}	2		-2
b_{2m}	1		1

Последнее свойство, указывающее на двойственность этих двух фильтров, заключается в том, что передаточную функцию одного из них можно получить, вычитая из единицы передаточную функцию другого. Интересно отметить, что причину этой двойственности нетрудно обнаружить, обращаясь к используемому конформному отображению

$$\operatorname{tg}(\pi f_m T) = \frac{1}{j} \frac{1 - z_m^{-1}}{1 + z_m^{-1}}.$$

Полоснопропускающие и заграждающие тангенсные фильтры. В качестве примера полоснопропускающего фильтра рассмотрим фильтр с передаточной функцией

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\cos(2\pi f T) - D}{E \sin(2\pi f T)} \right)^M},$$

где

$$D = \frac{\cos(2\pi f_c T)}{\cos(\pi B T)}, \quad E = \operatorname{tg}(\pi B T)$$

и M — четное число.

Эта функция имеет следующие характеристики:

f	$ H(f) ^2$	$10 \log_{10} H(f) ^2$
0	0	$-\infty$
$f_c - B/2$	$\frac{1}{2}$	-3.
f_{\max}	1	0
$f_c + B/2$	$\frac{1}{2}$	-3.
F	0	$-\infty$

Она принимает максимальное значение в частоте $f_{\max} = \arccos D$. Уравнение для полюсов сводится к соотношению

$$E \sin(2\pi f_m T) = s_m (\cos(2\pi f_m T) - D),$$

где, как и в случае синусного фильтра Баттеруорта, s_m — корень из единицы. Обе части этого уравнения можно возвести в квадрат и решить его относительно $\cos(2\pi f_m T)$:

$$\cos(2\pi f_m T) = \frac{D \pm E s_m \sqrt{1 + E^2 s_m^2 - D^2}}{1 + E^2 s_m^2}.$$

Напомним, что величины s_m комплексные. Поэтому до непосредственного перехода к решению уравнения необходимо сделать некоторые упрощения.

После упрощений уравнение принимает вид

$$\cos(2\pi f_m T) = a_m + j b_m.$$

Теперь можно обратиться к процедуре, которая использовалась для синусного фильтра Баттеруорта.

Заграждающий тангенсный фильтр получается так же, как и его синусный аналог. Для этого из единицы вычтем передаточную функцию полоснопропускающего тангенсного фильтра:

$$\begin{aligned} 1 - \frac{1}{1 + \left(\frac{\cos(2\pi f T) - D}{E \sin(2\pi f T)} \right)^M} &= \\ &= \left[1 + \left(\frac{\cos(2\pi f T) - D}{E \sin(2\pi f T)} \right)^M - 1 \right] \left[1 + \left(\frac{\cos(2\pi f T) - D}{E \sin(2\pi f T)} \right)^M \right]^{-1} = \\ &= \frac{(\cos(2\pi f T) - D)^M}{(E \sin(2\pi f T))^M + (\cos(2\pi f T) - D)^M}. \end{aligned}$$

Полюсы здесь остаются теми же, что и в полоснопропускающем случае. Перейдем к вычислению нулей. Ими служат решения уравнения

$$(\cos(2\pi f T) - D)^M = 0,$$

которое преобразуется к виду ($m = 1, \dots, M$)

$$\cos(2\pi f_m T) = D.$$

После некоторых стандартных операций получают коэффициенты для нерекурсивных членов (с точностью до постоянного множителя):

$$b_0 = 1, \quad b_1 = -2E, \quad b_2 = 1.$$

4.11. ДРУГИЕ РЕКУРСИВНЫЕ ФИЛЬТРЫ

Кроме базисных фильтров Баттеруорта можно реализовать и фильтры других типов, а именно фильтр Чебышева типа 1 (пульсация в полосе пропускания), фильтр Чебышева типа 2 (пульсация в полосе заграждения), фильтр эллиптического типа (пульсация и в полосе пропускания, и в полосе заграждения)

и фильтры Бесселева и Гауссова типов. Для каждого такого типа можно получить синусный и тангенсный варианты и выбрать один из четырех фильтров: низких частот, высоких частот, полоснопропускающий или заграждающий.

Аналоговые формы фильтров этих типов рассмотрены в работах Гийемана (1957) и Стоурера (1957). Соответствующие цифровые формы можно получить несколькими способами. В большинстве случаев эти авторы пользуются теми же методами, которые применялись для синусных фильтров, т. е. сначала находятся полюсы и нули, а затем с помощью комплексных отображений — синуса и косинуса — определяются коэффициенты.

В частности, ими были получены реализации для фильтров Чебышева и эллиптического типа. Сложность программ реализации фильтров Баттеруорта, Чебышева типа 1, Чебышева типа 2, эллиптического от фильтра к фильтру возрастает. Получить представление об этой сложности помогает таблица, в которой для каждого из четырех типов фильтров приведено число предложений Фортрана (включая комментарии), необходимое для реализации фильтров.

Тип фильтра	Число предложений
Баттеруорта	32
Чебышева (тип 1)	100
Чебышева (тип 2)	245
Эллиптический	238

Здесь не учтены еще 400 дополнительных предложений для синуса, тангенса и их вариантов, необходимых для получения фильтров низких и высоких частот, полоснопропускающих и заграждающих фильтров.

4.12. НЕРЕКУРСИВНЫЕ (ИКО) ФИЛЬТРЫ

Нерекурсивные фильтры (или, иначе, импульсные фильтры с конечной областью отклика — ИКО) были изучены Ормсби (1961). Предложенные им процедуры, которые ныне вытеснены более поздними конструкциями ИБО- (импульсные с бесконечной областью отклика) и ИКО-фильтров, долгое время применялись в промышленности и фактически до сих пор используются некоторыми предприятиями. Главной его заслугой было улучшение характеристик фильтров низких частот. В то время в качестве базовой конструкции применялся симметричный фильтр, для которого

$$y(i) = \sum_{k=-M}^M b_k x(i-k),$$

где

$$b_{-k} = b_k \quad \text{и} \quad b_k = \frac{\sin(2\pi BkT)}{\pi k}.$$

Симметричный фильтр такого типа имеет нулевой фазовый угол. С точки зрения электроники этот фильтр, разумеется, нереализуем, а в цифровом плане он реализуется только в ситуациях с «реальным временем» при помощи задержки. Теперь, по прошествии времени, понятно, что тот путь, который он предложил, отлично реализуется на ЭВМ.

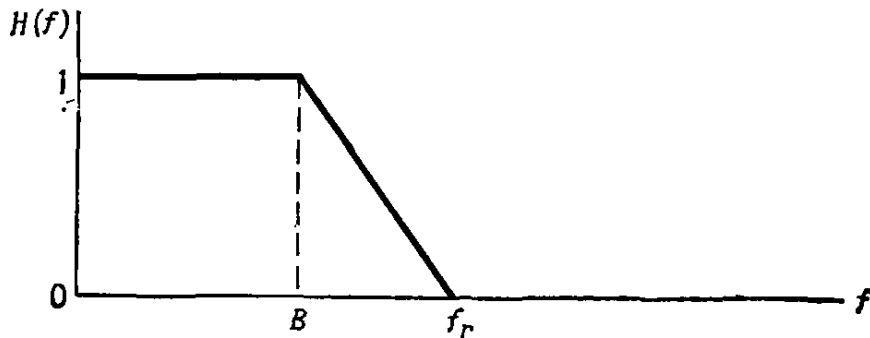


Рис. 4.16. Улучшенный фильтр Ормсби.

Ормсби заметил, что $|b_k| \leq 2BT/k$, и, следовательно, величина весов убывает слишком медленно. Поэтому в некоторых случаях для того, чтобы достичь хорошей фильтрации, значение M должно быть очень большим.

Значения весов получались с помощью обратного преобразования Фурье прямоугольного импульса

$$b_k = T \int_{-B}^B \exp\{j2\pi f k T\} df.$$

Множитель T перед интегралом обеспечивает нужную размерность. Его можно учесть также и после вычислений весов

Фильтр существенно улучшила предложенная Ормсби замена прямоугольного импульса трапециевидным (рис. 4.16). При этом из-за дополнительного треугольника, начало которого совпадает с точкой B (частотой отсечки), а конец — с точкой f_r (частотой отреза), веса убывают уже не как $1/k$, а как $1/k^2$, т. е. гораздо быстрее, чем в случае фильтра низких частот с прямоугольной передаточной функцией. Сейчас эти веса уже не используются, поскольку их вытеснили другие, предложенные позднее.

После Ормсби изучение ИКО-фильтров продвинулось далеко вперед. Пожалуй, наиболее значительным вкладом в этом вопросе являются работы Паркса и Мак-Клеллана (1972) и Мак-Клеллана, Паркса и Рабинера (1973). Для получения импульсных фильтров с конечной областью отклика широко используются программа

Паркса — Мак-Клеллана, распечатка которой приведена в их статье, и модификация алгоритма Ремеза. Мы применяли эти методы получения фильтров и склоняемся к тому мнению, что они в крайней степени полезны. Сейчас эти методы широко применяются в промышленности. В области ИКО-фильтров самым плодовитым автором является, пожалуй, Рабинер; см. работу Рабинера и Гоулда (1978), в которой достаточно подробно представлены его результаты.

Рабинер выделяет три основных метода получения весов ИКО-фильтров: 1) использование во временной области прямоугольного окна, 2) частотное ограничение, 3) разнообразные приемы, минимизирующие ошибки при некотором критерии. Последний из этих методов затрагивается в следующем параграфе.

Вместо пересказа работы Рабинера детально изложим один метод, не рассматривавшийся до сих пор в литературе. В нем используется окно, предложенное Поттером (1971). Авторам метод сообщили Бикфорд и Глейз, когда они оба работали в Time/Data Corrogation (Санта-Клара, Калифорния 95051).

Процедура получения весов в этом случае выглядит следующим образом. Сначала для промежутка $(-M, M)$ вычисляются веса прямоугольного импульса. (На самом деле вычисления производятся только для $k = 0, \dots, M$, для остальных k они излишни.) Затем полученные веса умножаются на функцию, определяющую окно. Эта функция улучшает характеристику фильтра за счет выхода за пределы полосы пропускания. Для этой цели можно использовать почти все окна, которые применяются при вычислениях плотности спектра мощности. Очень часто в этом случае выбираются окна Хэннинга и Хэмминга (см. гл. 8).

В упомянутой работе Поттера рассмотрено много примеров различных окон. Для реализации фильтра подходит предложенное им окно P310. Оно имеет вид

$$\omega_k = \frac{c_k}{\omega} \left[d_0 + 2 \sum_{p=-3}^3 d_p \cos \frac{\pi p k}{M} \right],$$

где

$$c_k = \begin{cases} 1/2 & \text{при } k = \pm M, \\ 1 & \text{в остальных случаях,} \end{cases}$$

$$\begin{aligned} d_0 &= 1, & d_{-1} &= d_1 = 0.684988, \\ d_{-2} &= d_2 = 0.202701, & d_{-3} &= d_3 = 0.0177127 \end{aligned}$$

и

$$\omega = d_0 + 2 \sum_{p=-3}^3 d_p = 2.8108034.$$

На рис. 4.17 представлена программа для получения весов этого фильтра. Последовательность команд вызова программы сводится к предложению

CALL LPSPBG(M, B, T, BK)

где M — размах фильтра; всего используется $2M + 1$ весов (в силу симметрии необходимо получить всего $M + 1$ весов); B — точка отсечки фильтра в герцах; T — выборочный интервал в секундах; BK — веса фильтра; соответствующий массив должен иметь размерность $M + 1$ или больше. Нормируем d коэффициентов (в программе $D0$, $D(1)$, $D(2)$ и $D(3)$) делением на 2.8108034.

```

SUBROUTINE LPSPBG (MM,B,T,BK)
  DIMENSION BK(1),D(3)
  DATA D0/0.35577019/,D(1)/0.2436983/,D(2)/0.07211497/,
  *D(3)/0.00630165/,PI/3.14159265/
  C   ПРОГРАММА ВЫЧИСЛЯЕТ ВЕСА НИЗКОПОЛОСНОГО ИКО-ФИЛЬТРА
  C   МЕТОД, ПРЕДЛОЖЕННЫЙ ПОТТЕРОМ, БИКФОРДОМ И ГЛЕЙЗОМ
  C   ВСЕГО ИМЕЕТСЯ 2M+1 ВЕСОВ... В LPSPBG ВЫЧИСЛЯЕТСЯ M+1
  C   T ЕСТЬ ИНТЕРВАЛ ВЫБОРКИ В СЕКУНДАХ
  C   B ЕСТЬ ПОЛОСА ПРОПУСКАНИЯ В ГЦ
  C   РЕЗУЛЬТАТЫ ЗАПОМИНАЮТСЯ В BK
  M=MM
  C   СНАЧАЛА ПОЛУЧАЮТ ВЕСА ПРЯМОУГОЛЬНОЙ ЧАСТИ
  FACT=2.*B*T
  BK(1)=FACT
  FACT=FACT*PI
  DO 5 I=1,M
  FI=I
  5   BK(I+1)=SIN(FACT*FI)/(PI*FI)
  C   ТРАПЕЦИЕВИДНОЕ СГЛАЖИВАНИЕ В КОНЦЕ
  BK(M+1)=BK(M+1)/2.
  C   ТЕПЕРЬ ПРИМЕНЯЕТСЯ ОКНО ПОТТЕРА P310
  SUMG=BK(1)
  DO 15 I=1,M
  SUM=DO
  FACT=PI*FLOAT(I)/FLOAT(M)
  DO 10 K=1,3
  10  SUM=SUM+2.*D(K)*COS(FACT*FLOAT(K))
  BK(I+1)=BK(I+1)*SUM
  15  SUMG=SUMG+2.*BK(I+1)
  M1=M+1
  DO 20 I=1,M1
  20  BK(I)=BK(I)/SUMG
  RETURN
  END

```

Рис. 4.17. Подпрограмма LPSPBG для получения весов низкочастотного ИКО-фильтра.

На рис. 4.18 показан результат применения окна Поттера P310 в случае прямоугольного фильтра низких частот. В этом примере $T = 1$ ($F = 0.5$), $B = 0.1245$, $M = 128$ (общий размах фильтра — 257 точек). Как видно из рисунка, в результате применения окна нижний уровень фильтра уменьшается в среднем примерно на 40 дБ; ширина полосы при этом возрастает очень незначительно. Это существенно лучше результатов, которые получаются при использовании так называемого окна Кайзера, рассмотренного Рабинером и Гоулдом (1978).

Эту программу нетрудно приспособить для фильтров высоких частот, заграждающих и полоснопропускающих фильтров. По сути дела, никаких изменений в программе производить не нужно: 1) для фильтра высоких частот с полосой пропускания от B до F

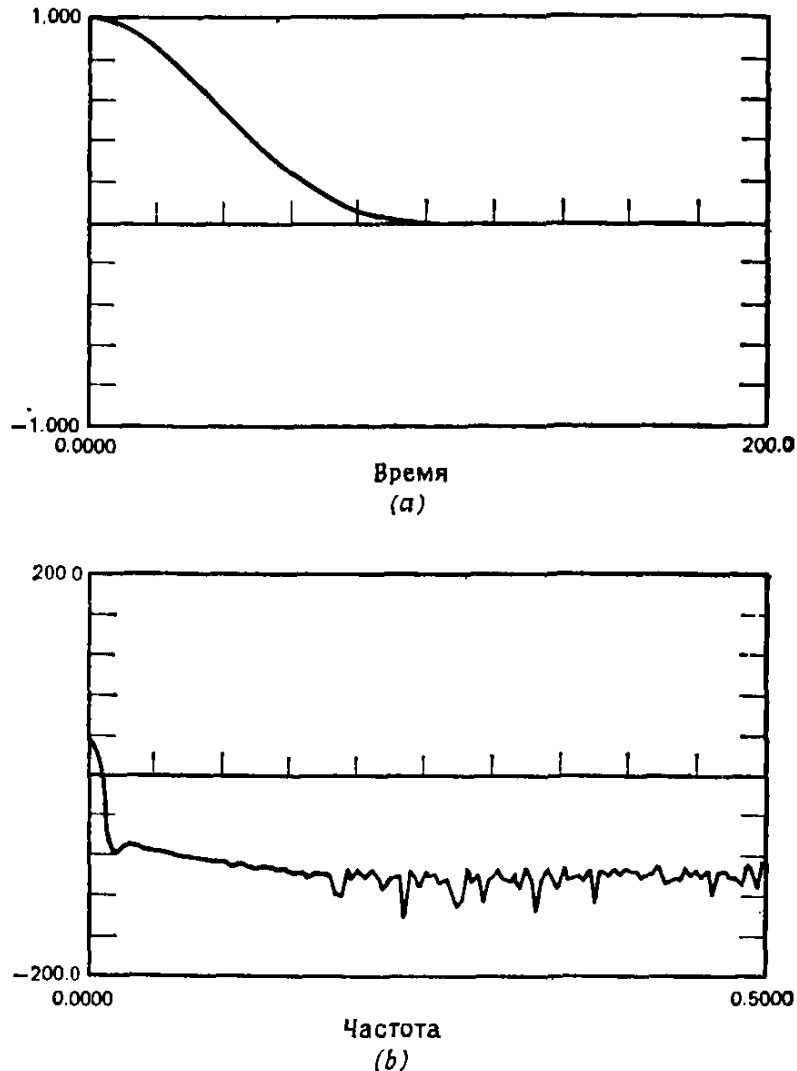
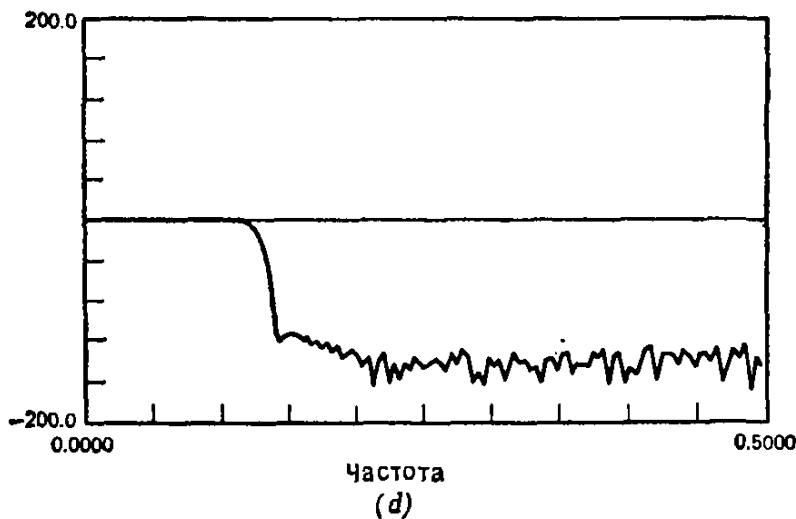
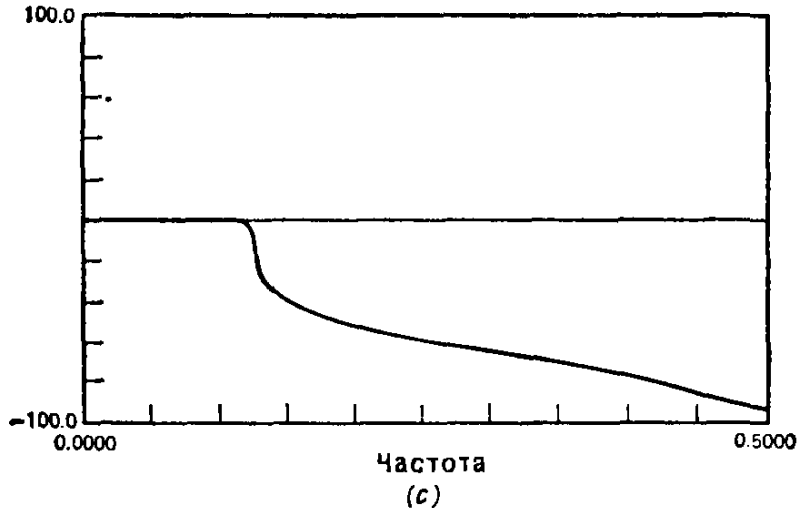


Рис. 4.18. Окно Поттера Р310 и низкочастотный фильтр до и после применения окна. a — окно Поттера Р310 во временной области; b — окно Поттера Р310 в области частот; c — низкочастотный ИКО-фильтр до применения окна; d — тот же

необходимо получить фильтр низких частот с полосой $0-B$, а затем вычесть центральный вес из единицы, а у остальных изменить знаки на противоположные, т. е. $b'_0 = 1 - b_0$, $b'_k = -b_k$, $k \neq 0$ (штрихованным весам соответствует случай высоких частот); 2) для полоснопропускающего фильтра с полосой пропускания от f_1 до f_2 Гц нужно получить два фильтра низких частот с полосами $0-f_1$ и $0-f_2$ соответственно, а затем получить веса, придерживаясь правила: (вес для фильтра высоких частот) = (вес для фильтра низких частот до f_2) - (вес для фильтра низких частот до f_1); 3) для заграждающего фильтра, не пропускающего информацию в интервале (f_1, f_2) , получить те же два

фильтра низких частот, что и в случае 2, и определить веса по правилу: (вес для заграждающего фильтра) = $1 - (\text{вес для фильтра низких частот до } f_2) + (\text{вес для фильтра низких частот до } f_1)$.



фильтр после применения окна Поттера. Обратите внимание на изменение масштаба. Здесь $T=1$, $B=0.1245$, $M=128$ (для общего числа точек, равного 257).

4.13. ПРИЕМЫ АППРОКСИМАЦИИ ФИЛЬТРОВ

В литературе описывается немало процедур, предназначенных для расчета цифровых фильтров. В них используются алгоритмы Флетчера — Пауэлла (Жюльен и Сид-Ахмед (1973)) и минимаксные процедуры или процедуры Чебышева (Хольц (1973)). Наиболее важными в этом вопросе являются уже упомянутые выше работы Паркса и Мак-Клеллана и их программы для ЭВМ. Используемый там метод включает в себя алгоритм Флетчера — Пауэлла (Флетчер и Пауэлл (1963)). Для этого алгоритма составлена программа на Фортране, которая имеется в различных источниках (IBM, 1968).

Предположим что ситуация выглядит следующим образом.

1. В виде таблицы неотрицательных чисел заданы значения $|H(f_i)|^2$ и f_i , $i = 1, \dots, N$.
2. По формуле $\lambda_i = 2\pi f_i T$ определены значения λ_i .
3. Фильтр должен быть рекурсивным и иметь M составляющих.
4. Аппроксимацией функции $|H(f_i)|^2$ служит функция $|\hat{H}(f_i)|^2$. Передаточную функцию каждой составляющей фильтра будем аппроксимировать отношением многочленов

$$\frac{\sin^4(\lambda_i/2) + 2p_m^2 \sin^2(\lambda_i/2) + (p_m^4 + q_m^4)}{\sin^4(\lambda_i/2) + 2r_m^2 \sin^2(\lambda_i/2) + (r_m^4 + s_m^4)}.$$

Выбор таких формул диктуется необходимостью гарантировать выполнение неравенства $|\hat{H}(f_i)|^2 \geq 0$ для любых i вне зависимости от знаков p_m и q_m . Кроме того, эти формулы гарантируют устойчивость фильтра.

Таким образом, передаточная функция фильтра имеет вид ($i = 1, \dots, N$)

$$|\hat{H}(f_i)|^2 = A \prod_{m=1}^M \frac{\sin^4(\lambda_i/2) + 2p_m^2 \sin^2(\lambda_i/2) + (p_m^4 + q_m^4)}{\sin^4(\lambda_i/2) + 2r_m^2 \sin^2(\lambda_i/2) + (r_m^4 + s_m^4)}.$$

Параметр A означает общий постоянный множитель.

Перечислим операции, предусмотренные в алгоритме.

1. Для величин p_m, q_m, r_m, s_m ($m = 1, \dots, M$) задаются начальные значения.

2. Для $m = 1, \dots, M$ и $i = 1, \dots, N$ вычисляются величины

$$g_{im} = \sin^4 \frac{\lambda_i}{2} + 2p_m^2 \sin^2 \frac{\lambda_i}{2} + (p_m^4 + q_m^4),$$

$$h_{im} = \sin^4 \frac{\lambda_i}{2} + 2r_m^2 \sin^2 \frac{\lambda_i}{2} + (r_m^4 + s_m^4).$$

Предварительно должны вычисляться и запоминаться значения $\sin^2(\lambda_i/2)$ и $\sin^4(\lambda_i/2)$.

3. Вычисляются значения

$$P_i = \prod_{m=1}^M \frac{g_{im}}{h_{im}}, \quad i = 1, \dots, N.$$

4. Вычисляется A^* (оценка для A):

$$A^* = \left(\sum_{i=1}^N |\hat{H}(f_i)|^2 \right) / \left(\sum_{i=1}^N P_i |\hat{H}(f_i)|^2 \right).$$

5. Для $i = 1, \dots, N$ вычисляются ошибки

$$E_i = A^* P_i - |H(f_i)|^2 = |\hat{H}(f_i)|^2 - |H(f_i)|^2.$$

6. Вычисляется среднеквадратичная ошибка

$$Q = N^{-1} \sum_{i=1}^N E_i^2.$$

7. По формулам для частных производных Q по φ

$$\frac{\partial Q}{\partial \varphi} = \frac{2}{N} \sum_{i=1}^N E_i \frac{\partial |\hat{H}(f_i)|^2}{\partial \varphi}$$

и формулам для частных производных $|\hat{H}(f_i)|^2$ по p_m , q_m , r_m и s_m ($i = 1, \dots, N$, $m = 1, \dots, M$)

$$\begin{aligned} \frac{\partial |\hat{H}(f_i)|^2}{\partial p_m} &= \frac{4A^*P_i}{g_{im}} \left(p_m \sin^2 \frac{\lambda_i}{2} + p_m^3 \right), \\ \frac{\partial |\hat{H}(f_i)|^2}{\partial q_m} &= \frac{4A^*P_i}{g_{im}} q_m^3, \\ \frac{\partial |\hat{H}(f_i)|^2}{\partial r_m} &= -\frac{4A^*P_i}{h_{im}} \left(r_m \sin^2 \frac{\lambda_i}{2} + r_m^3 \right), \\ \frac{\partial |\hat{H}(f_i)|^2}{\partial s_m} &= -\frac{4A^*P_i s_m^3}{h_{im}} \end{aligned}$$

вычисляются $\partial Q/\partial p_m$, $\partial Q/\partial q_m$, $\partial Q/\partial r_m$ и $\partial Q/\partial s_m$, где $m = 1, \dots, M$.

8. К полученным данным применяется алгоритм Флетчера — Пауэлла, который вырабатывает новые значения p_m , q_m , r_m и s_m ($m = 1, \dots, M$).

9. С этого пункта возвращаются на второй шаг, если значения не подходят, или приступают к дальнейшим действиям, если полученные результаты приемлемы.

10. После этого находят полюсы и нули. Для m -й составляющей фильтра нули определяют из уравнения

$$\sin^4 \frac{\lambda_m}{2} + 2p_m^2 \sin^2 \frac{\lambda_m}{2} + (p_m^4 + q_m^4) = 0,$$

которое решают как квадратное уравнение относительно $\sin^2 \frac{\lambda_m}{2}$:

$$\sin^2 \frac{\lambda_m}{2} = -\frac{2p_m^2 \pm \sqrt{4p_m^4 - 4p_m^4 - 4q_m^4}}{2} = -p_m^2 \pm jq_m^2.$$

В итоге получаются решения

$$f_{zm} = \frac{1}{\pi T} \arcsin (-p_m^2 \pm jq_m^2)^{1/2}.$$

Аналогичное уравнение для полюсов приводит к решениям

$$f_{pm} = \frac{1}{\pi T} \arcsin (-r_m^2 \pm js_m^2)^{1/2}.$$

Из них используются только M нулей и полюсов, которые соответствуют устойчивому фильтру.

11. Для вычисления коэффициентов a_{1m} , a_{2m} , b_{1m} и b_{2m} можно воспользоваться соответствующей частью видоизмененной программы получения синусного фильтра низких частот Баттеруорта (здесь $b_{0m} = 1$ для всех m). Последним вычисляется общий постоянный множитель для всех нерекурсивных членов. Например, если потребовать, чтобы

$$|H(f_0)|^2 = A,$$

то все b членов умножаются на множитель c , определяемый формулой

$$c = \left(A / \left| \prod_{m=1}^M \frac{1 + b_{1m}e^{-j2\pi f_0 T} + b_{2m}e^{-j4\pi f_0 T}}{1 + a_{1m}e^{-j2\pi f_0 T} + a_{2m}e^{-j4\pi f_0 T}} \right| \right)^{1/M}.$$

Этот алгоритм нетрудно приспособить для вычисления весов или нерекурсивных, или полностью рекурсивных фильтров, исключая или r_m и s_m , или p_m и q_m соответственно.

Упражнения

4.1. Покажите, что передаточная функция

$$|H(f)|^2 = \frac{1}{\omega_n^4 [(1 - (f/f_n)^2)^2 + (2\zeta f/f_n)^2]}$$

имеет максимум в частоте $f_p = f_n (1 - 2\zeta^2)^{1/2}$.

4.2. Получите передаточную функцию фильтра

$$y_i = 1/4 x_{i-1} + 1/2 x_i + 1/4 x_{i+1}.$$

Ответ: $H(f) = \cos^2(\pi f T)$.

4.3. Получите передаточную функцию фильтра

$$y_i = \frac{1}{2^N} \sum_{n=0}^N \binom{N}{n} x \left(i + n - \frac{N}{2} \right),$$

где N — четное и

$$\binom{N}{n} = \frac{N!}{(N-n)! n!}.$$

Ответ: $H(f) = \cos^N(\pi f T)$.

4.4. Приведите выражение для параметра α фильтра низких частот первого порядка с точкой половинной мощности, равной B Гц, т. е.

$$|H(B)|^2 = 1/2 \quad \text{и} \quad |H(0)|^2 = 1.$$

Ответ: $\alpha = 2A^2 + 1 - 2A\sqrt{A^2 + 1}$, где $A = \sin(\pi T B)$.

4.5. Как было показано, фильтр второго порядка с действительными полюсами может быть представлен в виде двух фильтров первого порядка. Покажите, что то же самое справедливо и для фильтра второго порядка с комплексными полюсами, полагая, что коэффициенты фильтра могут быть комплексными. Приведите выражение для коэффициентов через α и β .

4.6. Проверьте, что выражение, приводившееся для единичной импульсной функции отклика фильтра второго порядка с комплексными полюсами, правильно.

а) Используя разностное уравнение фильтра и полученное для $y(i)$ выражение, получите по пять первых значений ($i=0, \dots, 4$). Сравните полученные результаты.

б) Покажите, что решение удовлетворяет разностному уравнению.

Задачи для решения на ЭВМ. В некоторых задачах потребуется использование подпрограмм LPSB и TTRAN из приложения А. Для выведения результатов на дисплей полезно обращаться к подпрограмме PRPLOT.

4.7. Составьте программу для вычисления параметра α в случае фильтра низких частот первого порядка с характеристиками, перечисленными в упр. 4.4. Программа должна предусматривать следующие операции.

а) Вычисление параметра α для $T=0.005$ и $B=3.125, 6.25, 12.5, 25, 20$ и 100 Гц.

б) Вычисление $y(i)$ для $x(i)=1, 0, 0, \dots$, где $i=0, \dots, 25$.

с) Вычисление и построение графиков коэффициента усиления и фазы передаточной функции при $f=0, 2, 4, \dots, F$ Гц.

4.8. Составьте программу для вычисления коэффициентов a_1, a_2 и b_0 в общем случае комплексного фильтра второго порядка. Зафиксируйте в программе значение $T=0.005$ и определите коэффициенты для пар (f_n, ζ) в следующем порядке.

а) Зафиксируйте значение f_n равным 50 Гц, а ζ последовательно присваивайте значения 0.5, 0.25, 0.1, 0.05, 0.01, 0.001.

б) Зафиксируйте значение ζ равным 0.05, а f_n последовательно присваивайте значения 25, 10, 5, 2.5 Гц.

с) В перечисленных случаях вычислите такие значения коэффициента b_0 , чтобы в пиковой частоте коэффициент усиления фильтра был равным единице.

д) Для пар (50, 0.05) и (50, 0.01) получите отклик фильтра (численно и графически) на последовательность $x(i)=1, 0, 0, 0$. Согласуется ли полученная последовательность с теоретической?

е) Для первых четырех пар (f_n, ζ) , перечисленных в п. а), вычислите коэффициент усиления и фазу передаточной функции. Получите соответствующие графики. Прокомментируйте результаты с точки зрения ширины полосы пропускания.

4.9. Проведите трансляцию программы LP2 (рис. 4.4.) и составьте управляющую программу, которая производит следующие операции.

а) Вычисление по коэффициентам A1 и A2 (a_1 и a_2) значений f_n и ζ при $T=0.005$ и $BW=32, 16, 8, 4, 2, 1$ Гц. Прокомментируйте результаты.

б) Получение и построения графиков коэффициента усиления и фазы передаточной функции для $f=0, 1, \dots, F$ Гц. Прокомментируйте результаты.

с) Вычисление отклика фильтра на последовательность $x(i)=1, 0, 0, \dots, i=0, \dots, 25$. Прокомментируйте результаты.

4.10. Проведите трансляцию программы LPSB и составьте управляющую программу, которая производит следующие операции:

а) Дублирование (для проверки правильности работы программы) результатов, приведенных в табл. 5.3 для шестиполосного фильтра, или результатов, полученных в предыдущем упражнении для двухполосного фильтра.

б) Вычисление с помощью подпрограммы TTRAN коэффициента усиления и фазы передаточной функции шестиполосного фильтра при $T=0.005$ и $B=32, 16, 8, 4, 2, 1$ Гц. Прокомментируйте результаты.

с) Определение полюсов для каждой из 18 пар коэффициентов, полученных в п. б). Изобразите положение полюсов на комплексной плоскости и прокомментируйте результаты.

4.11. Описанным в § 4.7 способом переделайте для фильтра высоких частот Баттеруорта программу LPSB. Выполните с этой программой для фильтра

высоких частот упр. 4.10, используя значения $B = 68, 84, 92, 96, 98$ и 99 Гц. Сравните новые результаты с теми, которые были получены в упр. 4.10, и прокомментируйте возникшую картину.

4.12. Проведите трансляцию программы с рис. 4.13 для получения полоснопропускающего фильтра, имеющего шесть рекурсивных коэффициентов, полагая $T = 0.005$. Получите (численно и графически) значения коэффициента усиления и фазы в точках $f = 0, 1, \dots, F$ Гц, полагая

а) $f_c = 50$ Гц и $B = 1, 2, 5, 10, 25$ Гц;

б) $B = 2$ Гц и $f_c = 2, 4, 10$ и 20 Гц.

Прокомментируйте результаты.

4.13. Переделайте программу, использованную в упр. 4.12, для случая заграждающего фильтра и сделайте те же вычисления.

4.14. Проведите трансляцию программы LPTB с рис. 4.15 для вычисления коэффициентов фильтров низких частот Баттеруорта. Повторите упр. 4.10, используя вместо программы LPSB программу LPTB. Сравните результаты этих двух упражнений и прокомментируйте их.

4.15. а) Напишите программу вычисления коэффициента усиления и фазы передаточной функции симметричного ИКО-фильтра.

б) Напишите программу для вычисления весов фильтра низких частот, в которой

$$b_k = \frac{\sin(2\pi BkT)}{\pi k}, \quad k = 0, \dots, M.$$

с) Используя программы из пп. а) и б), составьте программу, которая строит графики коэффициента усиления, полагая $T = 0.005$, $B = 32, 16, 8, 4, 2, 1$ Гц и $M = 8, 16, 32, 64$. Прокомментируйте результаты. Сравните их также с результатами, полученными в упр. 4.10.

4.16. Программу LPSPBG с рис. 4.17, реализующую алгоритм Поттера, Бикфорда и Глейза, примените в упр. 4.10 и сравните результаты, которые получатся в обоих случаях.

Глава 5

ПРАКТИЧЕСКИЕ АСПЕКТЫ ЦИФРОВОЙ ФИЛЬТРАЦИИ

5.1. ВВЕДЕНИЕ

В этой главе затронут круг вопросов, связанных с использованием на практике методов цифровой фильтрации. Среди них будут рассмотрены шум и искажение, неустойчивость, различные виды программной записи алгоритмов фильтрации, сглаживание данных и децимация.

Выбор конкретного фильтра невозможно подчинить какому-либо общему правилу. Дело в том, что как необходимый тип фильтра (например, полоснопропускающий или низких частот), так и те условия, в которых должна проводиться фильтрация, в различных приложениях существенно отличаются.

С другой стороны, во многих приложениях требуется только низкочастотная фильтрация. В этом случае в качестве своего рода отправной точки выбора может служить шестиполосный фильтр низких частот Баттеруорта. Иначе говоря, правомерна такая постановка вопроса: «Нужно ли в данном приложении отказываться от шестиполосного фильтра Баттеруорта?» Нередко, если оптимальный выбор неосуществим, наиболее подходящим оказывается именно этот тип фильтра.

5.2. ШУМ И ИСКАЖЕНИЕ

Во многих вычислительных операциях, производимых ЭВМ, возникает шум. При использовании арифметики с фиксированной точкой это преимущественно бывает в результате умножения. После перемножения меньшая значащая половина произведения стирается, а бóльшая значащая половина (усеченная или округленная) запоминается. В данном случае *величину шума* определяет разница между истинным ответом и числом, которое получилось после усечения или округления. Уровень такого шума колеблется в пределах примерно половины последнего значащего бита.

При использовании арифметики с плавающей точкой ошибками такого рода дело не ограничивается. Здесь возникают новые погрешности, обусловленные антипереполнением при сложении. Когда складываются два числа с плавающей точкой, одно из которых значительно больше другого, все меньшее число или его часть могут быть потеряны.

Реализации фильтров, рассмотренные в гл. 4, не учитывали особенностей, связанных с конечным представлением чисел. Поэтому на практике характеристики реализаций фильтров отклоняются от характеристик теоретических реализаций, предполагающих использование идеальных арифметических операций.

С этим связано усложнение задачи реализации фильтров на ЭВМ, имеющих уменьшенное число битов. Например, применение для этой цели мини-ЭВМ с длиной слова, равной 12—16 битам, приводит к большему ограничению возможностей, чем применение больших ЭВМ со словами, длина которых равна 24—60 битов.

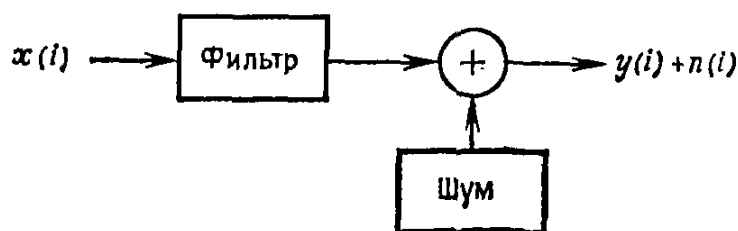


Рис. 5.1. Модель с шумом для идеального фильтра.

Для арифметики с конечным представлением чисел характерен ряд явлений, среди которых наиболее важными будут шум, искажение и неустойчивость. Они затронуты во многих работах. Назовем три сборника: первые два изданы под редакцией Рабинера и Рейдера (1972) и Оппенгейма и др. (1976), третий — под редакцией Лью (1975).

На рис. 5.1 изображена модель с шумом. К выходу идеального фильтра добавляются значения, порождаемые источником шума. В результате вместо последовательности значений $y(i)$ возникает последовательность $y(i) + n(i)$, где $n(i)$ — шум.

Характер шума в значительной степени зависит от используемой арифметики и от реализации фильтра.

Перечислим основные результаты, относящиеся к шуму.

1. Шум — функция числа битов, используемых в арифметических операциях; чем меньше это число, тем выше уровень шума.

2. Каскадные и параллельные¹⁾ реализации дают меньший по сравнению с комбинированными реализациями уровень шума. Если число полюсов или нулей равно четырем или больше четырех, то более предпочтительной будет каскадная форма фильтра. Это справедливо и для ИБО-, и для ИКО-фильтров.

3. Реализация фильтра в каскадной форме некоммутативна по отношению к выходу. Это означает следующее. При идеальных арифметических операциях выход фильтра не зависел бы от

¹⁾ Как правило, в дальнейшем параллельная и цепная реализации не будут затрагиваться. Хотя во многих случаях идеальные реализации таких типов имеют лучшие характеристики, чем другие, из-за большей сложности они оказываются непригодными для наших целей.

порядка расположения двухполюсных (двухнулевых) составляющих фильтра. В действительности порядок следования составляющих второго порядка в реализации фильтра влияет на шум. Поэтому возникает возможность выбора такой перестановки составляющих фильтров второго порядка, что полученный на выходе общего фильтра шум будет минимальным

Вычислительный шум включает в себя определенный «уровень» шума. Это нетрудно увидеть на примере идеализированной арифметической операции над числами с фиксированной точкой. Допустим, что для дробной части числа выделено m битов слова (само слово может иметь и более m битов). Тогда точно так же, как и при преобразовании в цифровую форму данных, при умножении каждый раз будет возникать ошибка, порядок которой равен примерно половине последнего значащего бита десятичной записи. Если предположить, что ошибки некоррелированы, то дисперсия ошибки после k последовательных умножений будет равна $k/12$ бит².

Таким образом, с учетом структуры слова получается уровень шума, равный $k2^{-2m}/12$ бит². Если перейти к децибелам, то

$$\begin{aligned} 10 \log_{10} (k2^{-2m}/12) &= 10 \log_{10} k - 20m \log_{10} 2 - 10 \log_{10} 12 \approx \\ &\approx 10 \log_{10} k - 10.8 - 6m. \end{aligned}$$

Следовательно, уровень шума имеет некоторую нижнюю границу, определенную числом битов, входящих в дробную часть слова ЭВМ. Уровень возрастает при увеличении числа умножений: чем больше фильтр, тем выше уровень шума.

Этот результат лишь приближенно отражает реальную картину. На практике, например, допущение о некоррелированности может не выполняться. Кроме того, полученный результат в каком-то смысле дает только нижнюю границу возможного шума; как будет показано, шум, порождаемый ИБО-фильтром, значительно возрастает, когда этот фильтр становится неустойчивым.

Одна из причин, по которой более предпочтительной является каскадная реализация фильтра, состоит в следующем: помимо того, что каждая каскадная составляющая вносит в данные свой собственный шум при фильтрации, она также уменьшает шум, появившийся на предшествующих стадиях каскадной фильтрации. Значит, в каскадном фильтре происходит и генерация, и сокращение шума. Это свойство в гораздо меньшей степени относится к комбинированным формам фильтров. Им более свойствен большой собственный шум, что, пожалуй, важно учитывать в приложениях.

Искажение. Искажение является результатом существенного расхождения передаточной функции фильтра в том виде, в ко-

тором он реализуется, с теоретической передаточной функцией. Обычно в качестве меры искажения берут среднеквадратичное значение

$$\epsilon^2 = \int_0^F |H(f) - \hat{H}(f)|^2 df,$$

где ϵ^2 — среднеквадратичное искажение, $H(f)$ — теоретическая передаточная функция, $\hat{H}(f)$ — передаточная функция фильтра, реализованного на ЭВМ с конечной длиной слова.

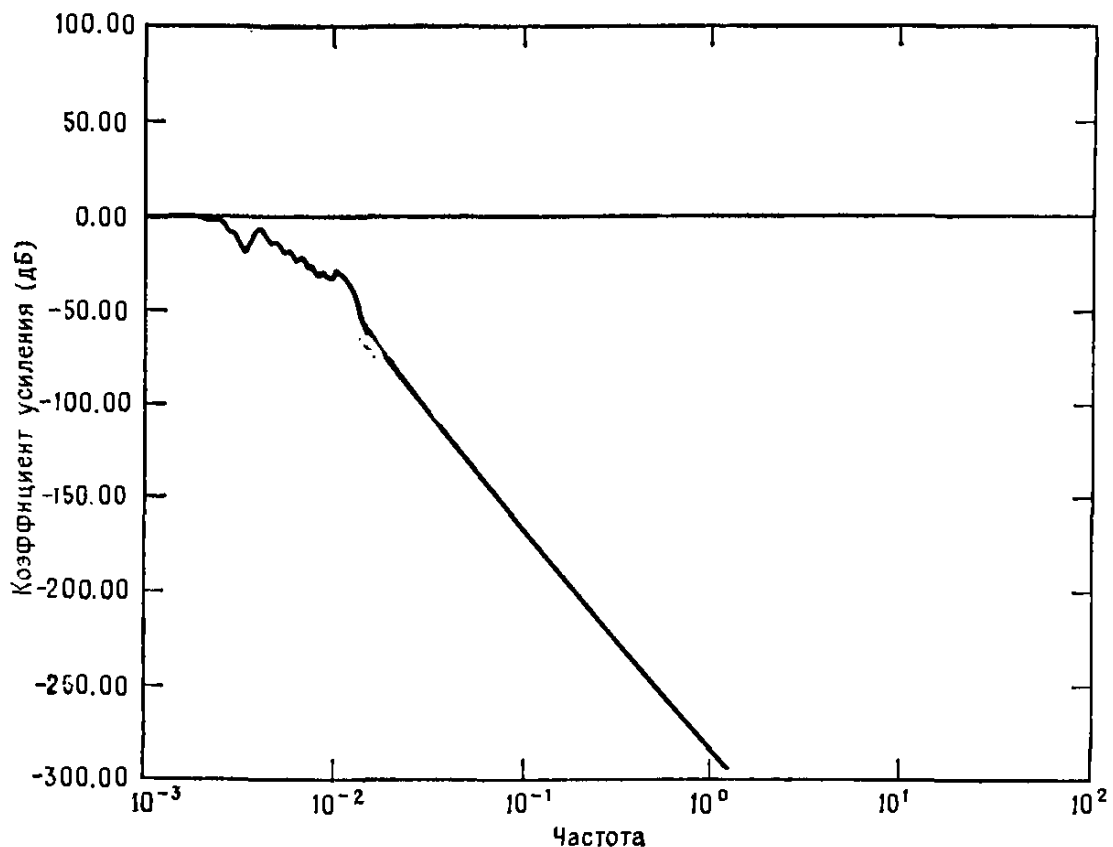


Рис. 5.2. Искажение передаточной функции, при котором она обретает очертания, свойственные низкочастотному фильтру.

Первое и наиболее явное проявление искажения состоит в погрешностях и отклонениях от нужного вида передаточной функции. При уменьшении числа битов передаточные функции фильтров заметно искажаются (рис. 5.2.). Такое уменьшение числа битов приводит к тому, что исходная форма передаточной функции может измениться до неузнаваемости.

Искажение играет, пожалуй, меньшую роль, чем шум. К тому моменту, когда искажение, измеренное в линейном масштабе, станет заметным, шум, измеренный в логарифмическом масштабе, т. е. в децибелах, может оказаться достаточно большим для того, чтобы характеристика фильтра стала неудовлетворительной.

Разумеется, определение искажения тоже можно дать в логарифмической записи, т. е. в децибелах:

$$\epsilon_{\text{дБ}}^2 = 10 \int_0^F \log_{10} \left| \frac{H(f)}{\hat{H}(f)} \right|^2 df.$$

И в этом случае уровень шума будет проявляться сильнее.

Измерение $\hat{H}(f)$. Передаточную функцию $\hat{H}(f)$ можно измерять по крайней мере двумя способами.

Первый метод. На той же ЭВМ и с той же арифметикой, с которой вычисляется функция $\hat{H}(f)$ как преобразование Фурье фильтра, вычислить

$$\hat{H}(f) = \prod_{m=1}^{M/2} \frac{b_{0m} + b_{1m} \exp(-j\omega T) + b_{2m} \exp(-2j\omega T)}{1 + a_{1m} \exp(-j\omega T) + a_{2m} \exp(-2j\omega T)}.$$

Второй метод. Для каждой частоты f_p , в которой следует определить $\hat{H}(f)$, получить последовательность $x_p(i) = \sin(2\pi f_p i T)$ и пропустить ее через фильтр. Выход фильтра обозначим $y_p(i)$. Накопив достаточно длинную последовательность для переменного отклика, вычислить величину

$$s_p^2 = \frac{1}{N_p} \sum_{i=q}^{N_p+q-1} y_p^2(i).$$

В идеале f_p и N_p должны выбираться так, чтобы числу точек N_p соответствовало бы целое число периодов синусоиды. Параметр q выбирается таким образом, чтобы при $i \geq q$ выход $y_p(i)$ оставался устойчивым. Далее, вычисляется последовательность

$$z_p(i) = x_p(i) y_p(i),$$

для которой определяется последовательность корреляционных параметров

$$\rho_p = \frac{1}{s_p 2^{-1/2} N_p} \sum_{i=q}^{N_p+q-1} z_p(i),$$

где $2^{-1/2}$ и s_p — стандартные отклонения входа и выхода, используемые для нормировки. Допустим, что установившийся выход фильтра $y_p(i)$ имеет вид

$$y_p(i) = A_p \sin(2\pi f_p i T + \varphi_p), \quad \text{где } i \geq q.$$

Тогда $s_p^2 = A_p^2/2$, и, значит,

$$\begin{aligned} z_p(i) &= \sin(2\pi f_p i T) A_p \sin(2\pi f_p i T + \varphi_p) = \\ &= \frac{A_p}{2} [\cos \varphi_p - \cos(4\pi f_p i T + \varphi_p)]. \end{aligned}$$

Суммирование на целом числе периодов членов последовательности $z_p(i)$ обращает в нуль крайнюю справа сумму косинусов, поэтому

$$\begin{aligned} \rho_p &= \frac{\sqrt{2}}{s_p N_p} \sum_{i=q}^{N_p - q + 1} [\cos \varphi_p - \cos (4\pi f_p i T + \varphi_p)] = \\ &= \left(\frac{\sqrt{2}}{(A_p / \sqrt{2}) N_p} \right) \frac{A_p}{2} N_p \cos \varphi_p = \cos \varphi_p. \end{aligned}$$

Таким образом, коэффициент усиления фильтра в точке f_p оказывается равным $A_p = \sqrt{2} s_p$, а его фаза φ_p — равной $\arccos(\rho_p)$. Разумеется, эти расчеты не застрахованы от некоторых ошибок, обусловленных тем, что в вычислениях использовано не целое число периодов синусоиды и что момент перехода q в установившийся режим определен неверно. Поэтому к всплескам в частотах следует относиться с известной осторожностью.

Более предпочтительным считается второй метод, поскольку в нем точнее отражаются стандартные операции для вычисления передаточной функции аналогового фильтра. Вместе с тем этот метод неудобно и невыгодно применять при вычислениях для больших наборов частот f_p , в особенности когда произведение BT невелико, где B — ширина полосы пропускания фильтра. Грубо говоря, для переменного отклика время достижения на выходе установившегося режима пропорционально $1/BT$. Поэтому при уменьшении значения BT требуется все большее время для получения установившейся части отклика.

Первая процедура, упомянутая выше, гораздо проще и удобнее для реализации. Однако ее нужно рассматривать скорее как указатель наличия шума и искажения, а не как средство для определения их силы.

5.3. УХУЖДЕНИЕ ФИЛЬТРА

Хотя шум и искажение играют важную роль сами по себе, главным показателем качества цифрового фильтра остается устойчивость. В сущности и шум, и искажение, и неустойчивость суть только разные стороны одного и того же явления — ухудшения характеристики фильтра, связанного с использованием вычислений ограниченной точности.

Под неустойчивостью подразумевается появление бесконечных значений на выходе фильтра при конечных значениях входа. Примером совершенно неустойчивого фильтра служит, например, фильтр с разностным уравнением

$$y(i) = 2y(i-1) + x(i),$$

который в ответ на последовательность $1, 0, 0, \dots$ выдает последовательность $1, 2, 4, \dots, 2^i, \dots$.

Условия устойчивости фильтра второго порядка получены в гл. 4. Область устойчивости снова изображена на рис. 5.3. На бесконечной плоскости всех возможных пар коэффициентов для фильтров второго порядка областью устойчивости фильтров служит небольшой треугольник, расположенный в начале коорди-

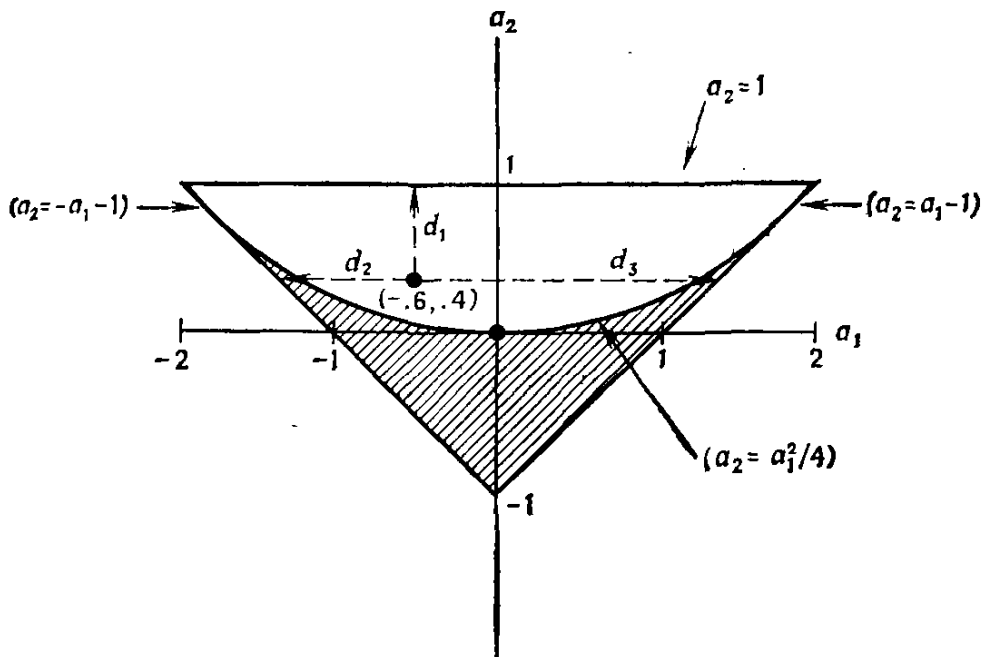


Рис. 5.3. Треугольник устойчивости для коэффициентов фильтра второго порядка. Каждая рекурсивная пара коэффициентов (a_1, a_2) определяет на плоскости одну точку. Если эта точка попадает в треугольник, то фильтр устойчив. Если она находится на границе или вне треугольника, то фильтр неустойчив.

нат. На рисунке показана точка $a_1 = -0.6$, $a_2 = 0.4$, которая попадает в треугольник коэффициентов устойчивых фильтров. Для оценки устойчивости нередко полезно обращаться к расстояниям d_1 , d_2 и d_3 . Чем ближе точка (a_1, a_2) расположена к сторонам треугольника, тем меньше оказывается одно из этих расстояний. В данном примере эти расстояния имеют следующие значения:

$$d_1 = 0.6, \quad d_2 = 0.8, \quad d_3 = 2.0.$$

Для исследования устойчивости фильтра применяется и другой способ, в котором определяется некоторое пороговое значение. Как было показано Отнесом и Мак-Нейми (1970), для фильтров низких частот можно указать такую точку на шкале частот, что при значениях ширины полосы пропускания, меньших этой точки, цифровые фильтры теряют нужный вид передаточной функции и становятся неустойчивыми. Такое пороговое значение

выражается формулой

$$B_{t, N} = \frac{2^{-(t-1-\beta_N)/N}}{2\pi T},$$

где N — число полюсов фильтра, t — общее число битов слова применяемой ЭВМ, β_N — такое целое число, что

$$2^{\beta_N - 1} < \binom{N}{\lfloor \frac{N}{2} \rfloor} \leq 2^{\beta_N}.$$

В табл. 5.1 приводятся значения порога для двух ЭВМ, работающих в режиме плавающей точки. Первая из них — IBM 704 (результаты для которой можно отнести также к ЭВМ IBM 709, 7090 и ЭВМ серии SRU 1106, 1107, 1108 и 1109), а вторая — IBM 360 (такой же формат используется в IBM 370 и ЭВМ серии Хегох 5, 6, 7 и т. д.). Значения, приведенные в таблице, относятся к каскадным реализациям фильтра, в комбинации составных частей которых входят фильтры низких частот Баттерворта синусного и тангенсного типа; $T = 0.005$, $N \geq 2$.

Кроме теоретических значений B для проверки на ЭВМ обоих типов были просчитаны $B_{\text{действ}}$ для реальных ситуаций (см. соответствующую колонку). Результаты почти одинаковы.

В качестве примера рассмотрим передаточные функции, изображенные на рис. 5.4. В этом примере для шестиполюсных фильтров значения B выбирались от 0.00031623 до 10 Гц. Расчеты производились на ЭВМ SRU 1108. При вычислениях передаточных функций использовался первый метод, в котором $\hat{H}(f)$

Таблица 5.1

Пороговые значения $B_{t, N}$ (в Гц) при нескольких значениях N для двух форматов слов ЭВМ с плавающей точкой

N	Формат IBM 704 ($t=36$ бит)		Формат IBM 360 ($t=32$ бит)	
	B	$B_{\text{действ}}$	B	$B_{\text{действ}}$
2	0.0054	—	0.032	—
3	0.13	0.10	0.32	0.15
4	0.6	0.4	1.0	1.0
5	1.3	1.3	1.0	2.0
6	2.4	2.4	3.2—5.0 ^{a)}	3.8

^{a)} При $N=6$ вычислить порог трудно из-за шестнадцатеричного (4-битового) округления, применяемого на IBM 360; это порождает для данного значения ряд проблем.

получаются по конечным представлениям весов фильтров с помощью ЭВМ, используемой при самой фильтрации. Для фильтров рассматривалась каскадная реализация. Как и в табл. 5.1, возьмем $T = 0.005$. Частота Найквиста равна в этом случае 100 Гц, что позволяет перейти к процентам на шкале частот. Действительные значения точек отсечки приводятся в табл. 5.2.

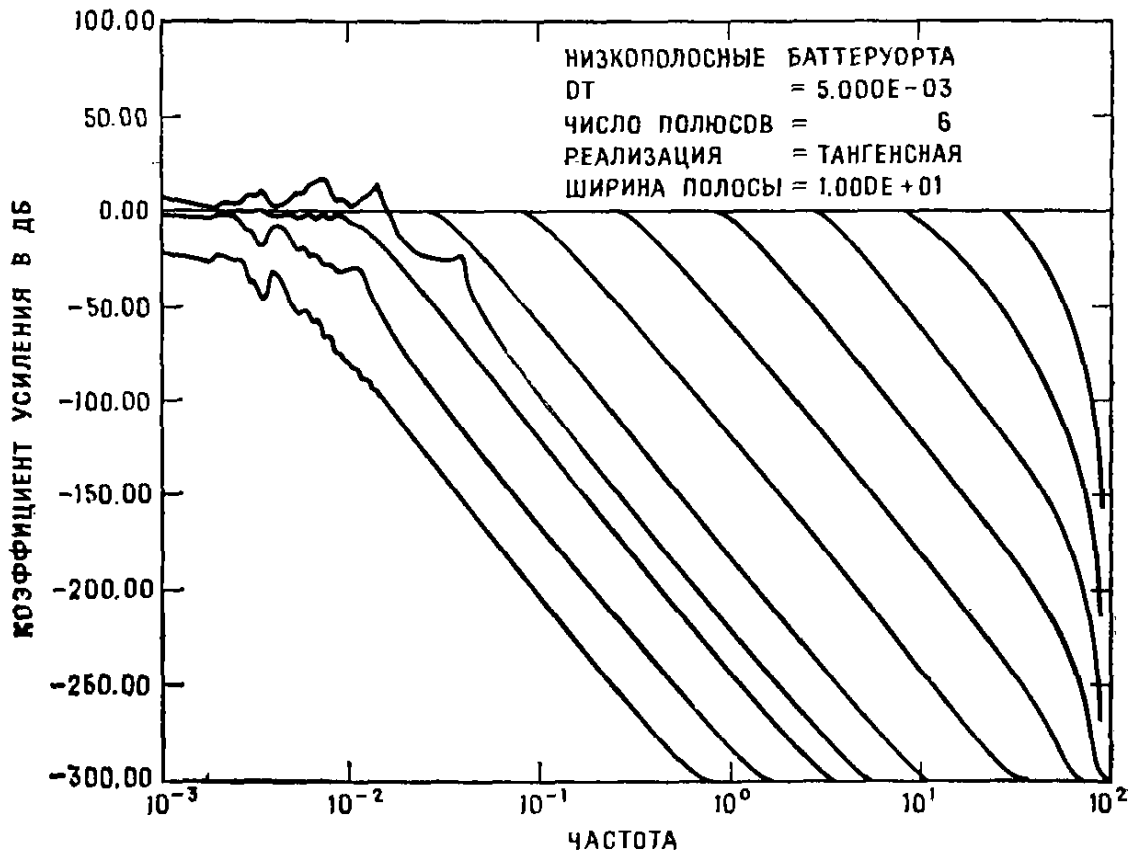


Рис. 5.4. Расчет десяти тангенсных низкочастотных фильтров Баттеруорта с разными точками отсечки, выполненный на ЭВМ SRU 1108. Здесь $T = 0.005$, $F = 100$ Гц. Для фильтра с наименьшей точкой отсечки смещение настолько сильно, что он попал между седьмым и восьмым фильтрами.

Соответствующие веса составляющих фильтров приводятся в табл. 5.3.

После шестого фильтра в этом ряду происходит ухудшение характеристик фильтров. Это ухудшение выражается не только в том, что в полосе пропускания коэффициент усиления испытывает заметные всплески, но и в положении, которое начинает занимать 3-децибеловая точка. Точки отсечки 7—10 фильтров смещены вправо от того положения, которое они должны были занимать теоретически. Каждый из этих последних фильтров или может оказаться, или оказывается неустойчивым. В частности, частота отсечки для наименьшего значения B сместилась настолько сильно, что попала между значениями точек отсечки 0.0032 и 0.01 Гц.

Понятие плоскости коэффициентов и ее области устойчивых коэффициентов без труда переносится на случай фильтров комбинированного типа более высоких порядков. Например, для фильтра третьего порядка с разностным уравнением

$$y(i) = b_0 x(i) - a_1 y(i-1) - a_2 y(i-2) - a_3 y(i-3)$$

Таблица 5.2

Частоты отсечки и поведение 10 фильтров, приведенных на рис. 5.4

Фильтр	Точка отсечки (Гц)	Поведение фильтра
1	10.	Хорошее
2	3.1625	»
3	1.	»
4	0.3162	»
5	0.1	»
6	0.0316	»
7	0.01	Пограничное или плохое
8	0.0032	Плохое
9	0.001	»
10	0.0003	»

Таблица 5.3

Веса шестиполосных тангенсных фильтров Баттеруорта (в каскадной реализации) с диапазоном частот отсечки от 0.01 Гц до 10 Гц (см. рис. 5.4)^{a)}

	B	b_0	a_{11}	a_{21}	a_{12}	a_{22}	a_{13}	a_{23}
1	10.00000	0.02046936		0.540254		0.641352		0.851887
2	3.16228	0.00231823	-1.464868	0.825146	-1.561018	0.868927	-1.761249	0.949944
3	1.00000	0.00024185	-1.816146	0.941106	-1.859712	0.956544	-1.940330	0.983972
4	0.31623	0.00002452	-1.940148	0.980991	-1.955578	0.986049	-1.982893	0.994871
5	0.10000	0.00000247	-1.980893	0.993949	-1.985951	0.995567	-1.994772	0.998375
6	0.03162	0.00000025	-1.993939	0.998083	-1.995557	0.998596	-1.998365	0.999486
7	0.01000	0.00000002	-1.998082	0.999393	-1.998595	0.998596	-1.999485	0.999486
			-1.999393		-1.999556	0.999556	-1.999837	0.999837

^{a)} Частота выборки — 200 выб./с. Для нерекурсивных членов указан только общий постоянный множитель; у каждой части фильтра имеется по три члена.

область устойчивых коэффициентов фильтров можно представлять как тетраэдр в пространстве троек всевозможных коэффициентов фильтров (a_1, a_2, a_3) . Все тройки коэффициентов, которые попадают внутрь этого тетраэдра в трехмерном пространстве, соответствуют устойчивым фильтрам. Разумеется, что плоскость, определенная уравнением $a_3 = 0$, дает ту же плоскость, что была и в двумерном случае.

Аналогично для фильтра с бесконечной областью отклика (ИБО-фильтра), имеющего M полюсов, областью устойчивости будет многогранник в M -мерном пространстве. Коэффициенты (a_1, \dots, a_M) , попадающие в него, соответствуют устойчивым фильтрам. Можно, конечно, и более подробно остановиться на этом случае, но мы не будем этого делать по той причине, что комбинированные формы фильтров при $M \geq 2$, вообще говоря, не рекомендуются. В этом случае обращаются к каскадным формам фильтров, которые могут быть устойчивыми даже тогда, когда соответствующий фильтр комбинированного вида будет неустойчивым. Короче говоря, каскадные реализации фильтров всегда имеют больший диапазон устойчивости, меньший уровень шума и искажений, чем комбинированные. Поэтому каскадные реализации предпочтительнее.

Разумеется, не обходится без исключений. В вычислительном плане каскадная реализация менее экономна, чем комбинированная. Так что для использования устойчивой (на данной ЭВМ) комбинированной формы фильтра при обработке больших объемов данных нет никаких препятствий.

Веса для комбинированных реализаций первых семи фильтров низких частот Баттеруорта синусного типа приводятся в табл. 5.4. Из этой таблицы видно, что веса устойчивых комбинированных фильтров имеют значительно большие абсолютные значения. Если для фильтров каскадного типа справедливы оценки

$$|a_1| < 2, \quad |a_2| < 1,$$

то соответствующие веса комбинированных фильтров оцениваются следующим образом:

$$|a_k| < \binom{M}{k} = \frac{M!}{k!(M-k)!}.$$

[Замечание. Эти оценки, дающие границы для коэффициентов, не могут служить условиями устойчивости. Хотя коэффициенты устойчивых фильтров обязательно удовлетворяют этим оценкам, можно привести и такие коэффициенты, лежащие в тех же границах, что соответствующие им фильтры будут неустойчивыми (например, $a_1 = 1$, $a_2 = -0.5$).]

Таким образом, комбинированная форма фильтра требует для реализации большего числа битов и в целой, и в дробной частях. Это приводит к тому, что общее число полезных битов при комбинированной реализации уменьшается.

Таблица 5.4

Веса шестиполосных синусных фильтров Баттеруорта
(в комбинированной реализации)
с диапазоном частот отсечки от 0.01 Гц до 10 Гц^{а)}

<i>B</i>	<i>b</i> ₀	<i>a</i> ₁	<i>a</i> ₂	<i>a</i> ₃	<i>a</i> ₄	<i>a</i> ₅	<i>a</i> ₆
1	10.000000	0.512018-03		09.6857		06.4863	0.298003
			-4.79661		-10.5252		-2.14663
2	3.162300	0.791620-06		13.1553		11.5766	0.581305
			-5.61649		-16.4476		-4.34912
3	1.000000	0.904631-09		14.4005		13.8295	0.885697
			-5.87863		-18.8154		-5.42168
4	0.316230	0.943210-12		14.8088		14.6205	0.962342
			-5.96162		-19.6191		-5.81099
5	0.100000	0.955728-15		14.9394		14.8791	0.987935
			-5.98786		-19.8789		-5.93960
6	0.031623	0.959648-18		14.9808		14.9617	0.996169
			-5.99616		-19.9616		-5.98083
7	0.010000	0.953604-21		14.9939		14.9879	0.998787
			-5.99879		-19.9878		-5.99393
Предел	0.000000-00	-6.00000	15.0000	-20.0000	15.0000	-6.00000	1.000000

^{а)} Скорость выборки — 200 выб./с.

Нерекурсивные коэффициенты как показатели устойчивости. В гл. 4 указывалось, что с целью нормировки фильтра обычно вводится постоянный множитель. При этом коэффициент усиления фильтра в полосе пропускания становится равным единице. Оказывается, этот постоянный множитель доставляет полезную информацию об устойчивости.

Поясним, как это получается. Предположим, что $\hat{H}(f)$ — передаточная функция до нормировки, f_{\max} — частота, в которой значение $|\hat{H}(f)|^2$ достигает наибольшей величины. Передаточную функцию $H(f)$ определим в этом случае посредством формулы

$$H(f) = \frac{\hat{H}(f)}{|\hat{H}(f_{\max})|}.$$

Эту нормировку обычно проводят, умножая все нерекурсивные члены на $1/|\hat{H}(f_{\max})|$. Обозначим символами $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_M$ нерекурсивные веса фильтра комбинированного вида до нормировки. Коэффициент \hat{b}_0 по определению полагается равным единице.

Тогда

$$b_k = \frac{\hat{b}_k}{|\hat{H}(f_{\max})|} = b\hat{b}_k,$$

где b — нормирующий множитель:

$$b = \frac{1}{|\hat{H}(f_{\max})|}.$$

Мы вскоре убедимся, что с уменьшением ширины полосы пропускания соответственно уменьшается значение b . В момент достижения такой ширины полосы пропускания, что соответствующее значение b окажется равным машинному нулю, возникают два взаимосвязанных явления. Легче всего их проследить на примере синусоиды $x(i)$ с максимальной частотой f_{\max} :

$$x(i) = \sin(2\pi f_{\max} iT), \quad i \geq 0.$$

Предположим, что i достаточно велико для того, чтобы был достигнут установившийся режим. Если ширина полосы пропускания стремится к нулю, то в частоте f_{\max} рекурсивные веса дают в пределе -1 (читатель может проверить это, полагая сначала $f_{\max} = 0$, а затем переходя к общему случаю). Поэтому в том особом случае, когда синусоида берется в частоте f_{\max} и достигается установившийся режим, рекурсивную составляющую фильтра можно записать в виде

$$-\sum_{k=1}^M a_k y(i-k) = \sin(2\pi f_{\max} iT + \varphi) - \varepsilon_i,$$

где ε_i — малый синусоидальный член (также имеющий частоту f_{\max}), а φ — фазовый угол, определяемый фильтром. Следовательно,

$$y_i = \sum_{k=0}^M b_k x(i-k) + \sin(2\pi f_{\max} iT + \varphi) - \varepsilon_i,$$

что равносильно соотношению

$$\varepsilon_i = b \sum_{k=0}^M \hat{b}_k x(i-k).$$

Таким образом, результатом воздействия нерекурсивной составляющей фильтра оказывается синусоида с относительно малой амплитудой, сокращающая член ε_i .

Первый случай (фиксированная точка). Если в этом случае b меньше 2^{-m} , где m — число битов, отведенных в слове ЭВМ для дробной части, то b оказывается равным нулю. Поэтому

на фильтр не подается никаких значений. Кроме того,

$$|\hat{H}(f_{\max})| = \frac{1}{b} = \frac{1}{0} = \infty,$$

и, значит, фильтр неустойчив.

Второй случай (плавающая точка). Этот случай чуть более сложен, но в основном сводится к тому же результату: хотя для чисел с плавающей точкой диапазон представления шире, при сложении одно число может выступать по отношению к другому в качестве нуля. Если слово с плавающей точкой имеет m битов для мантиссы и числа x и y связаны неравенством $|x| > |2^m y|$, то $x + y = x$. Таким образом, при суммировании в результате *антипереполнения* x полностью поглощает значение y . Значит, хотя y может оказаться и не настолько малой величиной, чтобы быть машинным нулем, при сложении в режиме плавающей точки она может выступить в качестве нуля.

Следовательно, если $b < 2^{-m}$, где m — число битов мантиссы, отведенных для мантиссы слова с плавающей точкой, то для данных, значения которых превосходят единицу, величина b может выступать в качестве нуля, и результат получится такой же, как и в первом случае.

Итак, множитель b позволяет сделать вывод об относительной устойчивости фильтра, реализованного в комбинированной форме.

То же относится и к постоянному множителю b^1 : $b^1 = b^{2/M}$ в случае каскадной реализации.

Нормирующий множитель b помогает также определить число битов, необходимых для реализации фильтра. Величина b должна быть такой, чтобы $b > 2^{-m}$, где m — число битов дробной части машинного слова, используемого при реализации. Если прологарифмировать это неравенство (основание логарифма возьмем равным 2), то получим оценку $m > -\log_2 b$. Таким образом, по известному значению b легко определить минимальное значение m . Оказывается, что такое соотношение между b и m накладывает определенные условия на $H(f)$. Предположим, что значение b неизвестно, а для $H(f)$ известен только общий вид. Спрашивается, можно ли получить отсюда информацию о множителе b и, значит, о числе битов m ? Ответ положительный.

Допустим, что передаточная функция $H(f)$ отвечает устойчивому фильтру и что ее нули расположены на верхней полуплоскости (на левой полуплоскости для переменной $j\omega$). Нетрудно установить, что ненормированная передаточная функция $\hat{H}(f)$ удовлетворяет в этом случае соотношению

$$\int_0^F \log_2 |\hat{H}(f)|^2 df = 0.$$

Как указано в работе Грэя и Мэркела (1974), из него вытекают некоторые важные следствия. В частности,

$$\begin{aligned} \int_0^F \log_2 |H(f)|^2 df &= \int_0^F \log_2 |b\hat{H}(f)|^2 df = \\ &= \int_0^F 2 \log_2 b df + \int_0^F \log_2 |\hat{H}(f)|^2 df = 2F \log_2 b. \end{aligned}$$

Следовательно,

$$\log_2 b = \frac{1}{2F} \int_0^F \log_2 |H(f)|^2 df.$$

Таким образом, число битов, необходимое для реализации данного нормированного фильтра с минимальной фазой, определяется формулой

$$m > -\frac{1}{2F} \int_0^F \log_2 |H(f)|^2 df.$$

Из этой оценки непосредственно следует, что фильтр не может быть реализован, если функция $|H(f)|^2$ равна 0 на каком-нибудь интервале частот f ненулевой длины.

Поясним на примерах, как используется эта формула.

1. Биномиальный фильтр M -го порядка (относящийся к ИКО-фильтрам), имеет веса

$$b_k = \binom{M}{k} 2^{-M}, \quad k = -M, \dots, M,$$

и такую передаточную функцию $H(f)$, что

$$|H(f)|^2 = \cos^{2M}(\pi fT).$$

Следовательно,

$$-\int_0^F \log_2 |H(f)|^2 df = -2M \int_0^F \log_2 (\cos(\pi fT)) df = M,$$

и для дробной части слова ЭВМ требуется M битов. В справедливости этого результата помогает убедиться тот факт, что коэффициенты

$$b_0 = b_M = 1/2^M$$

при числе битов, меньшем M , обратятся в нули.

2. Фильтр низких частот первого порядка

$$y(i) = (1 - \alpha)x(i) + \alpha y(i - 1), \quad 0 < \alpha < 1,$$

имеет коэффициент b , равный $1 - \alpha$. Поскольку $m > -\log_2 b$, при $(1 - \alpha) = 2^{-N}$ имеем $m > N$. Эти соотношения вместе с интегральной формулой, на которой они основываются, дают важную информацию о фильтрах.

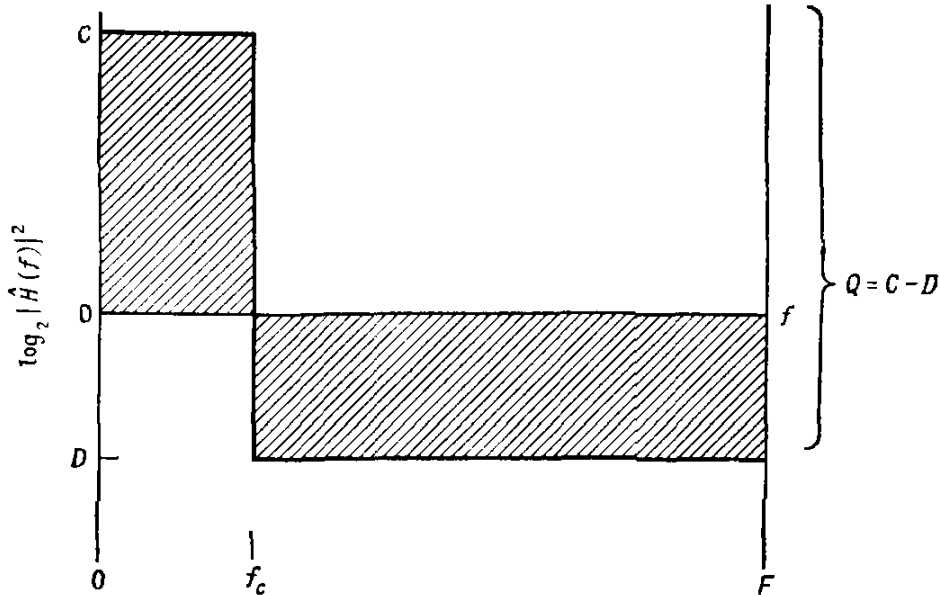


Рис. 5.5. Идеализированный низкочастотный фильтр.

Рассмотрим ненормированный фильтр низких частот, показанный на рис. 5.5. Согласно интегральному соотношению

$$\int_0^F \log_2 |\hat{H}(f)|^2 df = 0,$$

должно выполняться равенство

$$Cf_c = -D(F - f_c).$$

В левой части этой формулы стоит выражение для площади прямоугольника, расположенного над нулевым уровнем, а справа — выражение для площади прямоугольника, расположенного под нулевым уровнем.

Решим это уравнение относительно C :

$$C = -D \left(\frac{F - f_c}{f_c} \right).$$

Значение коэффициента b определяется формулой $b = 2^{-C/2}$. Для примера выберем значение $f_c = F/4$, полагая $C - D = 20$. Это соответствует интервалу 60 дБ. После несложных подсчетов можно получить значение $b = 1/181$.

На рис. 5.6 изображен полученный результат, сопоставленный с реальным фильтром. В качестве последнего выступает шестиполосный шестинулевой фильтр Чебышева типа 2 с гаше-

нием на 60 дБ в полосе заграждения и с частотой отсечки $f_c = F/4$. Нормирующий множитель оказывается равным $1/132$. Фильтр Чебышева имеет значительно менее четкие очертания, чем фильтр идеальной формы, представленный на рис. 5.5.

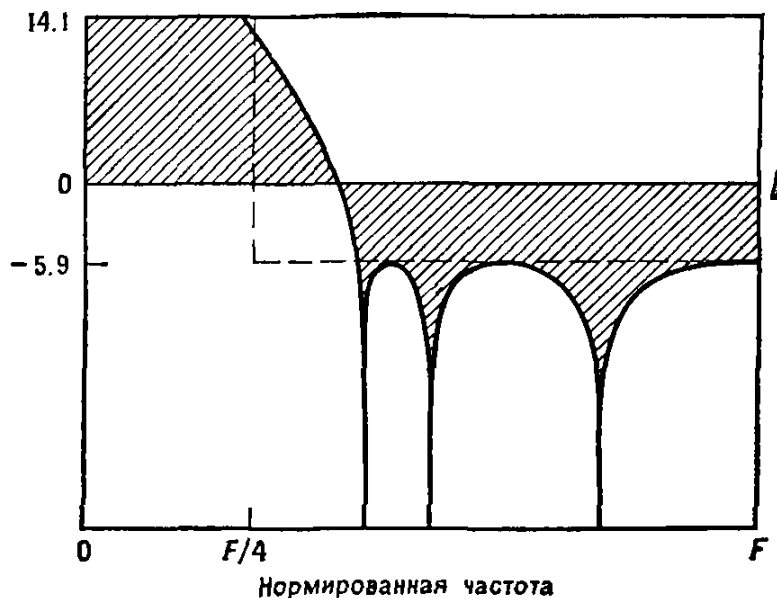


Рис. 5.6. Примерные очертания фильтра Чебышева второго типа. Интеграл для заштрихованной области равен нулю.

5.4. РЕАЛИЗАЦИЯ ФИЛЬТРА

В этом параграфе затронуты некоторые процедуры, применяемые в практической фильтрации.

На рис. 5.7 показана подпрограмма VFILT1. Она предназначена для получения одного значения выхода фильтра по одному значению на входе. Подпрограмма выглядит следующим образом.

1. Для связи с внешней программой FILTR1 в предложении COMMON перечисляются общие переменные: X — значение на входе (одно), Y — значение на выходе (тоже одно), B0 — единый для всех частей комбинированной реализации множитель, A1 — рекурсивные веса a_{1m} , A2 — рекурсивные веса a_{2m} , M — число полюсов фильтра.

2. Обращение к подпрограмме по основному названию VFILT1 происходит только один раз. При этом устанавливаются начальные значения, обнуляется память и т. д.

3. Вход в подпрограмму по названию FILT1 дает возможность получить одно новое значение выхода по одному новому значению входа. При этом значение входа размещается в области переменной X, затем совершается обращение к подпрограмме по названию FILT1 и значение выхода фильтра помещается в область переменной Y.

4. Веса фильтра следует поместить в область COMMON программы FILTR1.

```

SUBROUTINE BFILT1
COMMON /FILTR1/ X, Y, B0, A1(10), A2(10), M
DIMENSION Y0(11), Y1(11)
DO 10 K = 1, 11
Y0(K) = 0.
10 Y1(K) = 0.
M2 = M - M/2
RETURN
ENTRY FILT1
Y0(1) = X
DO 20 K = 1, M2
K1 = K + 1
YI = B0 * Y0(K) - A1(K) * Y0(K1) - A2(K) * Y1(K1)
Y1(K1) = Y0(K1)
20 Y0(K1) = YI
Y = YI
RETURN
END

```

Рис. 5.7. Подпрограмма BFILT1. Обратите внимание на то, что она имеет две точки входа. Вход BFILT1 используется лишь один раз — для начала операции фильтрации. Вход FILT1, следующий за ним, используется по одному разу для каждой точки данных, которую следует подвергнуть воздействию фильтра.

Поясним роль этой подпрограммы на примере программы для фильтрации. Допустим, что основная программа имеет следующий вид:

```

COMMON XX(1000), YY(1000) } задание областей COMMON
COMMON/FILTR1/X, Y, т. д. }
.
.
CALL BFILT1 } установка начальных значений для фильтрации
.
.
M = 6
T = 0.005
V = 25.
CALL LPSB(...) } получение весов шестипольного синусно-
                    го фильтра низких частот

```

·
·
·
(получение любых данных XX)
·
·
·

```
DO 100 I=1, 1000
X = XX(I)
CALL FILT1
YY(I) = Y
100 CONTINUE
```

} цикл, в котором непосредственно проводится фильтрация

·
·
·
(вывод на дисплей или на другое устройство)

Операцию фильтрации данных можно непосредственно вставить в основную программу. Однако ее выделение в подпрограмму имеет ряд преимуществ.

Предложения COMMON используются по двум причинам. Во-первых, отсутствие данных в операторах вызова сокращает время работы программы. Во-вторых, в предложении ENTRY не требуется перечисления переменных. Это позволяет пользоваться программой на тех машинах, на которых возможность такого перечисления не предусмотрена.

Программа VFILT1 предназначена только для таких фильтров, у которых уравнение для каждой части каскадной реализации имеет вид

$$y^{(m)}(i) = b_0 y^{(m-1)}(i) - a_{1m} y^{(m)}(i-1) - a_{2m} y^{(m)}(i-2),$$

где $y^{(m-1)}(i)$ — вход m -й составляющей фильтра, а $y^{(m)}(i)$ — соответствующий выход этой составляющей. Таким образом, в каждом уравнении присутствует только один нерекурсивный член, коэффициент при котором остается одним и тем же для всех составляющих фильтра.

В более общем случае соответствующее уравнение имеет вид

$$y^{(m)}(i) = b_{0m} y^{(m-1)}(i) + b_{1m} y^{(m-1)}(i-1) + \\ + b_{m2} y^{(m-1)}(i-2) - a_{1m} y^{(m)}(i-1) - a_{2m} y^{(m)}(i-2).$$

На рис. 5.8 представлена подпрограмма VFILT2, которая полностью аналогична подпрограмме VFILT1. Единственное отличие связано с увеличением до трех числа нерекурсивных членов для каждой составляющей фильтра. С этим, естественно, связано увеличение необходимого объема памяти и числа перемещений данных.

Обе подпрограммы имеют один недостаток. При их работе значительная часть времени уходит на сдвиги данных.

В подпрограмме BFILT3 (рис. 5.9) сделана попытка обойти этот недостаток заменой непосредственного перемещения данных

```

SUBROUTINE BFILT2
COMMON /FILTR2/ X, Y, B0(10), B1(10), B2(10),
*      A1(10), A2(10), M
DIMENSION Y0(11), Y1(11), Y2(11)
DO 10 K = 1, 11
Y0(K) = 0.
Y1(K) = 0.
10  Y2(K) = 0.
M2 = M - M/2
RETURN
ENTRY FILT2
Y0(1) = X
DO 20 K = 1, M2
K1 = K + 1
YI = B0(K) * Y0(K) + B1(K) * Y1(K) + B2(K) * Y2(K)
      - A1(K) * Y0(K1) - A2(K) * Y1(K1)
Y2(K1) = Y1(K1)
Y1(K1) = Y0(K)
20  Y0(K1) = YI
Y2(1) = Y1(1)
Y1(1) = Y0(1)
Y = YI
RETURN
END

```

Рис. 5.8. Подпрограмма BFILT2 аналогична подпрограмме BFILT1; исключение составляет более широкий набор рекурсивных коэффициентов.

перестановкой индексов. Эта подпрограмма работает очень хорошо, особенно в том случае, если она написана не на алгоритмическом, а на машинном языке. В то же время подпрограмма на Фортране не всегда обеспечивает сокращение времени работы, поскольку на разных ЭВМ компиляция происходит неодинаково. Некоторые компиляторы, плохо приспособленные к работе с индексами, не дают более эффективной машинной подпрограммы. Поставив во главу угла скорость, следует или написать подпрограмму на машинном языке, или поэкспериментировать с программами на

Фортране и определить самую быструю из них для имеющейся ЭВМ.

На рис. 5.10 показана подпрограмма, в которой реализована циклическая буферная схема, применяемая для ИКО-фильтров.

```

SUBROUTINE BFILT3
COMMON /FILTR3/ X, Y, B0(10), B1(10), B2(10),
      A1(10), A2(10), M
DIMENSION YM (11,3)
DO 10 K = 1, 11
DO 10 I = 1, 3
10 YM (K, I) = 0.
M2 = M - M/2
J1 = 1
J2 = 3
J3 = 2
RETURN
C
C
ENTRY FILT3
YM (1, J1) = X
DO 20 K = 1, M2
K1 = K + 1
YMM = B0(K) * YM(K,J1) + B1(K) * YM(K,J2)
*      + B2(K) * YM(K,J3) - A1(K) * YM(K1,J2)
*      - A2(K) * YM(K1,J3)
20 YM(K1,J1) = YMM
J4 = J1
J1 = J3
J3 = J2
J2 = J4
Y = YMM
RETURN
END

```

Рис. 5.9. Подпрограмма BFILT3. Применяется аналогично программе BFILT2. В подпрограмме BFILT3 предусмотрена перестановочная буферная схема, обеспечивающая циклический сдвиг.

В этой программе используется большой внутренний вектор, предназначенный для запоминания 1001 значения. В программе этот массив назван BUFFER. Предполагается, что разностное

уравнение фильтрации имеет вид

$$y(i) = \sum_{k=-M/2}^{M/2} b_k x(i-k),$$

где M — нечетное. Если дополнительно предположить, что фильтр симметричен, то можно это уравнение переписать в виде

$$y(i) = b_0 x(i) + \sum_{k=1}^{M/2} b_k [x(i-k) + x(i+k)].$$

Поскольку внутри буфера никаких перемещений данных не происходит, время, необходимое для поиска нужной части данных, увеличивается по сравнению с более простыми, нециклическими буферными схемами. Как и в случае рекурсивного

```

SUBROUTINE BFILT4
COMMON /FILTR4/ X, Y, B(1001), M
DIMENSION BUFFER (1001)
DO 5 I = 1, 1001
5  BUFFER (I) = 0.
LOC = 1
M2 = M/2
RETURN
ENTRY FILT4
BUFFER(LOC) = X
L0 = LOC - M2
IF (L0. LT. 1) L0 = L0 + M
Y = BUFFER (L0) * B(1)
DO 10 I = 1, M2
LL = L0 - I
IF (LL. LT. 1) LL = LL + M
LR = L0 + I.
IF (LR. GT. M) LR = LR - M
Y = Y+(BUFFER(LL)+BUFFER(LR)) * B(I + 1)
10 CONTINUE
LOC = LOC + 1
IF (LOC. GT. M) LOC = LOC - M
RETURN
END

```

Рис. 5.10. Циклическая буферная схема для симметричных ИКО-фильтров.

фильтра, можно сделать эту процедуру весьма эффективной, переходя к машинному языку, хотя соответствующая программа на Фортране могла быть довольно плохой.

5.5. ДЕЦИМАЦИЯ

Собственно говоря, слово «децимация» означает потерю одной десятой военных сил в ходе сражения или в результате наказания¹⁾. В обработке сигналов этим словом принято называть удаление ненужной или излишней информации и соответственное уплотнение данных.

Как правило, сокращение выборки децимацией необходимо в целях экономии. Допустим, что в течение 30 мин проводится тестовое испытание со скоростью выборки, равной 400 выб./с. В результате будет получено $30 \cdot 60 \cdot 400 = 720\,000$ значений данных. Если попытаться изобразить их на графике, длина которого равна 30 дюймам, то на каждый дюйм попадет 24 000 значений. На каждую точку, которую может отметить графопостроитель, имеющий интервал 0.01 дм, приходится по 240 точек выборки. В этом случае при децимации должна быть оставлена для дальнейшей обработки не более чем $1/240$ часть всей выборки.

В большинстве случаев децимации сопутствует необходимость низкочастотной фильтрации. Эта фильтрация должна предотвратить подмену частот, т.е. появление в низких частотах информации, относящейся к высоким частотам.

В качестве примера рассмотрим временную последовательность

$$x(i) = \cos\left(\frac{3\pi i}{4}\right).$$

Первые десять точек этой последовательности изображены в верхней части рис. 5.11. Все эти точки принадлежат синусоиде с частотой, составляющей 75% частоты Найквиста. Для наглядности точки нанесены на график косинуса. Если каждую вторую точку, начиная с $i=1$, выбросить, то получится последовательность, изображенная на нижней части рисунка. Точки этой последовательности можно принять за значения косинуса с частотой, равной 25% исходной частоты Найквиста.

Для предотвращения такой подмены частот следует применять низкочастотную фильтрацию. Точка отсечки определяется в этом случае следующим образом.

Допустим, что индекс децимации равен p (т.е. в новой последовательности останется только каждое p -е значение исходного ряда), а исходная частота Найквиста равна $F = S/2 = 1/(2T)$ Гц. Тогда новая частота Найквиста должна быть равной

¹⁾ В Древнем Риме, во времена Республики, децимация служила наказанием за проявленную в сражении трусость: провинившийся отряд обычно делили на группы по 10 солдат и в каждой группе бросался жребий. Тот, кому выпадал жребий, подвергался ужасной каре: с него сдирали кожу его же товарищи.

p -й части исходной частоты Найквиста, т. е. F/p Гц, поскольку всякая частота, попадающая в интервал $(F/p, F)$, после децимации может появиться в частотах, меньших новой частоты Найквиста. Итак, децимации должна предшествовать фильтрация данных с помощью фильтра низких частот, точка отсечки которого равна $F' = 1/(2pT)$ Гц. Формулы децимации сведены в табл. 5.5.

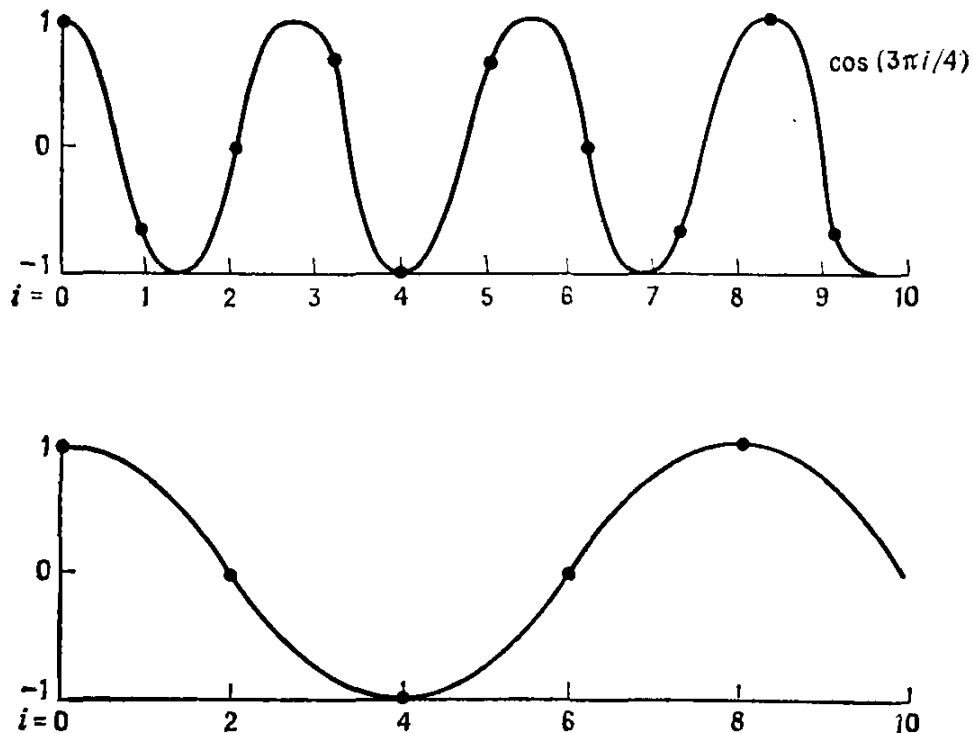


Рис. 5.11. Косинусоида до (вверху) и после (внизу) децимации, при которой выбирается каждая вторая точка.

Выбор фильтра. Точку отсечки фильтра, применяемого при децимации, можно выбрать несколькими способами. Одним таким стандартным способом предусматривается выбор точки отсечки

Таблица 5.5

Формулы децимации в терминах величин T , S и F

	Через T	Через S	Через F
Скорость выборки и интервал	$S = \frac{1}{T}$	$T = \frac{1}{S}$	$S = 2F$ $T = 1/(2F)$
Частота Найквиста	$F = \frac{1}{2T}$	$F = \frac{S}{2}$	F
Частота Найквиста после децимации	$F' = \frac{1}{2pT}$	$F' = \frac{S}{2p}$	$F' = \frac{F}{p}$
Новый интервал и скорость	$T' = pT$ $S' = 1/(pT)$	$T' = p/S$ $S' = S/p$	$T' = p/(2F)$ $S' = 2F/p$

фильтра в частоте F' , т. е. в 3-децибеловой точке. В этой точке мощность исходного сигнала падает наполовину, а соответствующая амплитуда уменьшается примерно на 30%. Примерно такое же уменьшение происходит для информации, которая имеет частоты чуть правее точки отсечки. Эта информация может появляться левее новой частоты Найквиста. Во втором способе предусматривается снижение амплитуды данных, которые могут вызывать подмену частот, до заданных пределов. Например, можно потребовать, чтобы в новой точке отсечки F' мощность сигнала падала бы на 60 дБ. Такое уменьшение мощности гарантирует применение фильтра низких частот с глушением на 60 дБ информации, расположенной правее точки F' . Данные с частотами, близкими к частоте Найквиста, в этом случае могут потеряться. Итак, при первом способе выбора 3-децибеловое падение мощности может вызывать подмену частот, что не позволяет делать выводы о низких частотах с достаточной уверенностью; во втором способе проблема подмены частот уже не играет такой существенной роли, но зато использование глушения на 60 дБ может приводить к потере полезных данных с частотами ниже F' .

Вслед за определением точки отсечки фильтра по величине F' возникает проблема выбора типа фильтра (рекурсивного — ИБО или нерекурсивного — ИКО). Помимо этого, нужно определить число коэффициентов конкретного типа фильтра. Преимущества и недостатки выбора того или иного фильтра сформулированы в табл. 5.6.

Как указано в таблице, при выборе нерекурсивных фильтров отпадает необходимость вычисления каждого значения выхода. В этом случае при децимации p к 1 нужно вычислять только каждое p -е значение выхода. В этом состоит главное преимущество выбора нерекурсивных фильтров.

Какого-то определенного правила выбора необходимого типа фильтра сформулировать нельзя. Помимо индекса децимации этот выбор диктуется многими другими факторами, в частности характером аппаратного оборудования программного обеспечения. Наша точка зрения по этому вопросу сводится к следующему.

Во-первых, мы считаем, что для децимации, проводимой на устройствах с фиксированной точкой (назовем здесь аппаратное оборудование специального назначения, мини- и микро-ЭВМ), более предпочтительным является выбор нерекурсивного фильтра.

Во-вторых, на ЭВМ широких возможностей или на вычислительных машинах для обработки больших массивов данных с плавающей точкой и объемом слова не меньше 32 битов более подходящими могут оказаться рекурсивные фильтры: исследователь заранее знает характер той информации, которую он может получить, и, значит, ему не нужно сначала анализировать фильтр и только потом применять его. Кроме того, важным преимуще-

Таблица 5.6

Преимущества и недостатки ИКО- и ИБО- фильтров при децимации

ИКО	ИБО
Преимущества	
<ol style="list-style-type: none"> 1. Необходимо вычислять только каждую p-ю точку выхода 2. Фильтр всегда устойчив 3. Легко проводить бесфазовую фильтрацию 4. Фильтрацию можно проводить в режиме фиксированной точки на минн-ЭВМ 	<ol style="list-style-type: none"> 1. Объем необходимой памяти как для весов, так и для данных сильно сокращается 2. Как правило, можно добиться, чтобы фильтр имел в полосе пропускания почти линейную фазу 3. Часто наиболее эффективный фильтр (в вычислительном плане)
Недостатки	
<ol style="list-style-type: none"> 1. Может потребоваться большое число весовых коэффициентов 2. Возможна зашумленность 3. Может потребоваться большой объем памяти для весовых коэффициентов и для данных 	<ol style="list-style-type: none"> 1. Должны вычисляться все значения на выходе. (Имеются формулировки, в которых не требуется, чтобы вычислялись все значения на выходе, однако они оказываются обычно менее эффективными) 2. Возможна зашумленность и даже неустойчивость 3. Бесфазовая фильтрация невозможна (если только данные не проходят через один и тот же фильтр дважды, каждый раз в противоположных направлениях) 4. Наилучшая реализация — в арифметике с плавающей точкой

ством такого выбора, особенно при обработке большого числа рядов, служит меньший объем памяти, необходимый для запоминания данных.

Децимацию может оказаться необходимым проводить в два и более приемов. Децимация в несколько этапов носит название *каскадной* децимации. Предположим, что $p = q^2$. Тогда процесс децимации естественно разбить на два шага, в каждом из которых проводится децимация q к 1. Например, допустим, что $f_s = 1\,000\,000$ и необходима децимация с индексом $p = 10\,000$ (т. е. 10 000 к 1). Для этого потребуется фильтр, точка отсечки которого располагается в частоте, равной $1/10\,000$ частоты Найквиста. На многих ЭВМ такого значения точки отсечки невозможно добиться. Вместе с тем разбиение процесса на два этапа с децимацией 100 к 1 на таких ЭВМ сделать нетрудно. Поэтому, если фильтр для одного этапа реализовать невозможно, прибегают к последовательной фильтрации.

По отношению к результатам фильтрации расположение фильтров в последовательной форме, вообще говоря, не коммутативно. Например, если требуется децимация 30 к 1, то с учетом того,

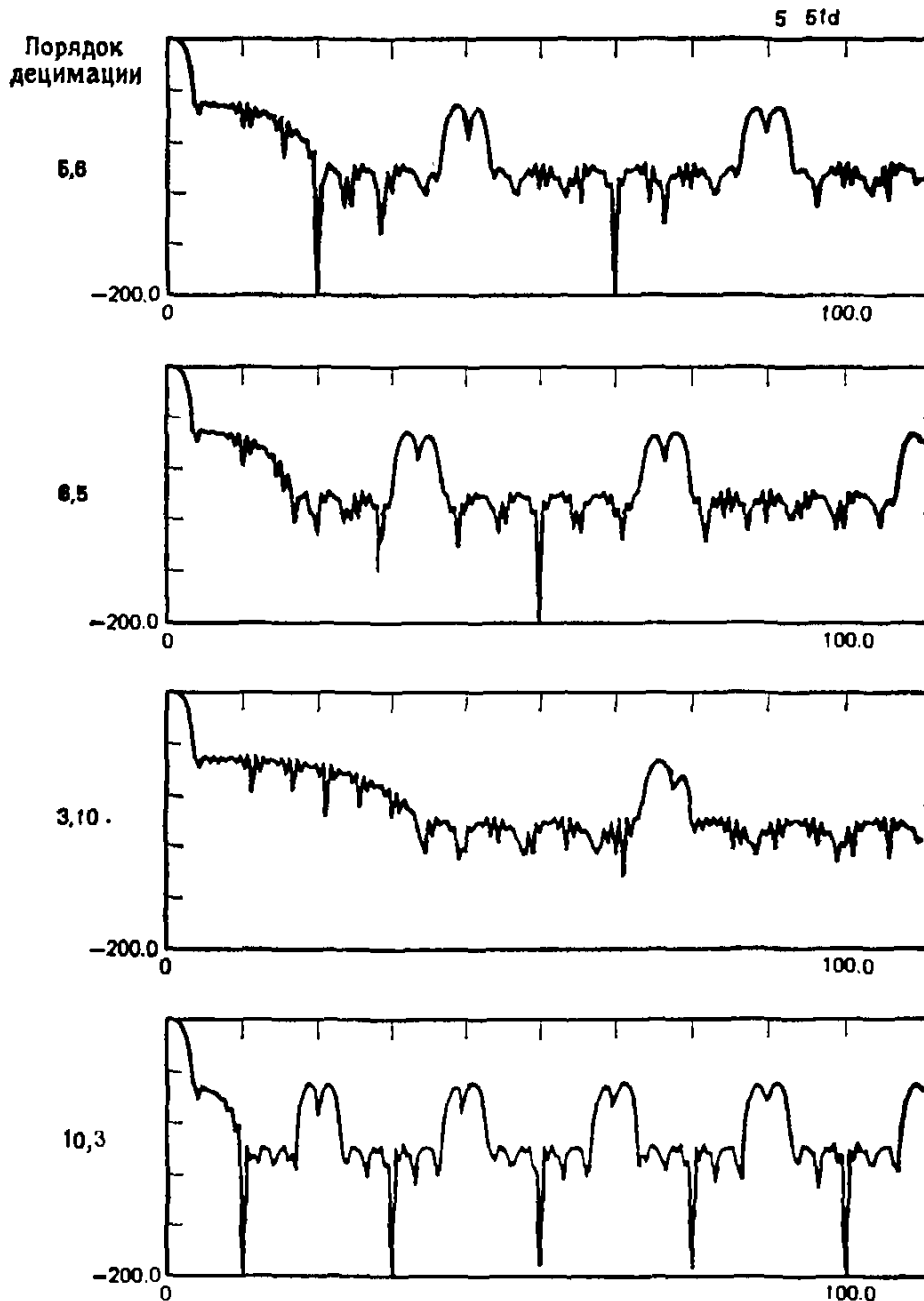


Рис. 5.12. Четыре варианта фильтров, осуществляющих децимацию 30 к 1.

что $30 = 2 \cdot 3 \cdot 5$, ее можно провести $3 \cdot 2! + 3! = 12$ различными способами (например, 5 к 1 и 6 к 1, 6 к 1 и 5 к 1, 3 к 1 и 10 к 1, 10 к 1 и 3 к 1 и т. д.).

На рис. 5.12 показаны передаточные функции для перечисленных в скобках вариантов каскадной децимации. Эти общие передаточные функции получены для нерекурсивных фильтров, реализованных с помощью процедуры Паркса — Мак-Клеллана,

рассмотренной в предыдущей главе. Параметры для фильтров, использованных перед децимацией, приведены в следующей таблице (частота Найквиста полагается равной 100 Гц):

Децимация	M	Частота отсечки (Гц)
3 к 1	18	13.3
5 к 1	30	8.
6 к 1	36	6.7
10 к 1	60	4.

Глушение в полосе заграждения этих фильтров равно примерно -46 дБ. При каскадной реализации этот уровень почти везде становится в два раза ниже, за исключением полос с подменной частот, где появились двойные выступы.

Хотя полоса пропускания во всех четырех показанных случаях и одинакова, полоса заграждения в каждой из них непохожа на другие. Например, полоса заграждения при децимации 5 к 1 и 6 к 1 выглядит иначе, чем в случае 6 к 1 и 5 к 1. Однако максимальные значения выступов в полосах заграждения нигде не превышают -46 дБ, что дает право говорить о том, что характеристика составного фильтра ничуть не хуже, чем исходного.

При фильтрации с помощью нерекурсивных фильтров, предотвращающей подмену частот, следует обратить внимание на следующее.

Во-первых, при фильтрации требуется только p -я часть выхода. Поэтому схему фильтрации необходимо изменить так, чтобы $p-1$ значений входа только запоминалось и никаких других операций с ними не проводилось. По p -й точке входа должно вычисляться одно значение выхода.

Во-вторых, значение M выбирается, как правило, от $5p$ до $10p$ или более в зависимости от необходимых характеристик в полосе заграждения (чем больше M , тем ниже значения коэффициента усиления в полосе заграждения).

В-третьих, вместо запоминания данных можно ограничиться только запоминанием *частных сумм*. Этот прием может сохранить значительный объем памяти.

Поясним, как это происходит. Главная экономия происходит за счет запоминания примерно M/p частных сумм вместо $5p-10p$ значений данных, которые запоминаются при обычных способах. Иначе говоря, если $M=rp$, то необходимо запоминание только r частных сумм. При этом для каждого нового значения входа выполняется r умножений этого значения на разные коэффициенты и каждое такое произведение добавляется к нужной частной сумме. Каждая частная сумма с p точками выступает в качестве значения выхода, и начинает вычисляться новая частная сумма.

Прием, в котором вычисляются частные суммы, позволяет сократить в p раз объем памяти, необходимый при других стандартных методах. Это имеет важное значение при обработке данных, поступающих по многим каналам.

5.6. ОБРАТНАЯ ДЕЦИМАЦИЯ

Иногда вместо уменьшения требуется увеличение скорости выборки. В частности, это может понадобиться для получения согласованных скоростей выборок двух несовпадающих временных рядов. Более подробно остановимся на этом в § 5.7. А сей-

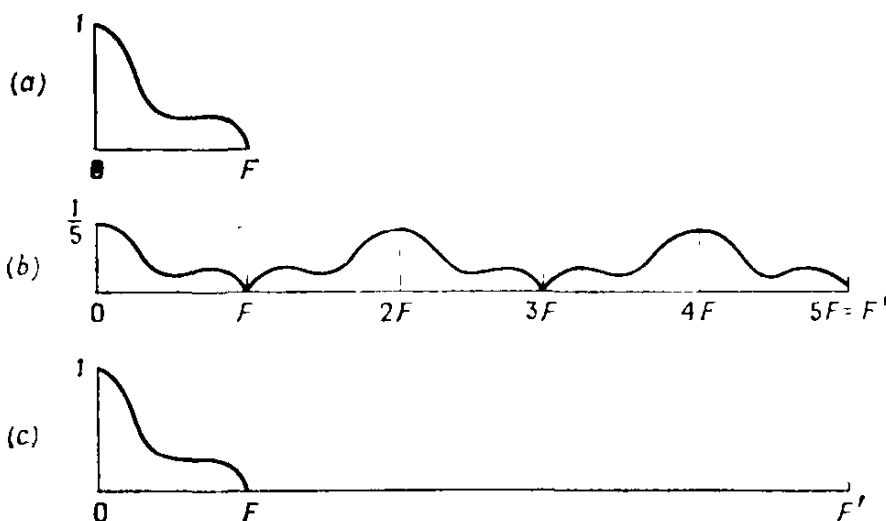


Рис. 5.13. Амплитуды спектров на разных стадиях обратной децимации функции данных.

час покажем, как производится обратная децимация — процесс, при котором получаются новые значения между старыми. Эта операция включает в себя интерполяцию и выглядит следующим образом.

Сначала между исходными данными вставляется необходимое число дополнительных нулей. При обратной децимации 1 к p между соседними данными вставляется $p - 1$ нулей. Затем исходные данные умножаются на p . Наконец, новая последовательность исходных данных (умноженных на p) и нулей пропускается через фильтр низких частот, точка отсечки которого совпадает с исходной частотой Найквиста.

Предположим, например, что скорость выборки исходных данных равнялась 200 выб./с и требуется увеличить эту скорость до 1000 выб./с. Этому соответствует обратная децимация 1 к 5. Если исходные данные обозначить $x(0), x(1), \dots$, то цифровая фильтрация будет проводиться с последовательностью $5x(0), 0, 0, 0, 0, 5x(1), 0, 0$ и т. д.

Фильтр низких частот производит интерполяцию и уменьшение амплитуды данных. Из-за этого понижения мощности в слу-

чае обратной децимации в отличие от прямой децимации приходится умножать данные на p .

Получающуюся картину помогает понять рис. 5.13. Предположим, что амплитуда преобразования Фурье данных имеет вид, представленный на рис. 5.13, *a*. Добавление между исходными данными нулей приводит к изменениям, отраженным на рис. 5.13, *b*. Общая высота графика понизилась, а мощность периодически

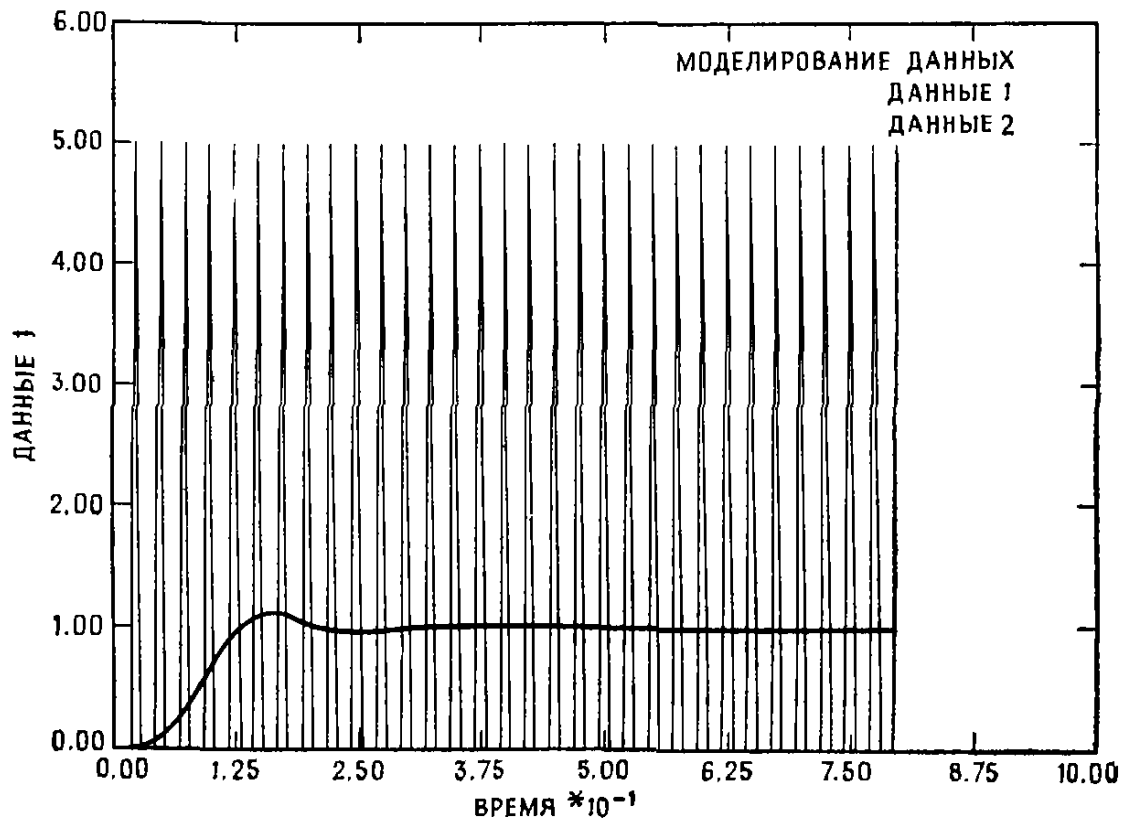


Рис. 5.14. Исходная функция и полученная из нее после обратной децимации 1 к 5. Функция в виде гребешка получена добавлением четырех нулей между единичными значениями исходной функции, умноженными на 5. Вторая функция получена после фильтрации с помощью подходящего цифрового фильтра.

рвспределилась на новом частотном диапазоне. Отметим, что на всех рисунках показан диапазон от 0 до частоты Найквиста. На рис. 5.13, *c* показан результат фильтрации, описанной выше. Вся мощность, располагавшаяся между F и F' , пропала, поэтому никаких изменений из-за подмены частот в амплитудах не происходит.

В качестве второго примера приведенной процедуры рассмотрим последовательность данных до и после фильтрации, изображенную на рис. 5.14. Исходный ряд представлял собой последовательность из 33 точек, каждый элемент которой равнялся единице. Как и в предыдущем случае, индекс обратной децимации выбран равным 5. Между каждыми двумя единичными значениями при этой обратной децимации 1 к 5 вставляется по 4 нуля, а

сами значения умножаются на 5. При этом получается функция, показанная на рис. 5.14 (она имеет форму гребешка). Вторая функция на этом рисунке есть результат фильтрации гребнеобразной функции с помощью шестиполосного фильтра низких частот.

Если исключить из рассмотрения начальный участок, соответствующий неустановившемуся режиму, то фильтрованные данные оказываются очень близкими к единице, т. е. данные снова имеют постоянное значение, равное единице.

5.7. СВЕДЕНИЕ К ОБЩЕЙ СКОРОСТИ ВЫБОРКИ

В некоторых задачах сравнения двух временных рядов (например, при вычислениях кросс-корреляции, кросс-спектра и передаточной функции) требуется получить для двух рядов одинаковую скорость выборки. Чтобы добиться такого согласования скоростей выборок, достаточно применить прямую или обратную децимации. Однако, как будет видно, следует соблюдать определенную осторожность, чтобы возникающие при этом потери информации не оказались слишком большими.

Предположим, что скорости выборок равны S_1 и S_2 соответственно. Рассмотрим два случая.

Первый случай. Пусть $S_1 = pS_2$, где p — целое число. В этом случае лучше всего провести децимацию первой последовательности с индексом p (т. е. оставить только p -ю часть исходной последовательности). Разумеется, можно сделать и иначе, т. е. провести обратную децимацию 1 к p второй последовательности. Однако этого не следует делать по двум причинам. Первая из них состоит в том, что число данных для обработки увеличится в p раз. Вторая причина связана с возможностью появления ложной информации о частотах выше $S_2/2$, не содержащейся в данных второй последовательности, если они были предварительно подвергнуты фильтрации.

Второй случай. В случае $S_1 \neq pS_2$ определяется наименьшее общее кратное S_1 и S_2 , т. е. такое наименьшее число S_3 , что и S_1 , и S_2 делят его без остатка. Иначе говоря, существуют такие целые числа p и q , что $S_1p = S_2q = S_3$. Если S_1 и S_2 взаимно просты, то $p = S_2$ и $q = S_1$ (предельный случай).

После этого нужно выбрать большее из чисел p и q . Допустим, что это большее число есть p . Тогда для второй последовательности проводится обратная децимация 1 к q , а затем прямая децимация p к 1. Общая скорость выборки будет равна

$$\frac{S_2q}{p} = \frac{S_3}{p} = \frac{S_1p}{p} = S_1.$$

Предложенная процедура дает наименьшую из возможных скоростей выборки, что всегда желательно. Например, положим $S_1 = 800 = 2^5 \cdot 5^2$, $S_2 = 1000 = 2^3 \cdot 5^3$. Наименьшим общим кратным этих чисел будет $S_3 = 2^5 \cdot 5^3 = 4000$, т. е. $p = 5$, $q = 4$. Поэтому для второй последовательности нужно сначала провести обратную децимацию 1 к 4 (в результате получится скорость выборки 4000 выб./с), а затем провести прямую децимацию 5 к 1, чтобы получить скорость выборки, которая использовалась для получения первой последовательности.

Добавим, что при обеих децимациях фильтрации приводят к изменению фазового угла или к увеличению задержки, отвечающей второй последовательности. Это следует учитывать в тех задачах анализа, где фаза и задержка могут иметь значение, например в задачах анализа кросс-спектра.

5.8. КОМПЛЕКСНАЯ ДЕМОДУЛЯЦИЯ

Понятие комплексной демодуляции уже упоминалось в § 4.8 в связи с полоснопропускающими фильтрами. Собственно говоря, это понятие относится к умножению последовательности на комплексную экспоненту $\exp(j2\pi f_c iT)$. В результате такого умножения происходит сдвиг $X(f)$ на частоту f_c . Покажем, как это происходит. Пусть

$$X(f) = T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi f iT).$$

Определим последовательность $y(i)$ формулой

$$y(i) = x(i) \exp(j2\pi f_c iT).$$

Это равносильно умножению $x(i)$ на синус и косинус с частотой f_c Гц. Преобразование Фурье последовательности $y(i)$ в этом случае имеет вид

$$\begin{aligned} Y(s) &= T \sum_{i=-\infty}^{\infty} y(i) \exp(-j2\pi siT) = \\ &= T \sum_{i=-\infty}^{\infty} x(i) \exp(j2\pi f_c iT) \exp(-j2\pi siT) = \\ &= T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi (s - f_c) iT). \end{aligned}$$

Если переменную $s - f_c$ заменить величиной f , т. е. сделать замену $f = s - f_c$ (следовательно, $s = f + f_c$), то уравнение примет вид

$$Y(f + f_c) = T \sum_{i=-\infty}^{\infty} x(i) \exp(-j2\pi f iT) = X(f).$$

Таким образом, комплексной демодуляции во временной области отвечает сдвиг в частотной области.

Комплексная демодуляция обычно проводится при помощи двух фильтров, как это показано на рис. 5.15. В прямоугольнике справа вычисляются амплитуда и фаза входящих сигналов.

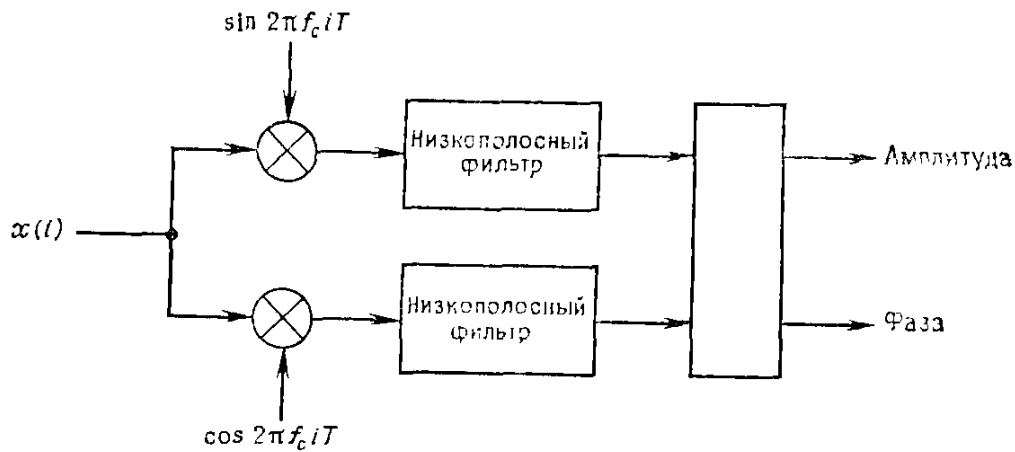


Рис. 5.15. Комплексная демодуляция.

Делается это следующим образом. Определим последовательности

$$x_s(i) = x(i) \sin(2\pi f_c iT), \quad x_c(i) = x(i) \cos(2\pi f_c iT)$$

и обозначим символом $y_s(i)$ последовательность значений, полученных фильтрацией из последовательности $x_s(i)$, а $y_c(i)$ — из последовательности $x_c(i)$. Последовательность значений амплитуды и фазы определим формулами

$$a(i) = \sqrt{2(y_s^2(i) + y_c^2(i))}, \quad \varphi(i) = \operatorname{arctg}\left(-\frac{y_s(i)}{y_c(i)}\right).$$

Действие комплексной демодуляции поясним на примере последовательности

$$x(i) = A \cos(2\pi f_c iT + \varphi_c),$$

где φ_c — произвольный фазовый угол. Тогда

$$\begin{aligned} x_s(i) &= A \cos(2\pi f_c iT + \varphi_c) \sin(2\pi f_c iT) = \\ &= \frac{A}{2} \{ \sin[2\pi(f_c + f_c) iT + \varphi_c] + \sin[2\pi(f_c - f_c) iT - \varphi_c] \} = \\ &= \frac{A}{2} [\sin(4\pi f_c iT + \varphi_c) - \sin \varphi_c]. \end{aligned}$$

Если пренебречь задержкой по фазе, то для последовательности $y_s(i)$ получим выражение

$$y_s(i) = -\frac{A}{2} \sin \varphi_c$$

(слагаемое, соответствующее высокой частоте, при фильтрации исчезает). Аналогично

$$\begin{aligned} x_c(i) &= A \cos(2\pi f_c iT + \varphi_c) \cos(2\pi f_c iT) = \\ &= \frac{A}{2} \{ \cos[2\pi(f_c + f_c) iT + \varphi_c] + \cos[2\pi(f_c - f_c) iT + \varphi_c] \} = \\ &= \frac{A}{2} [\cos(4\pi f_c iT + \varphi_c) + \cos \varphi_c]. \end{aligned}$$

По тем же причинам, что и в случае $y_s(i)$, имеем

$$\begin{aligned} y_c(i) &= \frac{A}{2} \cos \varphi_c, \\ y_s^2(i) + y_c^2(i) &= \frac{A^2}{4} \sin^2 \varphi_c + \frac{A^2}{4} \cos^2 \varphi_c = \frac{A^2}{2}. \end{aligned}$$

Поэтому

$$a(i) = \sqrt{2(y_s^2(i) + y_c^2(i))} = A$$

и

$$\varphi(i) = \operatorname{arctg} \left(-\frac{y_s(i)}{y_c(i)} \right) = \operatorname{arctg} \left(\frac{\sin \varphi_c}{\cos \varphi_c} \right) = \varphi_c.$$

Следует помнить, что задержка, которой пренебрегли, в фазовом угле присутствует. Несмотря на это, рассмотренная процедура используется в первую очередь для определения передаточных функций: на входе и выходе исследуемого устройства устанавливаются два одинаковых комплексных демодулятора. Коэффициент усиления определяется из отношения входной последовательности к выходной, а фаза — из разницы двух фазовых углов. Применение одинаковых фильтров в обоих комплексных демодуляторах вносит соответствующий фазовый угол.

Упражнения

5.1. Используя двойную точность в режиме плавающей точки, нетрудно смоделировать на Фортране операции, которые производятся при фильтрации в режиме фиксированной точки. Предположим, что числа с фиксированной точкой, для которых проводится моделирование, имеют m битов в дробной части. Тогда все вычисления, за исключением умножения и округления, могут непосредственно производиться с двойной точностью. Для указанных исключений нужно предусмотреть две подпрограммы.

Моделирование округления

```
SUBROUTINE SIMRND (AA,M)
DOUBLE PRECISION AA,A,FACT
FACT=2.D0**M
A=AA*FACT+0.5D0
A=A-DMOD(A,1.D0)
AA=A/FACT
RETURN
END
```

Моделирование умножения

```

SUBROUTINE SIMMLT (A,B,C,M)
DOUBLE PRECISION A,B,C,CC
CC=A*B/(2.D0**M)
CALL SIMRND (CC,M)
C=CC
RETURN
END

```

а) Моделируя вычисления с фиксированной точкой, при помощи приведенных подпрограмм повторите упр. 4.7 (b и c). Вычисления проведите при $M=16, 12, 8$ и 4 , $T=0.005$ и только при одном значении $B=4$ Гц. Для делений при вычислениях передаточной функции потребуется подпрограмма моделирования деления с фиксированной точкой SIMDIV. Напишите эту подпрограмму, просмотрев предварительно подпрограмму SIMMLT.

б) Параллельно вычислениям, моделирующим вычисления с фиксированной точкой, проведите те же вычисления с плавающей точкой. Используя результаты вычислений с плавающей точкой в качестве контрольных, определите погрешность. Для этого получите разности результатов для двух вариантов использованной арифметики, возведите эти разности в квадраты и осредните по всей длине вычислений. Прокомментируйте полученный для погрешности результат.

5.2. В этой задаче приводится способ, равнозначный методу измерения функции $\hat{H}(f)$, описанному под вторым номером в § 5.2. При решении этой задачи следует воспользоваться тем моделированием арифметики с фиксированной точкой, которое рассмотрено в предыдущем упражнении. Общая программа для решений должна включать в себя следующие элементы:

а) Подпрограмму SIMSIN для рекурсивного вычисления $\sin(2\pi f_p i T)$. При ее составлении нужно воспользоваться алгоритмом, описанным в § 4.3 в п. «Фильтр второго порядка как осциллятор». В этой программе следует использовать одинарную точность в режиме плавающей точки. Полученный выход должен преобразовываться в числа с двойной точностью и округляться при помощи подпрограммы SIMRND.

б) Подпрограмму SIMFIL для вычисления значений последовательности

$$y(i) = b_0 x(i) - a_1 y(i-1) - a_2 y(i-2)$$

моделируя действия с фиксированной точкой.

с) Управляющую программу, в которой предусмотрено следующее.

1) Ввод значения m или его задание в операторе данных.

2) Получение весов фильтра низких частот первого или второго порядка (в последнем случае обращением к LP2) и преобразование результатов с плавающей точкой в форму с фиксированной точкой, используя моделирование округления.

3) Использование подпрограммы SIMSIN для получения синусоиды заданной частоты и подпрограммы SIMFIL для фильтрации этой синусоиды.

4) Преобразование значений, полученных после достижения установившегося режима на выходе фильтра, в форму чисел с плавающей точкой и одинарной точностью; вычисление среднего квадрата мощности. Это дает коэффициент усиления и передаточную функцию фильтра в частоте, с которой генерировалась синусоида.

5) Вычисления по общей программе при значениях $T=0.005$, $B=4$, $f_p=2, 4, 8, 16, 32, 64$ Гц, $m=16, 12, 8, 4$.

Общую программу просчитайте и для фильтра первого, и для фильтра второго порядка. Объясните полученные результаты.

5.3. Устойчив ли фильтр

$$y(i) = 2x(i) - 4x(i-1) + 8x(i-2) + 2y(i-1) + 4y(i-2).$$

Объясните ответ. Бывают ли патологические случаи?

5.4. Веса биномиального фильтра M -го порядка (нерекурсивного) определяются формулой

$$b_k = \binom{M}{k} / 2^M, \quad \text{где } k = -M, \dots, M.$$

Покажите, что $|H(f)|^2 = \cos^{2M}(\pi fT)$. Получите выражение для 3-децибеловой точки (т. е. для такого значения f , что $10 \log_{10} |H(f)|^2 = 10 \log_{10} 1/2$) через частоту и M . Что напоминает каскадная версия этого фильтра? Каковы нули этого фильтра?

5.5. Перепишите программу VFILT1. Используя ее, подпрограммы PRPLOT, LPSB и управляющую программу, аналогичную описанной в упр. 5.2, выполните следующие действия.

а) Фильтрацию последовательности $x(i) = 1, 0, 0, \dots, i = 0, \dots, 100$. Получите шесть таких последовательностей фильтрованных данных (по 101 точке в каждой), полагая во всех случаях $T = 0.005$.

б) Применяемый в каждом из шести случаев фильтр должен иметь шесть рекурсивных весов (три каскадные части). Точки отсечки должны быть равными 32, 16, 8, 4, 2, 1 Гц.

с) Прокомментируйте полученные результаты.

5.6. Напишите тестовую программу для сравнения эффективности подпрограмм VFILT2 и VFILT3. *Замечание:* на ЭВМ, которая предназначена для такой проверки, можно оценивать время работы обеих подпрограмм по-разному. Можно применить такую процедуру: получить по 1000 точек синусоиды; время фильтрации при помощи подпрограммы VFILT2 и время фильтрации при помощи подпрограммы VFILT3 выдать на печать. Если будет время, попытайтесь продолжить это сравнение и напишите подпрограмму VFILT3 на машинном языке. Просчитайте временной тест в этом случае. Это не простая задача, поэтому решать ее необязательно, хотя и желательно.

5.7. Напишите и проверьте работу подпрограммы VFILT4 для биномиального фильтра.

а) Положите $M = 8, T = 0.005, N = 1016$.

б) Получите и отфильтруйте синусоиду длины N . Отбросив первые и последние M точек, вычислите среднее и дисперсию значений выхода.

с) Частоту синусоиды f возьмите равной 9.3, 18.6, 37.2, 74.4 Гц. Прокомментируйте результаты.

5.8. Используя веса, полученные с помощью подпрограммы LPSBVG (рис. 4.17), повторите упр. 5.7, полагая $M = 8, T = 0.005$ и $B = 18.6$. Прокомментируйте результаты. Далее, вычислите передаточные функции обоих фильтров и сравните их. Что можно взять в качестве альтернативы подпрограммы TTRAN, которую в данном случае применять нельзя? Можно ли в этом случае выбрать FFTRAN? Если да, то почему?

5.9. Напишите тестовую программу для изучения вопросов, связанных с децимацией. В ней должно быть предусмотрено следующее.

а) Получение с помощью подпрограммы TDRAND 1024 значений данных, имеющих равномерное распределение. Из каждого значения должно вычитаться 0.5, чтобы среднее значение было равно 0.

б) Фильтрация данных с помощью подпрограмм VFILT1 и LPSB. Предполагается, что $T = 0.005$, фильтр имеет шесть полюсов и $B = 20$ Гц.

с) Децимация исходных и фильтрованных данных 4 к 1 ($p = 4$).

д) Вычисление для обеих последовательностей преобразования Фурье. (*Замечание:* для этого отдельно к каждой последовательности примените подпрограмму FFTRAN, задавая в качестве действительной части члены последовательности, а в качестве мнимой — нули. Существует и более эффективный

способ такой операции, но он будет затронут только в следующей главе.) Затем — вычисление квадратов абсолютных величин полученных значений преобразований Фурье.

В каждом преобразовании Фурье должно оставаться по 129 точек. Почему? Получите графическое изображение результатов и сравните их.

Повторите действия при $B = 12.5$ и 25 Гц. Сравните полученные результаты со случаем 20 Гц.

5.10. Предположим, что требуется децимация 10 000 к 1. Попробуйте для этой цели воспользоваться подходящим шестиполосным фильтром, полученным при помощи подпрограммы LPSB. Удовлетворительны ли результаты? Что может служить разумной альтернативой?

5.11. Получите 64 точки равномерно распределенного псевдослучайного шума с помощью подпрограммы TDRAND. Проведите обратную децимацию 1 к 4, включающую необходимую операцию фильтрации. Прокомментируйте результаты. Вычислите и получите в виде графика квадраты абсолютных значений преобразования Фурье последовательности данных, подвергнутых обратной децимации, до и после фильтрации. Прокомментируйте результаты.

Глава 6

ПРЕОБРАЗОВАНИЯ ФУРЬЕ

6.1. ЭЛЕМЕНТЫ ТЕОРИИ

Значительная часть задач анализа временных рядов связана с преобразованием Фурье и методами его эффективного вычисления. В этих задачах преобразование Фурье играет важную роль не только само по себе, но и как необходимый промежуточный шаг в определении плотностей спектра мощности, кросс-спектральных плотностей, передаточных функций, сверток и корреляционных функций (табл. 6.1). Кроме того, преобразование Фурье можно использовать в задачах интерполяции значений. Эта глава содержит довольно краткий обзор теоретических сведений, поскольку главный упор сделан на те свойства преобразования, которые важны в приложениях и позволяют проводить вычисления разных функций наиболее быстро и экономно. Подробнее с преобразованием Фурье и его применениями читатель может ознакомиться по многим книгам. Назовем весьма полезные для наших целей монографии Колмогорова и Фомина (1976), Винера (1963).

Вопросы, связанные с алгоритмом быстрого преобразования Фурье, освещены здесь весьма скупо. Читателю, которому потребуется более широкое знакомство с этими вопросами, рекомендуем обратиться к книгам Отнеса и Эноксона (1972) и Ахо, Хопкрофта и Ульмана (1979).

Определение преобразования Фурье. Приведем сначала уже упомянутое нами раньше определение преобразования Фурье на бесконечной прямой:

$$X(f) = \int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft) dt, \quad -\infty < f < \infty.$$

В этом определении функция $x(t)$ обычно бывает временной траекторией (хотя и не всегда). В таких случаях переменной t служит время, а переменной f — частота. Вопрос существования преобразования Фурье $X(f)$ для случайной функции $x(t)$ с математической точки зрения довольно сложен. Однако конечная длина записи и ограниченные амплитуды временных траекторий снимают этот вопрос, поскольку для таких траекторий преобразование Фурье всегда существует.

Таблица 6.1

Функции, получаемые с помощью БПФ

1. Плотность спектра мощности

$$S_{xx}(k) = \frac{1}{P} \overline{X^*(k) X(k)} = \frac{1}{P} |\overline{X(k)}|^2$$

2. Плотность кросс-спектра

$$S_{xy}(k) = \frac{1}{P} \overline{X^*(k) Y(k)}$$

3. Передаточная функция

$$H_{xy}(k) = \frac{Y(k)}{X(k)}$$

4. Свертка ^{a)}

$$u(i) = \text{БПФ}^{-1}[U(k)], \quad U(k) = X(k) Y(k)$$

5. Ковариация ^{a)}

$$s_{xy}(i) = \frac{N}{N-i} \text{БПФ}^{-1}[S_{xy}(k)]$$

$$S_{xy}(k) = \frac{1}{P} \overline{X^*(k) Y(k)}$$

а) Для ковариации и свертки по крайней мере последние половины последовательностей $X(k)$ и $Y(k)$ могут быть нулями.

Величины $x(i)$, $y(i)$ и $u(i)$, где $i=0, 1, \dots, N-1$, суть конечные временные последовательности $X(k)$, $Y(k)$ и $U(k)$ — соответствующие БПФ ($k=0, 1, \dots, \frac{N}{2}$ при четном N , $k=0, 1, \dots, \frac{N-1}{2}$ при нечетном N). Черта означает осреднение

Обратимся теперь к преобразованию Фурье, определенному на конечном интервале дискретных значений времени. Это преобразование вычисляется, как правило, при помощи алгоритма быстрого преобразования Фурье. Определение преобразования для комплексной функции $z(i)$, $i=0, 1, \dots, N-1$, имеет в этом случае вид

$$Z(k) = T \sum_{i=0}^{N-1} z(i) \exp\left(-j \frac{2\pi}{N} ik\right), \quad \text{где } k=0, 1, \dots, N-1. \quad (6.1)$$

Напомним, что в этом определении для краткости вместо $z(iT)$ используется обозначение $z(i)$. Такое же сокращение записи от $Z(bk)$ к $Z(k)$ для области частот обусловлено конечной длиной записи $z(i)$, приводящей к дискретным значениям частот; величина b определяется из соотношений

$$b = \frac{1}{NT} = \frac{1}{P}.$$

Такое определение величины b позволяет в условиях применимости теоремы о выборке во временной области пользоваться вариантом этой теоремы для области частот и наоборот. Знание N значений в области частот позволяет восстановить все N значений во временной области. Более того, для ограниченных во временной и частотной областях функции и ее изображения все промежуточные значения можно восстановить, используя преобразования как интерполяционные операторы (см. теорему о вы-

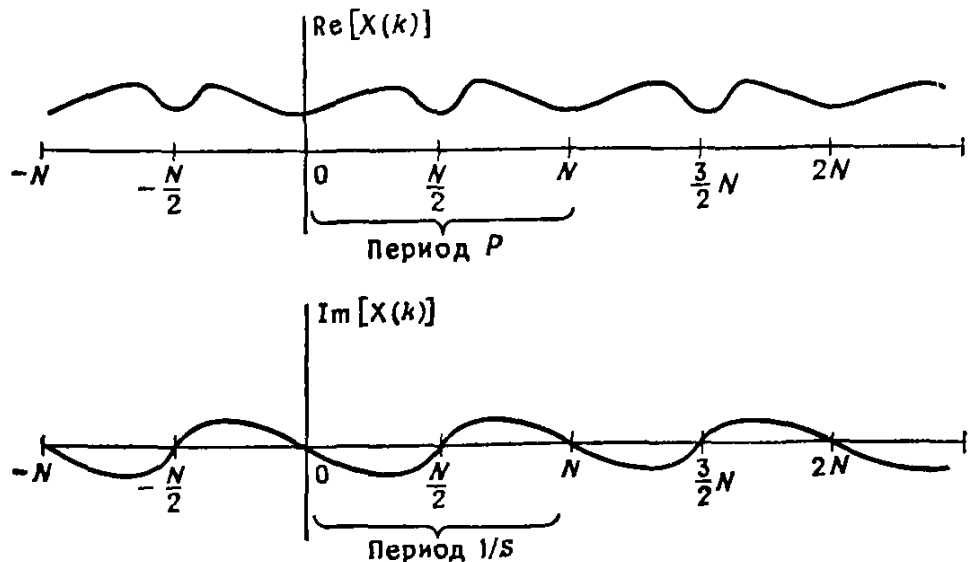


Рис. 6.1. Иллюстрация периодичности дискретного преобразования Фурье.

борке). Отметим периодичность $Z(k)$. Если попытаться продолжать вычисления при $k > N$, то полученные значения будут повторять уже имеющиеся значения $Z(k)$, т. е.

$$Z(N+k) = Z(k)$$

(рис. 6.1).

На практике в задачах анализа временных рядов обрабатываются не комплексные, а действительные величины. Поэтому определение (6.1) можно переписать в виде

$$X(k) = T \sum_{i=0}^{N-1} x(i) \exp\left(-j \frac{2\pi}{N} ik\right), \quad k = \begin{cases} 0, 1, \dots, \frac{N}{2} \\ \quad \quad \quad (N - \text{четное}), \\ 0, 1, \dots, \frac{N-1}{2} \\ \quad \quad \quad (N - \text{нечетное}). \end{cases} \quad (6.2)$$

Таким образом, для N значений данных получается примерно вдвое меньше значений частот. Если продолжать вычисления при следующих значениях k , то будут получаться числа, комплексно-

сопряженные уже имеющимся для первых значений k , а именно

$$X^*(N - k) = X(k) \quad (6.3)$$

(рис. 6.2).

Это непосредственно вытекает из формулы (6.2) и отражает тот факт, что косинус — четная, а синус — нечетная функции.

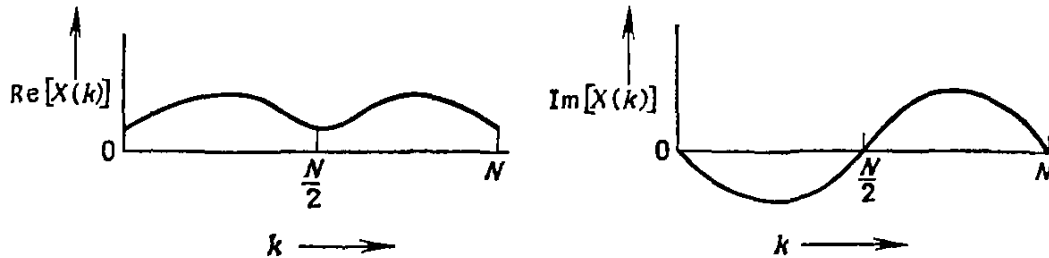


Рис. 6.2. Поведение, типичное для четной действительной и нечетной мнимой частей преобразования Фурье действительных функций.

Действительно,

$$\operatorname{Re}[X(k)] = T \sum_{i=0}^{N-1} x(i) \cos\left(\frac{i2\pi}{N} ik\right)$$

есть четная (симметрическая) функция (т. е. $\operatorname{Re} X(k) = \operatorname{Re} X(-k)$), а

$$\operatorname{Im}[X(k)] = -T \sum_{i=0}^{N-1} x(i) \sin\left(\frac{2\pi}{N} ik\right)$$

есть нечетная (антисимметрическая) функция ($\operatorname{Im} X(k) = -\operatorname{Im} X(-k)$). Доказательство предлагаем провести читателю (см. упр. 6.3).

Нетрудно видеть, что общее число действительных и мнимых частей частотной функции равно числу исходных данных временной функции. Отметим, что

$$X(0) = X^*(N) = X^*(0), \quad (6.4)$$

и для четных N

$$X\left(\frac{N}{2}\right) = X^*\left(N - \frac{N}{2}\right) = X^*\left(\frac{N}{2}\right). \quad (6.5)$$

Поэтому как в нулевой частоте, так и в частоте Найквиста значения преобразования оказываются действительными. В случае нечетного N величина $X([N-1]/2)$ не попадает точно в частоту Найквиста и, следовательно, имеет мнимую часть. Различие случаев четного и нечетного N отражено на рис. 6.3.

Заметим, что большинство современных мини-ЭВМ, используемых в системах для анализа Фурье, имеет алгоритм преобразования для последовательностей с числом членов, равным степени 2 (т. е. четным). Однако в больших вычислительных сис-

темах используются алгоритмы, в которых число данных может быть равным не только степени 2, но и 3, 4 и 5 (Синглтон (1969)). В этих случаях необходимо учитывать различия, связанные с четностью и нечетностью числа данных.

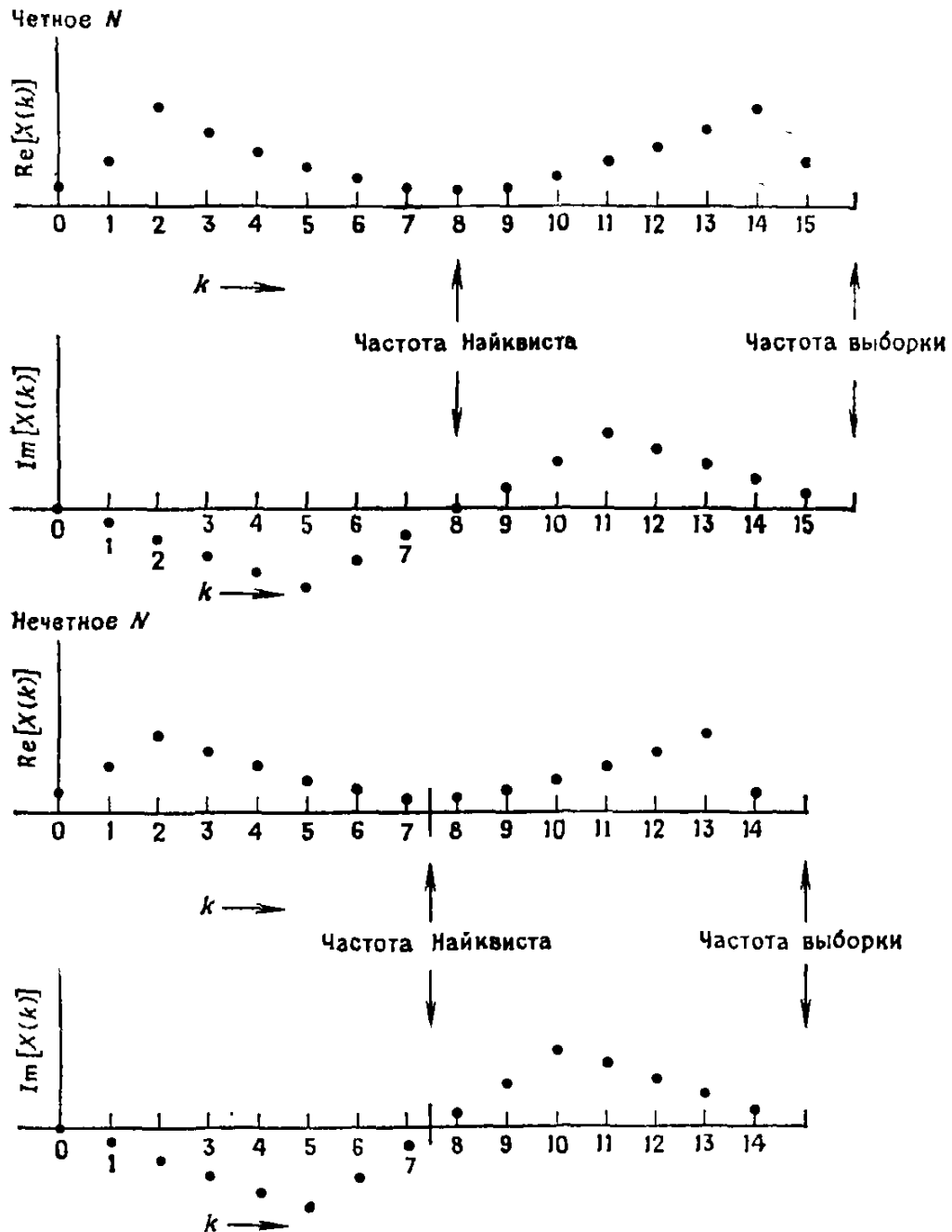


Рис. 6.3. Различия преобразований Фурье в случаях четного и нечетного числа точек.

Из формул (6.3)–(6.5) и четности действительной части преобразования Фурье вытекает, что N действительных компонент, полученных преобразованием из N действительных чисел, при *четном* N дают пару действительных (в частотах, соответствующих $k=0$ и $k=N/2$) и $(N-2)/2$ комплексных чисел (на дейст-

вительную и мнимую части которых уходит по две компоненты); при *нечетном* N только одна компонента ($k=0$) дает действительное число, а остальные $N-1$ уходят на $(N-1)/2$ комплексных чисел.

Следует отметить, то эти детали хотя и имеют определенное значение для правильного построения программ для ЭВМ, но в научно-технических задачах, связанных с преобразованием Фурье, не играют важной роли.

Полярные координаты (модуль и фаза). Значение преобразования Фурье в данной частоте — комплексное число — нередко записывают в полярных координатах

$$\begin{aligned} |X(k)| &= \{\operatorname{Re}^2[X(k)] + \operatorname{Im}^2[X(k)]\}^{1/2}, \\ \varphi_x(k) &= \operatorname{Arctg} \frac{\operatorname{Im}[X(k)]}{\operatorname{Re}[X(k)]}, \end{aligned} \quad (6.6)$$

где обозначение Arctg введено для главного значения.

Величину $|X(k)|$ называют по-разному — коэффициентом усиления или модулем. Величину $\varphi_x(k)$ называют фазой или аргументом. На рис. 6.4 в качестве иллюстрации приведены графики действительной части, мнимой части, модуля и фазы преобразования Фурье.

Обратное преобразование Фурье. Чтобы получить формулу, определяющую обратное преобразование Фурье, достаточно в определении (6.2) заменить знаки показателей экспонент на противоположные и поменять ролями $x(i)$ и $X(k)$, т. е.

$$x(i) = b \sum_{k=0}^{N-1} X(k) \exp \left[j2\pi \frac{ik}{N} \right], \quad \text{где } i=0, 1, \dots, N-1. \quad (6.7)$$

Отметим, что пределы суммирования в этой формуле требуют знания N значений $X(k)$. Поскольку величины $X(k)$ определялись только для k от 0 до $N/2+1$ или $(N-1)/2$, недостающие величины нужно получить, исходя из симметрии по Эрмиту значений преобразования Фурье действительной последовательности, т. е. по формуле (6.3).

Например, действительная и мнимая части преобразования Фурье действительной функции могут выглядеть так, как это показано на рис. 6.1. Неявно предполагается, что для этого преобразования N — четное число. Случай нечетного N будет рассмотрен ниже.

Нетрудно понять, что комплексно-сопряженную симметричную часть функции $X(k)$ совсем необязательно непосредственно вводить в правую часть (6.7). В пункте, посвященном специальному обратному преобразованию Фурье, рассмотрен метод, в котором предусмотрена такая возможность удаления лишних значений.

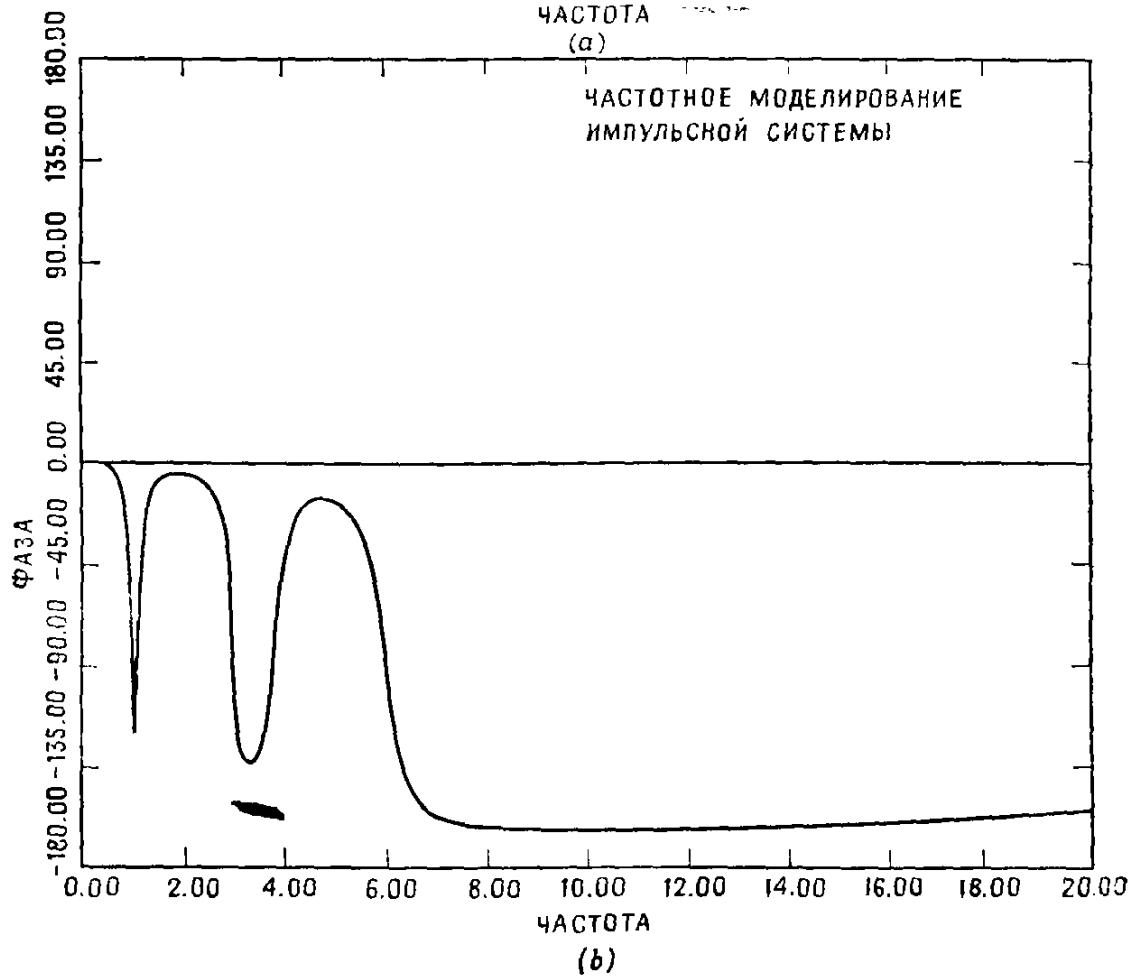
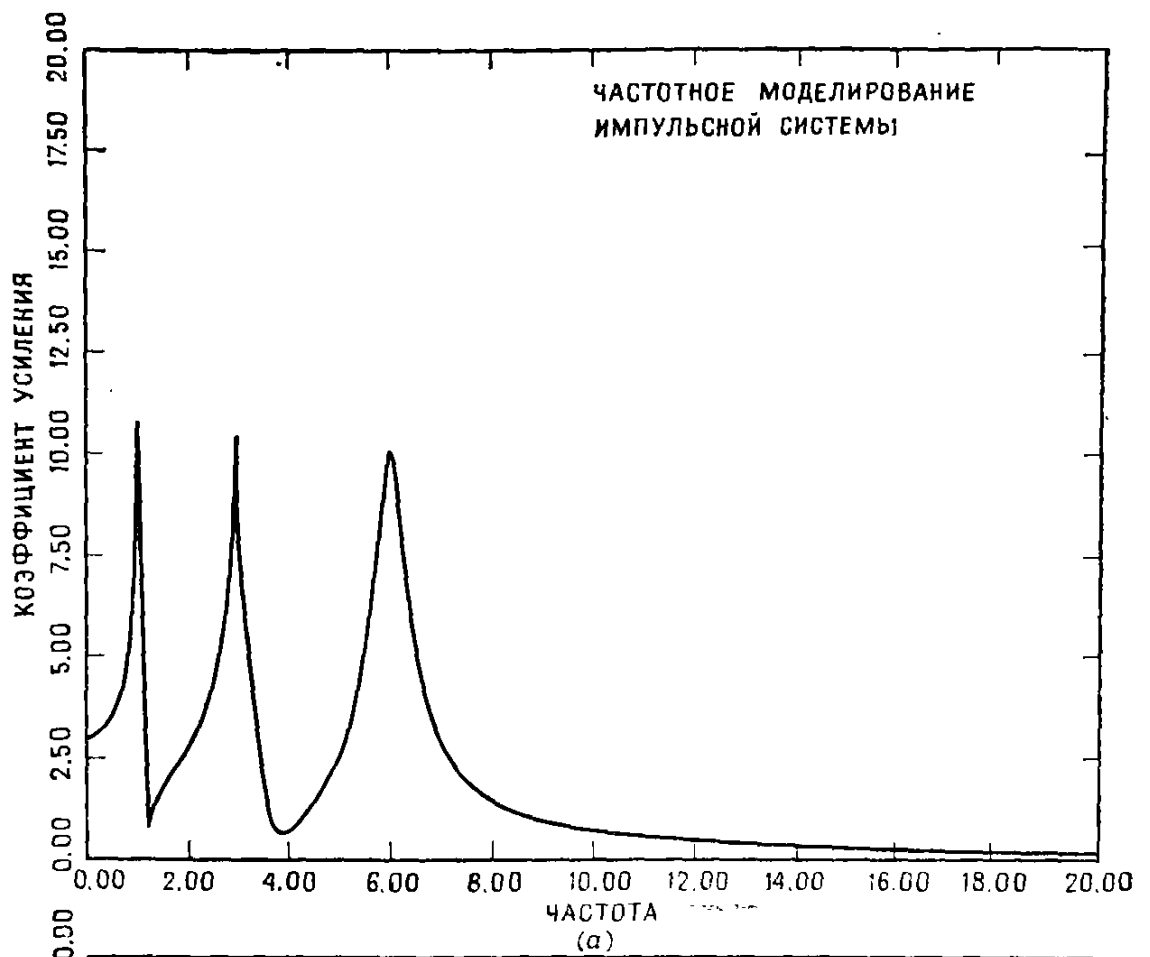
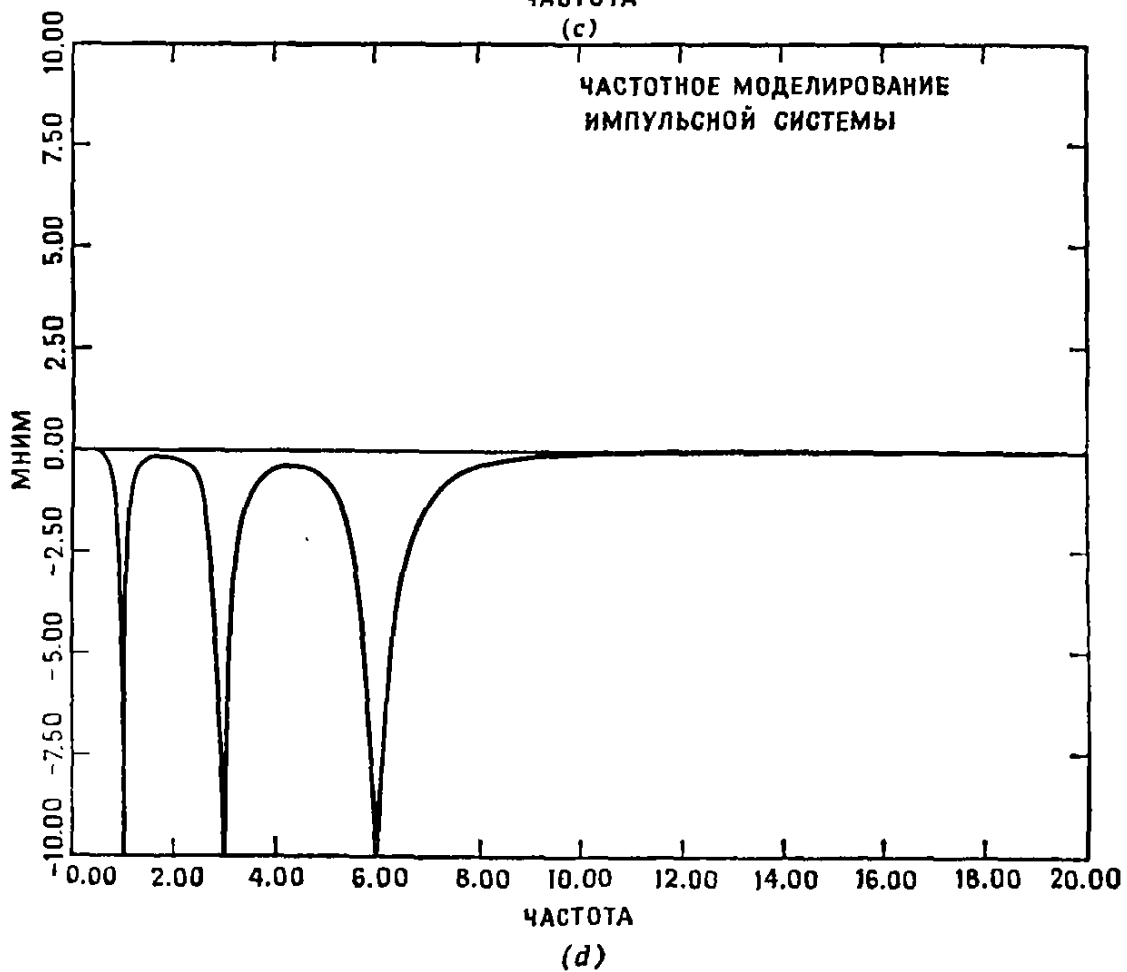
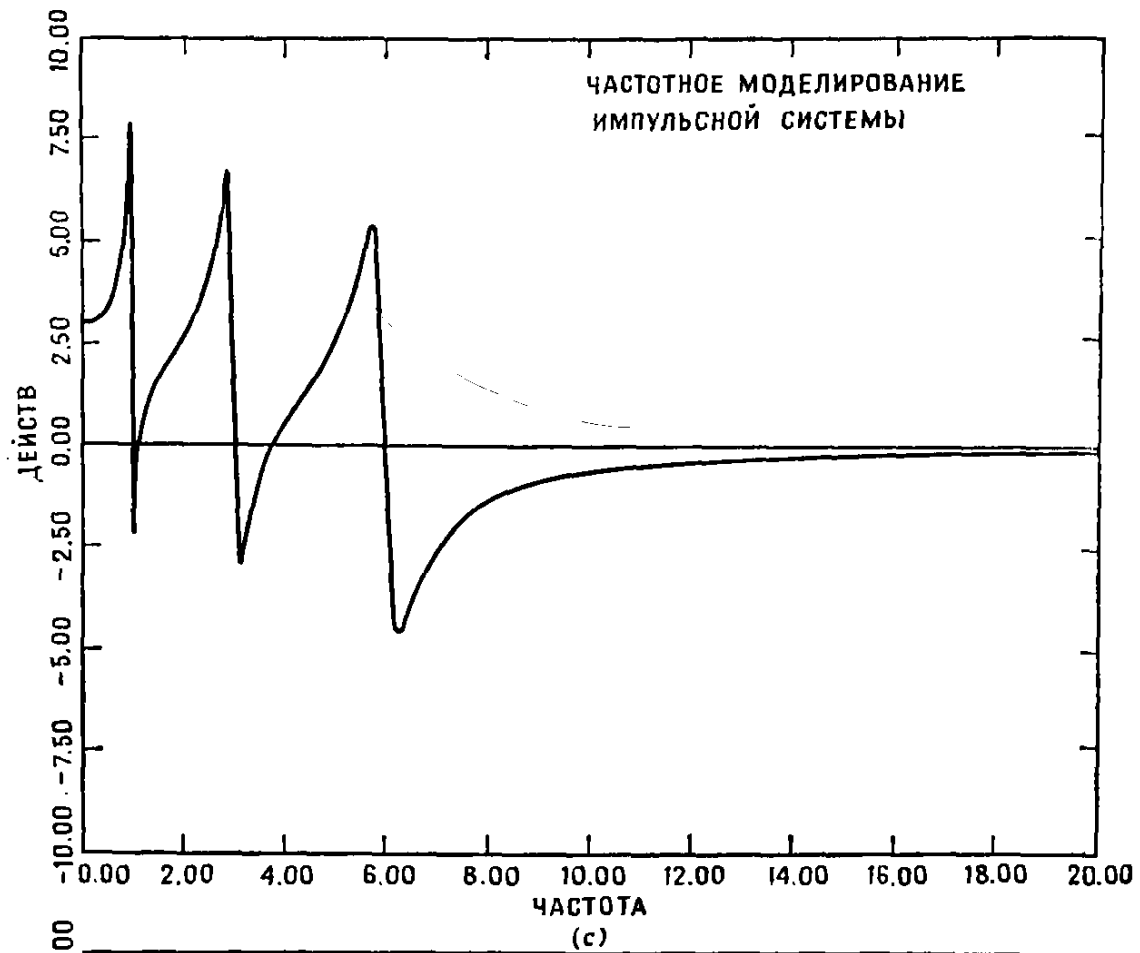


Рис. 6.4. Сопоставление декартовых и полярных координат. *a* — абсолютные



величины; b — фаза; c — действительная часть; d — мнимая часть.

Далее, при вычислении преобразования Фурье действительной последовательности с помощью алгоритма комплексного преобразования Фурье необязательно заполнять нулями весь массив, отведенный под мнимые части. В этом случае действительную последовательность лучше разнести по обоим массивам, отведенным под действительные и мнимые части.

Отметим, что в определениях прямого и обратного дискретных преобразований Фурье в явном виде присутствуют временной и частотный интервалы выборок. Во многих учебниках и статьях приводится несколько измененный вариант определения преобразования, соответствующий случаю нормированного временного интервала T , равного 1. При этом шаг b , с которым чередуются частоты, равен

$$P^{-1} = (NT)^{-1} = N^{-1}.$$

Формулы для преобразований имеют в этом случае вид

$$X(k) = \sum_{i=0}^{N-1} x(i) \exp \left[-j2\pi \frac{ik}{N} \right],$$

$$x(i) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \exp \left[j2\pi \frac{ik}{N} \right].$$

Теперь рассмотрим несколько соотношений, полезных при вычислениях преобразования Фурье. Одно из них — формула преобразования «двойной длины» — служит основой алгоритма быстрого преобразования Фурье с числом данных, равным степени 2.

Спаренное преобразование. Обычные алгоритмы быстрого преобразования Фурье предназначены для последовательностей комплексных данных, хотя на практике временные ряды представляют собой последовательности действительных величин. К счастью, алгоритмы для комплексных последовательностей нетрудно приспособить для вычисления преобразований действительных последовательностей. Допустим, что имеются две временные последовательности $x(i)$ и $y(i)$, $i = 0, 1, \dots, N-1$. Представим эти две последовательности в качестве действительной и мнимой частей комплексной последовательности $z(i)$. Определив преобразование Фурье

$$Z(k) = T \sum_{i=0}^{N-1} z(i) \exp \left[-j2\pi \frac{ik}{N} \right], \quad \text{где } k = 0, 1, \dots, N-1,$$

можно получить из него преобразования $X(k)$ и $Y(k)$ по формулам

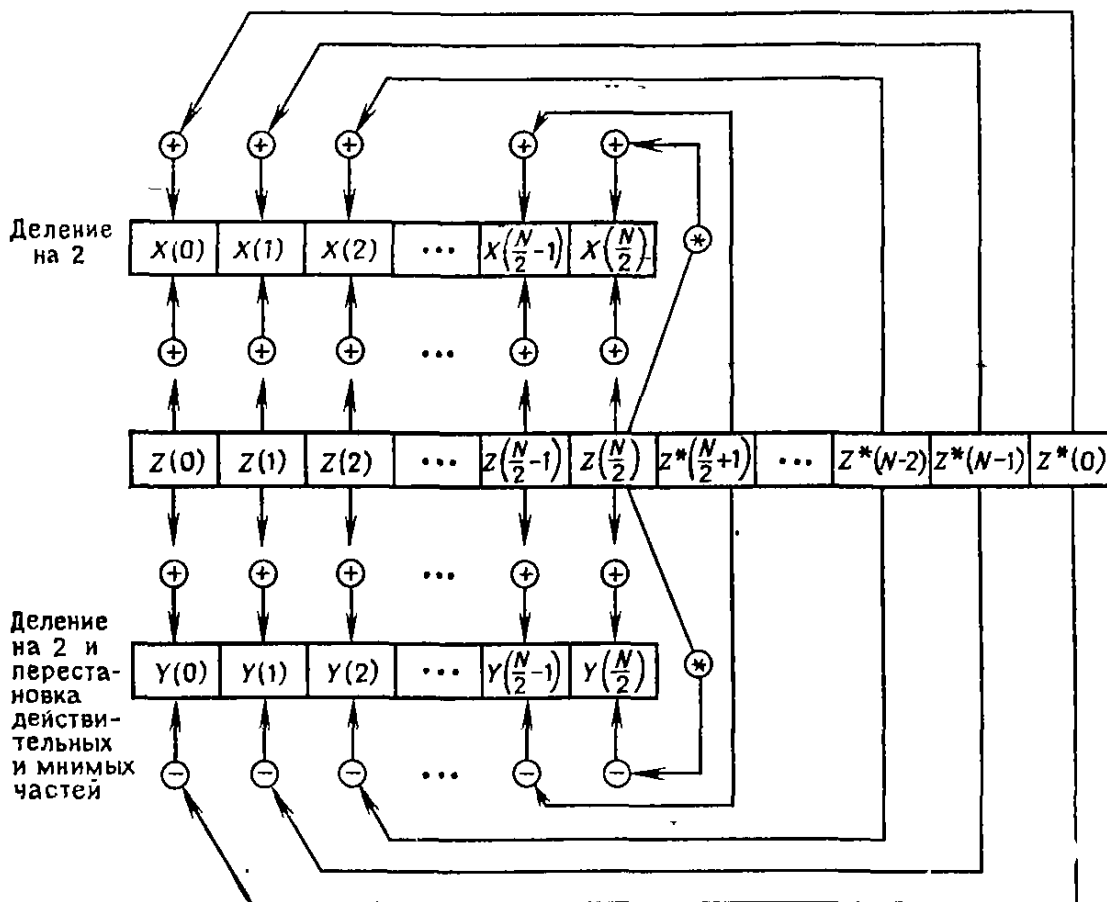
$$X(k) = \frac{Z(k) + Z^*(N-k)}{2}, \quad (6.8)$$

$$Y(k) = \frac{Z(k) - Z^*(N-k)}{2j}, \quad (6.9)$$

где

$$k = \begin{cases} 0, 1, \dots, \frac{N}{2}, & \text{если } N \text{ — четное,} \\ 0, 1, \dots, \frac{N-1}{2}, & \text{если } N \text{ — нечетное.} \end{cases}$$

При этом получается по $N/2$ значений преобразования. Как отмечалось выше, для действительных последовательностей этого достаточно. Схема, иллюстрирующая вычисление преобразований по формулам (6.8) и (6.9), изображена на рис. 6.5.



6.5. Диаграмма вычисления спаренного преобразования расщеплением.

Преобразование двойной длины. Так называется метод получения преобразования Фурье (комплексной) последовательности $z(i)$, где $i = 0, 1, \dots, 2N-1$, по преобразованиям Фурье пары последовательностей вдвое меньшей длины. В этом методе по последовательности $z(i)$ строятся последовательности

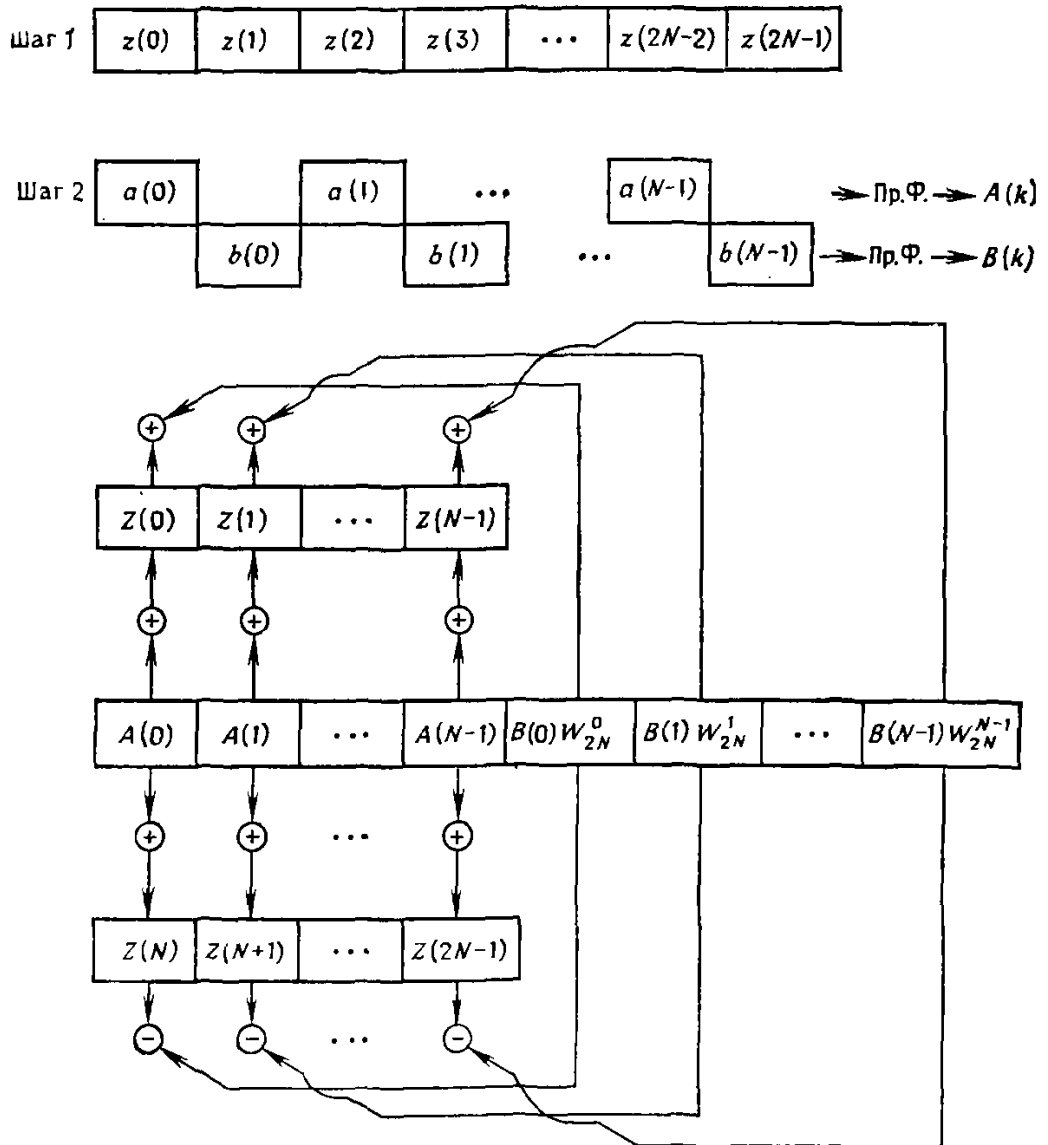
$$a(i) = z(2i), \quad b(i) = z(2i+1),$$

где $i = 0, 1, \dots, N-1$. После этого вычисляются преобразования

$$A(k) = T \sum_{i=0}^{N-1} a(i) \exp\left(-j2\pi \frac{ik}{N}\right)$$

и

$$B(k) = T \sum_{i=0}^{N-1} b(i) \exp\left(-j2\pi \frac{ik}{N}\right),$$

где $k = 0, 1, \dots, N-1$.

6.6. Диаграмма вычисления преобразования двойной длины.

Нужную последовательность значений преобразования Фурье получают по формулам

$$Z(k) = A(k) + B(k) W_{2N}^k, \quad Z(N+k) = A(k) - B(k) W_{2N}^k, \quad (6.10)$$

где $k = 0, 1, \dots, N-1$ и

$$W_{2N} = \exp\left(-j \frac{2\pi}{2N}\right) = \exp\left(-j \frac{\pi}{N}\right).$$

Заметим, что для получения преобразования временной последовательности длины N можно воспользоваться формулами (6.10) для комплексного преобразования длины $N/2$. Схема вычисления по формулам (6.10) изображена на рис. 6.6.

Ряды Фурье. Понятие ряда Фурье (см. Колмогоров и Фомин (1976), с. 406)

$$x(t) = \frac{1}{2} a_0 + \sum_{k=1}^{\infty} (a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t)),$$

где $\omega_0 = 2\pi/P$, тесно примыкает к комплексному преобразованию Фурье. Дискретным вариантом ряда Фурье служит разложение

$$x(i) = \left[\frac{1}{2} a_0 + \sum_{k=1}^{N/2} \left(a_k \cos \frac{2\pi i k}{N} + b_k \sin \frac{2\pi i k}{N} \right) \right],$$

где

$$a_k = 2T \sum_{i=0}^{N-1} x(i) \cos \frac{2\pi i k}{N}, \quad b_k = 2T \sum_{i=0}^{N-1} x(i) \sin \frac{2\pi i k}{N}.$$

Заметим, что

$$a_k = X(k) + X^*(k), \quad b_k = -[X(k) - X^*(k)],$$

т. е. классический ряд Фурье периодической функции можно получить непосредственно из комплексного преобразования Фурье.

Однако следует отметить, что с математической точки зрения ряды и преобразования Фурье имеют существенное различие. Для исследования преобразования Фурье случайной функции требуются более сложные математические построения, чем для рядов Фурье детерминированных функций. Главное различие связано с вопросами существования или сходимости интегралов, входящих в определения. Впрочем, в практическом анализе данных редко приходится сталкиваться с этими тонкостями. Более того, те вычислительные алгоритмы, которые используются в анализе данных, равно хорошо применимы как для вычисления разложения в ряд Фурье детерминированной периодической функции, так и для получения конечного преобразования Фурье цифровой выборки случайного процесса на бесконечном интервале.

Вычисление обратного преобразования. Как уже было отмечено, преобразование Фурье действительной последовательности требует определения только $N/2$ (или $(N-1)/2$) комплексных величин. Это обусловлено комплексной сопряженностью значений преобразования действительной последовательности (см. рис. 6.1). Наоборот, такая комплексная сопряженность симметричных относительно середины последовательности значений

преобразования приводит при обратном преобразовании Фурье к действительной последовательности. Очевидно, что в этом случае вторая половина значений прямого преобразования Фурье не несет никакой дополнительной информации и может быть отброшена. На этом соображении основано вычисление обратного преобразования. Сначала вычисляется последовательность значений $U(k)$ по формуле

$$U(k) = Z(k) + Z^*(M-k) + j[Z(k) - Z^*(M-k)]W_M^{-k/2}, \quad (6.11)$$

где $k = 0, 1, \dots, M-1$, $M = N/2$ или $(N-1)/2$, а $Z(k)$ — последовательность длины N значений преобразования Фурье последовательности $z(i)$. Из общего числа значений для вычислений по формуле (6.11) требуется только $N/2$ или $(N-1)/2$, что и позволяет не учитывать комплексно-сопряженной части преобразования последовательности. Само вычисление обратного преобразования Фурье проводится по формуле

$$\begin{aligned} z(2i) &= \operatorname{Re} u(i), \\ z(2i+1) &= \operatorname{Im} u(i), \quad i = 0, 1, \dots, M-1, \end{aligned} \quad (6.12)$$

в которой действительные и мнимые части чисел $u(i)$ определяются из обратного преобразования Фурье последовательности $U(k)$:

$$u(i) = b \sum_{k=0}^{M-1} U(k) \exp \left[j \frac{2\pi ik}{M} \right].$$

Преимущество способа (6.12) расположения результатов связано с тем, что при работе на Фортране именно таким образом в памяти ЭВМ на сердечниках располагаются соответствующие комплексные числа. Схема вычисления показана на рис. 6.7.

Чтобы получить обратное преобразование Фурье двух функций одновременно, можно, обратив соотношения (6.8) и (6.9), т. е. записав их в виде

$$Z(k) + Z^*(N-k) = 2X(k), \quad Z(k) - Z^*(N-k) = 2jY(k),$$

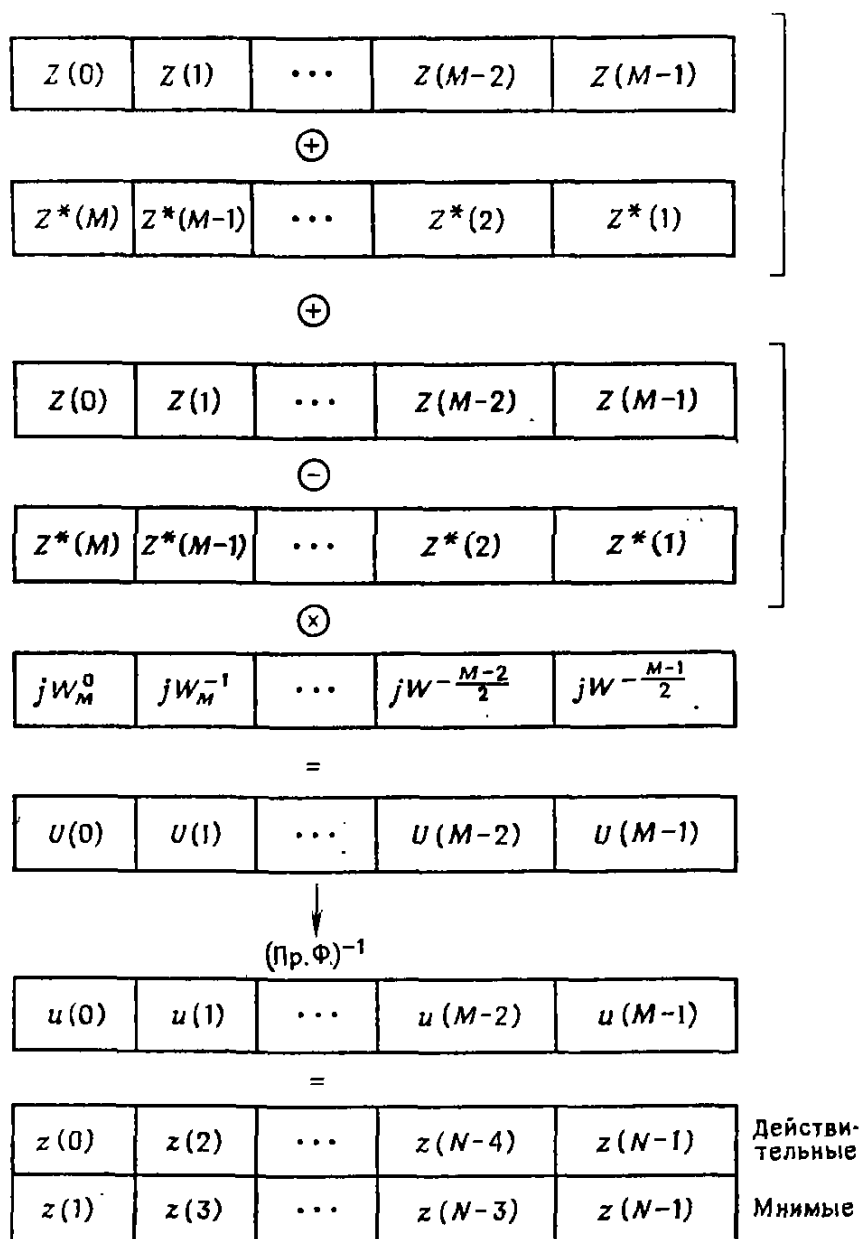
получить значения преобразования $Z(k)$ и $Z^*(N-k)$. Для этого нужно сложить оба равенства и вычесть второе из первого:

$$\begin{aligned} Z(k) &= X(k) + jY(k), \\ Z^*(N-k) &= X(k) - jY(k). \end{aligned}$$

Здесь

$$k = \begin{cases} 0, 1, \dots, \frac{N}{2}, & \text{если } N \text{ — четное,} \\ 0, 1, \dots, \frac{N-1}{2}, & \text{если } N \text{ — нечетное.} \end{cases}$$

Получив таким образом N значений, следует применить обратное преобразование Фурье. В результате действительные последовательности $x(i)$ и $y(i)$ будут получены как действительные и мнимые части последовательности $z(i)$ соответственно.



6.7. Иллюстрация вычисления обратного преобразования по формулам (6.11) и (6.12).

Шаг в области частот как функция длины записи. В следующем параграфе мы остановимся на двух алгоритмах быстрого преобразования Фурье (БПФ) для последовательностей, число членов которых равно степени 2 или, в более общем случае, произведению степеней 2, 3, 4 и 5 (т. е. $N = 2^n 3^m 4^p 5^q$). Если число членов последовательности не совпадает в точности с одним из этих чисел, то прибегают к увеличению длины последова-

тельности за счет добавления нулей в начале и конце. При этом изменяется шаг в области частот. Если исходная длина равна N , то при получении преобразования Фурье шаг в области частот будет равен

$$b = \frac{1}{NT} = \frac{S}{N} = \frac{F}{N/2}.$$

Напомним, что получающиеся значения $X(k)$, где

$$k = \begin{cases} 0, 1, \dots, \frac{N}{2}, & \text{если } N \text{ — четное,} \\ 0, 1, \dots, \frac{N-1}{2}, & \text{если } N \text{ — нечетное.} \end{cases}$$

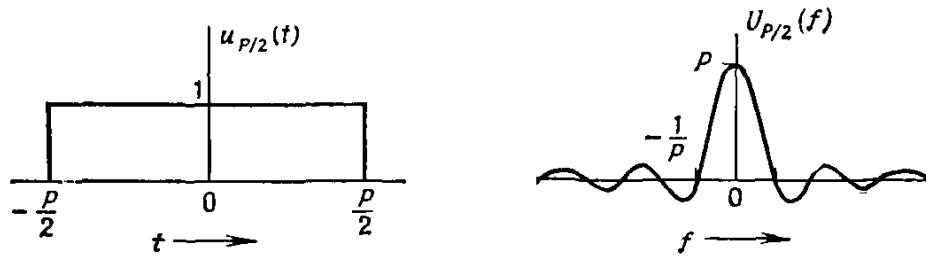
располагаются в частотах kb , т. е. $X(k) = X(bk)$. Первое из значений располагается в частоте $f=0$ и равно

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

Для последнего значения, расположенного в случае четного N в частоте $f=F$, справедлива формула

$$X\left(\frac{N}{2}\right) = \sum_{i=0}^{N/2-1} [x(2i) - x(2i+1)].$$

Добавление нулей. При получении длины последовательности N' за счет добавления ΔN нулей к исходной последовательности



6.8. Прямоугольный импульс и его преобразование.

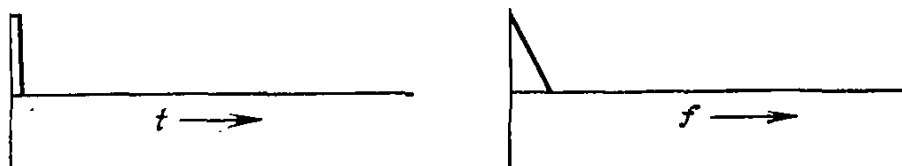
сти длины N шаг в частотной области изменяется следующим образом:

$$b' = \frac{1}{(N + \Delta N)T} = \frac{1}{N'T}.$$

При N , близких к N' , т. е. при малых относительно N' значениях ΔN , выполняется приближенное равенство $b \approx b'$. При малых значениях N' разница между b и b' может оказаться довольно существенной. В качестве иллюстрации такого изменения шага в области частот рассмотрим преобразование Фурье $U_{P/2}(f)$ прямоугольного импульса $u_{P/2}(t)$ с длиной носителя P . Эти функции изображены на рис. 6.8.

Результат применения быстрого преобразования Фурье без добавления нулей при $NT = P$ изображен на рис. 6.9.

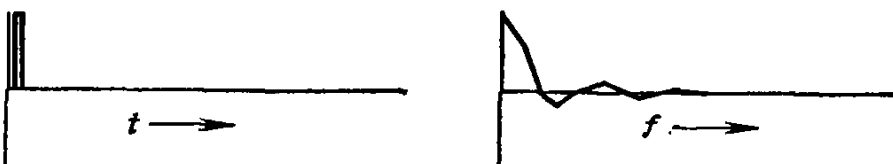
Если до преобразования добавлено N нулей (по $N/2$ нулей перед прямоугольным импульсом и после него), то в результате



6.9. БПФ без добавления нулей.

применения БПФ к последовательности с $2N$ членами произойдет изменение, показанное на рис. 6.10.

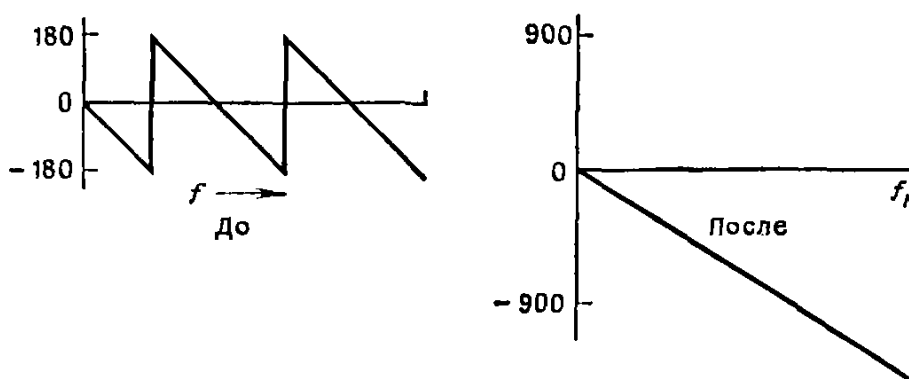
Таким образом, при добавлении нулей происходит интерполяция значений функции $U_{P/2}(f)$ тем точнее, чем больше нулей

6.10. БПФ с добавлением N нулей.

добавляется. Такого же рода интерполяция происходит для всякой функции, к которой добавляются нули.

Позднее будет показано, что, хотя шаг в области частот b изменяется, эффективная разрешающая полоса B_e остается неизменной вне зависимости от числа добавленных нулей.

Длина записи. В анализе цифровых временных рядов число данных, преобразованных в цифровую форму, ограничено воз-



6.11. Продолжение фазы.

можностями ЭВМ. Длина выборки зависит от скорости преобразования в цифровую форму и ограничена величиной

$$P = NT = \frac{N}{S}.$$

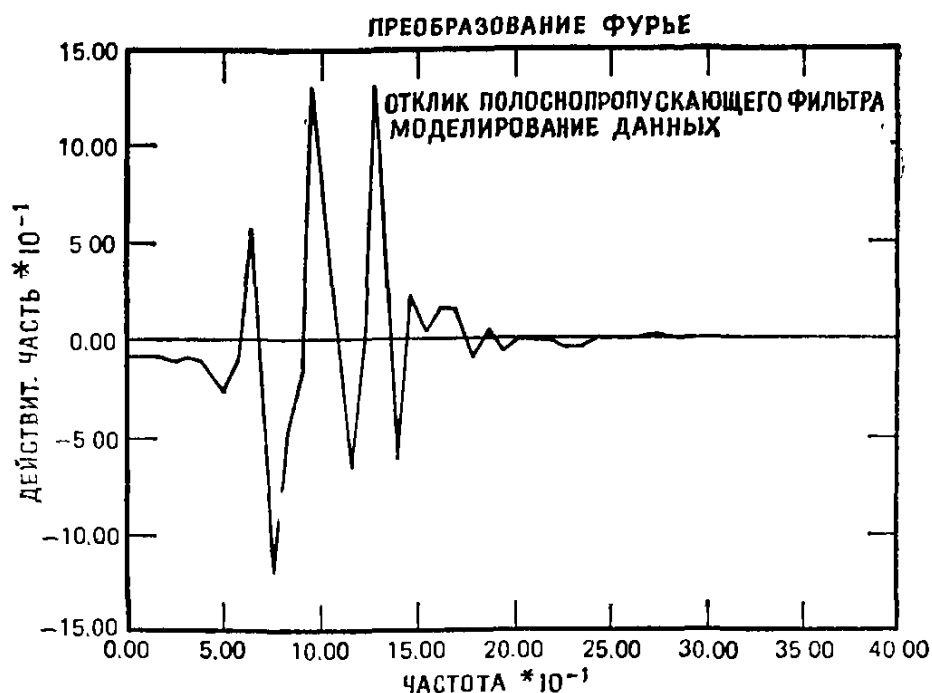
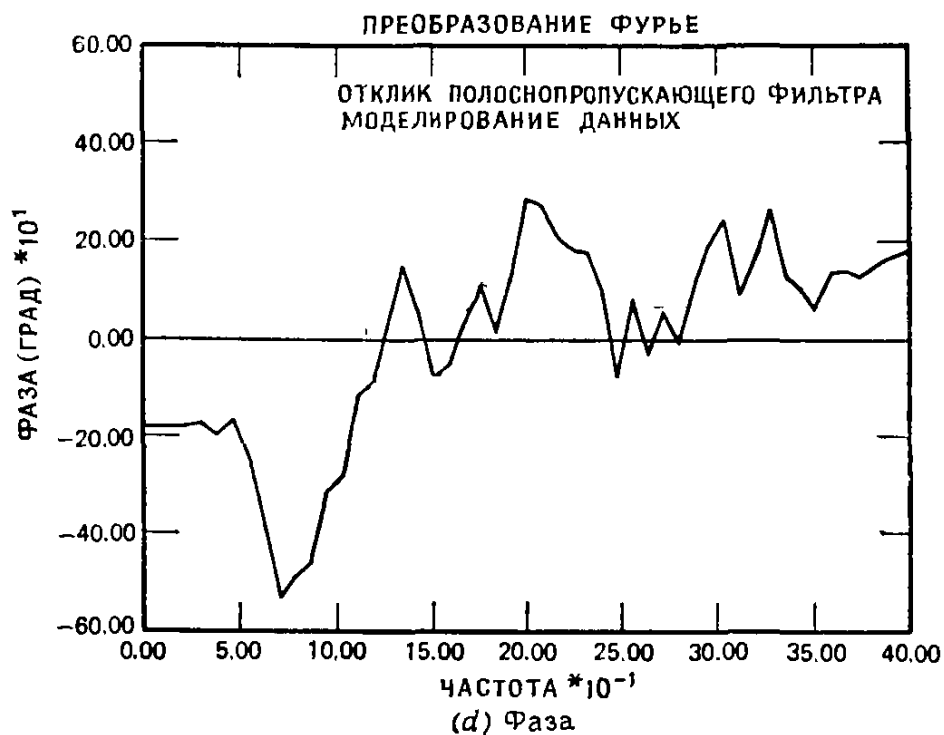
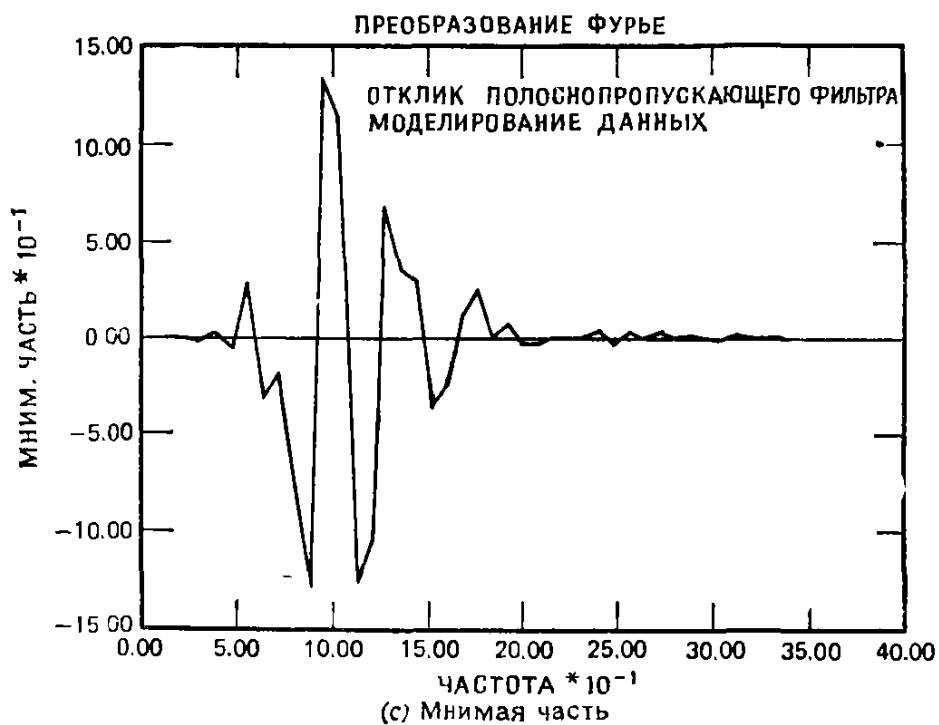


Рис. 6.12. График преобразования Фурье отклика

Для мини-ЭВМ максимальная величина N колеблется обычно в пределах от 1024 (2^{10}) до 8192 (2^{13}). Как правило, в этих случаях выбирается алгоритм быстрого преобразования Фурье для N , равного степени 2. Для ЭВМ средних и широких возможностей величина N колеблется от 1024 до 32 768 (2^{15}), и для ЭВМ широких возможностей чаще всего выбирается алгоритм БПФ с N , равным произведению степеней 2, 3, 4 и 5.



узкополосного фильтра (узкополосного шума).

В обоих случаях нередко требуются длины записей, большие, чем максимально возможная длина выборки $P_s = N_s T$. Такое ограничение обходят с помощью двух методов. Во-первых, для получения точного преобразования Фурье на более длинных интервалах выборки используют алгоритм двойной длины или его варианты. Во-вторых, часто вычисляют преобразование, усредненное по нескольким частям выборки.

В первом случае получается лучшее разрешение для спектра, равное $b = 1/NT$. Во втором случае разрешение ограничено величиной возможной выборки и равно $b_s = 1/N_s T$. В восьмой главе будут рассмотрены вопросы связи шума со степенями свободы и соотношением этих двух методов.

Продолжение фазы. Фигурирующее в формуле (6.6) главное значение арктангенса заключено в границах от -180° до $+180^\circ$.

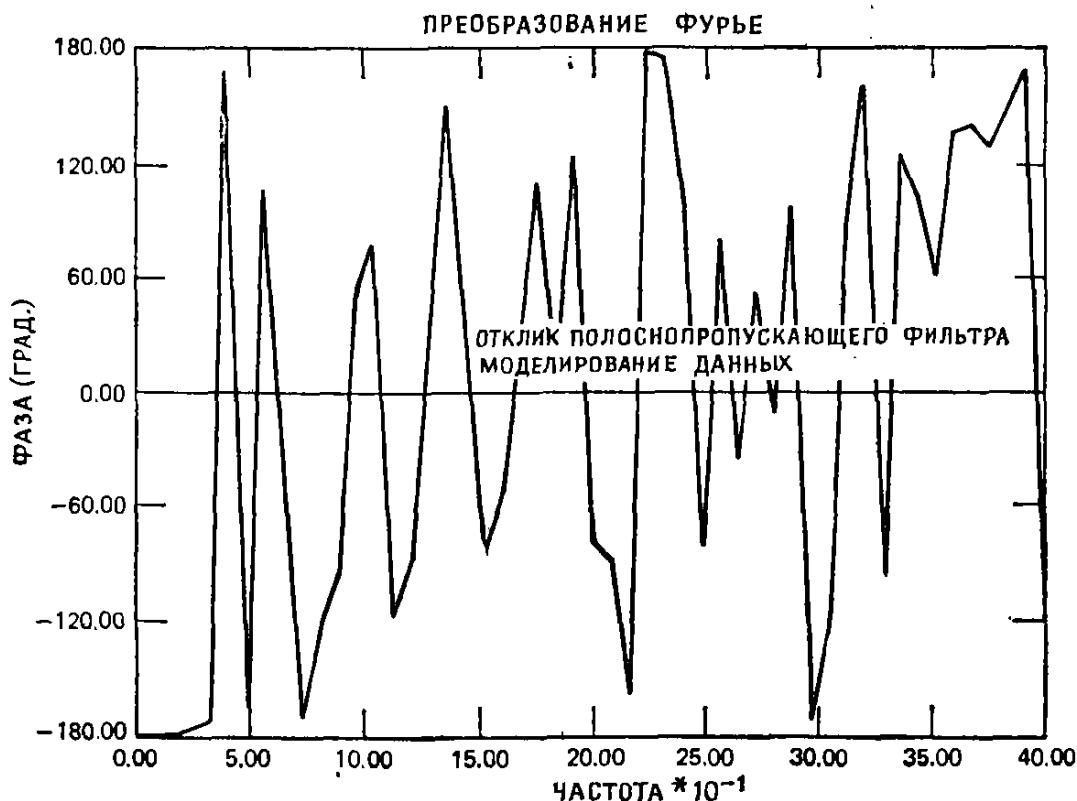


Рис. 6.13. График фазы с рис. 6.12, а без продолжения фазы.

«Продолжением фазы» мы называем процедуру, позволяющую устранять разрывы при переходах через 180° . При этом предполагается, что фазовая функция — «гладкая» функция частоты. Проиллюстрируем на примере большой скачок значения фазы, который может произойти при переходе от $+(180-\epsilon)^\circ$ к $-(180-\epsilon)^\circ$. Величина этого скачка достигает $\Delta\phi = (180-\epsilon)^\circ + (180-\epsilon)^\circ = (360-2\epsilon)^\circ$, хотя сам скачок может быть равным всего 2ϵ , поскольку $-(180-\epsilon)^\circ = (180+\epsilon)^\circ$. Выбор последнего варианта, отвечающего непрерывному случаю, предусмотрен в подпрограмме TTRAN с рис. А.7 (в приложении А).

Эффект продолжения фазы демонстрируется на рис. 6.11. На нем приведены графики линейной фазы (соответствующей случаю простого временного сдвига) до и после продолжения фазы. Второй из них — типичный пример графика, полученного на ЭВМ при продолжении фазы. На рис. 6.12 показаны действительная и мнимая части, коэффициент усиления и продолжен-

ная фаза псевдослучайного белого шума, подвергнутого узкополосной фильтрации. На рис. 6.13 приведен график той же фазы, что и на предыдущем рисунке, но на этот раз без продолжения фазы, т. е. на графике приведены соответствующие величины главного значения арктангенса. На рис. 6.13 хорошо прослеживается нерегулярность графика, полученного без продолжения фазы.

6.2. АЛГОРИТМ БЫСТРОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ

В этом параграфе нами будет дан краткий вывод формул для быстрого преобразования Фурье. При этом мы будем следовать подходу Кули—Тьюки (Кули и Тьюки (1965)), основанному на умножении матриц. Кроме того, будет рассмотрен второй вариант алгоритма БПФ, называемый алгоритмом Сэнда—Тьюки (Джентльмен и Сэнд (1966)). Оба рассматриваемых алгоритма можно назвать каноническими в том смысле, что все остальные методы получения БПФ сводятся в конечном счете к одному из них. По причинам, которые станут яснее позже, вариант Кули—Тьюки принято называть «децимацией по времени», а вариант Сэнда—Тьюки—«децимацией по частотам».

Мы обсудим три программы для ЭВМ. Первые две из них входят в обзор Мэйнэрда (1973) под вторым и первым номерами соответственно. (Всего в этот обзор включено 10 программ.) Первая рассмотренная здесь нами программа (Отнес, Натанс и Эноксон (1969)) имеет очень простую запись и предназначена для решения читателем вычислительных задач этой главы. Эта программа составлена для последовательностей, длина которых равна степени 2. Более общий случай последовательностей, длина которых равна произведению степеней 2, 3, 4 и 5, предусмотрен во второй программе (Синглтон (1969)). Третья программа, предназначенная для вычисления преобразований последовательностей с длиной, превышающей внутренний (на сердечниках) объем памяти ЭВМ, также принадлежит Синглтону (1967).

Матричная запись алгоритма БПФ. Напомним уже встречавшееся раньше обозначение для комплексной экспоненты

$$W_N = \exp \left[-j \frac{2\pi}{N} \right] = \cos \frac{2\pi}{N} - j \sin \frac{2\pi}{N}.$$

Заметим, что, используя это обозначение, экспоненты, входящие в формулу преобразования Фурье, можно переписать в виде

$$W_N^{ik} = \exp \left[-j \frac{2\pi ik}{N} \right] = \cos \left(\frac{2\pi}{N} ik \right) - j \sin \left(\frac{2\pi}{N} ik \right). \quad (6.13)$$

Расположение величин W_N на комплексной плоскости показано на рис. 6.14. Эти величины в теории функций комплексного

переменного называют *корнями N -й степени из единицы* (в данном случае $N = 16$). Из формулы (6.13) можно выводить некоторые довольно сложные тригонометрические соотношения; два из них приведены в упр. 6.11.

Матричный подход для объяснения БПФ был предложен МакКованом (1966). Дальнейшее развитие этот подход получил в работе Тейлхеймера (1969). Мы в качестве пояснения матричного

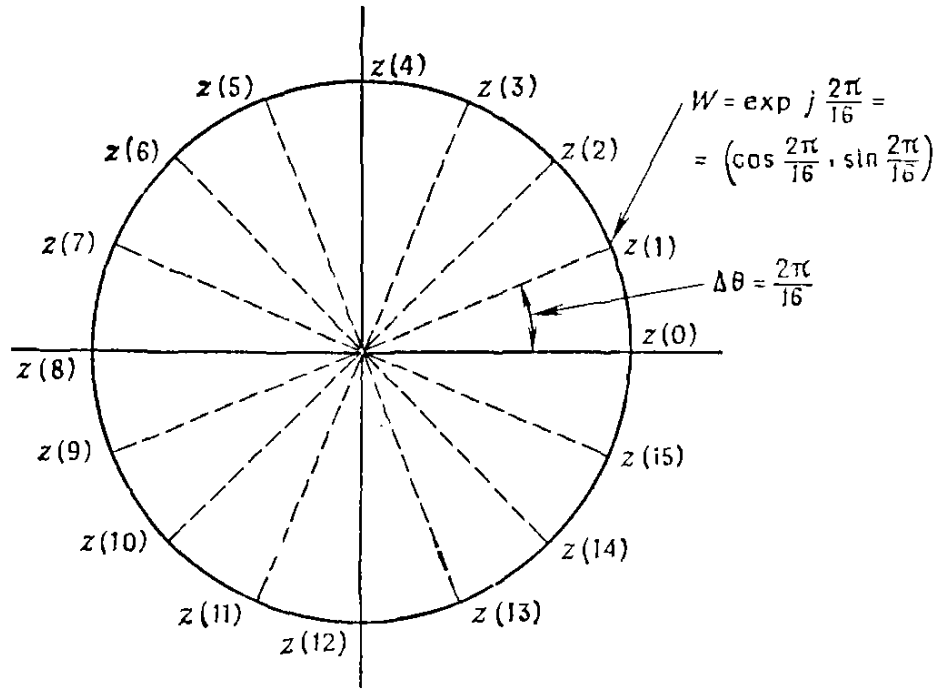


Рис. 6.14. Точки данных, расположенные на единичной окружности комплексной области.

подхода рассмотрим пример, в котором маневр самолета раскладывается на составляющие компоненты — отклонение от курса, наклон относительно поперечной оси и бортовой наклон. В матричном виде поворот плоскости координат на угол θ можно записать следующим образом:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

В трехмерном пространстве всякий поворот можно свести к поворотам в трех координатных двумерных плоскостях (см. рис. 6.15) — повороту в координатной плоскости (y, z) (бортовой наклон), повороту в координатной плоскости (x, z) (наклон относительно поперечной оси) и повороту в плоскости (x, y) (отклонение от курса).

Таким образом, полный поворот описывается произведением трех матриц, соответствующих двумерным преобразованиям:

$$T(\theta_1, \theta_2, \theta_3) = \begin{pmatrix} \cos \theta_3 & -\sin \theta_3 & 0 \\ \sin \theta_3 & \cos \theta_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta_2 & 0 & -\sin \theta_2 \\ 0 & 1 & 0 \\ \sin \theta_2 & 0 & \cos \theta_2 \end{pmatrix} \times \\ \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_1 & -\sin \theta_1 \\ 0 & \sin \theta_1 & \cos \theta_1 \end{pmatrix}.$$

отклонение от курса наклон относительно поперечной оси

бортовой наклон

Поэтому вычисления, необходимые для получения преобразования координат, можно свести к последовательному вычислению

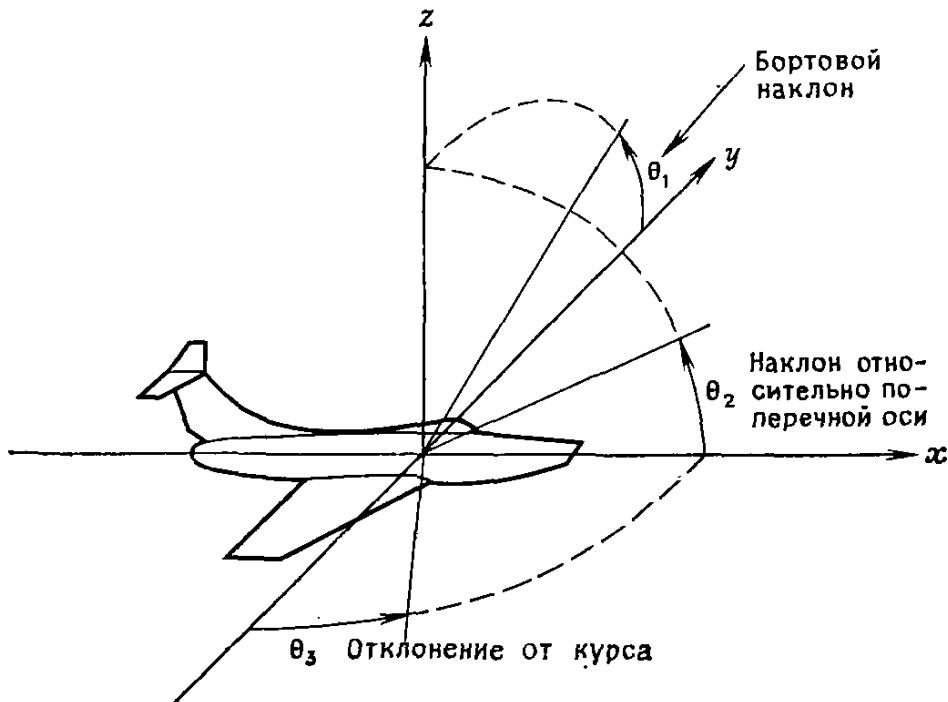


Рис. 6.15. Иллюстрация компонент маневра — отклонения от курса, наклона относительно поперечной оси и бортового наклона.

трех двумерных преобразований. Если такого разбиения вычисления на этапы не проводить, то потребуется $3^2 = 9$ операций (для умножения матрицы 3×3 на трехмерный вектор). При разбиении число операций равно $3 \times 2^2 = 12$. В четырехмерном случае потребуется $4^2 = 16$ и $4 \times 2^2 = 16$ операций соответственно (здесь вращение можно свести к четырем вращением двумерных плоскостей). Уже в пятимерном случае, где прямое вычисление требует $5^2 = 25$, а с разбиением — только $5 \times 2^2 = 20$ операций, видна возможность сокращения необходимого для выполнения операций

времени. Это уменьшение становится все более заметным с ростом размерности пространства и показывает преимущество второго метода вычисления. По сути дела, то же самое происходит и при вычислении при помощи алгоритмов БПФ. Дискретное конечное преобразование Фурье N данных можно представлять как вращение в N -мерном пространстве. (В линейной алгебре такие вращения называют ортогональными преобразованиями.) В матрице, соответствующей этому преобразованию, показатель величины W определяется произведением номеров строки и столбца:

$$\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} W^{0 \cdot 0} & W^{0 \cdot 1} & W^{0 \cdot 2} & \dots & W^{0 \cdot (N-1)} \\ W^{1 \cdot 0} & W^{1 \cdot 1} & W^{1 \cdot 2} & \dots & W^{1 \cdot (N-1)} \\ W^{2 \cdot 0} & W^{2 \cdot 1} & W^{2 \cdot 2} & \dots & W^{2 \cdot (N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ W^{(N-1) \cdot 0} & W^{(N-1) \cdot 1} & W^{(N-1) \cdot 2} & \dots & W^{(N-1) \cdot (N-1)} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{bmatrix},$$

или, в более короткой матричной записи, $X = Wx$.

Приведем явный вид матрицы в частном случае $N = 2^p = 8$, $p = 3$:

$$W = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & W^1 & W^2 & W^3 & W^4 & W^5 & W^6 & W^7 \\ 1 & W^2 & W^4 & W^6 & W^8 & W^{10} & W^{12} & W^{14} \\ 1 & W^3 & W^6 & W^9 & W^{12} & W^{15} & W^{18} & W^{21} \\ 1 & W^4 & W^8 & W^{12} & W^{16} & W^{20} & W^{24} & W^{28} \\ 1 & W^5 & W^{10} & W^{15} & W^{20} & W^{25} & W^{30} & W^{35} \\ 1 & W^6 & W^{12} & W^{18} & W^{24} & W^{30} & W^{36} & W^{42} \\ 1 & W^7 & W^{14} & W^{21} & W^{28} & W^{35} & W^{42} & W^{49} \end{bmatrix}.$$

Эту матрицу можно разложить в произведение $(p+1=4)$ матриц:

$$W = W_3 W_2 W_1 W_0$$

или, в более общем случае, в произведение

$$W = W_p W_{p-1} \dots W_1 W_0. \quad (6.14)$$

Последняя матрица W_p — матрица перестановки, переупорядочивающая полученный на последнем шаге вектор перестановкой битов индекса каждого его элемента в обратном порядке.

Правило получения представления W вида (6.14) сформулировано Тейлхеймером (1969). Для алгоритма Кули — Тьюки, в кото-

ром $N = 2^p$, представление вида (6.14) получить довольно просто. Такое представление в частном случае $N = 2^3$ было получено Мак-Кованом (1966):

$$W_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & W^0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & W^0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & W^0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & W^0 \\ 1 & 0 & 0 & 0 & W^4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & W^4 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & W^4 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & W^4 \end{bmatrix},$$

$$W_1 = \begin{bmatrix} 1 & 0 & W^0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & W^0 & 0 & 0 & 0 & 0 \\ 1 & 0 & W^4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & W^4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & W^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & W^2 \\ 0 & 0 & 0 & 0 & 1 & 0 & W^6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & W^6 \end{bmatrix},$$

$$W_2 = \begin{bmatrix} 1 & W^0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & W^4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & W^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & W^6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & W^1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & W^5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & W^3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & W^7 \end{bmatrix}.$$

Обозначение W^0 для единицы оставлено только из соображений симметрии. Учет при программировании особенностей вида этих матриц (нет необходимости производить умножения на нули и единицы) позволяет сократить время работы программы.

Вариант Кули — Тьюки. Обозначим символом A_l вектор данных, полученный после l -го умножения матриц. Вариант Кули — Тьюки алгоритма преобразования Фурье основан на вычислении

этого вектора данных по формулам

$$\begin{aligned}
 & A_l(k_0, k_1, \dots, k_{l-2}, \overset{k_{l-1}}{\downarrow} 0, i_{p-l-1}, \dots, i_0) = \\
 & = A_{l-1}(k_0, k_1, \dots, k_{l-2}, \overset{-i_{p-l}}{\downarrow} 0, i_{p-l-1}, \dots, i_0) + \\
 & + A_{l-1}(k_0, k_1, \dots, k_{l-2}, 1, i_{p-l-1}, \dots, i_0) \times \\
 & \cdot X W^{(k_{l-2}2^{l-2} + \dots + k_0) \downarrow 1 \cdot 2^{p-l}}, \quad l=1, 2, \dots, p,
 \end{aligned} \tag{6.15a}$$

```

SUBROUTINE FFTRAN (SIGN,T,X,NPOW)                                90010
C                                                                 90020
C   МЕТОД КУЛИ-ТЮКИ ВЫЧИСЛЕНИЯ ПРЕОБРАЗОВАНИЯ ФУРЬЕ          90030
C   ВКЛЮЧАЕТ ВЫЧИСЛЕНИЯ СИНУСОВ, КОСИНУСОВ И ПЕРЕУПОРЯДОЧЕНИЕ 90040
C   ДАННЫХ ЧЕРЕЗ ОБРАТНУЮ ПЕРЕСТАНОВКУ БИТОВ ИНДЕКСОВ        90050
C                                                                 90060
C   SIGN = УКАЗАТЕЛЬ НАПРАВЛЕНИЯ ПРЕОБРАЗОВАНИЯ ФУРЬЕ       90070
C   = -1. ДЛЯ ПРЯМОГО ПРЕОБРАЗОВАНИЯ                          90080
C   = 1. ДЛЯ ОБРАТНОГО ПРЕОБРАЗОВАНИЯ                         90090
C   T     = ШАГ ПО ВРЕМЕНИ                                     90100
C   X     = МАССИВ ДЛЯ ПРЕОБРАЗОВАНИЯ ФУРЬЕ                  90110
C   NPOW  = СТЕПЕНЬ 2 (ОБЪЕМ БЛОКА = 2 ** NPOW)              90120
C                                                                 90130
C   DIMENSION X(1), CS(2), MSK(13)                             90140
C   COMPLEX X,CXCS,HOLD,XA                                     90150
C   EQUIVALENCE (CXCS,CS)                                     90160
C   NMAX=2**NPOW                                             90170
C   ZZ=6.283185306*SIGN/FLOAT(NMAX)                          90180
C   DELTA=T                                                  90190
C   IF (SIGN) 10,10,5                                         90200
5   DELTA=1./(T*FLOAT(NMAX))                                  90210
10  MSK(1)=NMAX/2                                             90220
DO 15 I=2,NPOW                                              90230
15  MSK(I)=MSK(I-1)/2                                         90240
    NN=NMAX                                                    90250
    MM=2                                                         90260
C                                                                 90270
C   ВНЕШНИЙ ЦИКЛ ДЛЯ СЛОЕВ NPOW                               90280
C                                                                 90290
C   DO 45 LAYER=1,NPOW                                         90300
C   NN=NN/2                                                     90310
C   NW=0                                                         90320
C   DO 40 I=1,MM,2                                             90330
C   II=NN*I                                                     90340
C                                                                 90350
C   CXCS = CEXP(2*PI*NW*SIGN/NMAX)                             90360
C                                                                 90370
C   W=FLOAT(NW)*ZZ                                             90380
C   CS(1)=COS(W)                                               90390
C   CS(2)=SIN(W)                                               90400
C                                                                 90410
C   ВЫЧИСЛЕНИЕ ЭЛЕМЕНТОВ ДЛЯ ОБЕИХ ПОЛОВИН КАЖДОГО БЛОКА  90420
C                                                                 90430
C   DO 20 J=1,NN                                               90440

```

Рис. 6.16а. Программа быстрого преобразования Фурье, часть I. Программа написана Л. Д. Эноксоном, Ховардом Натансом и Р. К. Отнесом.

$$\begin{aligned}
 A_l(k_0, k_1, \dots, k_{l-2}, \overset{k_{l-1}}{\downarrow} 1, i_{p-l-1}, \dots, i_0) = \\
 = A_{l-1}(k_0, k_1, \dots, k_{l-2}, \overset{i_{p-l}}{\downarrow} 0, i_{p-l-1}, \dots, i_0) - \\
 - A_{l-1}(k_0, k_1, \dots, k_{l-2}, 1, i_{p-l-1}, \dots, i_0) \times \\
 \times W^{(k_{l-2}2^{l-2} + \dots + k_0) \overset{i_{p-l}}{\downarrow} 1 \cdot 2^{p-l}}, \quad l = 1, 2, \dots, p.
 \end{aligned}
 \tag{6.15b}$$

```

II=II+1                                90450
IJ=II-NN                                90460
XA=CXCS*X(II)                           90470
X(II)=X(IJ)-XA                           90480
X(IJ)=X(IJ)+XA                           90490
20 C                                     90500
C                                         90510
C                                         90520
C      ВЫЧИСЛЕНИЕ ОБРАТНЫХ АДРЕСОВ      90530
C                                         90540
DO 25 LOC=2,NPROW                         90550
LL=NW-MSK(LOC)                            90560
IF (LL) 30,35,25                          90570
25  NW=LL                                  90580
30  NW=MSK(LOC)+NW                        90590
    GO TO 40                               90600
35  NW=MSK(LOC+1)                         90610
40  CONTINUE                              90620
C                                         90630
C      ПРОВОДИТСЯ ЗАКЛЮЧИТЕЛЬНОЕ ПЕРЕУПОРЯДОЧЕНИЕ
C      А ТАКЖЕ УМНОЖЕНИЕ НА DELTA        90640
C                                         90650
45  MM=MM*2                               90660
    NW=0                                   90670
    DO 80 I=1,NMAX                         90680
    NW1=NW+1                               90690
    HOLD=X(NW1)                            90700
    IF (NW1-I) 60,55,50                   90710
50  X(NW1)=X(I)*DELTA                     90720
55  X(I)=HOLD*DELTA                       90730
C                                         90740
C                                         90750
C                                         90760
C      ВЫЧИСЛЕНИЕ ОБРАТНОГО АДРЕСА     90770
C                                         90780
60  DO 65 LOC=1,NPROW                     90790
    LL=NW-MSK(LOC)                         90800
    IF (LL) 70,75,65                       90810
65  NW=LL                                  90820
70  NW=MSK(LOC)+NW                        90830
    GO TO 80                               90840
75  NW=MSK(LOC+1)                         90850
80  CONTINUE                              90860
    RETURN                                  90870
    END                                    90880

```

Рис. 6.16b. Программа быстрого преобразования Фурье, часть II.

Значительно проще эти соотношения выглядят в программе на Фортране (рис. 6.16). Им соответствуют предложения 48 и 49 во внутреннем цикле, имеющие вид

$$\begin{aligned} X(IJ) &= X(IJ) + X(II) * CXCS \\ X(II) &= X(IJ) - X(II) * CXCS \end{aligned} \quad (6.16)$$

где

$$\begin{aligned} CXCS &= W^{(k_{l-2}2^{l-2} + \dots + k_0)2^{p-1}} \\ IJ &= (k_0, k_1, \dots, k_{l-2}, 0, i_{p-l-1}, \dots, i_0) \\ II &= (k_0, k_1, \dots, k_{l-2}, 1, i_{p-l-1}, \dots, i_0) \end{aligned}$$

Отметим, что (6.16) — формулы для двухточечного преобразования Фурье (см. упр. 6.5), в которых дополнительно производится умножение на экспоненту. Читателю следует обратить внимание на следующие особенности алгоритма.

1. Внутренний цикл, в котором вычисляется двухточечное преобразование Фурье с дополнительным умножением на экспоненту, соответствует формулам (6.10) для преобразования двойной длины.

2. Исходные данные — массив A_0 и временной индекс — раскладываются на двоичные компоненты. На каждом шаге рекурсии вычисляется преобразование для одной компоненты.

3. Рекурсия проводится, начиная со старшего бита временного индекса (i_{p-1}), который преобразуется в младший бит частотного индекса (k_0).

4. В результате перестановки компонент индекса в обратном порядке возникает следующее соотношение для элементов массива X :

$$A_l(k_{p-1}, k_{p-2}, \dots, k_0) = X(k_0, k_1, \dots, k_{p-1}).$$

Следовательно, для того чтобы частоты в распечатке или графике имели естественный порядок, следует переупорядочить элементы полученного массива в соответствии с обратной перестановкой битов индексов.

В программе БПФ с рис. 6.16 для обратной перестановки битов индексов предусмотрена такая процедура. Все биты предыдущего индекса вплоть до первого, содержащего 0, полагаются равными нулю. Заменяя в этом первом бите с нулем нуль на единицу, получают новый индекс. Например, при $p=3$ начинают с $000_2 = 0_{10}$ и получают $100_2 = 4_{10}$ (т. е. индекс, порядок битов в котором противоположен порядку битов индекса $001_2 = 1_{10}$). Следующий индекс получается равным $010_2 = 2_{10}$ и т. д. (Нам

сообщали, что для этой цели существуют машинные команды, правда, не на всех ЭВМ.) Как в варианте Кули—Тьюки, так и в варианте Сэнда—Тьюки предусмотрена обратная перестановка битов и переупорядочение (программно) конечных результатов в естественную очередность. Отметим, что некоторыми уловками, например предварительной перестановкой исходных данных, можно достичь того же результата.

Процедура перестановки встречается в программе два раза: в строках 55—61 и в строках 79—85 (рис. 6.16). Оператор IF в строке 57 предназначен для проверки очередного бита. Первый раз процедура используется для вычисления аргументов синусов и косинусов, а во второй — для окончательного упорядочения данных.

Сделаем еще два замечания относительно программы БПФ с рис. 6.16. Вычисление синусов и косинусов комплексных экспонент проводится через библиотеку подпрограмм Фортрана (строки 39 и 40). Далее в вычислениях предусмотрены шаги T и b , соответствующие dt и df в интеграле Фурье. Эти величины из формул (6.1) и (6.7) входят в произведения на строках 73 и 74.

Этим алгоритмом следует пользоваться во всех упражнениях с четными номерами в конце главы.

В другом варианте алгоритма БПФ — Сэнда и Тьюки — рекурсивные соотношения выглядят следующим образом (см., например, книгу Отнеса и Эноксона (1972), в которой приведены соответствующие выкладки):

$$\begin{aligned}
 & \hat{A}_l(k_0, k_1, \dots, k_{l-2}, \overset{k_{l-1}}{\downarrow} 0, i_{p-l-1}, \dots, i_0) = \\
 & = [\hat{A}_{l-1}(k_0, k_1, \dots, k_{l-2}, \overset{i_{p-l}}{\downarrow} 0, i_{p-l-1}, \dots, i_0) + \\
 & + \hat{A}_{l-1}(k_0, k_1, \dots, k_{l-2}, 1, i_{p-l-1}, \dots, i_0)] \times \\
 & \times W(\overset{k_{l-1}}{\downarrow} (i_{p-l-1}2^{p-l-1} + \dots + i_0) 2^{l-1}), \\
 & \hat{A}_l(k_0, k_1, \dots, k_{l-2}, \overset{k_{l-1}}{\downarrow} 1, i_{p-l-1}, \dots, i_0) = \\
 & = [\hat{A}_{l-1}(k_0, k_1, \dots, k_{l-2}, \overset{i_{p-l}}{\downarrow} 0, i_{p-l-1}, \dots, i_0) -
 \end{aligned}
 \tag{6.17a}$$

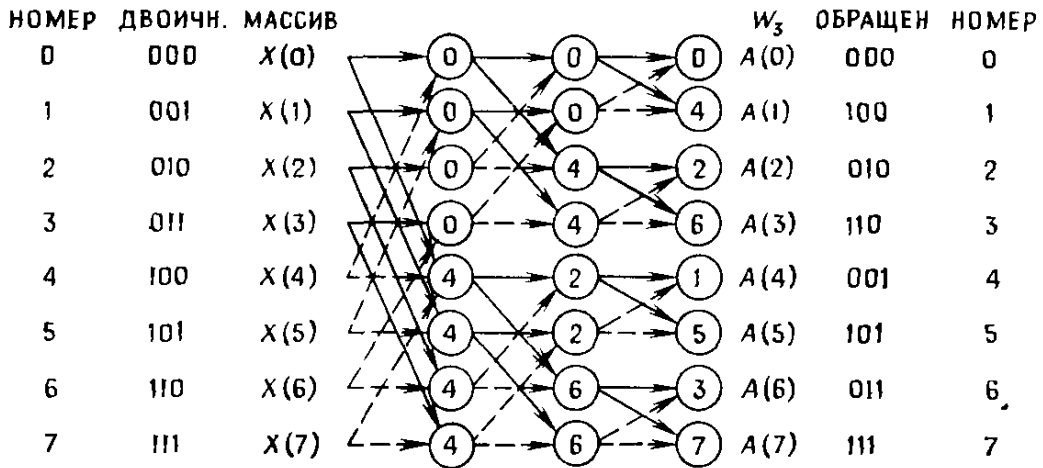


Рис. 6.17. Диаграмма алгоритма Кули — Тьюки, $N=8$, $p=3$ (Мак-Кован (1966)).

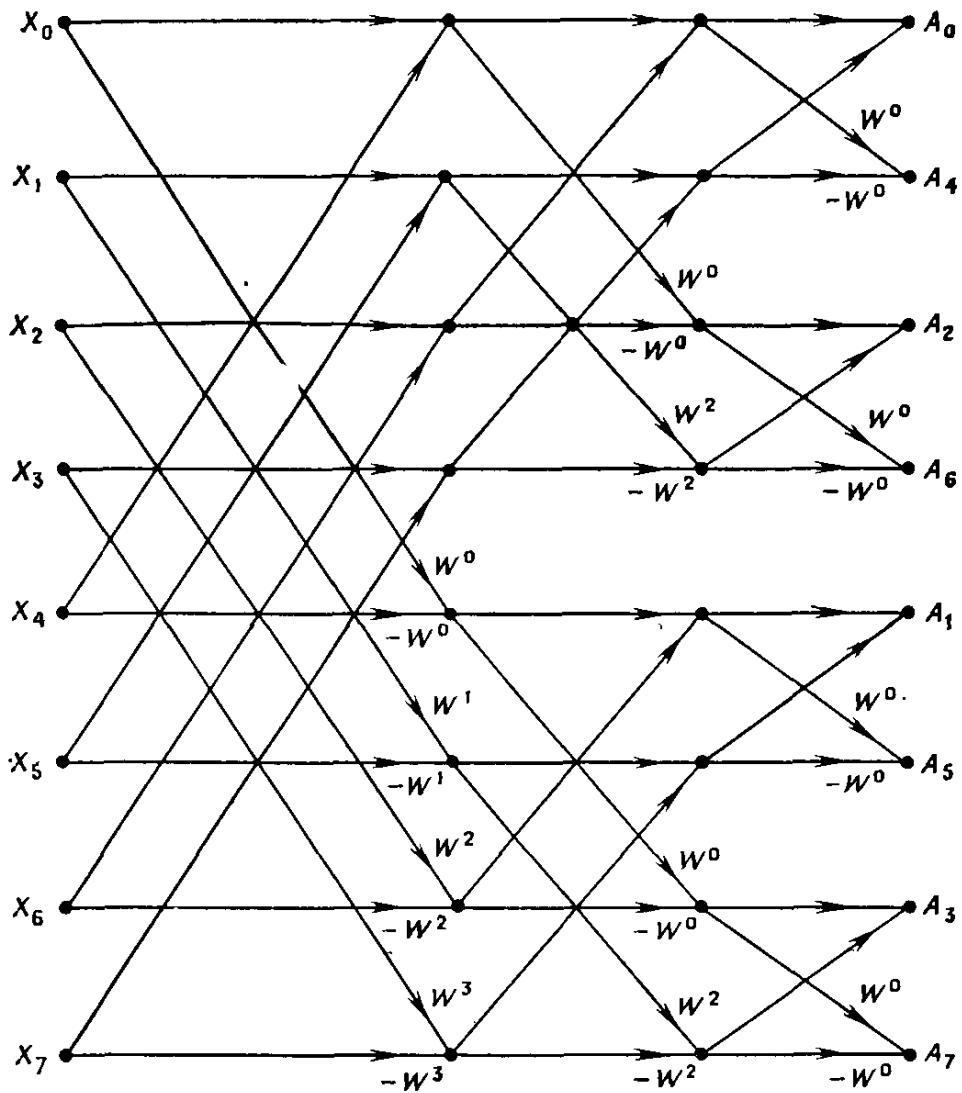


Рис. 6.18. Граф потока сигналов Сэнда — Тьюки (из статьи What Is the Fast Fourier Transform, IEEE Transactions on Audio Electroacoustics, Vol. AU-15, No. 2, pp. 45—55, 1969).

по частотам». Как видно из рисунка, на последней стадии алгоритма величины комплексных экспонент также разделяются $N/2$ позициями.

Пояснения к алгоритму. Предложенная на рис. 6.16 Фортран-программа — наиболее простой вариант для алгоритма К — Т, предназначенный для обработки исходных данных, расположенных в их естественном порядке, и для переупорядочения получающихся результатов в порядке, соответствующем исходному. Эта программа позволяет получать и прямое, и обратное преобразование Фурье последовательностей длины $N = 2^p$ с комплексными членами. Помимо перестановки получающихся результатов и выбора направления преобразования в программе предусмотрено задание величин T и p . Максимальное значение, которое может принимать p , выбрано в этой программе равным 13.

Массивы исходных данных определяют предложения

```

15      DIMENSION X(1), CS(2), MSK(13)
16      COMPLEX X, CXCS, HOLD, XA
17      EQUIVALENCE (CXCS, CS)

```

В предложении

```

18      NMAX=2**NPOW

```

вычисляется значение $N = 2^p$, где $p \leq 13$. Затем вычисляется величина $ZZ = 2\pi/N$, необходимая для определения аргументов комплексных экспонент:

```

19      ZZ=6.283185306*SIGN/FLOAT(NMAX)

```

Для прямого преобразования Фурье выбирается значение $DELTA = T$, а для обратного — $DELTA = b = 1/NT$:

```

20      DELTA=T
21      IF (SIGN) 10,10.5
22      5    DELTA=1./(T*FLOAT(NMAX))

```

После этого формируется последовательность двоичных векторов, используемых в переупорядочении:

```

23      10      MSK(1)=NMAX/2
24              DO 15 I=2,NPOW
25      15      MSK(I)=MSK(I-1)/2

```

В результате получается массив

$$\begin{aligned}
 \text{MSK}(1) &= 10\dots 0 \\
 \text{MSK}(2) &= 010\dots 0 \\
 \text{MSK}(p) &= \underbrace{000\dots 01}_{p \text{ битов}}
 \end{aligned}$$

Внутренний цикл вычисления БПФ составляют команды

```

44          DO 20 J=1,NN
45          II=II+1
46          IJ=II-NN
47          XA=CXCS*X(II)
48          X(II)=X(IJ)-XA
49      20    X(IJ)=X(IJ)+XA

```

Как было указано выше, CXCS есть комплексная экспонента

$$W^{(k_{l-2} 2^{l-2} + \dots + k_0) 2^{p-l}}$$

В строках 48 и 49 производится вычисление двухточечного преобразования Фурье, в которое дополнительно введен множитель поворота, т. е. экспонента. Отметим, что адреса данных разделены расстоянием, равным множителю NN (биты i_{p-l} или k_{l-1}); этот множитель определяет, кроме того, число проходов внутреннего цикла. Предложение $II = II + 1$, включенное в этот цикл, проводит значения индекса Фортран-программы от 1 до N , что отвечает пробеганию индексов нашего уравнения от 0 до $N - 1$.

В последовательности предложений, которая сейчас будет приведена, предусмотрены установка начального значения аргумента комплексной экспоненты ($NW = 0$), вычисление комплексной экспоненты, определение необходимого для аргумента комплексной экспоненты адреса с обращенным порядком битов и проверка для выхода из цикла по большей значащей части индекса данных. Поэтому индексация I отвечает первой части (6.18),

в то время как индекс J внутреннего цикла — второй части (6.18):

```

31      NN=NN/2
32      NW=0
33      DO 40 I=1,MM,2
34      II=NN*I
35      C
36      C          CXCS=CEXP(2*PI*NW*SIGN/NMAX)
37      C
38      W=FLOAT(NW)*ZZ
39      CS(1)=COS(W)
40      CS(2)=SIN(W)
:      :
50      C          :
51      C          .
52      C
53      C      ВЫЧИСЛЕНИЕ ОБРАТНОГО АДРЕСА
54      C
55      DO 25 LOC=2,NPOW
56      LL=NW-MSK(LOC)
57      IF (LL) 30,35,25
58      25      NW=LL
59      30      NW=MSK(LOC)+NW
60      GO TO 40
61      35      NW=MSK(LOC+1)
62      40      CONTINUE

```

Как следует из формулы (6.15), аргумент для вычисления синусов и косинусов получается умножением на 2^{p-l} , т. е. смещением расположения битов на $p-l$ позиций влево. Мы избегаем непосредственного умножения, получая нужный индекс из большей значащей части битов, и уменьшаем тем самым объем работы.

Новые адреса с обратным порядком следования битов NW определяются по правилу, которое уже выше упоминалось. Напомним, что по этому правилу происходит проверка старших битов предыдущего адреса с обратным порядком битов NW . Для этой цели служит массив MSK .

Заметим, что переменная NN , определяющая число прохождений внутреннего цикла, каждый раз уменьшается вдвое (см. предложение 31). Следовательно, в первый раз цикл проходит $N/2$ раз. Последний раз этот внутренний цикл проходится только один раз.

Теперь обратимся к последовательности предложений, в которой задаются пределы изменения индексов: NN для внутреннего цикла, MM для среднего цикла. Индекс внешнего цикла LL, который соответствует нижнему индексу l в формуле (6.15), пробегает значения от 1 до p :

```

26          NN=MAX
27          MM=2
28      C
29      C          ВНЕШНИЙ ЦИКЛ ДЛЯ СЛОЕВ NPOW
30          DO 45 LAYER = 1, NPOW
   ⋮          ⋮
63      45      MM=MM*2

```

Отсюда видно, что число MM, определяющее число проходов среднего цикла, сначала равно 2, а затем каждый раз удваивается. Если учесть, что вычисления в среднем цикле проводятся с шагом 2, то для него по формуле $\frac{MM}{2} \times \frac{NN}{2}$ можно определить число выполнений операций в строках 47—49. Для внешнего цикла это число будет равно (после p проходов)

$$\frac{2}{2} \frac{N}{2} + \frac{4}{2} \frac{N/2}{2} + \dots + \frac{N}{2} 1 = p \frac{N}{2}.$$

Следовательно, в центральной части БПФ производится примерно $Np/2$ комплексных умножений, комплексных сложений и комплексных вычитаний.

На последнем этапе работы проводится переупорядочение данных, соответствующее адресации с обратным порядком битов, и умножение результатов на T или b . Этим операциям отвечают строки программы с 68 по 86:

```

64      C
65      C          ПРОВОДИТСЯ ЗАКЛЮЧИТЕЛЬНОЕ УПОРЯДОЧЕНИЕ
66      C          А ТАКЖЕ УМНОЖЕНИЕ НА DELTA
67      C

```

```

68      NW=0
69      DO 80 I=1,NMAX
70      NW1=NW+1
71      HOLD=X(NW1)
72      IF(NW1-I) 60,55,50
73      50  X(NW1)=X(I)*DELTA
74      55  X(I)=HOLD*DELTA
75      C
76      C
77      C          ВЫЧИСЛЕНИЕ ОБРАТНОГО АДРЕСА
78      C
79      60  DO 65 LOC=1,NPOW
80      LL=NW-MSK(LOC)
81      IF (LL) 70,75,65
82      65  NW=LL
83      70  NW=MSK(LOC)+NW
84      GO TO 80
85      75  NW=MSK(LOC+1)
86      80  CONTINUE

```

Вся эта небольшая программа довольно проста. Ее работу можно несколько убыстрить, предварительно переупорядочив массив исходных данных. Это позволяет избежать тех переадресаций, которые необходимы при получении аргументов для синусов и косинусов. Скорость работы программы можно увеличить также рекурсивным получением синусов и косинусов. Однако, добровольно жертвуя памятью для увеличения скорости за счет выбора значений синусов и косинусов из таблицы, приходится отказываться от предварительного переупорядочения данных, поскольку оно не приносит желаемой выгоды из-за того, что соответствующее переупорядочение можно провести и непосредственно в таблицах синусов и косинусов.

В программу можно внести и другие усовершенствования. Например, при первом прохождении среднего цикла (т. е. при $NW=0$) нет необходимости умножений на $CXCS$, что позволяет оставить только сложения и вычитания (см. упр. 6.13). Внимательный учет такого рода особенностей алгоритма дает бесчисленное множество различных возможностей повышения эффективности программы БПФ. Количество таких усовершенствований диктуется лишь объемом и сложностью соответствующих изменений в программе. Никаких общих рекомендаций на этот счет привести нельзя, поскольку в каждом частном случае может оказаться наилучшим свой вариант.

Алгоритм Синглтона для произведения степеней 2, 3, 4 и 5. Исследованию программ БПФ для последовательностей с длиной, равной произведению степеней произвольных целых чисел,

посвящена работа Синглтона (1969). Наименее эффективной из этих программ оказалась программа для произведений степеней 2, 3, 4 и 5. Наиболее удобный и эффективный вариант этого случая предложен Мэйнердом (1973). Основным недостатком его программы — это сложность, поскольку программа в себя почти 400 предложений Фортрана. Мы не будем приводить текста программы, которую читатель может взять у Мэйнерда, и остановимся лишь на особенностях алгоритма. В табл. 6.2 приведены числа от 2 до 5000, которые можно представить в виде произведения степеней 2, 3 и 5. Эти числа гораздо плотнее, чем степени 2, заполняют выбранный промежуток.

Синглтон определил прямое преобразование формулой

$$\alpha_k = \sum_{i=0}^{N-1} x_i \exp \left[\frac{j2\pi ik}{N} \right].$$

Его определение соответствует нашему определению обратного преобразования, правда, без множителя b . При использовании его программы это расхождение нетрудно ликвидировать.

БПФ со смешанной базой $N = N_1 N_2 \dots N_p$ основано на представлении преобразования Фурье в матричном виде

$$\mathbf{X} = \mathbf{W}\mathbf{x}, \quad (6.19)$$

где матрицу \mathbf{W} можно разложить в произведение

$$\mathbf{W} = \mathbf{P}\mathbf{W}_p\mathbf{W}_{p-1} \dots \mathbf{W}_2\mathbf{W}_1,$$

в котором матрицы \mathbf{W}_i соответствуют сомножителям N_i , а \mathbf{P} — матрица перестановки. Каждую из матриц \mathbf{W}_i можно разбить в свою очередь на N_i/N_i подматриц размерности N_i . Далее, матрицы \mathbf{W}_i можно представить в виде произведения

$$\mathbf{W}_i = \mathbf{R}_i\mathbf{T}_i,$$

где \mathbf{R}_i — диагональная матрица вращения, а матрица \mathbf{T}_i распадается на N_i/N_i одинаковых квадратных матриц.

Сомножители в алгоритме Синглтона располагаются симметрично. Например, для $N = 270$ множители располагаются в таком порядке: $3 \times 2 \times 3 \times 5 \times 3$. После этого выбирается перестановка, соответствующая такому расположению сомножителей алгоритма.

Алгоритм Синглтона, использующий внешние запоминающие устройства. В другой известной работе Синглтона (1967) рассмотрено несколько способов вычисления БПФ последовательностей, длина которых не позволяет поместить все данные в быстродействующее запоминающее устройство на магнитных сердечниках. Мы рассмотрим алгоритм менее общего вида, чем

Таблица 6.2

Целые числа от 1 до 5000, представимые в виде $N = 2^i 3^j 5^k$

N	i	j	k	N	i	j	k	N	i	j	k
2	1	0	0	243	0	5	0	1458	1	6	0
3	0	1	0	250	1	0	3	1500	2	1	3
4	2	0	0	256	8	0	0	1536	9	1	0
5	0	0	1	270	1	3	1	1600	6	0	2
6	1	1	0	288	5	2	0	1620	2	4	1
8	3	0	0	300	2	1	2	1728	6	3	0
9	0	2	0	320	6	0	1	1800	3	2	2
10	1	0	1	324	2	4	0	1875	0	1	4
12	2	1	0	360	3	2	1	1980	7	1	1
15	0	1	1	375	0	1	3	1944	3	5	0
16	4	0	0	384	7	1	0	2000	4	0	3
18	1	2	0	400	4	0	2	2025	0	4	2
20	2	0	1	405	0	4	1	2048	11	0	0
24	3	1	0	432	4	3	0	2160	4	3	1
25	0	0	2	450	1	2	2	2187	0	7	0
27	0	3	0	480	5	1	1	2250	1	2	3
30	1	1	1	486	1	5	0	2304	8	2	0
32	5	0	0	500	2	0	3	2400	5	1	2
36	2	2	0	512	9	0	0	2430	1	5	1
40	3	0	1	540	2	3	1	2500	2	0	4
45	0	2	1	576	6	2	0	2560	9	0	1
48	4	1	0	600	3	1	2	2592	5	4	0
50	1	0	2	625	0	0	4	2700	2	3	2
54	1	3	0	640	7	0	1	2880	6	2	1
60	2	1	1	648	3	4	0	2916	2	6	0
64	6	0	0	675	0	3	2	3000	3	1	3
72	3	2	0	720	4	2	1	3072	10	1	0
75	0	1	2	729	0	6	0	3125	0	0	5
80	4	0	1	750	1	1	3	3200	7	0	2
81	0	4	0	768	8	1	0	3240	3	4	1
90	1	2	1	800	5	0	2	3375	0	3	3
96	5	1	0	810	1	4	1	3456	7	3	0
100	2	0	2	864	5	3	0	3600	4	2	2
108	2	3	0	900	2	2	2	3645	0	6	1
120	3	1	1	960	6	1	1	3750	1	1	4
125	0	0	3	972	2	5	0	3840	8	1	1
128	7	0	0	1000	3	0	3	3888	4	5	0
135	0	3	1	1024	10	0	0	4000	5	0	3
144	4	2	0	1080	3	3	1	4050	1	4	2
150	1	1	2	1125	0	2	3	4096	12	0	0
160	5	0	1	1152	7	2	0	4320	5	3	1
162	1	4	0	1200	4	1	2	4374	1	7	0
180	2	2	1	1215	0	5	1	4500	2	2	3
192	6	1	0	1250	1	0	4	4608	9	2	0
200	3	0	2	1280	8	0	1	4800	6	1	2
216	3	3	0	1296	4	4	0	4860	2	5	1
225	0	2	2	1350	1	3	2	5000	3	0	4
240	4	1	1	1440	5	2	1				

тот, который был предложен Синглтоном, и ограничимся следующими случаями.

Во-первых, будем считать, что алгоритм предназначен для последовательностей с длиной, равной степени 2.

Во-вторых, предположим, что в нашем распоряжении имеются либо четыре запоминающих устройства на ленте, либо один или более дисков, которые можно использовать как четыре устройства ввода-вывода.

Говоря о четырех запоминающих устройствах на ленте, которые встречаются в описании алгоритма, мы должны помнить, что для тех же целей можно воспользоваться одним запоминающим устройством на диске, в котором имеется 4 соответствующих области запоминания.

Допустим, что ряд, для которого нужно получить преобразование Фурье, содержит $N = 2^p$ точек (комплексных величин) и что данные записаны блоками по 2^r комплексных значений в каждом блоке, причем 2^r значительно меньше 2^p . Оказывается, для того чтобы получить преобразование этого ряда и переставить в нужном порядке данные, требуется $2p - r - 2$ проходов. Эти проходы распадаются на два этапа: в первом производится $p - r$ вычислительных проходов и $p - r - 2$ перестановочных проходов; во втором этапе производится r вычислительных проходов.

Например, если $N = 2^{15}$ и $2^r = 256 = 2^8$ (на один блок данных приходится, следовательно, по 512 слов для действительных чисел), то первый этап сведется к 7 вычислительным и 5 перестановочным проходам, а второй — к 8 вычислительным.

Вычисление начинается с записи первых $N/2$ комплексных величин, для которых нужно получить преобразование, на первую ленту и остальных $N/2$ комплексных величин на вторую ленту. Все четыре ленты перематываются к началу.

На первом этапе два блока данных, по одному с каждой из двух первых лент, записываются в память ЭВМ на сердечниках как отдельные массивы. Из каждого массива последовательно берутся комплексные величины, преобразуются и затем помещаются на место исходных величин. Преобразование проводится по формулам

$$Y(k) = X(k) + X(k + N/2) \exp \left[-j\pi \left(\frac{2(k/(p-q))}{2^{q-1}} \right) \right],$$

$$Y\left(k + \frac{N}{2}\right) = X(k) - X(k + N/2) \exp \left[-j\pi \left(\frac{2(k/(p-q))}{2^{q-1}} \right) \right],$$

где $X(k)$ есть k -е значение на ленте 1, $k = 0, \dots, N/2$; $X(k + N/2)$ есть k -е значение на ленте 2, $k = 0, \dots, N/2$; $Y(k)$ — замена для $X(k)$, а $Y(k + N/2)$ — замена для $X(k + N/2)$; q — номер вычислительного прохода, $q = 1, 2, \dots, m$. Вычисления с величинами

в круглых скобках должны проводиться в целочисленной арифметике.

Такие же операции проводятся на всех m вычислительных проходах.

После обработки 2^r величин каждого блока результаты записываются на ленты 3 и 4. Те результаты, которые получены из данных, находящихся на первой ленте, записываются на третью, а те, которые получены из данных со второй ленты, — на четвертую.

Вслед за первыми блоками с первой и второй лент считываются и обрабатываются вторые. Так продолжается до тех пор, пока вся информация с 1-й и 2-й лент не будет считана, обработана и записана на ленты 3 и 4. Все эти операции составляют первый вычислительный проход, входящий в первый этап получения преобразования.

На перестановочных проходах 1-го этапа проводятся обратная перемотка всех четырех лент и перестановка данных. При этом начинают с третьей ленты: блоки считываются с первой ее половины и записываются на 1-ю ленту, со второй половины блоки записываются на 2-ю ленту. После того как вся третья лента переписана, та же процедура повторяется для четвертой ленты.

Так производится перестановка при $q=1$. На следующих перестановочных проходах, по одному на каждый вычислительный проход, число блоков, последовательно записываемых с лент 3 и 4, равно $2^{p-r-1-q}$. Следовательно, при $q=1$ число блоков, последовательно выбираемых и записываемых на первые две ленты, равно 2^{p-r-2} , при $q=2$ число таких блоков равно 2^{p-r-3} и т. д. При $q=p-r-2$ последовательно берется по два блока, а при $q=p-r-1$ и $q=p-r$ никаких перестановок не производится. Например, для $N=2^{15}$, $r=8$ и $p=15$ число последовательно записываемых блоков при $q=1$ равно 32, при $q=2$ равно 16 и т. д. При $q=5$ последовательно записываются пары блоков, а при $q=6$ и $q=7$ перестановочных проходов нет. Таким образом, для двух последних вычислительных проходов первого этапа перестановочные проходы отсутствуют.

Вообще говоря, при $q=p-r-1$ (в нашем примере при $q=6$) должен быть один перестановочный проход. Но соответствующая перестановка учитывается на последнем вычислительном этапе, т. е. при $q=p-r$, следующим образом: при обработке берутся не пары блоков по одному с каждой ленты, как это было на предыдущих вычислительных проходах, а пары блоков с одной ленты (сначала с первой, а затем, после того как все данные с нее будут обработаны, со второй). Обработка данных этих блоков проводится так же, как и прежде. Такой ход вычислений используется только один раз, при $q=p-r$, т. е. на пос-

леднем вычислительном проходе 1-го этапа вычисления преобразования.

На втором этапе, включающем в себя r проходов, вычисления и перестановки объединены вместе. На каждом из этих проходов с данными поступают следующим образом. На вход подаются блски по одному, начиная с первого блока первой ленты входа. После окончания первой ленты на вход подаются блоки данных со второй ленты, начиная с последнего блока. В отличие от первого этапа теперь требуется 3 массива данных объемом 2^r (комплексных) — на один больше, чем раньше. Один из этих массивов предназначен для записи данных обрабатываемого блока. В двух других помещаются результаты для вывода на ленты.

Перестановка на q -м вычислительном шаге, где теперь $p - r - 1 \leq q < p$, производится следующим образом. Для обработки берутся пары комплексных значений блока, разделенные 2^{p-q} величинами. В нашем примере при $q = 8$ на первой стадии вычислений расстояние между парой обрабатываемых величин равно $2^{15-8} = 2^7 = 128$, поэтому вычисления на втором этапе начинаются для 1-й и 129-й величин, затем обрабатывается пара из 2-й и 130-й величин и т. д. Формулы для вычислений остаются теми же, что и на первом этапе. Результаты последовательно запоминаются в выходные массивы: величина $Y(k)$ в первый массив, а $Y(k + N/2)$ — во второй. После того как оба массива окажутся заполненными, они записываются на первую и вторую ленты выхода. Поскольку данные для обработки берутся из входного массива парами, для двух выходных массивов требуется два входных массива.

После каждого такого вычислительного прохода 2-го этапа ленты перематываются и меняются ролями: вход становится выходом и наоборот.

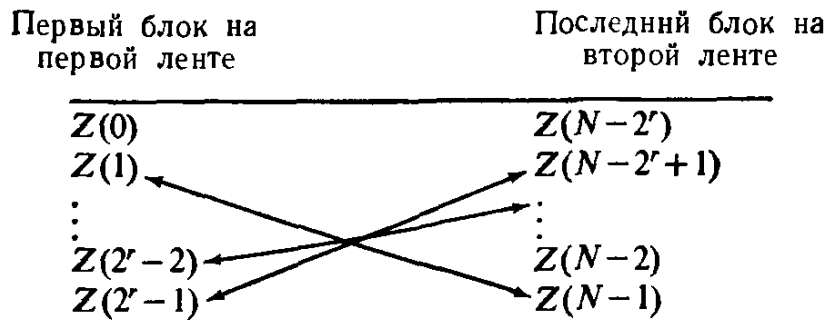
На последнем r -м проходе 2-го этапа на выходе получается искомое преобразование Фурье данных. Первая половина преобразования записана на первой ленте выхода, а вторая — на второй ленте.

Если исходная функция была действительной, т. е. все мнимые части были равны 0, то в результате получится нужное преобразование, и никаких дополнительных действий проводить не нужно.

Если же проводится спаренное преобразование Фурье, т. е. проводится оновременное определение преобразований двух действительных функций, скажем $x(i)$ и $y(i)$, которые выступают в качестве действительной и мнимой частей исходного ряда данных соответственно, то полученный результат еще необходимо разъединить. При этом используется алгоритм, основу которого составляют формулы (6.8) и (6.9).

Для этой цели сначала берутся первый блок результатов с первой ленты и последний блок результатов со второй ленты, затем обрабатываются второй и предпоследний блоки с первой и второй лент соответственно и т. д. Иначе говоря, необходимо чтение с конца второй ленты и возвращение ленты к началу предыдущего блока.

Следует учесть, что блоки результатов при такой процедуре накладываются не совсем точно, и одно комплексное значение должно запоминаться для следующей пары блоков. Как это получается, нетрудно понять, взглянув на следующую диаграмму:



Стрелки на ней указывают, как выбираются пары для получения отдельных преобразований $X(k)$ и $Y(k)$. Первая величина первого блока берется при этом безо всяких изменений, а первый элемент второго блока, как и говорилось, необходимо запоминать и использовать при обработке второй пары блоков. Это, естественно, несколько увеличивает сложность алгоритма.

6.3. ПРИМЕРЫ

В этом параграфе мы рассмотрим в качестве иллюстраций применения преобразования Фурье несколько примеров; в частности, вычисления преобразований: суммы синусов и косинусов, стационарного случайного процесса, переходного сигнала и временного ряда вибрационных данных, полученных с помощью акселерометра, как мы надеемся, дадут читателю возможность «почувствовать» преобразование Фурье и представить его вид в наиболее характерных ситуациях. Рассмотренные примеры помогут понять, как преобразование Фурье используется для интерполяции числовых функций, как проявляется эффект добавления к данным нулей и чем отличается рассмотренный выше случай широкополосного случайного шума от случая узкополосного шума; эти примеры помогут также оценить значение преобразования Фурье для практических задач и покажут те искажения, которые проистекают из типичных ошибок в данных.

Суммы синусов и косинусов. Напомним, что преобразование Фурье синусоиды частоты f_0 — пара дельта-функций, расположен-

ных в точках $\pm f_0$ (см. гл. 1). В силу линейности для суммы синусоид преобразованием Фурье будет служить сумма соответствующих пар дельта-функций. Далее, напомним, что конечное дискретное преобразование Фурье суммы синусоид будет суммой не пар дельта-функций, а суммой заменяющих их функций $(\sin x)/x$. Рассмотрим сумму трех дискретных синусоид:

$$x(i) = \sum_{k=1}^3 a_k \cos(2\pi i T f_k + \varphi_k), \quad (6.20)$$

где $i = 0, 1, \dots, N-1$. Эти синусоиды имеют следующие частоты и фазы:

k	a_k	f_k	φ_k
1	0.986	$F/16 = 1/8T$	$\pi/3 = 60^\circ$
2	1.225	$F/8 = 1/4T$	$\pi/5 = 36^\circ$
3	0.707	$3F/16 = 3/8T$	$\pi/4 = 45^\circ$

Последовательность значений из 128 точек, полученных на ЭВМ по слегка измененной формуле

$$x(i) = \sum_{k=1}^3 a(k) \cos\left(\frac{2\pi k}{16} i + \varphi_k\right), \quad (6.21)$$

где $i = 0, 1, \dots, 127$, изображена на рис. 6.19. На рис. 6.20 изображены графики преобразования Фурье этой последователь-

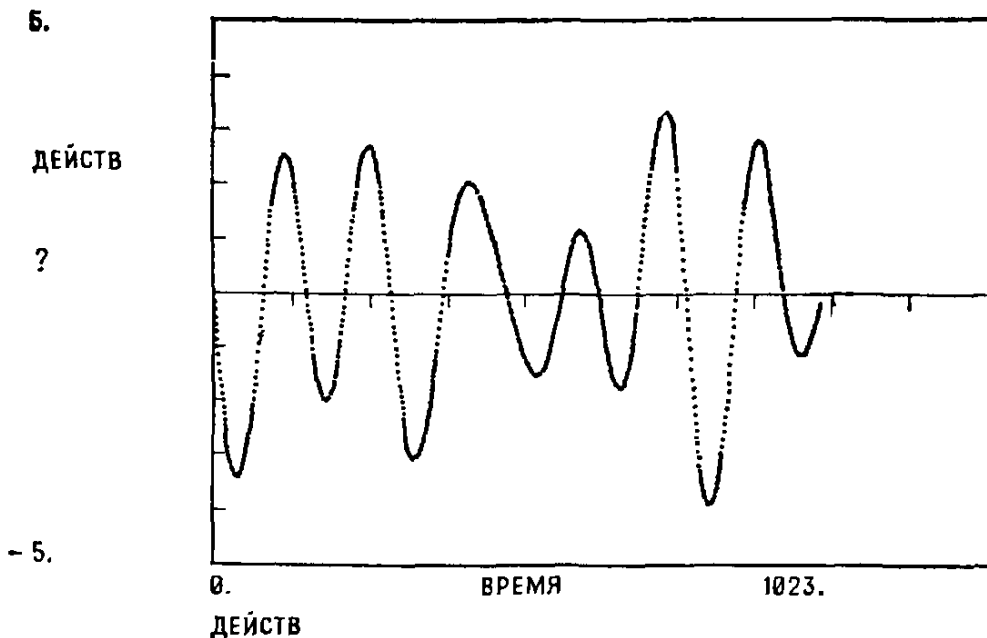
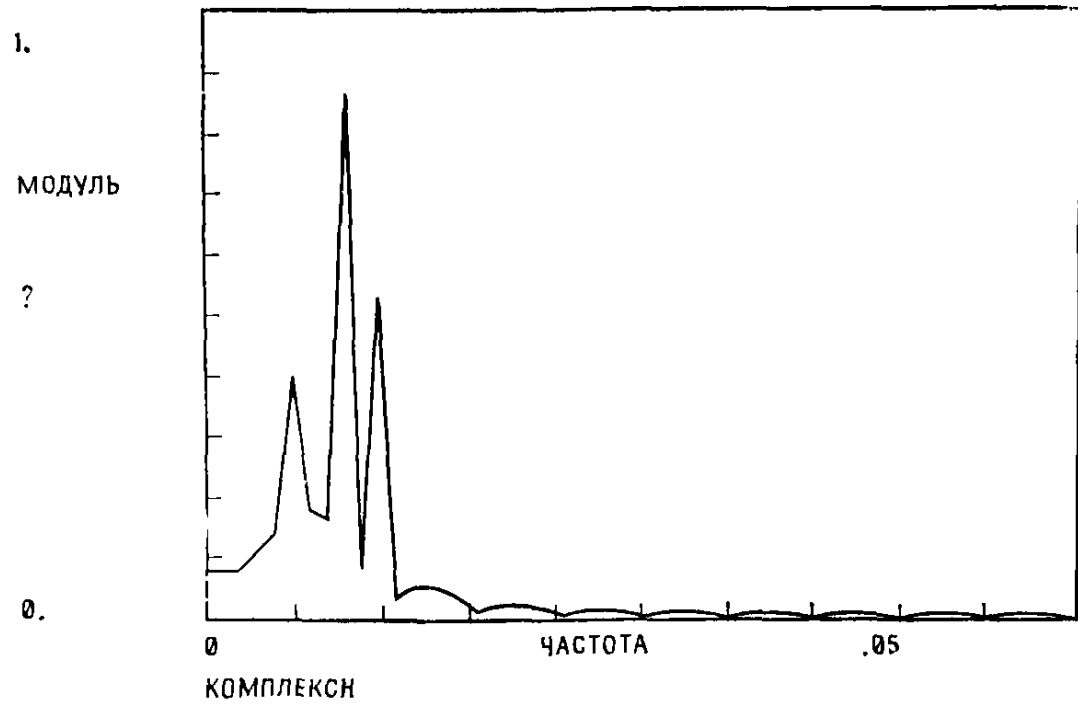
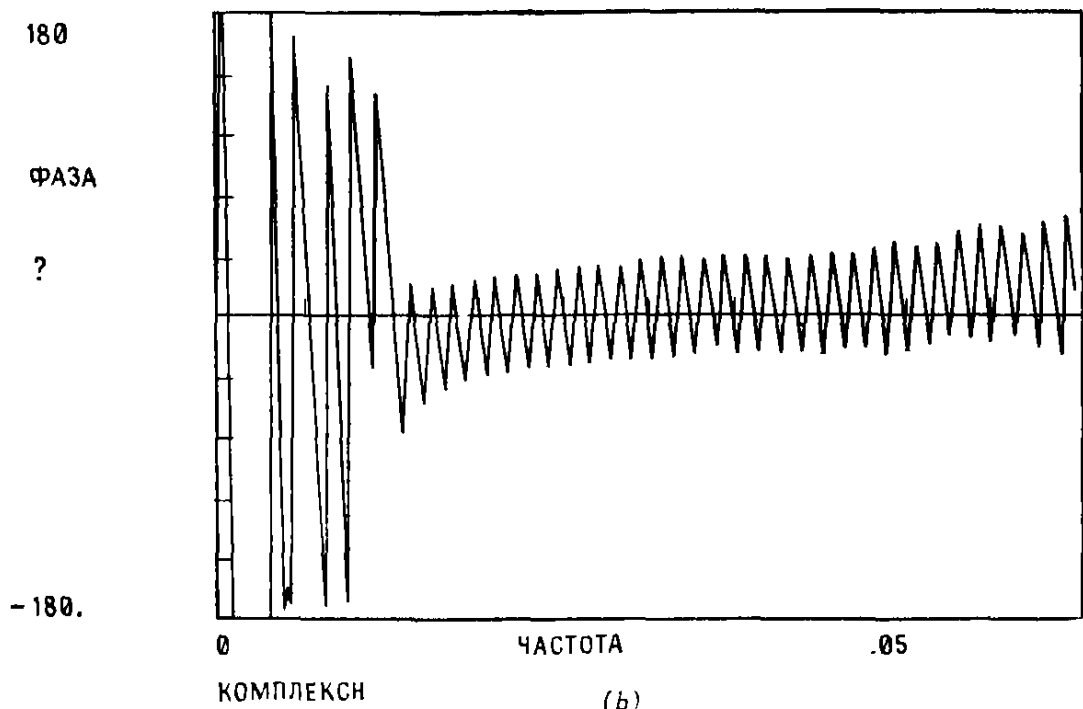


Рис. 6.19. График для трех синусоид без фаз.

ности в декартовой и в полярной системах координат. Рис. 6.21, на котором в полярных координатах изображено 1024-точечное



(a)

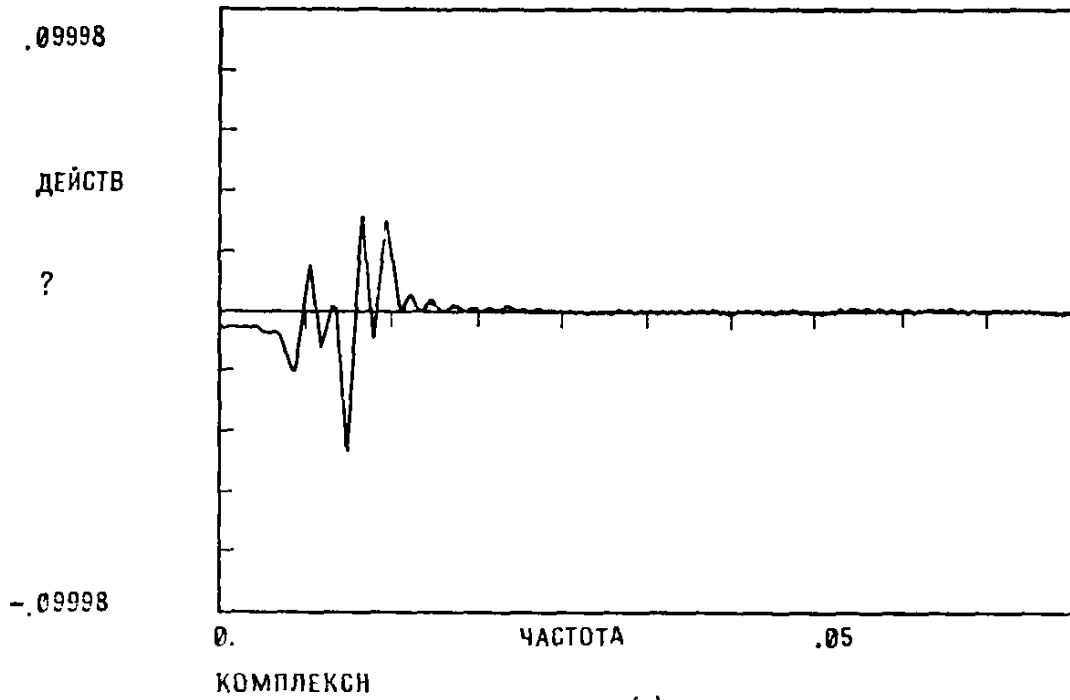


(b)

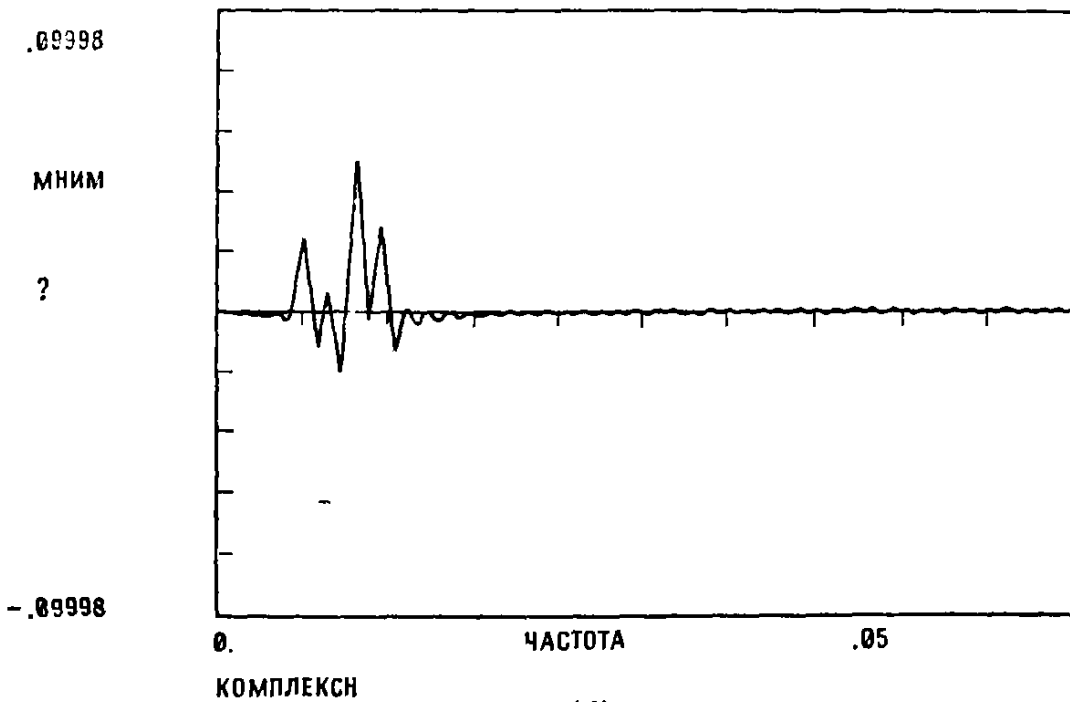
Рис. 6.20. Преобразование Фурье временной

преобразование Фурье той же последовательности, дает возможность проследить интерполяцию в частотной области.

Стационарный случайный шум. Рассмотрим преобразование Фурье стационарного случайного процесса. Для этой цели восполь-



(c)



(d)

последовательности, изображенной на рис. 6.19.

зуюемся генератором аналогового гауссовского белого случайного шума, выход которого преобразуется в цифровую форму. По 1024 точкам во временной области получим 513 точек в области частот. Для случайного шума $\mu_x = 0$, $\sigma_x^2 = 1$ и значения данных почти независимы друг от друга. Действительная и мнимая части

преобразования Фурье имеют гауссовское распределение с параметрами

$$\begin{aligned} E[X_R(k)] &= E[X_I(k)] = 0, \\ D[X_R(k)] &= D[X_I(k)] = \frac{1}{2NT}. \end{aligned} \quad (6.22)$$

На рис. 6.22 показана сама временная последовательность, преобразованная в цифровую форму, а на рис. 6.23 — ее преоб-

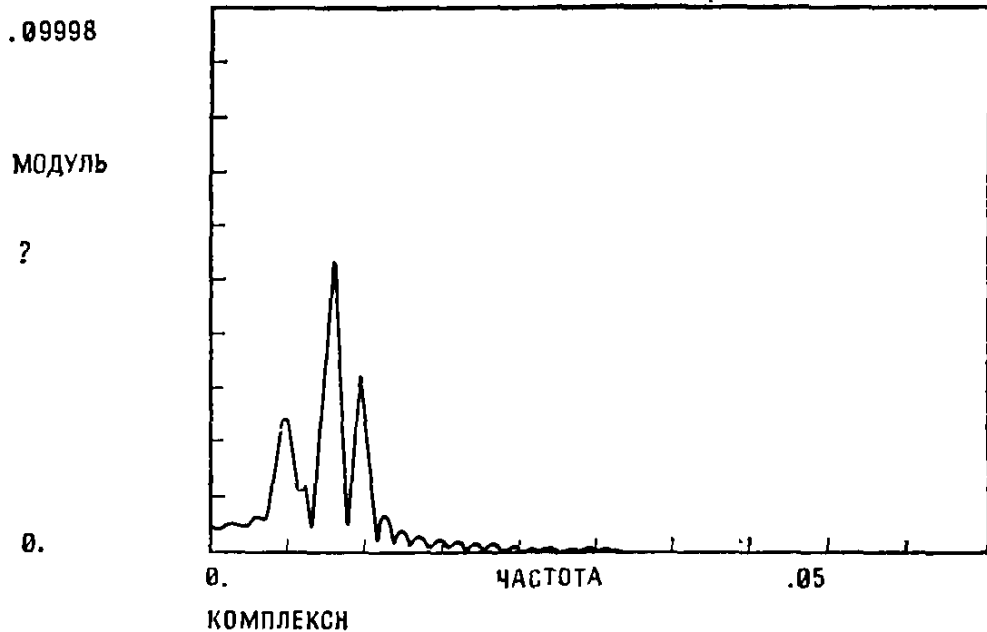


Рис. 6.21. Преобразование Фурье последовательности (6.21), дополненной 896 нулями.

разование Фурье. Амплитуда имеет распределение $[\chi(2)]$ с двумя степенями свободы, где переменная $\chi(n)$ есть квадратный корень из величины $\chi^2(n)$. Это распределение называют обычно распределением Релея.

Переходные данные, полученные при помощи акселерометра. Присоединим акселерометр к бойку небольшого ударника, который используется для удара по лопастям турбины. Это позволит определить импульсную функцию отклика при импульсной нагрузке (т. е. дельта-функции). Подробнее этот вопрос будет затронут в гл. 9. Сейчас мы хотим только показать вид преобразования Фурье для переходной (нестационарной) временной последовательности.

Данные преобразуются в цифровую форму со скоростью 4000 выб./с. График полученной временной последовательности приведен на рис. 6.24. Преобразование Фурье этой последовательности показано на рис. 6.25. Статистические результаты, которые справедливы для стационарных процессов, в данном случае неприменимы. Для переходных процессов такого рода

возникают многочисленные статистические трудности, обусловленные тем, что между полосами частот уже нет статистической независимости.

Интерполяция функций на ограниченной полосе частот. В этом пункте мы покажем, как проводится интерполяция значений дискретной функции на ограниченной полосе частот на примере распределения напряжения, полученного при структурном анализе строительных панелей.

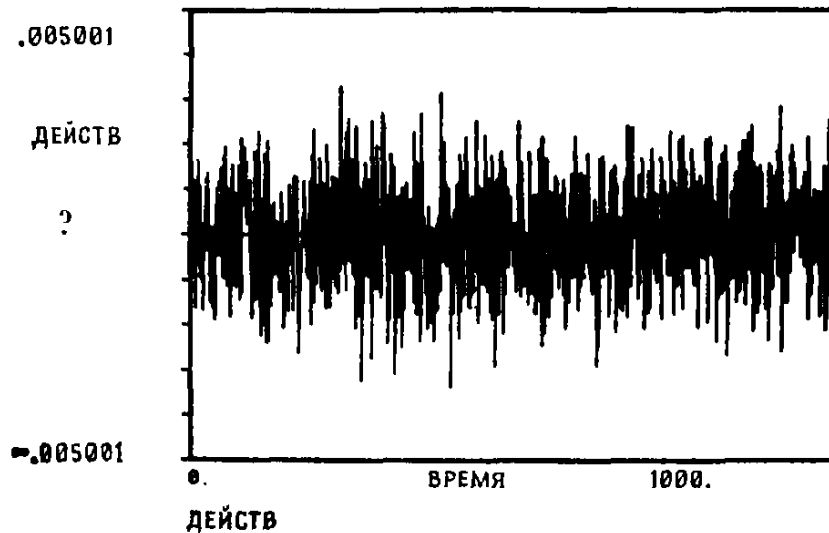


Рис. 6.22. Выход генератора гауссовского белого случайного шума, преобразованный в цифровую форму.

На рис. 6.26 показано, как определяется напряжение в двумерной плоскости при действии динамических точечных нагрузок. Вертикальное напряжение определяется при помощи конечного числа элементов, поэтому напряжение можно получить только в отдельных точках. Такой способ анализа дает общее представление о характере распределения напряжения, но сам по себе недостаточен для определения непрерывного распределения напряжения. На рис. 6.27 показано распределение напряжения на уровне, расположенном на 1.5 дюйма ниже плоскости нагрузки. Такая дискретная картина не дает точного представления о пиковых значениях напряжения под нагрузкой на 1.5-дюймовом уровне, поэтому ищут такие методы, чтобы оценить их более точно.

На рис. 6.28 приведен результат интерполяции имеющейся последовательности значений при помощи БПФ. Бросается в глаза, что пиковые значения напряжения интерполированной функции оказываются по величине значительно больше, чем у исходной функции. В этом и состоит основной результат проведенного анализа.

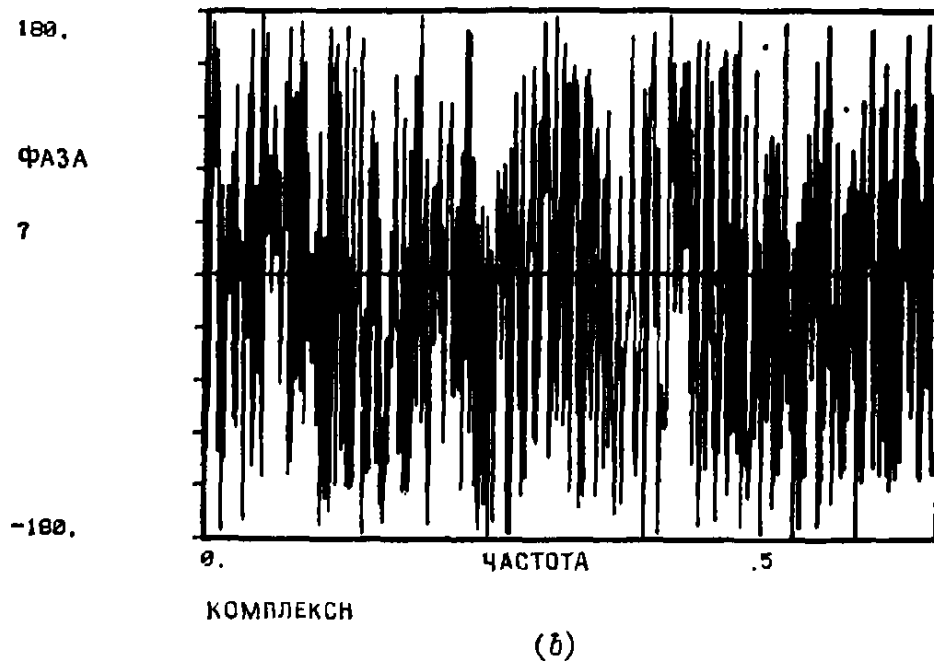
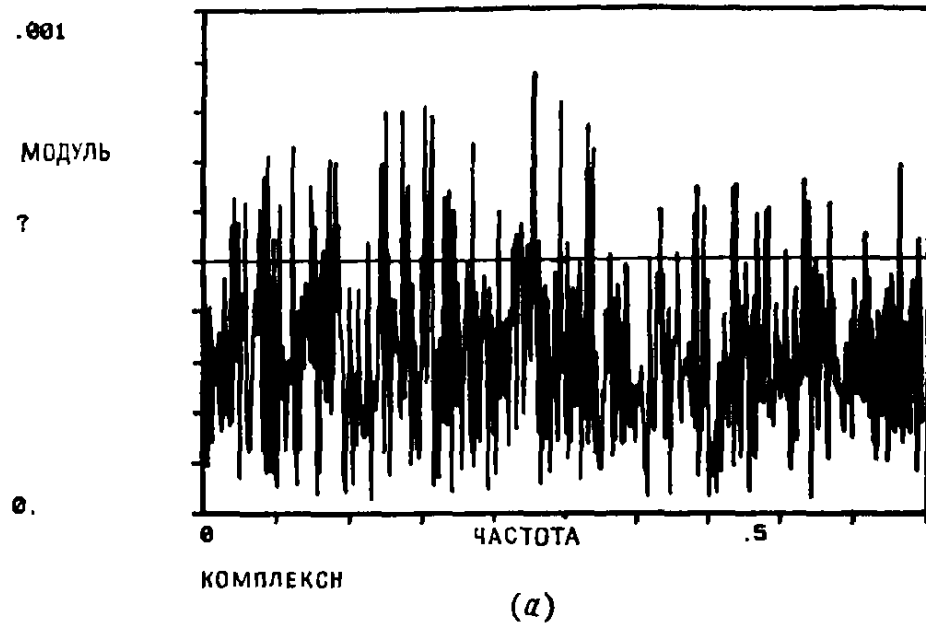
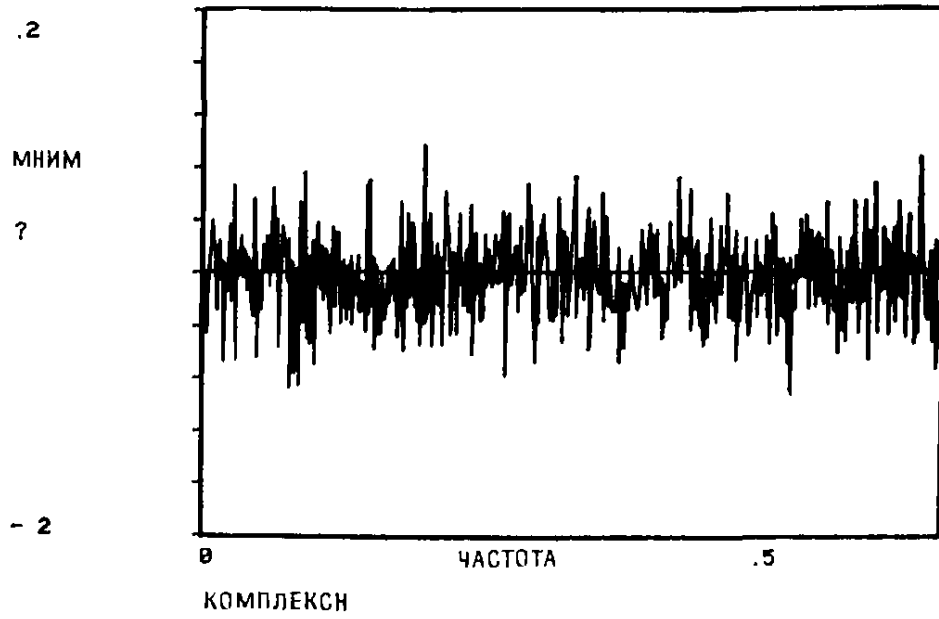
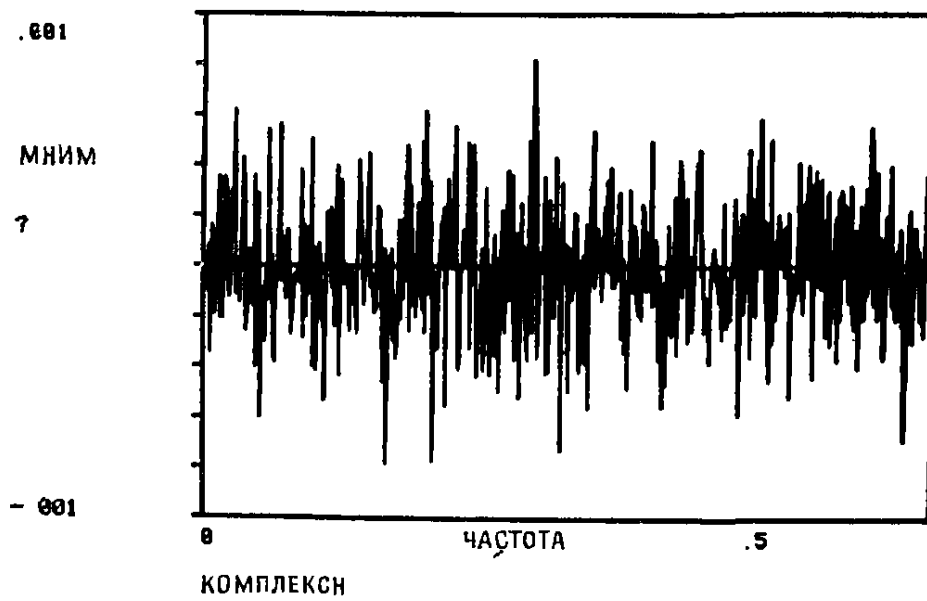


Рис. 6.23. Преобразование Фурье



(c)



(d)

стационарного белого шума.

При интерполяции к исходной последовательности сначала были добавлены нули, а затем для новой последовательности вычислялось преобразование Фурье. В исходной последователь-

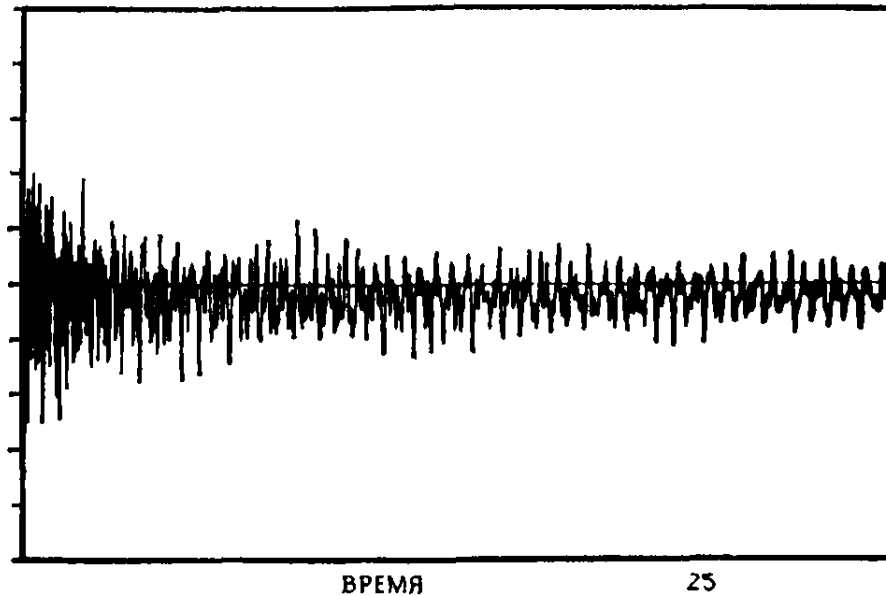


Рис. 6.24. Временная последовательность переходных данных, полученных при помощи акселерометра.

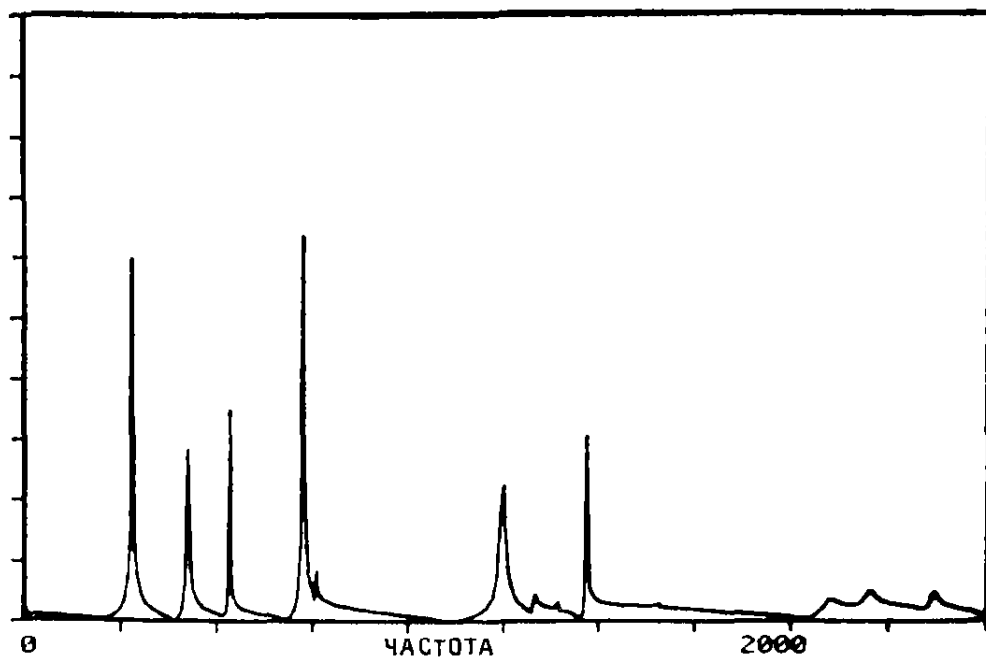


Рис. 6.25. Абсолютная величина преобразования Фурье переходных данных, приведенных на рис. 6.24.

ности было 16 точек, а в новой — 128 точек, т. е. на одно значение добавлялось по семь нулей.

Узкополосный случайный шум. На рис. 6.29, *a—e* показана картина, которую дает узкополосный шум. График на рис. 6.29

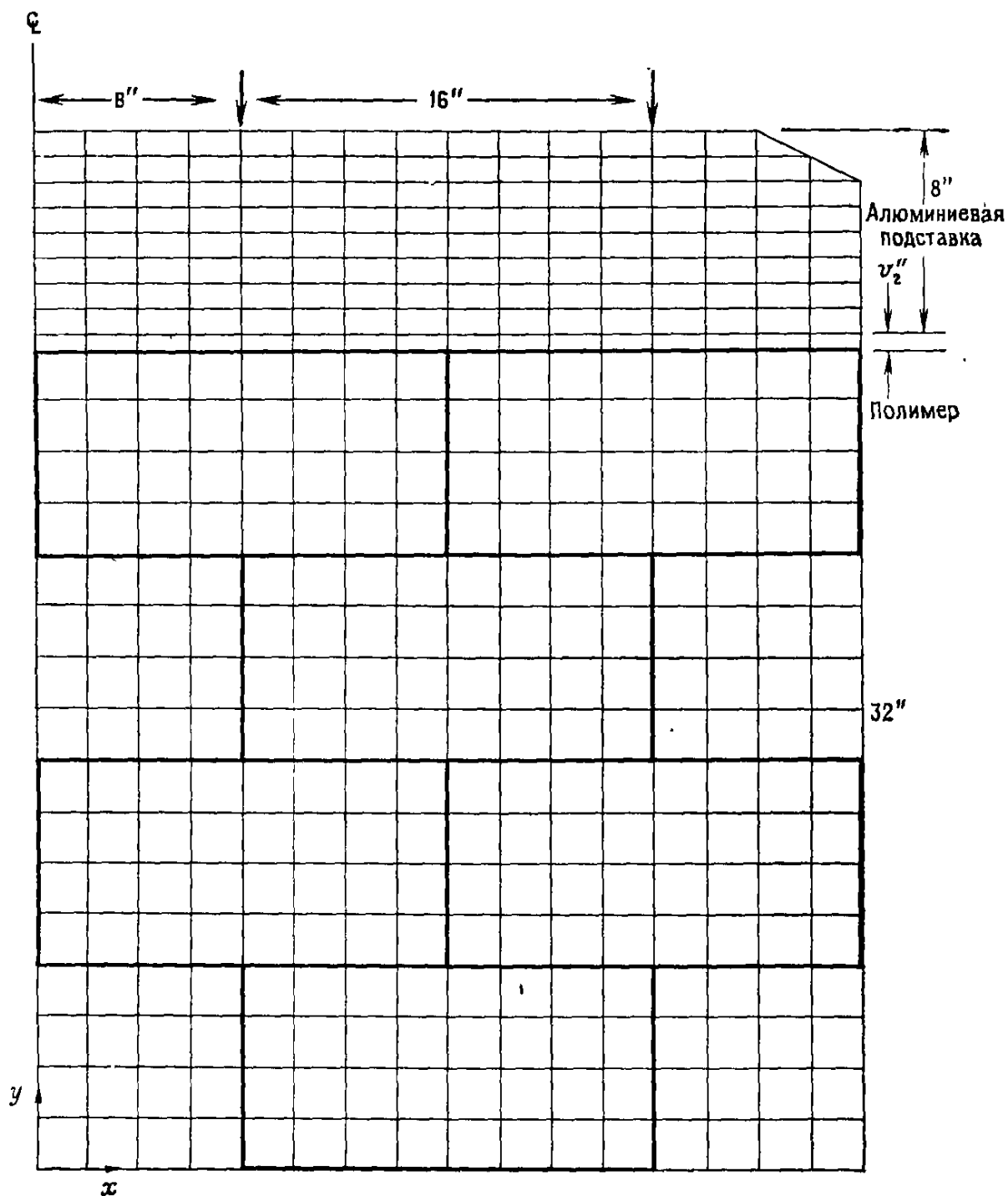


Рис. 6.26. Структурный анализ строительных панелей.

имеет известное сходство с модулированной синусоидой. Вместе с тем эти случаи имеют существенные различия, поскольку последовательность для шума дает узкую полосу в спектре и имеет случайную фазу.

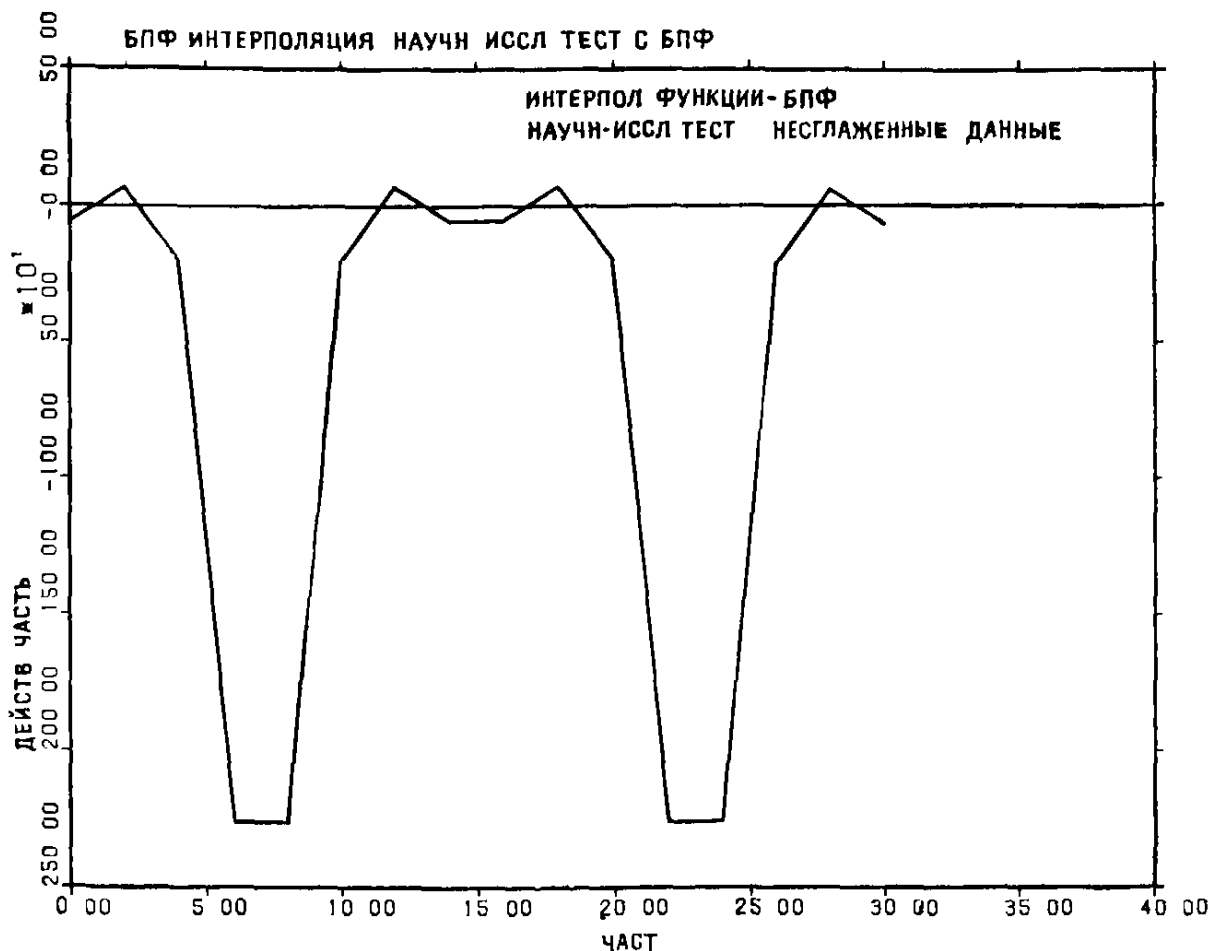


Рис 6 27 Исходные дискретные значения функции напряжения

Упражнения

6 1 Докажите что $Z(N+k) = Z(k)$ где

$$Z(k) = T \sum_{l=0}^{V-1} z(l) \exp\left(-j \frac{2\pi}{N} lk\right)$$

Ответ поскольку $\exp(-j2\pi l) = 1$ для всех l , имеем

$$\begin{aligned} z(l) \exp\left(-j \frac{2\pi}{V} 2[k+N]\right) &= z(l) \exp\left(-j \frac{2\pi}{N} lk\right) \exp\left(-j \frac{2\pi}{N} lN\right) = \\ &= z(l) \exp\left(-j \frac{2\pi}{N} lk\right), \end{aligned}$$

следовательно

$$\begin{aligned} Z(k+N) &= T \sum z(l) \exp\left(-j \frac{2\pi}{N} l(k+N)\right) = \\ &= T \sum z(l) \exp\left(-j \frac{2\pi}{N} lk\right) \end{aligned}$$

6 2 Вычислите преобразование для последовательности действительных величин из 128 точек, используя только один 64-точечный комплексный массив

6.3. Докажите, что $X(k) = X^*(N-k)$, если $x(t)$ — действительные числа

6.4. Вычислите обратное преобразование Фурье последовательности, полученной в упр 6.2, также используя один 64-точечный массив

6.5. Покажите, что преобразование Фурье последовательности, состоящей из двух величин $(x(0), x(1))$, дает величины $X(0) = (x(0) + x(1))T$ и $X(1) = (x(0) - x(1))T$

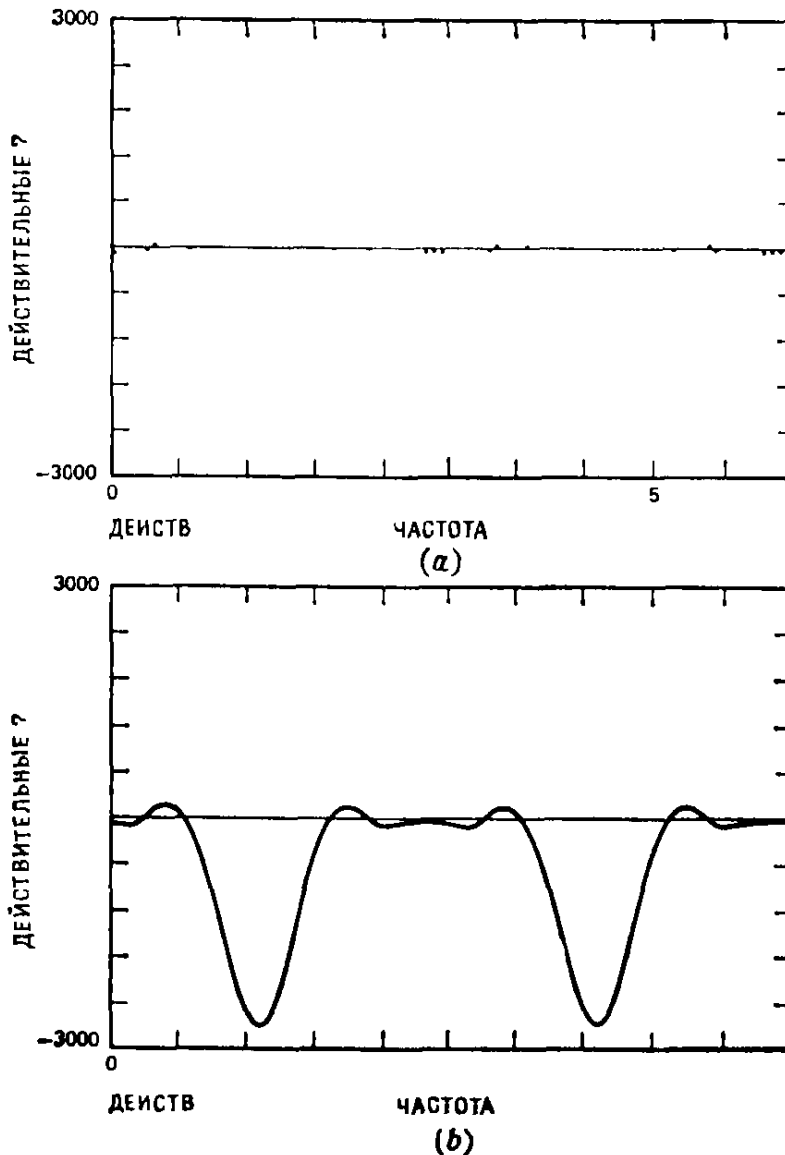


Рис 6.28 Результат интерполяции функции напряжения посредством БПФ с добавлением нулей 8 к 1

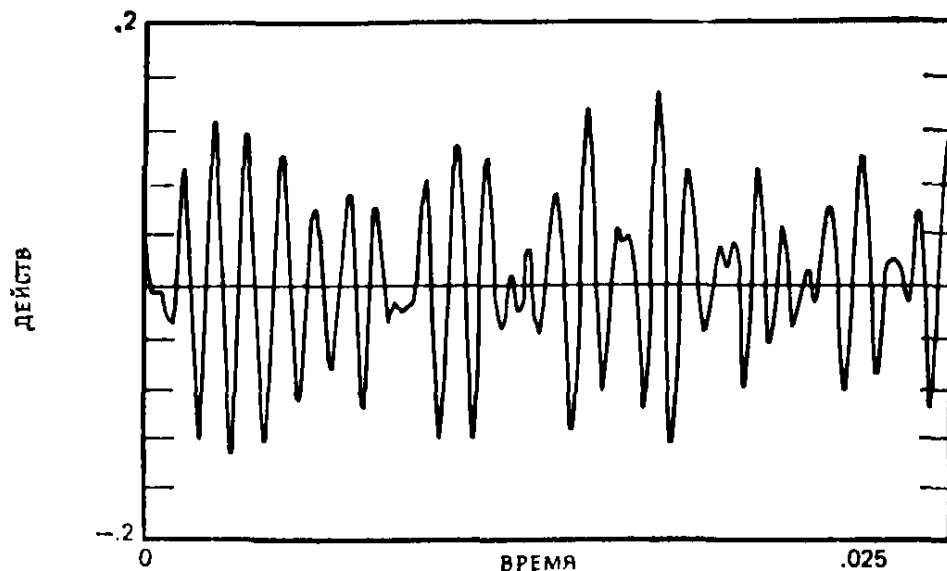
6.6. Получите преобразование, которое требуется в упр 6.2, используя 32-точечный массив

6.7. Докажите, что $X(0) = T \sum_{t=0}^{N-1} x(t)$.

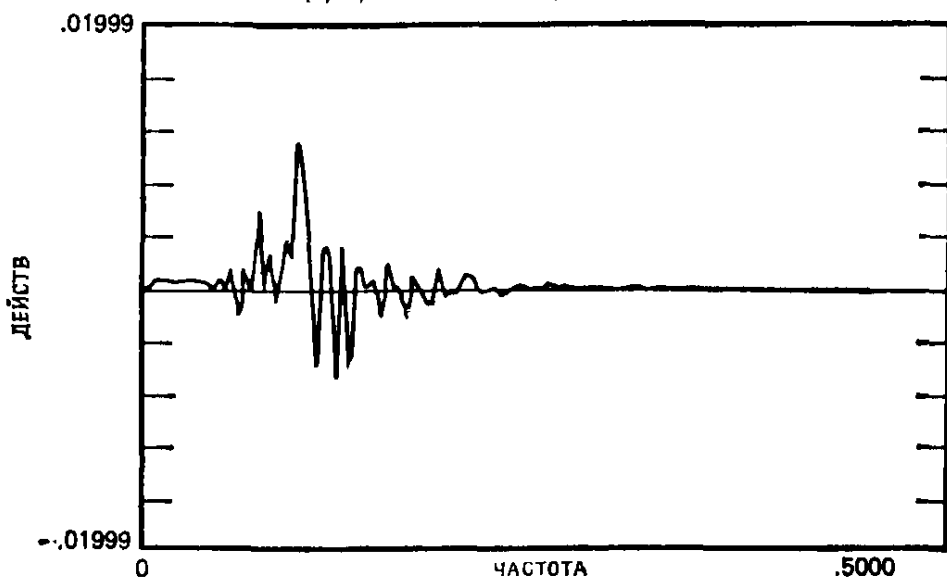
6.8. Переведите результаты упр. 6.2 в полярные координаты, т. е. выразите их через абсолютную величину и фазу

6.9. Докажите, что при четном N выполняется соотношение

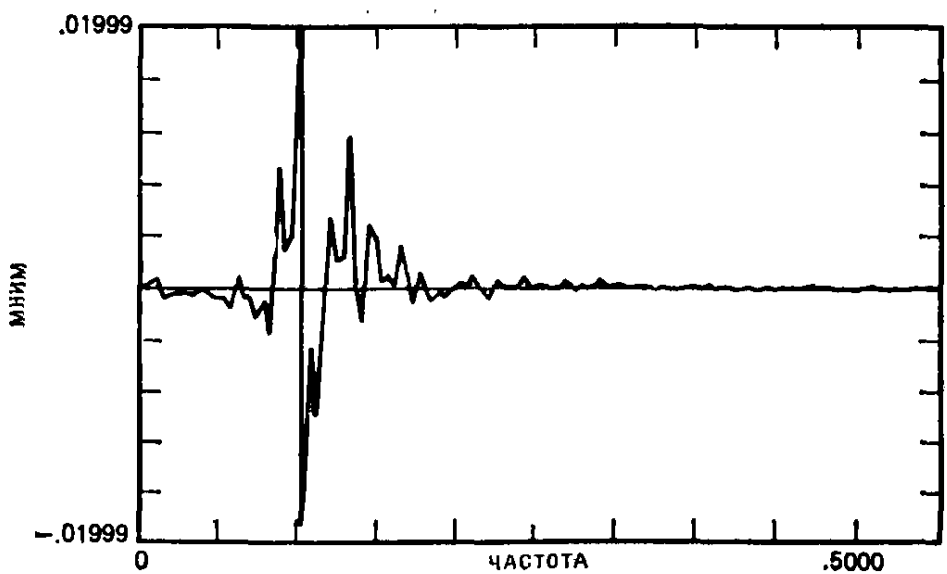
$$X(N/2) = \sum_{t=0}^{N-1} x(2t) - \sum_{t=0}^{N-1} x(2t+1)$$



(а) Временная последовательность



(б) Действительная часть преобразования Фурье



(с) Мнимая часть преобразования Фурье

Рис. 6,29, Узкополосный случайный процесс

6.10. Проведите продолжение фазы для последовательности, полученной в упр. 6.8.

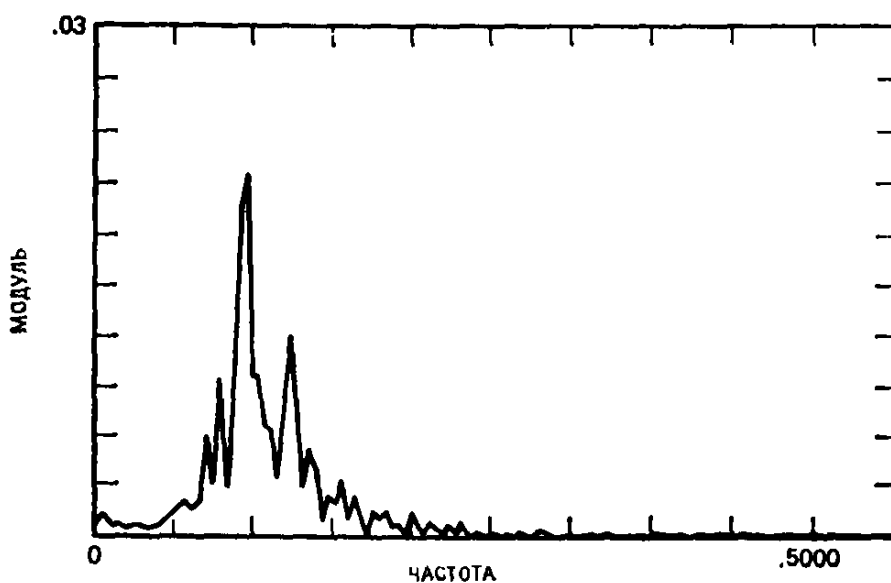
6.11. Приведите вывод формул Эйлера

$$\begin{aligned}\sin(n\alpha) &= 2 \sin((n-1)\alpha) \cos \alpha - \sin((n-2)\alpha), \\ \cos(n\alpha) &= 2 \cos((n-1)\alpha) \cos \alpha - \cos((n-2)\alpha).\end{aligned}$$

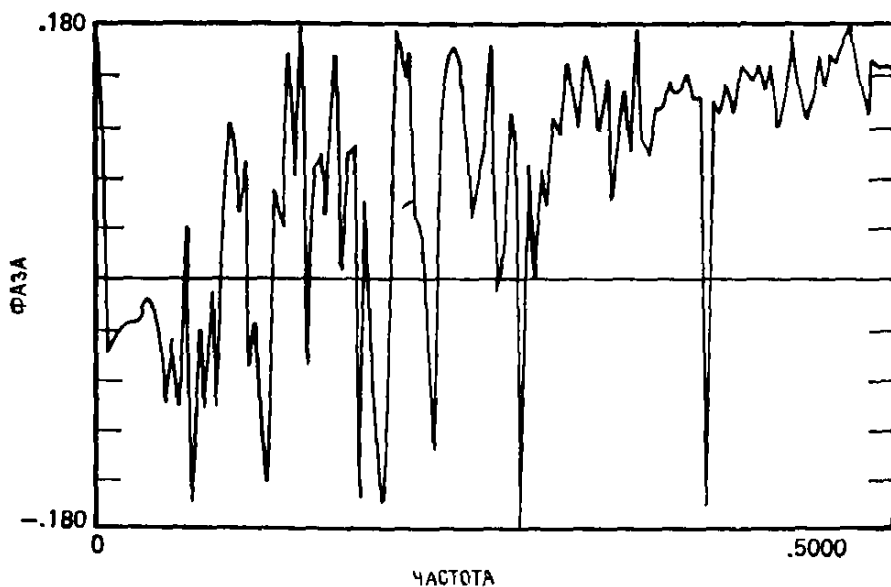
6.12. Получите выборку псевдослучайного белого шума, вычислите ее преобразование Фурье и изобразите результаты графически, воспользовавшись подпрограммой TRAND или какой-нибудь другой эквивалентной подпрограммой.

6.13. Перепишите подпрограмму FFT, в которой предусмотрено, что в первом проходе внешнего цикла умножения не нужны. Сколько операций сохранит такое усовершенствование?

6.14. Получите узкополосный процесс, изменив преобразование Фурье, полученное в упр. 6.12, и вычислите обратное преобразование при помощи



(d) Абсолютная величина преобразования Фурье



(e) Фаза преобразования Фурье

и его преобразование Фурье.

обратного БПФ. Приведите график для полученной временной последовательности.

6.15. Получите формулу (6.21) из определения (6.20).

Ответ: поскольку $f_k = k/(16T)$, для суммы

$$\sum a_k \cos(2\pi i T f_k + \varphi_k)$$

имеем

$$\sum a_k \cos\left(2\pi i T \frac{k}{16T} + \varphi_k\right) = \sum a_k \cos\left(2\pi k \frac{i}{16} + \varphi_k\right) = \sum a_k \cos\left(\frac{2\pi k}{16} i + \varphi_k\right).$$

6.16. Повторите графики, изображенные на рис. 6.29, используя результаты упр. 6.14.

6.17. Объясните, как получаются равенства (6.22).

Ответ: а) в силу того, что преобразование Фурье линейно, справедливы равенства $E[L(x)] = LE(x) = 0$;

б) интеграл квадрата преобразования Фурье должен быть равным σ_x^2 , поскольку

$$\Delta f = b = \frac{1}{NT}, \quad E|X^2| \frac{1}{NT} = \sigma_x^2$$

и

$$E \frac{X_R^2}{NT} + E \frac{X_I^2}{NT} = \sigma_x^2;$$

так как $EX_R^2 = EX_I^2$, то $2E[X_R^2/NT] = \sigma_x^2$, что и требовалось доказать.

6.18. Докажите, что модуль $|S(f) + a \exp[-j2\pi ft]|$ содержит синусоидальную составляющую.

Ответ:

$$\begin{aligned} |S(f) + a e^{-j\omega t}|^2 &= |S(f)|^2 + a^2 + 2\operatorname{Re}[a S(f) e^{-j\omega t}] = \\ &= |S(f)|^2 + a^2 + 2a \{\operatorname{Re}[S(f) \cos(\omega t)] + \operatorname{Im}[S(f) \sin(\omega t)]\}. \end{aligned}$$

Глава 7

КОВАРИАЦИЯ И СВЕРТКА

7.1. ЭЛЕМЕНТЫ ТЕОРИИ

Содержание этой главы составляют вопросы, связанные с ковариационными функциями и сверткой. Несмотря на то что во многих приложениях ковариация и свертка играют различные роли, мы рассмотрим их вместе, поскольку в вычислительном плане получение их мало чем отличается.

Хотя в анализе временных рядов термины «корреляционная функция» и «ковариационная функция» зачастую выступают как синонимы, в этой главе мы будем их разграничивать и пользоваться более устоявшимся в статистике термином ковариация, оставляя за корреляцией понятие нормированной величины со значениями в отрезке $[-1, 1]$. Далее будем придерживаться статистической терминологии.

Ковариационные (корреляционные) функции находят широкое применение в анализе вибраций и ударов, в задачах, связанных с передачей информации, радарным обнаружением, системами управления, а также в анализе данных океанографических исследований, словом, во всех областях, где уместно применение соотношений спектрального анализа, использующих преобразования Фурье. Если в задаче для анализа используются спектр мощности и кросс-спектр, то соответствующие корреляционные функции в принципе также могут применяться, поскольку они тесно связаны со своими преобразованиями Фурье. На практике могут существовать веские причины для выбора той или иной возможности. Например, время относительного смещения для двух сигналов можно определять по фазе кросс-спектра, но нередко более целесообразным в этом случае оказывается применение (кросс-) ковариационной функции. Уточнение «кросс» для ковариации излишне, однако мы часто прибегаем к этому, чтобы не путать ее с автоковариацией.

Можно говорить и о столь же широком применении функции свертки. Следует, впрочем, отметить, что большую часть таких применений можно отнести к цифровой фильтрации, поскольку свертка временной последовательности с импульсной функцией отклика линейного фильтра дает отклик (выход) фильтра. Поэтому функция свертки полезна при решении таких задач, как сглаживание (низкополосная фильтрация) данных и генерирование

случайного процесса с заданной формой спектра при моделировании и выделении сигналов (подгоночный фильтр). Ряд таких применений будет продемонстрирован в § 7.9.

Ковариация. Ковариационная функция имеет двойное применение. Во-первых, с ее помощью определяется степень «сходства» двух временных последовательностей при различных значениях сдвига по времени (кросс-ковариация). Во-вторых, она используется для определения степени корреляции временной функции с самой собой в зависимости от величины времени задержки (автоковариация). В качестве оценки кросс-ковариации для последовательностей $x(i)$ и $y(i)$ ($i = 0, 1, \dots, N-1$) используется величина

$$s_{xy}(i) = \frac{1}{N-i} \sum_{p=0}^{N-i-1} x(p) y(p+i),$$

где $i = -m, \dots, -1, 0, 1, \dots, m$. При этом всегда предполагается, что до вычисления оценки из значений последовательностей вычитаются соответствующие выборочные средние. Индекс i называют «запаздыванием», а величину m — «максимальным запаздыванием». Если T — интервал, с которым производится выборка, то величина $\tau_{\max} = mT$ будет соответствующим максимальным запаздыванием (или, иначе, задержкой) по времени. На практике, как правило, редко используются значения задержки, превышающие 10% всей длины записи. Если $y(i) = x(i)$, то $s_{xx}(i)$ называют автоковариационной функцией и, в силу ее симметричности относительно 0, вычисляют лишь для положительных значений запаздывания. Приведем несколько элементарных фактов, относящихся к ковариационной функции. Справедливы равенства

$$s_{xy}(-i) = s_{yx}(i), \quad s_{xx}(-i) = s_{xx}(i), \quad s_{xx}(0) = s_x.$$

Функция

$$s_{xy}(i) + s_{xy}(-i) = A_{xy}(i) \quad (7.1)$$

является четной, а функция $s_{xy}(i) - s_{xy}(-i) = B_{xy}(i)$ — нечетной. Для случайных данных $s_{xx}(i) \rightarrow 0$ и $s_{xy}(i) \rightarrow 0$ при $i \rightarrow \infty$.

Напомним, что для корреляций выполняется неравенство

$$|\rho_{xy}(i)| = \left| \frac{\sigma_{xy}(i)}{\sigma_x \sigma_y} \right| \leq 1, \quad (7.2)$$

где $\sigma_{xy}(i)$ есть «истинная» ковариационная функция. Важно отметить, что для ограниченной длины наблюдаемого ряда выборочная корреляция

$$r_{xy}(i) = \frac{s_{xy}(i)}{s_x s_y}$$

не обязательно ограничена единицей, как этого можно было бы ожидать из (7.2). Особенно вероятно нарушение такого ограничения при $i > N/10$. Другим несколько измененным определением выборочной кросс-корреляционной функции служит следующее:

$$s'_{xy}(i) = \frac{1}{N} \sum_{p=0}^{N-1} x(p) y(p+i). \quad (7.3)$$

Эта оценка кросс-ковариации $\sigma_{xy}(i)$ имеет несколько меньшую, чем $s_{xy}(i)$, дисперсию (согласно Парзену (1961)) и всегда удовлетворяет неравенству (7.2). Вместе с тем это смещенная оценка, причем смещение может достигать неприемлемо больших значений при умеренно больших величинах задержки i . Вследствие несмещенности оценка $s_{xy}(i)$ более предпочтительна, так как при разумных значениях i ее дисперсия ненамного больше, чем у (7.3). Как будет показано ниже, это очень важно для детерминированных данных (см. § 7.9).

Получение корреляционных функций стало экономичным с вычислительной точки зрения после открытия БПФ. Из дискретного варианта теоремы Винера — Хинчина вытекает, что

$$S_{xy}(k) = T \sum_{i=0}^{N-1} s_{xy}(i) \exp\left(-j \frac{2\pi ik}{N}\right), \quad k = 0, 1, \dots, N-1,$$

$$s_{xy}(i) = b \sum_{k=0}^{N-1} S_{xy}(k) \exp\left(j \frac{2\pi ik}{N}\right), \quad i = 0, 1, \dots, N-1.$$

Следовательно, кросс-ковариацию (автоковариацию) можно определить по кросс-спектру (спектру мощности), вычисленному приближенно. Этот вопрос будет затронут в § 7.4.

Почти во всех приложениях кросс-ковариационной функции определяется время относительной задержки. Это связано с тем, что для одинаковых последовательностей $y(i)$ и $x(i)$, сдвинутых друг относительно друга на \hat{i} единиц, оценка $s_{xy}(i)$ достигает максимума при значении \hat{i} , равном величине сдвига. Оказывается, что в некоторых случаях лучшие оценки для времени относительной задержки дает функция свертки. Позже мы еще вернемся к этому вопросу.

Допустим, что последовательность $y(i)$ есть запаздывающая версия последовательности $x(i)$ плюс посторонний шум, т. е.

$$y(i) = x(i+k) + n(i).$$

Тогда ковариационная функция имеет вид

$$s_{xy}(i) = \frac{1}{N-i} \sum_{p=0}^{N-1-i} [x(p) x(p+k+i) + x(p) n(p+i)] =$$

$$= s_{xx}(k+i) + s_{xn}(i).$$

Поскольку среднее оценки $s_{xn}(i)$ равно 0, кросс-ковариация есть сдвинутая по времени автоковариация, которая искажена кросс-ковариацией последовательности $x(i)$ и шума $n(i)$. Эта кросс-ковариация при больших N стремится к нулю. Поэтому, исследуя $s_{xy}(i)$ на максимум, можно получить значение запаздывания по времени, соответствующее величине задержки. В § 7.9 мы увидим, какие практические задачи возникают в этом случае.

Помимо задач определения задержки кросс-корреляционные (или ковариационные) функции находят применение в задачах установления эквивалентности временных последовательностей. Максимум значений нормированной кросс-корреляционной функции дает количественный показатель, отражающий степень «сходства» двух рядов данных. При этом для двух одинаковых временных функций, сдвинутых друг относительно друга, максимальное значение $\rho_{xy}(i)$ равно единице. Для функций, совершенно независимых друг от друга (некоррелированных), значение $\rho_{xy}(i)$ равно 0. Величина $s_e = \sqrt{1 - \rho_{xy}^2(i)_{\max}}$ служит мерой «согласия» между $x(i)$ и $y(i)$ и представляет собой среднеквадратичную величину разности, которую обычно называют «стандартной ошибкой» или «стандартной ошибкой оценки».

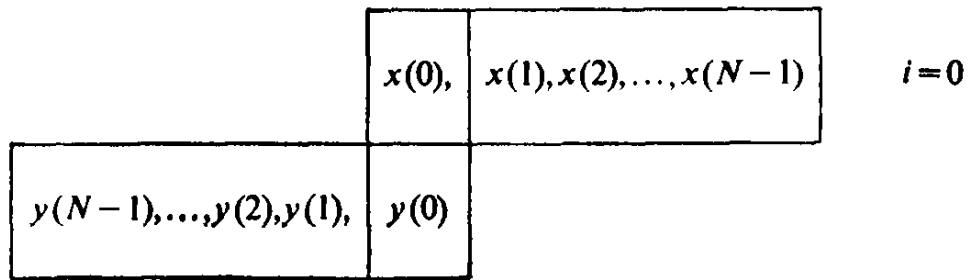
Автоковариационные функции и их преобразование Фурье, т. е. спектры мощности, тесно связаны и несут в принципе одну и ту же информацию. От плотности спектра мощности, допускающей, как правило, простую интерпретацию в инженерных приложениях, иногда, стремясь получить более наглядное графическое изображение результатов, переходят к автокорреляции. Например, к автокорреляционной функции обращаются, когда нужно определить коэффициенты для уравнения прогнозирования (т. е. для экстраполяции будущих значений временных последовательностей по их прошлым значениям).

Свертка. Формула для дискретной быстрой свертки¹⁾ отличается от формулы для ковариации очень немногим. Все различия сводятся лишь к отсутствию усреднения и изменению знака в одном временном индексе:

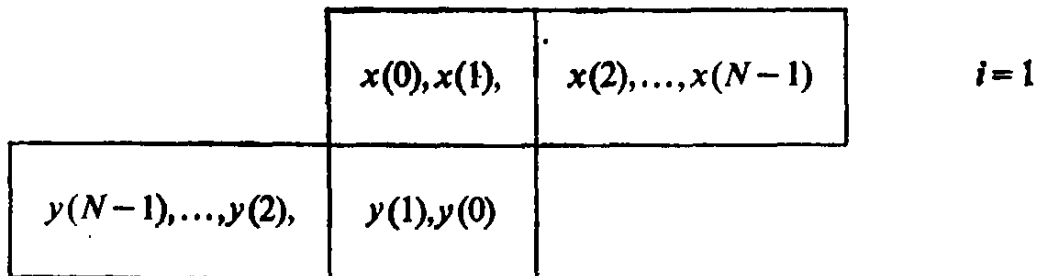
$$c_{xy}^l(i) = \begin{cases} T \sum_{p=0}^i x(p) y(i-p), & i=0, 1, \dots, N-1, \\ T \sum_{p=1}^{N-l-1} x(p+l) y(N-p), & i=l+N, l=0, 1, \dots, N-1. \end{cases} \quad (7.4)$$

¹⁾ Мы ввели этот термин (transient), чтобы подчеркнуть различие с круговой сверткой. Это понадобится в § 7.5.

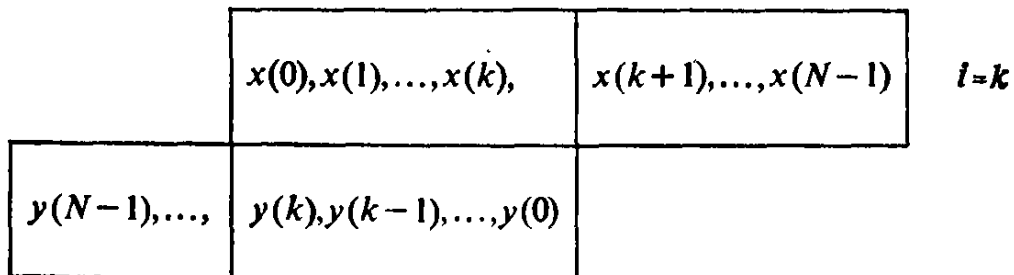
Проиллюстрируем эту формулу диаграммами



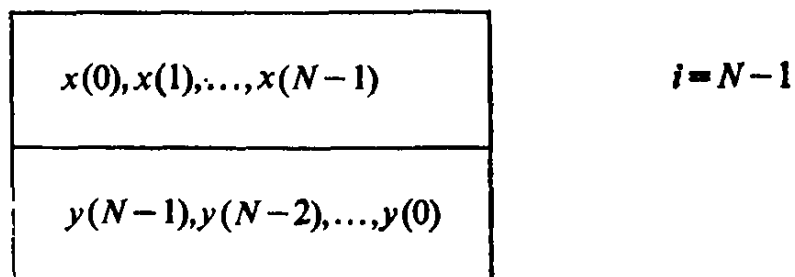
$$c'_{xy}(0) = x(0)y(0)$$



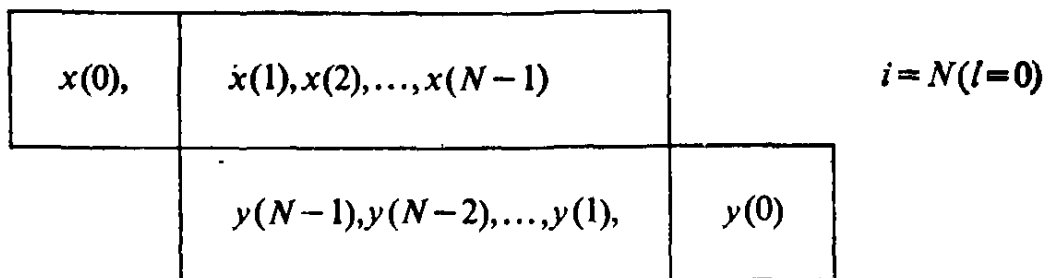
$$c'_{xy}(1) = x(0)y(1) + x(1)y(0)$$



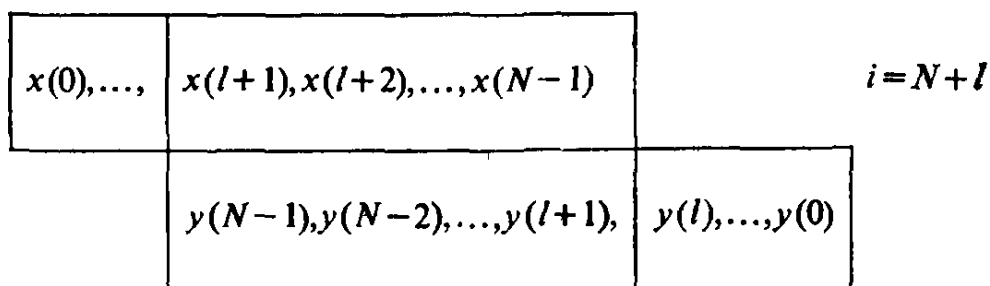
$$c'_{xy}(k) = x(0)y(k) + x(1)y(k-1) + \dots + x(k)y(0)$$



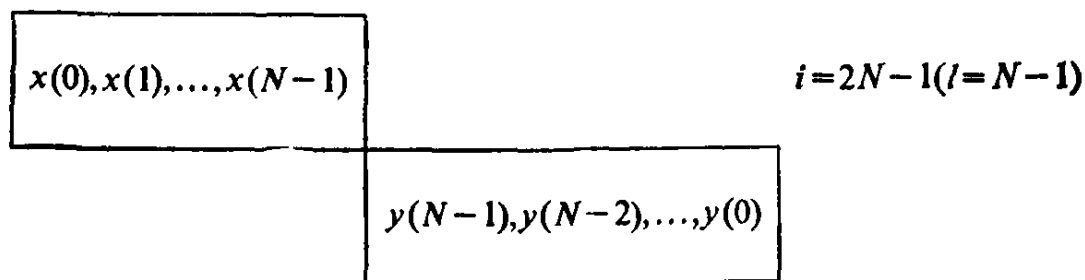
$$c'_{xy}(N-1) = x(0)y(N-1) + x(1)y(N-2) + \dots + x(N-1)y(0)$$



$$c_{xy}^l(N) = x(1)y(N-1) + x(2)y(N-2) + \dots + x(N-1)y(1)$$



$$c_{xy}^l(N+l) = x(l+1)y(N-1) + x(l+2)y(N-2) + \dots + x(N-1)y(l+1)$$



$$c_{xy}^l(2N-1) = 0$$

На свертку можно смотреть как на перестановку значений функции y в обратном порядке в последующем вычислении кросс-ковариации.

Наличие индекса l в формуле для *быстрой* свертки отражает тот факт, что появление быстрого преобразования Фурье очень ускорило вычисление дискретной свертки. Метод ее вычисления основан на *теореме о свертке* (входящей в теорию преобразования Фурье) и сводится к вычислению быстрого обратного преобразования Фурье произведения быстрых преобразований Фурье исходных последовательностей.

Для анализа данных временных рядов свертка имеет фундаментальное значение. Это связано в первую очередь с тем, что любая линейная система полностью характеризуется своей импульсной функцией отклика (см. гл. 1). Более того, отклик линейной системы при любом входе определяется сверткой с импульсной функ-

цией отклика. Поскольку цифровые фильтры представляют собой линейные системы, цифровую фильтрацию можно осуществлять при помощи свертки. В гл. 4 уже затрагивался вопрос о реализации нерекурсивных (ИКО) фильтров с помощью свертки.

Любую операцию сглаживания данных можно представлять как низкочастотную фильтрацию и осуществлять ее при помощи свертки. Скользящее усреднение, часто применяемое для сглаживания, задается формулой

$$\bar{x}(i) = \frac{1}{2M+1} \sum_{j=-M}^M x(i-j), \quad i = M, M+1, \dots, N-M-1. \quad (7.5)$$

(Замечание: значения выражения (7.5) в первых M и последних M точках иногда определяются в предположении, что значения данных, лежащих вне отрезка $[0, N-1]$, равны 0.)

Отметим, что величину $\bar{x}(i)$ можно рассматривать как свертку последовательности $x(i)$ с константой $a(i) = (2M+1)^{-1}$ для всех i . Придерживаясь такого подхода, можно дать следующее обобщение операции скользящего усреднения:

$$\bar{x}(i) = T \sum_{j=-M}^M a(j) x(i-j), \quad (7.6)$$

где

$$T \sum_{j=-M}^M a(j) = 1.$$

В § 8.8 для операции, заданной формулой (7.5), будет получена частотная функция отклика; это позволит оценить недостатки такого усреднения.

Соотношение (7.6) приводит к свертке специального вида. Ее называют *симметричной сверткой* и определяют формулами

$$\begin{aligned} c_{ax}(i) &= T \sum_{j=-M}^M a(j) x(i-j) = \\ &= Ta(0)x(i) + T \sum_{j=1}^M a(j)[x(i-j) + x(i+j)], \end{aligned} \quad (7.7)$$

где $i = 0, 1, \dots, N-1$ и

$$a(j) = a(-j), \quad j = 0, 1, \dots, M.$$

Значения данных, попадающих при суммировании по формуле (7.7) за пределы отрезка $[0, N-1]$, полагаются равными нулю. Это позволяет расширить область определения симметричной свертки на точки $i = -M, -(M-1), \dots$. Поэтому в качестве области определения можно взять множество

$$i = -M, -(M-1), \dots, -1, 0, 1, \dots, N-1, N, N+1, \dots, N+M-1.$$

Вообще, если в свертке участвуют N_1 -точечная и N_2 -точечная функции, то результат свертки определяется на отрезке с $N_1 + N_2$ точками.

В гл. 8 нам встретится еще одно применение свертки. Оказывается, свертка в области частот позволяет получить такие модификации преобразований Фурье, которые ослабляют эффект утечки. Вообще говоря, свертка прочно вплетена в ткань большинства задач анализа цифровых временных рядов. Следует, впрочем, отметить, что свертка как объект интерпретации редко представляет самостоятельный интерес и, как правило, отдельно не вычисляется.

7.2. РАЗЛИЧИЯ МЕЖДУ КОВАРИАЦИЕЙ И СВЕРТКОЙ

Вычисление ковариации включает операцию усреднения, поскольку ковариация представляет собой среднее кросс-произведения. Операция свертки усреднения не требует. Так как дискретные варианты ковариации и свертки должны быть аналогами соответствующих интегралов, следует позаботиться о тех постоянных множителях, которые будут играть ту же роль, что и dt в интегралах. Для ковариации этот вопрос, в силу усреднения, решается автоматически, поскольку $T/(NT) = 1/N$. Что же касается дискретной свертки, то в ее уравнение приходится специально вводить T — величину интервала, с которым производится выборка.

Хотя оба кросс-произведения для свертки и ковариации выглядят почти одинаково, они имеют одно коренное отличие. Временной индекс со знаком минус для $y(i)$ в верхней строке формулы (7.4) соответствует перестановке значений в обратном порядке.

Таким образом, если при подходящем масштабном множителе переставить значения $y(i)$ в обратном порядке и вычислить ковариационную функцию двух временных последовательностей, то получится свертка этих временных последовательностей. Введем обозначение $y'(i) = y(-i)$ и запишем соответствующие уравнения:

$$s_{xy'}(i) = \begin{cases} \frac{1}{N-i} \sum_{p=0}^{N-1} x(p) y'(i+p), & \text{положительные запаздывания} \\ & i = 0, 1, \dots, N-1, \\ \frac{1}{N-i} c_{xy}^i(i), & i = N-1, N-2, \dots, 0, \end{cases}$$

$$s_{y'x}(i) = \begin{cases} \frac{1}{N-i} \sum_{p=0}^{N-1} y'(p) x(i+p), & \text{отрицательные запаздывания} \\ & i = 1, 2, \dots, N-1, \\ \frac{1}{N-i} c_{xy}^i(i), & i = N, N+1, \dots, 2N-1. \end{cases}$$

Заметим, что получено выражение быстрой свертки, состоящее из двух частей, причем первое множество значений имеет обратный порядок, а второе — прямой. Кроме того, в отрицательные запаздывания не включен нуль, чтобы не повторялось значение свертки, расположенное в середине. Позднее, в § 7.5 будет показано, как упрощаются соответствующие выражения при переходе к круговой свертке с добавлением нулей.

7.3. ДЛИННЫЕ РЯДЫ И ОСНОВНАЯ ВЫЧИСЛИТЕЛЬНАЯ ПРОЦЕДУРА ДЛЯ КОВАРИАЦИИ

Уравнение ковариации обычно записывают отдельно для положительных и отрицательных запаздываний:

$$s_{xy}(i) = \frac{1}{N-i} \sum_{p=0}^{N-1} x(p) y(p+i), \quad s_{yx}(i) = \frac{1}{N-i} \sum_{p=0}^{N-1} y(p) x(p+i),$$

где $i = 0, 1, \dots, m$. Отметим, что нулевое запаздывание во второе уравнение можно не включать.

Выражение для ковариации хорошо приспособлено для расчетов при больших значениях N . Общий объем памяти, нужный для вычисления свертки, составляет $4(m+1)$ ячеек для исходных данных и $2(m+1)$ ячеек для накопления кросс-произведений. Резервируем четыре блока памяти по $m+1$ ячеек в каждом. Последовательность действий при вычислениях будет следующей.

1. Считывание $m+1$ начальных значений данных каждой из последовательностей $x(i)$ и $y(i)$ в два первых блока.

2. Накопление (т. е. суммирование) $m+1$ кросс-произведений для запаздывания 0, $m+1-1$ для запаздывания 1, ..., и, наконец, 1 для запаздывания m .

3. Считывание следующих $m+1$ значений данных в два оставшихся блока.

4. Накопление $m+1$ кросс-произведений для всех значений запаздывания.

5. Считывание следующих $m+1$ значений данных в первую пару блоков.

6. Накопление $m+1$ кросс-произведений для всех значений запаздывания.

7. Повторение шагов 3—6 до тех пор, пока не будет считан последний блок.

8. Накопление одного кросс-произведения для запаздывания 0, двух для запаздывания 1, ..., и, наконец, $m+1$ для запаздывания m .

9. Деление накопленных сумм кросс-произведений на $N-i$. Этим вычисления завершаются.

Однобитовое квантование или метод срезания экстремальных значений. Определим по стандартному гауссовскому процессу $x(t)$ с нулевым средним процесс

$$y(t) = \operatorname{sgn} x(t) = \begin{cases} 1, & x(t) \geq 0, \\ -1, & x(t) < 0. \end{cases}$$

Коэффициент корреляции $\rho_y(\tau)$ процесса $y(t)$, полученного срезанием экстремальных значений исходного процесса, связан с коэффициентом корреляции $\rho_x(\tau)$ процесса $x(t)$ соотношением

$$\rho_x(\tau) = \sin \left[\frac{\pi}{2} \rho_y(\tau) \right], \quad (7.8)$$

где $\rho_x(\tau)$ и $\rho_y(\tau)$ определяются отношениями

$$\rho_x(\tau) = \frac{s_x(\tau)}{s_x(0)}, \quad \rho_y(\tau) = \frac{s_y(\tau)}{s_y(0)}.$$

Вывод соотношения (7.8) можно найти в работе Уэйнреба (1963); теоретическим вопросам оценивания спектра посвящена работа Хинича (1967). Впрочем, начало исследований по этим вопросам следует отнести, согласно Кендаллу и Стьюарту (1973), к 1898 г. Очевидно, однобитовое квантование сильно упрощает необходимые при умножениях операции и сводит их к присвоению единице знаков плюс и минус. Такое упрощение вычислений нередко приносит существенную выгоду.

С применением метода срезания для вычисления корреляций связаны определенные проблемы. Одной из таких проблем является, например, сохранение неизменной статистической точности. Чтобы обеспечить постоянную точность, приходится ускорение вычислений компенсировать за счет роста длины ряда. Как показал Уэйнреб (1963), срезанию сопутствует рост дисперсии оценки корреляционной функции примерно в 2.5 раза ($\approx \pi^2/4$). Эта дисперсия пропорциональна P , поэтому чтобы дисперсия оставалась постоянной, нужно увеличить длину ряда в два с половиной раза по сравнению с тем случаем, когда используется вся информация, содержащаяся в данных. В работе Хинича (1967) получено обобщение этого результата, позволяющее применить его для вычислений спектра. Как показывают теоретические результаты Хинича, наихудшим является случай широкополосных

данных (некоррелированных данных). Для узкополосных данных наблюдается меньший рост дисперсии. В типичных случаях упомянутый множитель роста уменьшается с 2.5 до 1.1. Это вполне объяснимо, так как узкополосные данные имеют большое сходство с синусоидой. Крайний случай — синусоида — дает при вычислении ковариации нулевую статистическую изменчивость.

Вообще говоря, столь незначительный рост длины рядов при срезании демонстрирует хорошо известный специалистам по те-

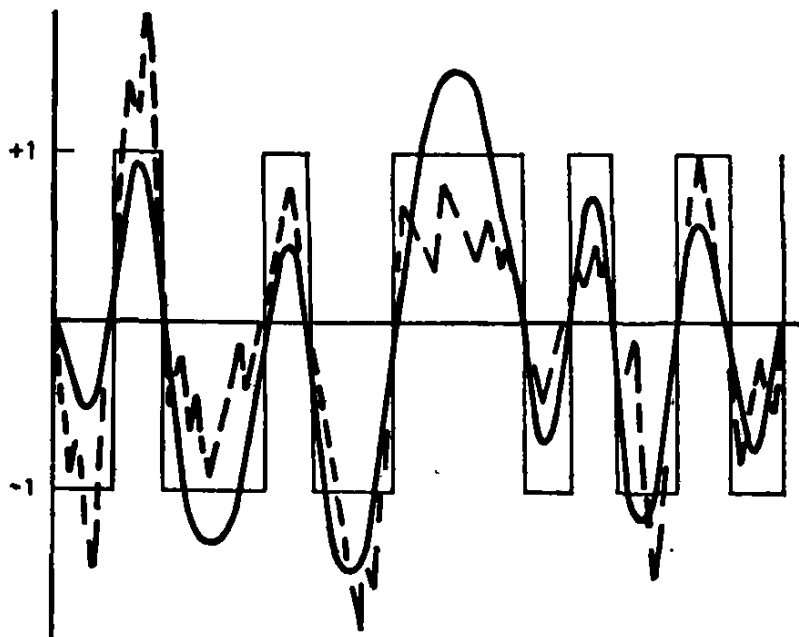


Рис. 7.1. Два временных ряда, имеющих одну и ту же версию, полученную методом срезания экстремальных значений.

ории случайных процессов факт: пересечения случайным процессом нулевого уровня несут огромное количество информации. Подчеркнем, что метод жесткого срезания применяется в трех случаях — гауссовского процесса, синусоиды и синусоиды, искаженной гауссовским шумом; для них справедливо соотношение (7.8). Заметим, что две функции (например, синусоида и ломаная), одинаково проходящие через нули, будут давать одну и ту же корреляционную функцию срезки (рис. 7.1). Поэтому, обращаясь к методу срезки для данных разных типов, следует проявлять определенную осторожность.

Методом жесткого срезания, как это вытекает из соотношения (7.8), можно получить непосредственно только нормированные корреляционные функции, поэтому для определения соответствующих масштабных множителей необходимо вычислять отдельно среднеквадратичное значение. Кроме того, при этом методе полностью исчезает возможность выявления нелинейных эффектов, из-за которых формы функций плотности вероятностей перестают отвечать гауссовскому случаю.

Сумма квадратов и метод четверти квадратов. В работе Шмида (1965) предложен технический прием, основанный на представлении произведения в виде суммы квадратов. В применении к нашим целям его способ дает возможность записывать кросс-произведения в виде линейной комбинации квадратов сомножителей

$$xy = 1/2 [(x + y)^2 - x^2 - y^2]. \quad (7.9)$$

Это соотношение очень похоже на достаточно широко известный метод «четверти квадратов», применяемый при построении аналоговых умножителей (Келли и др. (1966)). Работа этих устройств основана на соотношении

$$xy = 1/4 [(x + y)^2 - (x - y)^2]. \quad (7.10)$$

Приведенные соотношения используются в упр. 7.1 и 7.2, где с их помощью для ковариационных функций получена иная запись.

Использование таких соотношений имеет то преимущество, что у нас появляется возможность вычислять произведения с помощью таблицы, хранящейся в памяти ЭВМ. Таблица для квадратов двух множеств данных требует гораздо меньшего объема памяти, чем таблица для кросс-произведений того же числа данных. Если, например, при 10-битовом квантовании для таблицы всевозможных кросс-произведений потребовалось бы 2^{20} ячеек, то таблица для квадратов и сумм данных потребовала бы всего 2^{11} ячеек; разумеется, второй вариант более предпочтителен.

7.4. ВЫЧИСЛЕНИЕ КОВАРИАЦИИ И СВЕРТКИ ПРИ ПОМОЩИ БПФ

Функции ковариации и свертки можно вычислять по функции плотности кросс-спектра. Такой подход основан на соотношениях Винера—Хинчина. Поскольку для вычисления спектра необходимо определить преобразование Фурье временной последовательности, обратившись к БПФ, можно получить корреляционные функции. Оказывается, такой обходной путь приводит к цели в 5—100 раз быстрее в зависимости от того, какова максимальная величина запаздывания. Получить значения ковариационной функции для N запаздываний можно почти так же быстро, как и для числа запаздываний m , значительно меньшего N . Сэнд (1965) весьма подробно рассмотрел особенности этой процедуры.

Основной метод. Суть основного метода вычисления автоковариации сводится к следующему.

1. Сначала вычисляется преобразование $X(k)$ для временного ряда $x(i)$ ($i, k = 0, 1, \dots, N-1$).

2. После этого вычисляется «предварительный» спектр $\hat{S}_x(k) = (1/P) |X(k)|^2$.

3. Наконец, определяется автоковариационная функция

$$s_x(i) = \mathcal{F}^{-1}[\hat{S}_x(k)],$$

где \mathcal{F}^{-1} — быстрое обратное преобразование Фурье.

С этой процедурой связан целый ряд вопросов. Часть из них будет рассмотрена в настоящем параграфе. Подробное описание

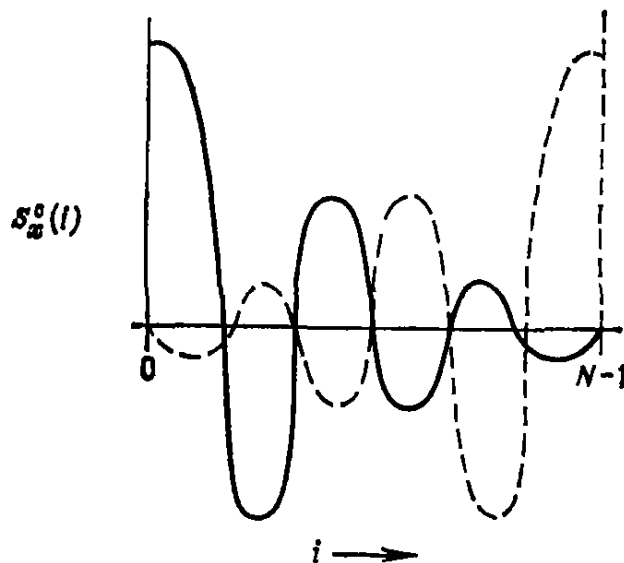


Рис. 7.2. Две части ковариационной функции, полученные без добавления нулей.

отдельных эффектов отложим до следующего параграфа. Все, что будет говориться о ковариационных функциях, в полной мере относится к тесно связанным с ними функциям свертки.

Начнем с наименее очевидной особенности предложенной процедуры. Как показал Сэнд (1965), обычно на 3-м шаге процедуры получается не ковариационная функция, а «круговая» автоковариационная функция $s_x^c(i)$, определяемая соотношением

$$s_x^c(i) = \frac{N-i}{N} s_x(i) + \frac{i}{N} s_x(N-i); \quad (7.11)$$

слагаемые правой части (без множителей) изображены на рис. 7.2.

Если к исходному ряду предварительно добавлять нули, то части круговой свертки будут расходиться. При добавлении N нулей они займут положение, показанное на рис. 7.3.

Последовательность действий процедуры нетрудно приспособить для вычисления некруговой ковариационной функции. Если исходная последовательность данных имеет число членов, равное степени 2, т. е. $N = 2^p$, то значения такой функции для всех N

запаздываний можно получить, добавляя N нулей. Если число членов исходной последовательности не равно степени 2, то дополнение ряда N_z нулями до первой степени 2 даст N_z неискаженных ковариаций для начальных запаздываний. В случае необходимости длину последовательности еще раз увеличивают, добавляя 2^p нулей. Приведем последовательность действий вы-

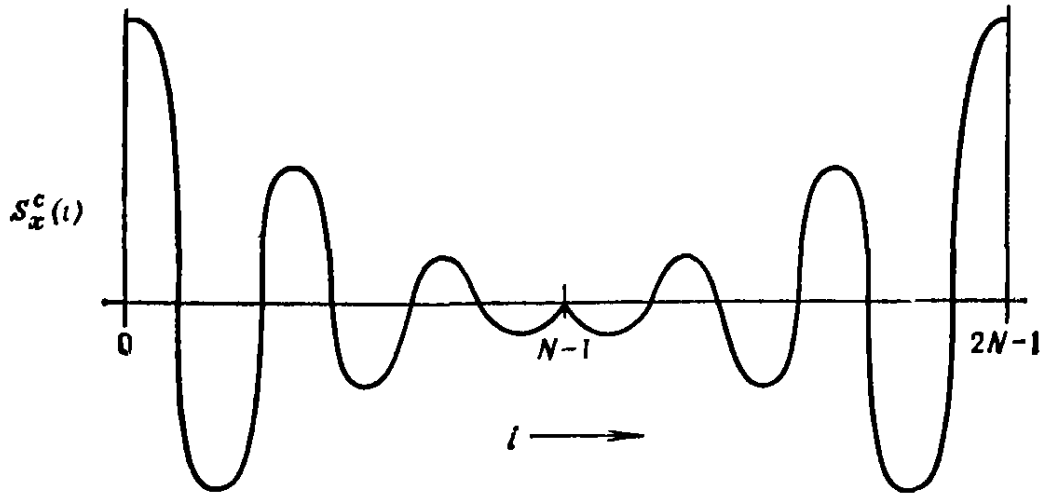


Рис. 7.3. Влияние добавления N нулей на круговую ковариационную функцию.

числения ковариационной функции, учитывающую рассмотренный эффект.

1. Исходный временной ряд $x'(i)$, $i = 0, 1, \dots, N-1$, дополняется N нулями. Новую последовательность будем обозначать $x(i)$, $i = 0, 1, \dots, 2N-1$.

2. Вычисляется $2N$ -точечное БПФ $X(k)$, $k = 0, 1, \dots, 2N-1$.

3. Для $k = 0, 1, \dots, 2N-1$ вычисляется «предварительный» спектр $\hat{S}_x(k) = (T/N) |X(k)|^2$.

4. Вычисляется быстрое обратное преобразование Фурье спектра $\hat{S}_x(k)$; чтобы получить правильные значения ковариации, первые N величин умножаются на $N/(N-i)$:

$$s_x(i) = \frac{N}{N-i} \mathcal{F}^{-1}[S_x(k)],$$

где $i = 0, 1, \dots, N-1$.

Для ковариационных функций общего вида вычисления проводятся так же, как и для автоковариации. В этом случае аналогом соотношения (7.11) служит формула

$$s_{xy}^c(i) = \frac{N-i}{N} s_{xy}(i) + \frac{i}{N} s_{xy}(i).$$

Подробное изложение метода для вычисления ковариации и свертки. При вычислении ковариационной функции и функции свертки используется произведение двух преобразований Фурье.

Поэтому в отличие от основного метода БПФ вычисляется для двух рядов. Вычислительная процедура состоит из следующих шагов.

1. Ряд $x(i)$ запоминается в массиве, отведенном для действительных частей, а ряд $y(i)$ — для мнимых, т. е. $z(i) = x(i) + jy(i)$, где $i = 0, 1, \dots, N-1$.

2. Вычисляется N -точечное быстрое преобразование Фурье ряда $z(i)$ (преобразование обозначим $Z^e(k)$):

$$Z^e(k) = \text{БПФ} [z(i)], \quad k = 0, 1, \dots, N-1.$$

3. Величины $z(i)$, $i = 0, 1, \dots, N-1$, считываются и умножаются на $W_{2N}^{i/2}$.

4. Вычисляется N -точечное БПФ последовательности $z(i)W_{2N}^{i/2}$ (обозначим это преобразование $Z^o(k)$):

$$Z^o(k) = \text{БПФ} [z(i) W_{2N}^{i/2}], \quad k = 0, 1, \dots, N-1.$$

5. Полное $2N$ -точечное БПФ получается по формулам

$$Z(2k) = Z^e(k), \quad Z(2k+1) = Z^o(k),$$

где $k = 0, 1, \dots, N-1$.

6. Для ковариационной функции вычисляется «предварительный» кросс-спектр ($k = 0, 1, \dots, N$)

$$\begin{aligned} \bar{S}_{xy}(k) &= \frac{1}{P} X^*(k) Y(k) = \\ &= \frac{1}{P} \left[\frac{Z(k) + Z^*(2N-k)}{2} \right]^* \left[\frac{Z(k) - Z^*(2N-k)}{2j} \right]. \end{aligned}$$

Для функции свертки вычисляется произведение двух преобразований Фурье ($k = 0, 1, \dots, N$)

$$\begin{aligned} C_{xy}(k) &= X(k) Y(k) = \\ &= \left[\frac{Z(k) + Z^*(2N-k)}{2j} \right] \left[\frac{Z(k) - Z^*(2N-k)}{2j} \right]. \end{aligned}$$

7. Первая половина значений кросс-спектра — функции, симметричной по Эрмиту, — определяется формулами

$$\bar{S}_{xy}(2k) = S_{xy}^e(k), \quad \bar{S}_{xy}(2k+1) = S_{xy}^o(k),$$

где $k = 0, 1, \dots, (N/2) - 1$.

8. Используя обратное преобразование, введенное в гл. 6, для N -точечной некруговой свертки или для ковариационной функции, вычисляют величины

$$S'_{xy}(k) = \bar{S}_{xy}(k) + \bar{S}_{xy}^*(N-k) + j [\bar{S}_{xy}(k) - \bar{S}_{xy}^*(N-k)] W_{2N}^{-k/2},$$

где $k = 0, 1, \dots, N-1$. Для экономии памяти вычисления необходимо проводить одновременно для k и $(N-k)$.

9. Вычисляется N -точечное быстрое обратное преобразование Фурье последовательности $S'_{xy}(k)$:

$$s'_{xy}(i) = \text{БПФ}^{-1}[S'_{xy}(k)], \quad i = 0, 1, \dots, N-1.$$

10. На этом последнем шаге определяется $2N$ -точечная некруговая свертка или некруговая ковариационная функция:

$$s_{xy}(2i) = \frac{N}{N-i} \operatorname{Re}[s'_{xy}(i)], \quad s_{xy}(2i+1) = \frac{N}{N-i} \operatorname{Im}[s'_{xy}(i)].$$

Отметим, что расположение комплексных чисел при использовании программы на Фортране позволяет при вычислении $s_{xy}(i)$ избегать какого-либо переупорядочения в памяти ЭВМ.

Только что предложенный метод требует объема памяти, необходимого для запоминания входной последовательности. Этот метод можно упростить, дополняя последовательность данных нулями и используя прямое и обратное преобразования двойной длины. Однако это потребует больших затрат времени и памяти.

При добавлении в последовательности нулей можно расположить данные в ином порядке. Например, располагая данные в начале каждой последовательности, т. е.

$$\begin{aligned} &x(0), x(1), \dots, x(N-1), 0(N), 0(N+1), \dots, 0(2N-1), \\ &y(0), y(1), \dots, y(N-1), 0(N), 0(N+1), \dots, 0(2N-1), \end{aligned}$$

для ковариации, вычисленной при помощи быстрых преобразований Фурье, получим последовательность

$$s_{xy}(0), s_{xy}(1), \dots, s_{xy}(N-1), s_{yx}(N), s_{yx}(N-1), \dots, s_{xy}(1).$$

Расположив те же данные так, чтобы последовательность x шла с начала, а последовательность y — с середины дополненной последовательности, т. е.

$$\begin{aligned} &x(0), x(1), \dots, x(N-1), 0(N), 0(N+1), \dots, 0(2N+1), \\ &0(0), 0(1), \dots, 0(N-1), y(0), y(1), \dots, y(N-1), \end{aligned}$$

получим для ковариации последовательность значений

$$s_{yx}(N), s_{yx}(N-1), \dots, s_{yx}(1), s_{xy}(0), s_{xy}(1), \dots, s_{xy}(N-1),$$

которая позволяет сразу получить ковариационные функции для i от $-N$ до $N-1$, поскольку $s_{yx}(i) = s_{xy}(-i)$.

Такого рода временные сдвиги можно получать циклической перестановкой в области частот, что достигается умножением $S_{xy}(k)$ на экспоненту

$$W_N^{kpN} = \exp\left[-j \frac{2\pi k}{N} (pN)\right] = \exp(-j2\pi kp),$$

где $p = i/N$ — доля общего числа запаздываний, для которых строится график. Если $p = 1$ ($i = N$), то на графике сначала

откладываются значения $s_{xy}(i)$ для отрицательных запаздываний, а затем — значения для положительных запаздываний; величина $s_{xy}(0)$ при этом окажется в центре $((N+1)$ -я точка). Если $\rho = 0.1$, то сначала появятся $0.1N$ значений для отрицательных запаздываний, потом — все значения для положительных запаздываний, а за ними последуют оставшиеся $0.9N$ значений для отрицательных запаздываний. Величину $\rho = i/N$ следует выби-

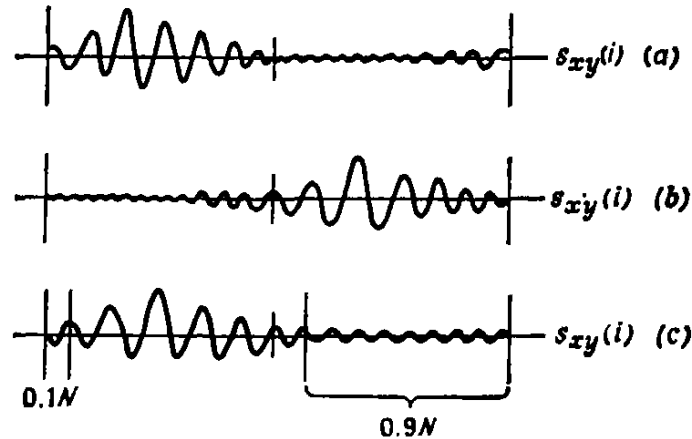


Рис. 7.4. Иллюстрация перестановки значений ковариационной функции для графического изображения. *a* — результаты в естественном порядке; *b* — после циклического сдвига на N точек (умножение $S_{xy}(k)$ на W_N^{-kN}); *c* — после циклического сдвига на $0.1N$ точек (умножение $S_{xy}(k)$ на $W_N^{-k0.1N}$).

рать так чтобы она была целой. На рис. 7.4 и 7.5 показаны циклические перемещения значений ковариаций при $\rho = 1$ и $\rho = 0.1$.

Вкратце способ получения свертки и ковариации при помощи преобразования Фурье можно сформулировать следующим образом.

Для свертки вычисления проводятся по формулам

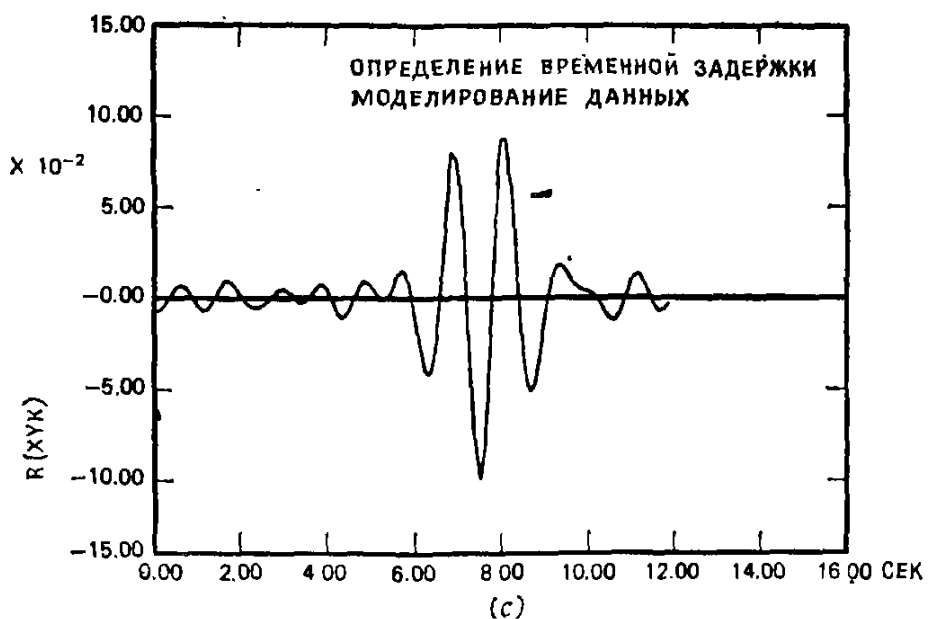
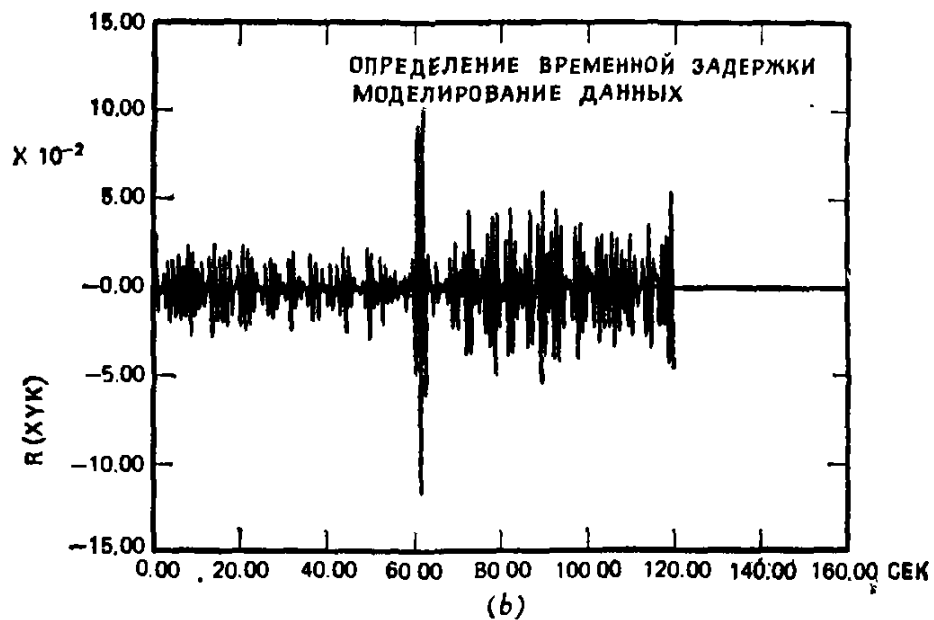
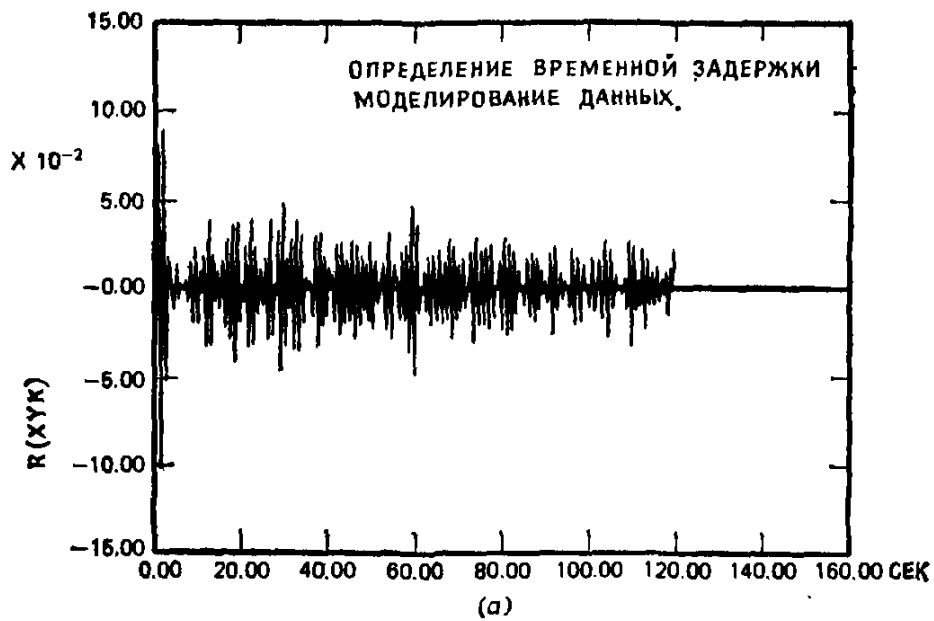
$$c_{xy}(i) = T \sum_{p=0}^{N-1} x(p) y(p-i), \quad i = 0, 1, \dots, N-1,$$

$$c_{xy}(i) = \text{БПФ}^{-1} [C_{xy}(k) = X(k) Y(k)],$$

а для ковариации — по формулам

$$s_{xy}(i) = \frac{1}{N-i} \sum_{p=0}^{N-i-1} x(p) y(p+i), \quad i = 0, 1, \dots, N-1,$$

$$s_{xy}(i) = \begin{cases} \frac{N}{N-i} \text{БПФ}^{-1} \left[S_{xy}(k) = \frac{1}{P} X^*(k) Y(k) \right], & i = 0, 1, \dots, N-1, \\ \frac{N}{i-N} \text{БПФ}^{-1} \left[S_{xy}(k) = \frac{1}{P} X^*(k) Y(k) \right], & i = N+1, \dots \\ & \dots, 2N-1. \end{cases}$$



7.5. Иллюстрация циклических сдвигов свертки. *a* — последовательность в естественном порядке; *b* — циклический сдвиг на N точек; *c* — циклический сдвиг на $0,1 N$ точек (первые $0,2 N$ точек графика).

В обоих случаях предполагается, что в конце временных последовательностей добавлены нули; это должно защищать от круговых эффектов, не позволяющих получать верные значения для быстрой ковариационной функции и свертки.

7.5. ЭФФЕКТЫ ПОДМЕНЫ И НАЛОЖЕНИЯ

В настоящее время, когда достигнуто полное понимание механизма возникновения кругового эффекта при вычислении свертки посредством БПФ, интересно вернуться к тому времени, когда с этим эффектом столкнулись впервые — при использовании аналоговых корреляционных анализаторов поворотного-барабанных типов. С их помощью получалась свертка вида

$$c_{xy}(i) = \sum_{p=0}^{N-1} x(p) y[(i-p) \bmod N], \quad i = 0, 1, \dots, N-1, \quad (7.12)$$

где запись $(i-p) \bmod N$ означает, что значение индекса $(i-p)$ берется по модулю N (т. е. представляет собой остаток деления $(i-p)$ на N). Свертка (7.12) включает в себя быструю свертку $c_{xy}^t(i)$; оставшуюся часть суммы назовем круговой частью свертки (7.12) и обозначим $c_{xy}^c(i)$ ¹). Теперь свертку (7.12) можно записать в виде

$$c_{xy}(i) = c_{xy}^t(i) + c_{xy}^c(i), \quad i = 0, 1, \dots, N-1, \quad (7.13)$$

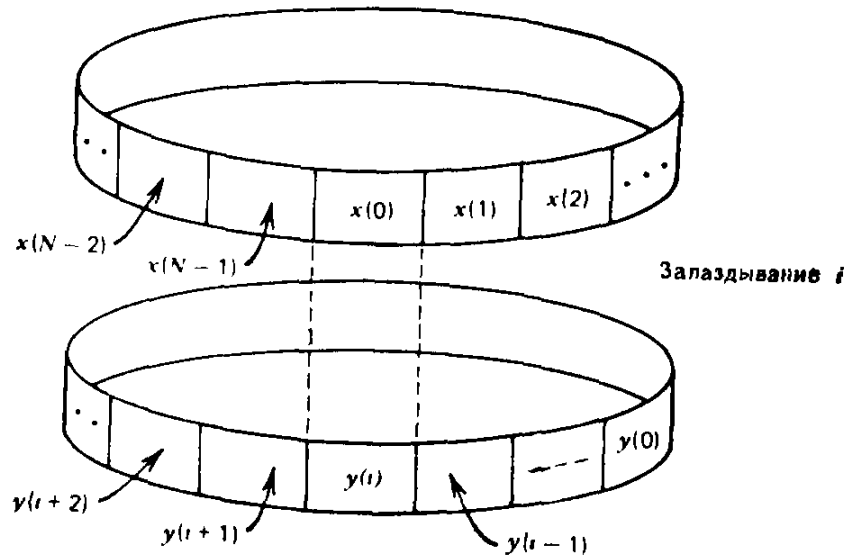
т. е.

$$c_{xy}(i) = \sum_{p=0}^i x(p) y(i-p) + \sum_{p=0}^{N-i-2} x(p+i+1) y(N-p-1),$$

где $i = 0, 1, \dots, N-1$.

Поясним, почему свертку $c_{xy}^c(i)$ называют круговой. Мысленно можно представлять себе, что данные, по которым вычисляется свертка (7.12), расположены на барабанах. С подобной ситуацией мы уже встречались в гл. 1. Расположение данных на барабанах

¹) Заметим, что $c_{xy}^c(i) = c_{xy}^t(N+i)$; это видно из формулы (7.4) при $T=1$. — Прим. перев.



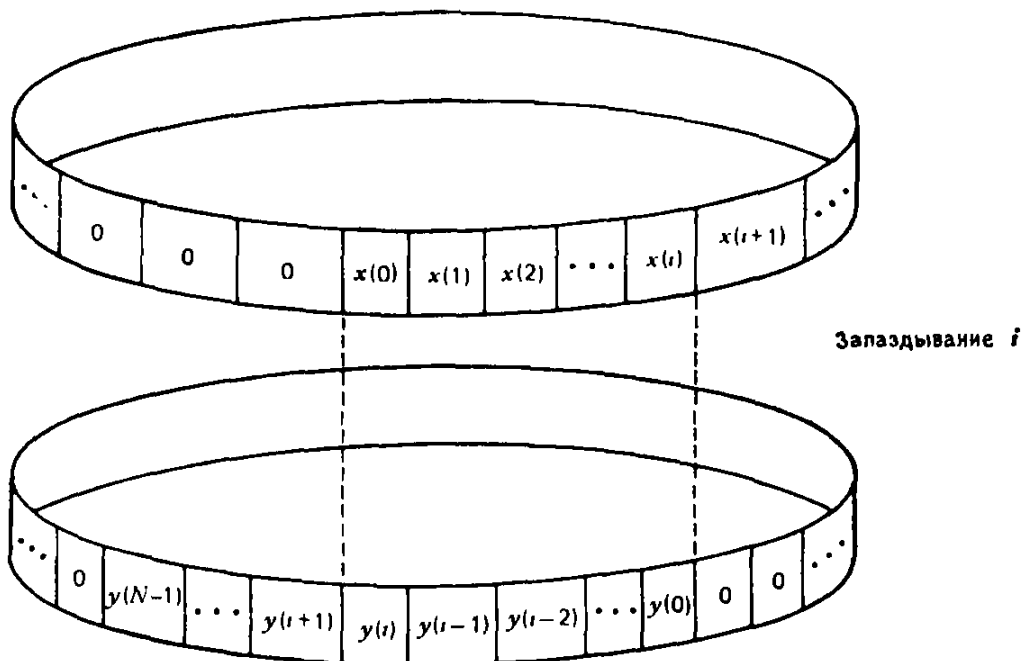
нагляднее показывает, как сумма (7.13) разбивается на сегменты

$$x(0)y(i) + x(1)y(i-1) + \dots + x(i)y(0)$$

и

$$x(i+1)y(N-1) + \dots + x(N-1)y(i+1).$$

От кругового эффекта можно избавиться, определяя две новые временные последовательности. Эти последовательности в отличие от исходных дополняются N нулями каждая. Свертка (7.12) для таких $2N$ -точечных последовательностей будет иметь в разложении (7.13) нулевую круговую часть. Проиллюстрируем этот эффект диаграммой



Как отмечалось в гл. 6, для функций в области частот добавление нулей эквивалентно интерполяции рядами Фурье (гар-

монической интерполяции). Поэтому при попытке определить свертку через функции в частотной области, значения которых предварительно должным образом не интерполируются, будет возникать круговой эффект и появятся искажения результатов. Мы еще коснемся вкратце этого вопроса.

Переходя к вопросу о соотношении свертки и корреляции, сделаем два замечания, касающихся временного индекса. Предположим, что ряды имеют нечетное число членов, равное $2M + 1$, где M — четное число, и нулевому значению временного индекса соответствуют значения рядов, расположенные посередине; таким образом, ряды $x(i)$ и $y(i)$ определены для $i = -M, \dots, -1, 0, 1, \dots, M$. После перестановки членов ряда $y(i)$ в обратном порядке положительные и отрицательные запаздывания для корреляции обретают следующий смысл:

$$x(-M), \dots, x(-1), x(0), x(1), \dots, x(M)$$

$$y(M), \dots, y(1), y(0), y(-1), \dots, y(-M)$$

$$\begin{aligned} & x(-M)y(M) + \dots + x(-1)y(1) + x(0)y(0) + x(1)y(-1) + \dots + x(M)y(-M) && \text{Зап. 0} \\ \oplus & x(-M+1)y(M) + \dots + x(-1)y(2) + x(0)y(1) + x(1)y(0) + \dots + x(M)y(-M+1) && \text{Зап. 1} \\ \ominus & x(-M)y(M-1) + \dots + x(-1)y(0) + x(0)y(-1) + x(1)y(-2) + \dots + x(M-1)y(-M) \\ \oplus & x(-M+k)y(M) + \dots + x(-1)y(k+1) + x(0)y(k) + x(1)y(k-1) + \dots + x(M)y(-M+k) && \text{Зап. } k \\ \ominus & x(-M)y(M-k) + \dots + x(-1)y(-k+1) + x(0)y(-k) + x(1)y(-k-1) + \dots + x(M-k)y(-M) \end{aligned}$$

Таким образом, получаются значения свертки (после умножения на T)

$$\begin{aligned} c_{xy}^t(i) &= T \sum_{p=-M+i}^M x(p) y(i-p), \quad i = 0, 1, \dots, 2M, \\ c_{xy}^t(i) &= T \sum_{p=-M}^{M+i} x(p) y(i-p), \quad i = -1, -2, \dots, -2M. \end{aligned}$$

Наоборот, если рассматривать свертку как ковариацию ряда $x(i)$ с обращенной во времени версией ряда $y(i)$, то для рядов $x(i)$ и $y(i)$, $i = 0, 1, \dots, N-1$, получим соотношения

$$y(N-i) = z(i),$$

$$Ns_{xz}^t(i) = \sum_{p=0}^{N-i-1} x(p) z(p+i), \quad i = -N, \dots, -1, 0, 1, \dots, N-1,$$

$$s_{xz}(i) = s_{xz}^t(i) + s_{xz}^t(N-i), \quad i = 0, 1, \dots, N-1. \quad (7.14)$$

В этом случае отрицательным запаздываниям для ковариации соответствуют положительные запаздывания для свертки. Как следует из формулы (7.14), для круговой свертки получаются

два слагаемых, причем первым N значениям свертки соответствуют обращенные во времени положительные запаздывания для ковариации. Если соотношение (7.14) записать для свертки, то получится сумма

$$c_{xy}(i-1) = c_{xu}^i(N-i-1) + c_{xu}^i(i-1), \quad i=1, 2, \dots, N,$$

где

$$\begin{aligned} c_{xy}^i(N-i-1) &= Ns_{xz}^i(i), \quad i=1, \dots, N-1, \\ c_{xy}^i(i-1) &= Ns_{xz}^i(N-i), \quad i=1, 2, \dots, N, \end{aligned}$$

причем для $i=N$

$$c_{xu}^i(-1) = Ns_{xz}^i(N) = 0.$$

Итак, кросс-корреляция $x(i)$ с обращенной во времени версией $y(i)$ дает свертку $c_{xy}(i)$, сдвинутую на одну позицию.

Наложение как подмена. При определении значений свертки (или ковариации) получается временная последовательность, содержащая $2N$ точек; эта последовательность получается из N -точечной временной последовательности. Из теоремы о выборке для функций в области частот вытекает, что подмены можно избежать, если интервал, с которым производится выборка, равен

$$b' = \frac{1}{2NT}.$$

Однако преобразования Фурье $x(i)$ и $y(i)$ дают вдвое большее значение

$$b = 2b' = \frac{1}{NT}.$$

Поэтому при вычислении для функции свертки обратного преобразования Фурье произведения

$$C_{xy}(k) = X(k)Y(k)$$

возникнет явление подмены. Этот эффект аналогичен подмене частот, рассмотренной раньше; роль частот во временной области играют суммы интервалов времени длины T . Можно показать, что ошибка, связанная с подменой во временной области, оказывается точно такой же, как ошибка, обусловленная наложением.

В этом случае имеется один выход — получить преобразования $X(k)$ и $Y(k)$ в точках, разделенных вдвое меньшим интервалом b' . Поэтому нужно либо получить недостающие значения преобразований Фурье интерполяцией, либо, что эквивалентно, до вычисления преобразований дополнить нулями последовательности $x(i)$ и $y(i)$. Отметим, что значения функции $C_{xy}(k)$ интер-

полировать в данном случае нельзя, поскольку

$$(X \widetilde{Y}(k)) \neq \widetilde{X}(k) \widetilde{Y}(k),$$

где тильдой обозначены значения, полученные интерполяцией.

Фильтрация посредством свертки. Первый фильтр, с которым мы встретились в этой книге, имел вид свертки. В непрерывном случае отклик $y(t)$ линейной системы (фильтра) для всякого входа $x(t)$ определенного вида задается интегралом свертки

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h(t-\tau) d\tau,$$

где $h(\tau)$ — импульсная функция отклика линейной системы с постоянными параметрами.

Конечным дискретным аналогом такого интеграла служит свертка

$$y^c(i) = T \sum_{p=0}^{N-1} x(p) h[(i-p) \bmod N], \quad i=0, 1, \dots, N-1,$$

которую мы называем «круговой». Как правило, предпочтение отдается другой свертке, а именно «быстрой» свертке, определенной соотношениями (7.4).

Одним из способов получения свертки (7.4) служит метод дополнения исходных последовательностей N нулями. При этом все $2N$ значений сверток $y^c(i)$ и $y(i)$ одинаковы. Такой подход требует, по существу, вдвое большего времени для вычислений. Кроме того, требуется вдвое больший объем памяти. Время вычислений и необходимый объем памяти можно сократить, применив специальные приемы, рассмотренные в гл. 6. Если имеется комплексная последовательность $z(i)$, $i=0, 1, \dots, N-1$, для которой мы хотим получить быстрое преобразование Фурье, соответствующее последовательности $x(i)$ длины $2N$ (дополненной), вся вторая половина которой суть нули, то

$$\begin{aligned} Z(2k) &= \text{БПФ} \{z(i)\}, \\ Z(2k+1) &= \text{БПФ} \{z(i) W_N^{i/2}\}, \quad k=0, 1, \dots, N-1. \end{aligned}$$

Далее, как показано в гл. 6, последовательность $y(i)$, $i=0, 1, \dots, 2N-1$, можно получить N -точечным обратным БПФ последовательности

$$U(k) = [V(k) + V^*(N-k)] + j[V(k) - V^*(N-k)] W_N^{-k/2},$$

где $k=0, 1, \dots, N-1$, и

$$V(k) = X(k) H(k), \quad k=0, 1, \dots, N-1,$$

а преобразования $X(k)$ и $H(k)$ получены расщеплением преобразования $Z(k)$, определенного соотношением

$$Z(k) = \text{БПФ} [z(i) = x(i) + jh(i)], \quad k = 0, 1, \dots, N-1.$$

Окончательный результат, т. е. последовательность $y(i)$, получается как чередование действительных частей $u(i)$, дающих члены последовательности с четными индексами, и мнимых частей $v(i)$, дающих члены с нечетными индексами. Иначе говоря,

$$\begin{aligned} y(2i) &= \text{Re БПФ}^{-1} [U(k)], \\ y(2i+1) &= \text{Im БПФ}^{-1} [U(k)], \quad i = 0, 1, \dots, N-1. \end{aligned}$$

Таким образом, последовательность $y(i)$ можно получить при помощи трех (комплексных) N -точечных БПФ вместо двух $2N$ -точечных (комплексных) БПФ. Поскольку число арифметических операций (при $N = 2^p$) пропорционально величине $N \log_2 N$, то, как показывает отношение

$$\frac{3N \log_2 N}{2 \times 2N \log_2 2N} = \frac{3p}{4(p+1)},$$

время вычислений сокращается примерно на четверть; необходимый объем памяти сокращается вдвое.

В гл. 3. были рассмотрены методы реализации фильтров с импульсными функциями отклика, имевшими конечные носители. Для таких фильтров метод, в котором используются БПФ и теорема о свертке, обычно оказывается с вычислительной точки зрения довольно выгодным. Граница, определяющая выбор одного из методов (прямого или использующего БПФ), лежит между значениями числа точек носителя импульсной функции отклика, равными 50 и 100¹⁾. К такому заключению приводят следующие соображения.

Число операций, необходимое для прямого метода вычисления быстрой свертки по формулам (7.4), равно

$$N_d = 2(1 + 2 + 3 + \dots + N) = 2 \frac{N(N+1)}{2} = N(N+1).$$

Число операций, которое требуется в методе, использующем БПФ, равно (предполагается, что $N = 2^p$; на каждое БПФ уходит по $4Nr$ операций)

$$N_{\text{БПФ}} = \underbrace{2(4Nr)}_{\substack{1 \text{ комплексное БПФ} \\ \text{плюс } 1 \text{ дополни-} \\ \text{тельное БПФ для} \\ \text{интерполяции зна-} \\ \text{чений}}} + \underbrace{4Nr}_{\substack{1 \text{ обратное БПФ} \\ \text{при вычислении} \\ \text{которого применя-} \\ \text{ется специальный} \\ \text{метод}}} = 12Nr.$$

¹⁾ Без предварительной децимации.

Здесь не учтены операции, необходимые при вычислении произведения преобразований, а также дополнительные операции, входящие в специальное обратное преобразование. Пренебрегая ими, получим отношение

$$\frac{N_d}{N_{\text{БПФ}}} = \frac{N(N+1)}{12Nr} = \frac{N+1}{12r}.$$

При $N = 2^7 = 128$ это отношение больше единицы, так как

$$\frac{N_d}{N_{\text{БПФ}}} = \frac{129}{84} > 1;$$

при $N = 2^6 = 64$ это отношение меньше единицы:

$$\frac{N_d}{N_{\text{БПФ}}} = \frac{65}{72} < 1.$$

Конкретное число необходимых для вычислений операций в определенной степени зависит от метода программирования, поэтому сделанный выше вывод о границах применимости методов вычислений требует определенных поправок. Кроме того, отдельные операции, которые требуются в методе, использующем БПФ, мы не учитывали. Тем не менее указанные границы служат хорошим ориентиром.

Для «короткой» импульсной функции отклика и временной последовательности, имеющей произвольно большую длину, можно применить метод, известный под названием «перекрытие при сложении». Можно показать (см. Стокхэм (1966)), что, разбивая последовательность $x(i)$ на сегменты длины, равной длине $h(i)$, и используя только что рассмотренный метод вычисления, можно получить точную свертку $x(i)$ и $h(i)$, запоминая каждый раз последнюю половину получающихся сегментов последовательности $y(i)$ и прибавляя их (поточечно) к первой половине следующего полученного сегмента $y(i)$. Эта процедура, схематически показанная на рис. 7.6, оказывается очень практичным и полезным методом фильтрации посредством свертки. Разумеется, было бы чрезвычайно неудобно ограничиваться только теми последовательностями, длина которых лимитируется объемом быстроедействующего запоминающего устройства на сердечниках. Очевидно, что в методе «перекрытия при сложении» вычисление БПФ импульсной функции отклика нужно проводить только один раз. В результате отношение, характеризующее скорость вычислений, примет вид

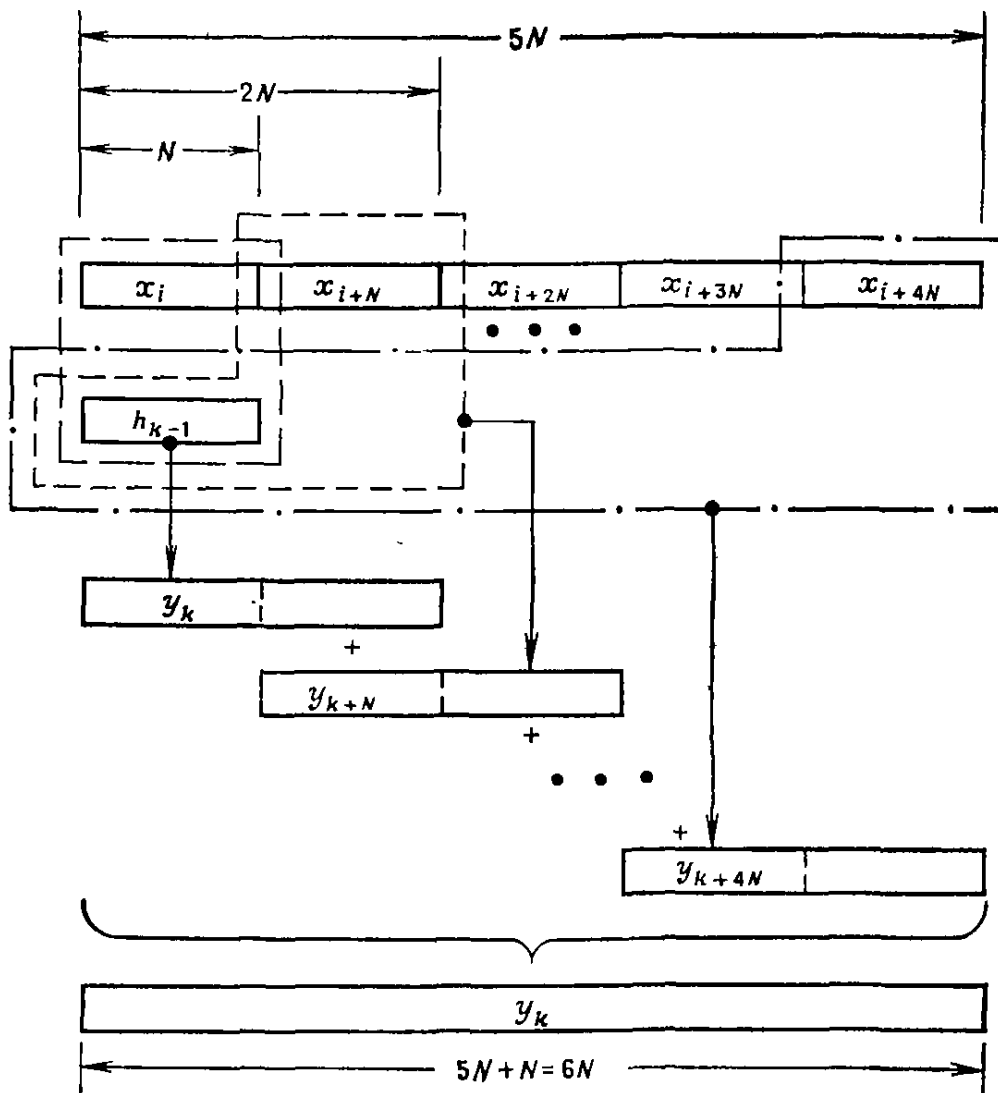
$$\frac{N_d}{N_{\text{БПФ}}} = \frac{N+1}{8r},$$

поскольку одно прямое БПФ теперь не нужно.

Для уже рассматривавшегося случая $N = 2^6 = 64$ это отношение станет равным

$$\frac{N_d}{N_{\text{БПФ}}} = \frac{65}{48},$$

поэтому нижняя граница для метода с БПФ уходит, пожалуй, даже левее значения 50.



7.6. Иллюстрация метода перекрытия при сложении для фильтрации посредством свертки на примере последовательности x , длина которой в шесть раз превосходит длину h .

7.6. КАК ВЫЧИСЛЯТЬ ФУНКЦИИ КОВАРИАЦИИ И СВЕРТКИ

Длина записи и число запаздываний для корреляционной функции. Привести какие-либо точные количественные оценки точности ковариационных функций для данных, включающих шум, чрезвычайно сложно. Ситуация здесь прямо противоположна той, которая возникает для функций плотности спектра; там нужные количественные оценки получают относительно

просто. Причина в том, что значения ковариационных функций во временной области сильно коррелированы (исключение составляет чистый белый шум), в то время как значения спектра мощности от точки к точке почти не коррелируют. Всякая попытка получить какие-то статистические результаты, относящиеся к ковариационной функции, неизбежно приводит к тому, что результат содержит истинную функцию ковариации, т. е. нужно а priori знать результат, который пытаются оценивать.

Для отдельной точки корреляционной функции можно воспользоваться весьма полезным результатом, в котором фигурирует преобразование, известное в статистике как «Z-преобразование Фишера». Так называют преобразование, осуществляемое посредством функции, обратной гиперболическому тангенсу. Можно показать, что значения этого преобразования на величинах $\hat{r}(i)$, т. е.

$$\hat{z} = \text{arth}[\hat{r}(i)] = \frac{1}{2} \log \frac{1 + \hat{r}(i)}{1 - \hat{r}(i)},$$

приближенно имеют нормальное распределение со средним

$$\mu_{\hat{z}} = \text{arth}[\rho(i)]$$

и дисперсией

$$\sigma_{\hat{z}}^2 = \frac{1}{N-2-i}.$$

Следовательно, для отдельных значений корреляционных функций можно строить доверительные границы. К сожалению, для всего множества значений оцениваемой функции эти границы оказываются неверными.

В практических ситуациях рекомендуется избегать значений корреляционной функции для запаздываний $i \geq 0.5N$ (см. по этому поводу работу Акайка (1962)). Показать, что поведение корреляционной функции при больших запаздываниях неустойчиво, довольно нетрудно. Что же касается других запаздываний, то нужно следовать хорошему правилу — не вычислять значений корреляционных функций для запаздываний, превышающих 10% длины всего ряда, т. е. вычислять корреляционные функции для $i \leq 0.1N$. Поскольку в большинстве случаев корреляционные функции используются для определения времени относительного сдвига двух рядов, следует позаботиться о том, чтобы длина ряда была по крайней мере в 10 раз больше ожидаемого значения относительного сдвига.

7.7. ОБЛАСТЬ ИМПУЛЬСНОГО ОТКЛИКА И ПОЛОСА ЧАСТОТ ПРИ ФИЛЬТРАЦИИ ПОСРЕДСТВОМ СВЕРТКИ

Для фильтрации посредством свертки существуют два метода, в которых используется БПФ.

Первый метод. Временная последовательность $x(i)$ и импульсный отклик $h(i)$.

1. Вводится импульсный отклик $h(i)$. Если это необходимо, то область импульсного отклика, содержащую M_1 точек, расширяют за счет добавления нулей таким образом, чтобы новое число точек M стало равным степени 2, т. е. $M = 2^p$. Затем добавляют еще M нулей и вычисляют $2M$ -точечное БПФ. Поскольку $h(i)$ — последовательность действительных значений, результат однозначно определяется $M+1$ значениями преобразования (остальные значения будут комплексно-сопряженными). Введем для полученного преобразования обозначение $H(k)$.

2. Считываются M значений ряда $x(i)$. К ним добавляются еще M нулевых значений, после чего вычисляется $2M$ -точечное БПФ. Как и в п. 1, запоминаются только $M+1$ значений результата. Последовательность этих значений обозначим $X^l(k)$.

3. Вычисляется произведение $H(k) X^l(k) = Y^l(k)$.

4. После получения $M-1$ комплексно-сопряженных значений вычисляется обратное $2M$ -точечное БПФ для $Y^l(k)$. Эти значения образуют последовательность $y^l(i)$.

5. Полученная временная последовательность действительных величин $y^l(i)$ содержит $2M$ членов. Первые M из них становятся первыми M величинами выходной временной последовательности $y(i)$. Вторая половина значений, содержащая M значений, запоминается.

6. Повторяются шаги 2—4. Новую последовательность обозначим $y^l(i)$. Следующие M значений выхода $y(i)$ получаются почленным сложением первых M значений новой последовательности $y^l(i)$ с M значениями последовательности $y^{l-1}(i)$, находящейся в памяти. Вторая половина значений новой последовательности $y^l(i)$ запоминается.

7. Шестой шаг повторяется до тех пор, пока на вход не поступит последняя часть входной последовательности $x(i)$. Если число членов в последней части меньше M , то длину последовательности увеличивают до M , считая недостающие значения нулями. Повторяются шаги 2—4; в результате получается завершающая часть последовательности $y^L(i)$. Первые M значений этой последовательности почленно складываются со значениями $y^{L-1}(i)$, находящимися в памяти. Остальные M значений последовательности непосредственно подаются на выход; этим завершаются вычисления.

Отметим, что число точек области импульсного отклика, равное M_1 , накладывает известные ограничения на действие фильтра. Именно, ширина эффективной полосы частот фильтра, которую обозначим B_e , ограничена снизу:

$$B_e \geq \frac{1}{M_1 T}.$$

На практике величина B_e может оказаться значительно больше; вычислительная полоса b равна

$$b = \frac{1}{M_1 T}.$$

Второй метод. Временная последовательность $x(i)$ и частотный отклик $H(k)$.

Метод, который используется в этом случае, полностью аналогичен методу, рассмотренному в первом случае. Исключение составляет только первый шаг, поскольку $H(k)$ задается непосредственно. (Если значения $H(k)$ заданы через коэффициент усиления и фазу, то предварительно их выражают через действительную и мнимую части. Если фазы отсутствуют, то их полагают равными нулю.)

Остановимся на некоторых требованиях, предъявляемых к последовательности $H(k)$.

1. Эта последовательность должна содержать $M+1$ значений. Если число значений M_1 меньше M , то последовательность дополняется нулями таким образом, чтобы число ее членов оказалось равным $M+1$ (где M — наименьшая степень 2, превосходящая M_1).

2. Преобразование отрезков последовательности $x(i)$, содержащих M значений ($P = MT$), дополняется нулями, поэтому интервал в области частот оказывается равным

$$b = \frac{1}{2P} = \frac{1}{2MT} = \frac{S}{2M} = \frac{F}{M},$$

где S — скорость выборки, а F — частота Найквиста.

3. Всякая последовательность $H(k)$, предназначенная для фильтрации, должна быть определена в частотах, разделенных интервалом длины $b = F/M = S/(2M)$.

4. Более того, всякая используемая для этой цели последовательность $H(k)$, $k = 0, 1, \dots, M$, должна представлять собой дискретное преобразование; значения этого преобразования, имеющие нечетные индексы, совпадают с теми интерполированными значениями, которые получаются при вычислении преобразования для M -точечной последовательности значений данных, дополненной M нулями. Сопоставьте это преобразование с обычным M -точечным преобразованием действительной последо-

вательности M значений, дающим последовательность M величин, $M/2 - 1$ из которых повторяют значения, уже имеющиеся в преобразовании. Таким образом, в обычном случае частоты разделяются интервалом длины $b = F/(M/2) = 2F/M = S/M$.

Обращаясь к фильтрации посредством свертки с функцией, определенной в частотной области, основное внимание следует уделять последнему п. 4. Интерполированные значения можно получать при помощи процедуры «обратной децимации», рассмотренной в § 5 гл. 5. Эти значения можно получить и иначе: сначала вычислить обратное преобразование Фурье, затем дополнить полученную последовательность нулями и вычислить преобразование Фурье для новой последовательности. Если не обеспечивать требований 2—4 (т. е. не подгонять к нужному значению разрешение b), то при фильтрации возникнут сильные ошибки и искажения

7.8. НОРМИРОВКА И УДАЛЕНИЕ СРЕДНЕГО ПРИ ВЫЧИСЛЕНИИ КОВАРИАЦИЙ

При вычислении ковариаций нормировка может осуществляться так, что она послужит причиной явных парадоксов. Для нормировки можно пользоваться значениями соответствующих выборочных автокорреляционных функций для нулевых смещений, поскольку

$$s_x(0) = s_x^2, \quad s_y(0) = s.$$

В этом случае корреляционную функцию можно записать в виде

$$r_{xy}(i) = \frac{s_{xy}(i)}{\sqrt{s_x(0)s_y(0)}}, \quad i = 0, 1, \dots, m.$$

Представим, что по недосмотру для временных последовательностей не было проведено удаление средних. В этом случае для нулевых запаздываний автоковариационных функций будут получены значения

$$s_x(0) = s_x^2 + \bar{x}^2, \quad s_y(0) = s + \bar{y}^2,$$

которые внесут при нормировке ошибку.

Рассмотрим, какие эффекты связаны с удалением среднего и нормировкой. Допустим, что

$$x(i) = \Delta x(i) + \bar{x}, \quad y(i) = \Delta y(i) + \bar{y}.$$

Тогда ковариационную функцию можно записать в виде

$$s_{xy}(i) = \frac{1}{N-i} \left[\sum_{p=0}^{N-i-1} \Delta x(p) \Delta y(p+i) + \bar{y} \sum_{p=0}^{N-i-1} \Delta x(p) + \bar{x} \sum_{p=0}^{N-i-1} \Delta y(p+i) + (N-i) \bar{x} \bar{y} \right] = \\ = \hat{\text{cov}}(x, y) + \bar{x} \bar{y} + \bar{x} \Delta x_f + \bar{y} \Delta y_l,$$

где Δx_f и Δy_l обозначают выборочные средние первых и последних $N-i$ значений соответствующих временных последовательностей с удаленными средними. После нормировки получим корреляционную функцию

$$r_{xy}(i) = \frac{s_{xy}(i)}{\sqrt{s_{xx}(0) s_{yy}(0)}} = \frac{\hat{\text{cov}}(x, y) + \bar{x} \bar{y} + \bar{x} \Delta x_f + \bar{y} \Delta y_l}{\sqrt{(s_x^2 + \bar{x}^2)(s_y^2 + \bar{y}^2)}} = \\ = \frac{\hat{\text{cov}}(x, y)}{s_x s_y} \frac{1}{A} + \frac{\bar{x} \bar{y}}{s_x s_y} \frac{1}{A} + K,$$

поэтому

$$\frac{\hat{\text{cov}}(x, y)}{s_x s_y} = r_{xy}(i) A - \frac{\bar{x} \bar{y}}{s_x s_y} - KA.$$

Величина A всегда меньше единицы, следовательно, данные с удаленным средним дают меньшее значение корреляции. Если средние имеют одинаковые знаки, то корреляция данных с удаленным средним окажется еще меньше. Но если средние имеют разные знаки, то смещение может пойти в другом направлении. Вообще говоря, кросс-корреляция временных последовательностей, из значений которых не удалены средние, может достигать произвольно больших величин. Однако если знаки выборочных средних противоположны, то это оказывается несправедливым утверждением.

При нормировке с помощью дисперсий, вычисленных на меньшем числе данных, о характере поведения кросс-корреляции никаких определенных выводов сделать нельзя. Выборочные дисперсии, вычисленные на укороченных отрезках последовательности данных, могут оказаться и больше, и меньше значения автокорреляционной функции при нулевом запаздывании.

7.9. ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ КОВАРИАЦИИ И СВЕРТКИ

Определение временных задержек. В приложениях наиболее удобна такая нормировка кросс-корреляционных функций, при которой их значения по абсолютной величине не превосходят

единицы. Почти все приложения кросс-корреляционных функций тем или иным образом сводятся к определению временных задержек. Что же касается автокорреляционных функций, то когда мы имеем дело с приложениями, более удобными оказываются их аналоги для области частот — функции плотности спектра мощности.

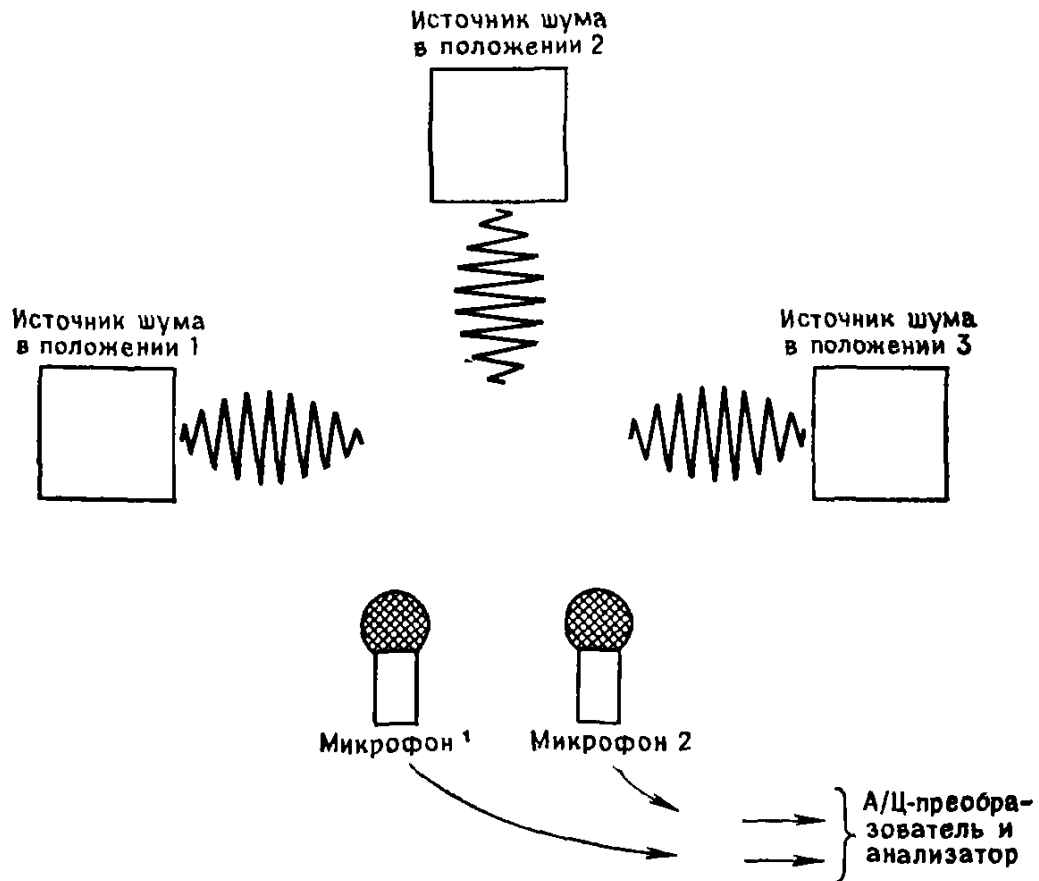


Рис. 7.7. Схема эксперимента для определения кросс-корреляций акустических источников.

Применение кросс-корреляционной функции мы проиллюстрируем на примере обнаружения источников шума. Проведем эксперимент, в котором источник шума располагается в трех различных положениях относительно пары микрофонов (рис. 7.7). Три соответствующие кросс-корреляционные функции приведены на рис. 7.8. Нетрудно видеть, что при расположении источника шума слева от микрофонов сигнал $x(t)$ опережает сигнал $y(t)$; среднему положению источника соответствует всплеск, по существу, при нулевой задержке; при расположении источника справа сигнал $y(t)$ опережает $x(t)$, на что указывает пик при отрицательном запаздывании. Учитывая, что скорость звука в данном случае равна 1100 фут/с и микрофоны установлены на расстоянии примерно 1 фут, можно прикинуть, что время задержек должно быть равно примерно 1 мкс. В данном случае

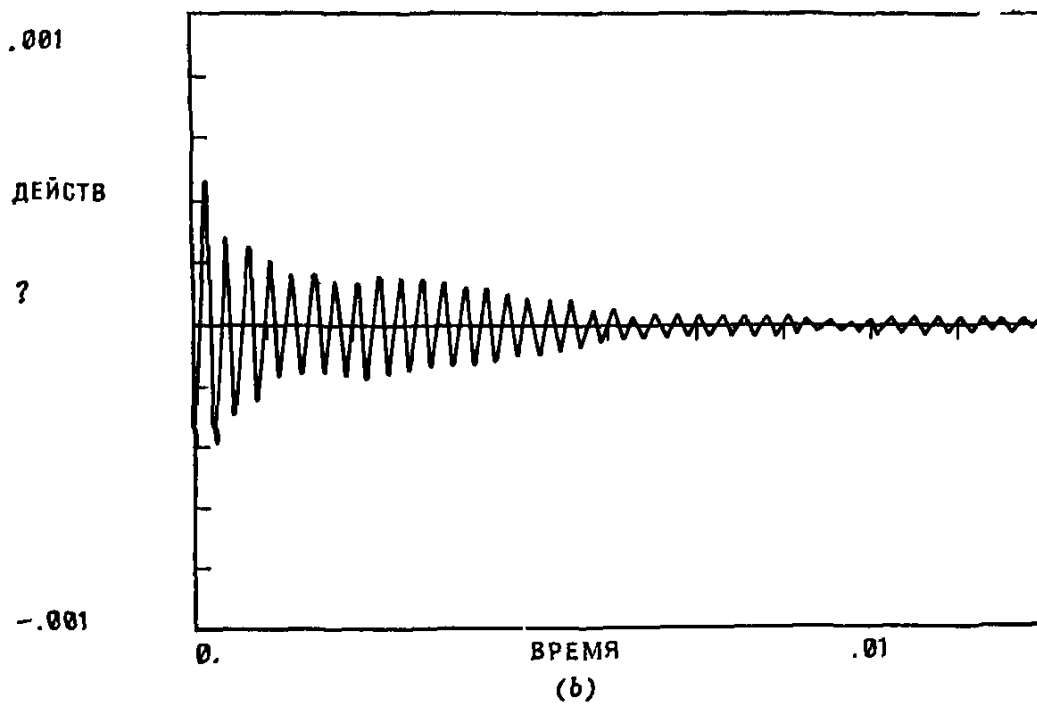
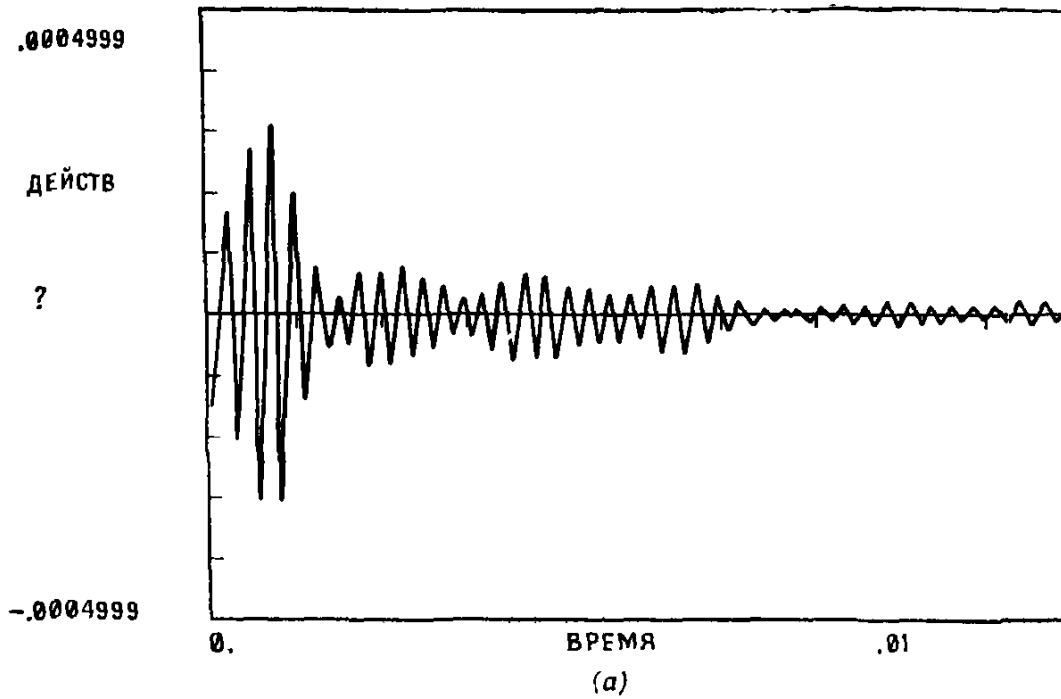


Рис. 7.8. Три кросс-корреляционные функции, полученные для трех положений источников шума. *a* — источник шума слева от микрофона 1; *b* — источник шума между микрофонами; *c* — источник шума справа от микрофона 2.

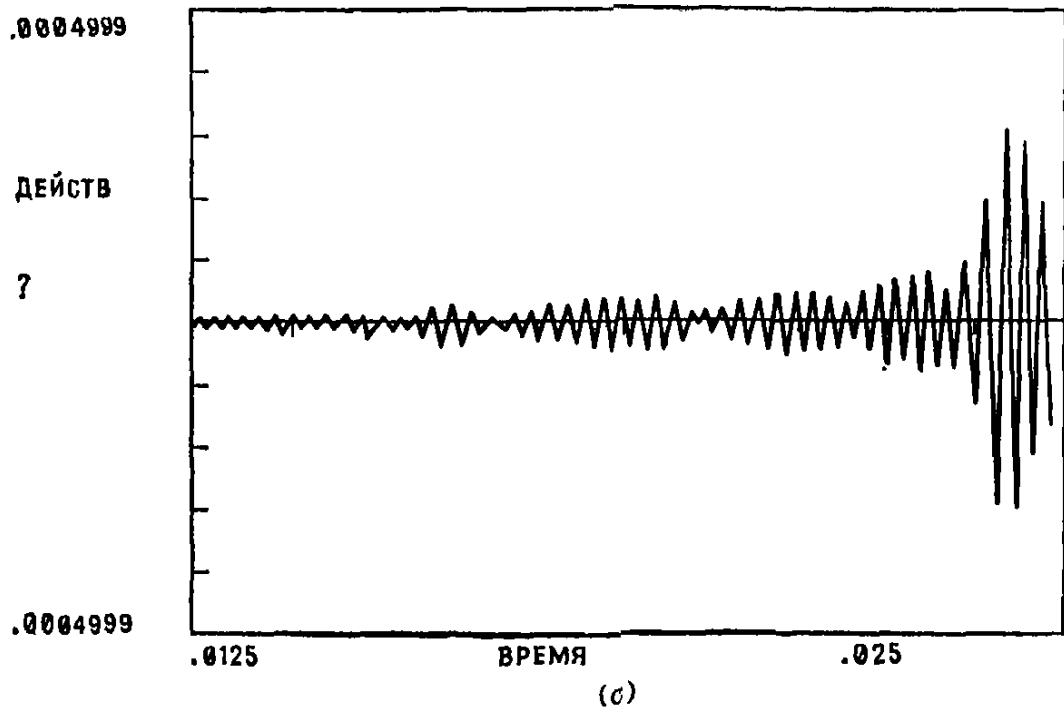


Рис. 7.8. Продолжение.

$S = 50\,000$ выб./с и $T = 0.020$ мкс, поэтому максимальные значения кросс-корреляции должны появляться при $i = 50$; с этим значением хорошо увязан объем выборки (вспомним рекомендацию, по которой N должно выбираться так, чтобы $i \leq 0.1N$).

Моделирование отклика линейных систем (фильтрации посредством свертки). Если мы можем оценивать функцию частотного отклика чувствительных линейных систем (к ним можно отнести большинство структур), то мы имеем возможность с достаточной степенью точности моделировать отклик таких систем. Предположим, что $\hat{H}_{xy}(k)$ — оценка частотного отклика и $x(i)$ — входной сигнал, для которого нужно определить отклик системы. Тогда оценкой отклика служит последовательность ($i = 0, 1, \dots, N-1$)

$$\hat{y}(i) = \mathcal{F}^{-1}[X(k) \hat{H}_{xy}(k)] \quad (7.15)$$

(предполагается, что против возможного кругового эффекта приняты соответствующие меры). Определение величин по формуле (7.15) эквивалентно свертке

$$\hat{y}(i) = T \sum_{p=0}^{N-1} x(p) h(p-i), \quad i = 0, 1, \dots, N-1, \quad (7.16)$$

где индекс $(p-i)$ берется по модулю N .

В структурном анализе непрерывный аналог свертки (7.16) обычно называют интегралом Дюамеля.

Моделирование такого рода представляет интерес во многих приложениях. Предположим, что имеется укрытие для чувствительного электронного оборудования радарной станции. Далее, допустим, что мы можем определить функции частотного отклика при воздействии извне.

Рассмотрим задачу о «прочности», в которой требуется определить, что произойдет внутри укрытия, если его наружная часть подвергается повышенному атмосферному давлению в результате воздушного взрыва, как это происходит при близком разрыве бомбы. Имея в распоряжении типичную временную последовательность для ожидаемой при взрыве бомбы ударной волны, можно оценить отклик при помощи формулы (7.15). Полнее с затронутым вопросом читатель может ознакомиться по работам Кеннеди и Сэффорда (1974) и Масри и Сэффорда (1976).

Иллюстрация наложения (подмена во временной области).

Для периодических (детерминированных) и случайных данных эффект наложения проявляется совершенно различным образом. Серьезные расхождения в результатах возникают также в зависимости от выбора того или иного делителя ($N-i$ или N).

В качестве примера рассмотрим автокорреляционную функцию синусоиды, вычисленную тремя различными способами (рис. 7.9): a — без добавления нулей и с делением на N , b — с добавлением нулей и делением на N , c — с добавлением нулей и делением на $N-i$. Отметим, что на рис. 7.9, a и b заметны противоположные по характеру искажения. В силу периодичности функции наложения (подмены) не происходит, поскольку та информация, которая накладывается друг на друга, оказывается правильной. Поэтому деление на N не дает никаких искажений результата — все N кросс-произведений суммы правильны. Другая картина возникает при добавлении нулей (рис. 7.9, b и c). В этих случаях предусмотрена возможность наложения, и, следовательно, в сумму входят только $N-i$ произведений. При этом деление на N вызывает заметное искажение — уменьшение истинных значений таким образом, что они попадают в треугольные области; при делении на $N-i$ существенных искажений не возникает.

Хотя деление на $N-i$ дает несмещенные результаты, нельзя не отметить того, что деление на N удерживает величины в интервале $[-1, 1]$. Мы придерживаемся того мнения, что в ситуациях со случайными данными при значениях задержки i , ограниченной величиной $0.1N$, чрезмерной статистической изменчивости можно избежать, добавляя нули, проводя деление на $N-i$ и имея достаточный объем выборки. Обратим внимание на ту проблему, которая связана с делением на величины, близкие к нулю, при значениях запаздывания, близких к N .

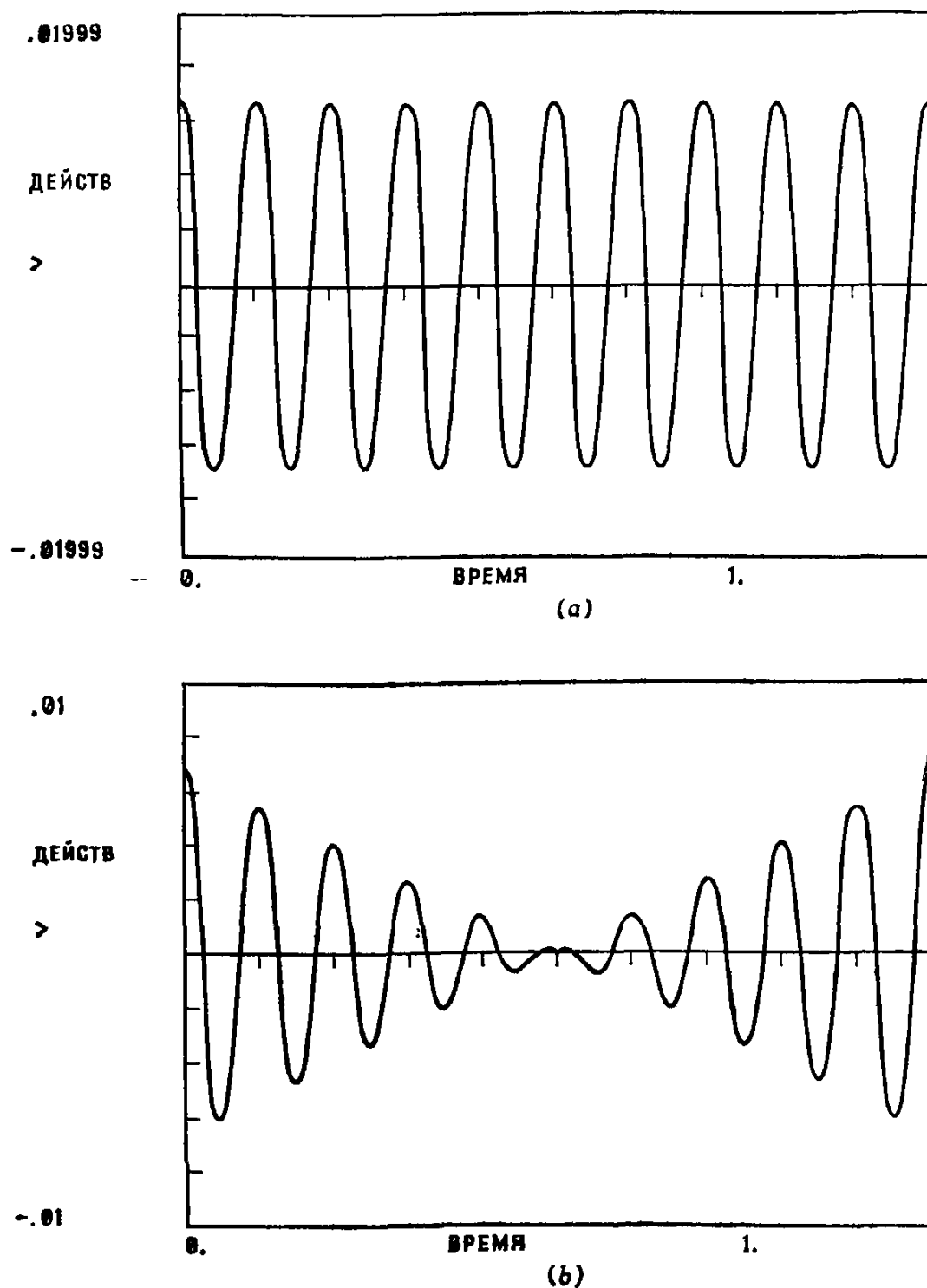


Рис. 7.9. Иллюстрация наложения, возникающего при вычислении кросс-корреляционных функций для синусоиды. *a* — наложение, делитель N ; *b* — наложения нет, делитель N ; *c* — наложения нет, делитель $N - 1$.

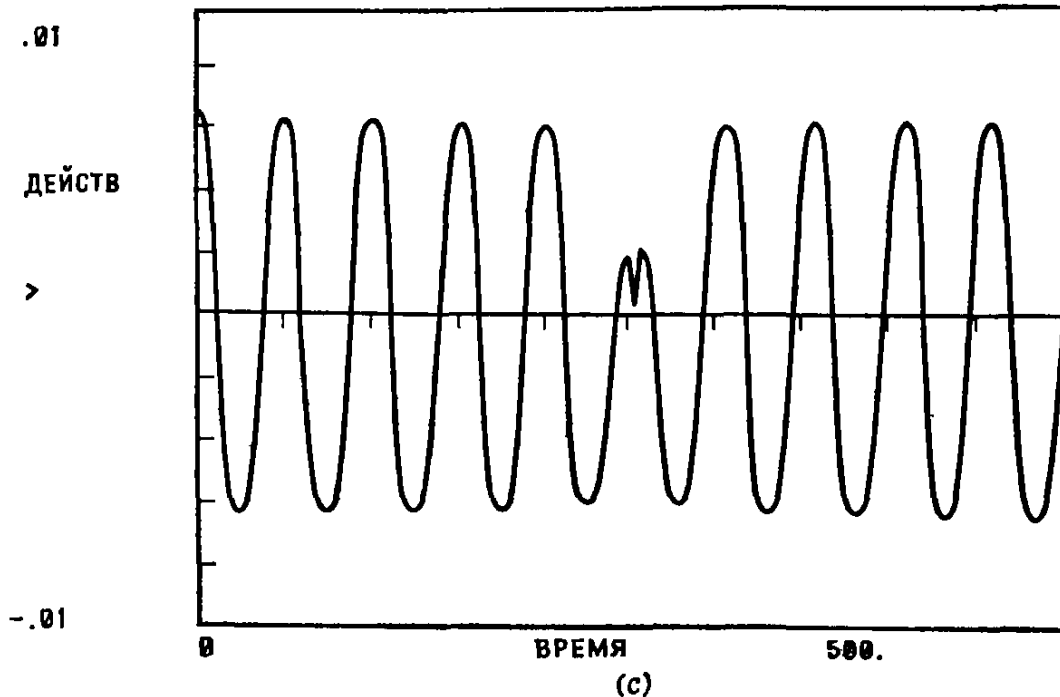


Рис. 7.9. Продолжение.

Хотя данные, изображенные на графиках, были подобраны специально, они вполне точно отражают те эффекты, которые наблюдаются на практике.

Упражнения

7.1. Используя формулу (7.9), перепишите определение свертки.

Ответ:

$$\hat{s}_{xy}(i) = \frac{1}{2N} \left(\sum_{p=1}^N [x(p) + y(p+i)]^2 - \sum_{p=1}^N x^2(p) - \sum_{p=1}^N y^2(p+i) \right).$$

7.2. Используя формулу (7.10), перепишите определение свертки.

Ответ: см. упр. 7.1.

7.3 Докажите, что $\mathfrak{F}[s_{xy}(\tau)] = \mathfrak{F}[X^*(t)Y(t)]$.

Ответ:

$$\mathfrak{F}[s_{xy}(\tau)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t) y(t+\tau) e^{-i2\pi f\tau} dt d\tau;$$

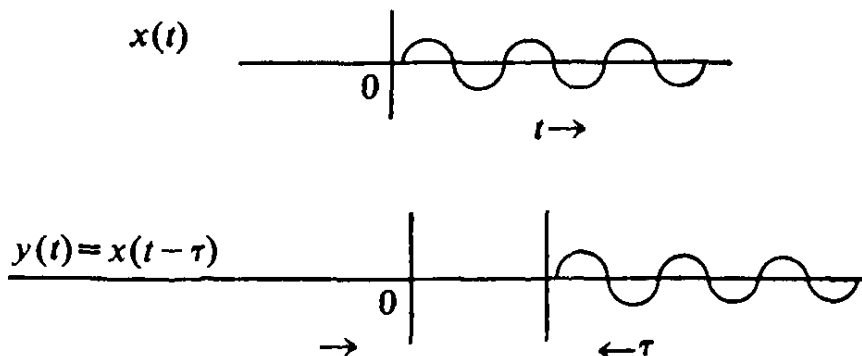
после замены переменных

$$u = t + \tau, \quad \tau = u - t, \quad d\tau = du$$

получим

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t) y(u) e^{-i2\pi f(u-t)} dt du &= \\ &= \int_{-\infty}^{\infty} x(t) e^{i2\pi ft} dt \int_{-\infty}^{\infty} y(u) e^{-i2\pi fu} du = X(f)^* Y(f). \end{aligned}$$

7.4. Докажите, что сдвигу графика функции отвечает линейная убывающая функция фазы с отрицательными значениями; сдвиг определяется как такое значение τ , что при $t = \tau$ функция $y(t)$, соответствующая перемещению $x(t)$ на τ , принимает значение $x(0)$ ($x(\tau - \tau) = x(0)$):



Ответ: поскольку $y(\tau) = x(0)$ и, значит, при $t = \tau$

$$y(t) = x(t - \tau) = x(0),$$

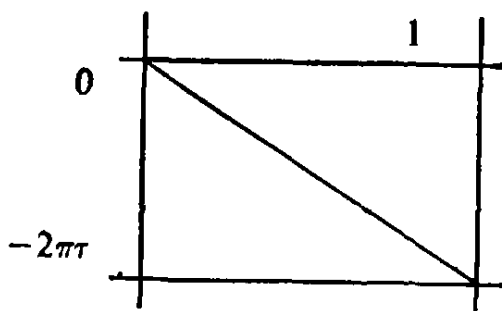
после замены переменных

$$u = t - \tau, \quad t = u + \tau, \quad dt = du$$

получим

$$\int x(t - \tau) e^{-j2\pi f t} dt = \int x(u) e^{-j2\pi f (u + \tau)} du = e^{-j2\pi f \tau} X(f),$$

Для $\varphi = -2\pi f \tau$ имеем



7.5. Напишите программу, в которой предусмотрены следующие действия.
 а) Получение функции, значения которой при $i = 0, 1, \dots, 2^p - 1$ равны 0.1 и при $i = 2^p, \dots, 127$ — нулю. Параметр p пробегает значения $1, \dots, 7$.
 б) Вычисление для этой последовательности во всех семи случаях функции круговой свертки и изображение результатов с помощью программы для графического изображения данных.
 Прокомментируйте результаты.

7.6. Повторите упр. 7.5, но программу измените так, чтобы к данным добавлялся гауссовский шум с нулевым средним и единичной дисперсией (получите этот шум сложением 12 идущих подряд значений, полученных с помощью подпрограммы TDRAND, и последующим вычитанием из суммы б). Поясните результаты.

7.7. Повторите упр. 7.6 для кросс-корреляции: получите с одинаковым шагом функции, используя для каждой из этих двух функций разные значения гауссовского шума. Поясните результаты.

Глава 8

ПЛОТНОСТИ КРОСС-СПЕКТРА И СПЕКТРА МОЩНОСТИ

8.1. ОБЗОР ОСНОВНЫХ МЕТОДОВ ВЫЧИСЛЕНИЯ СПЕКТРОВ

Функция плотности спектра мощности (ПСМ) определяет распределение дисперсии (среднеквадратичного значения) временной последовательности по частотам. Своим происхождением термин обязан электротехнике. Дело в том, что рассеяние мощности в электрической цепи пропорционально среднему квадрату поданного напряжения. Этим объясняется, почему в термин вошло слово мощность. Слово спектр подчеркивает, что функция ПСМ есть функция частот. В ПСМ значения мощности на бесконечно малых частотных интервалах нормируются длиной такого интервала, чтобы получалась функция плотности. Понятие ПСМ настолько далеко шагнуло за пределы электротехники, что в некоторых отраслях науки все чаще используется термин *автоспектральная плотность*.

Понятие ПСМ аналогично понятию дисперсии в классической статистике. Продолжая эту аналогию, понятию ковариации (как функции частот) сопоставляют функцию плотности кросс-спектра (ПКС). Понятие ПКС определяется для пары временных последовательностей. Как правило, в качестве этих последовательностей выступают вход и выход системы. ПКС — последний шаг, подводящий к понятиям передаточных функций и функций когерентности.

Возвращаясь к аналогии с классической статистикой, отметим, что передаточная функция соответствует регрессионному коэффициенту, а когерентность — квадрату коэффициента корреляции.

Вычислительные определения. Для вычисления функций плотности спектра существуют три способа. Каждый из этих способов основан на своем вычислительном определении ПСМ; асимптотически все три определения дают один и тот же результат. Во всех остальных способах вычислений используется операция приведения к белому шуму. Она будет затронута только во втором томе, где помимо прочего будут рассмотрены также вопросы применения для оценивания ПСМ линейно упреждающего кодирования (КЛУ), метода Прони, авторегрессионного скользящего усреднения (УАС) и спектрального анализа максимума энтропии (САМЭ). А сейчас обратимся к следующим методам.

Метод преобразования Фурье

$$\text{ПСМ} \quad \hat{S}_{xx}(f) = \frac{1}{P} \overline{|X(f)|^2}; \quad (8.1)$$

$$\text{ПКС} \quad \hat{S}_{xy}(f) = \frac{1}{P} \overline{X^*(f) Y(f)}.$$

Метод корреляционной функции (метод Блэкмена—Тьюки (Б—Т))

$$\text{ПСМ} \quad \hat{S}_{xx}(f) = \int_{-P/2}^{P/2} \hat{R}_{xx}(\tau) \cos(2\pi f\tau) d\tau; \quad (8.2)$$

$$\text{ПКС} \quad \hat{S}_{xy}(f) = \int_{-P/2}^{P/2} \hat{R}_{xy}(\tau) \exp(-j2\pi f\tau) d\tau.$$

Метод полоснопропускающего фильтра

$$\text{ПСМ} \quad \hat{S}_{xx}(f) = \frac{1}{B_e} E[x^2(t, f, B_e)]; \quad (8.3)$$

$$\text{ПКС} \quad \text{Re}[\hat{S}_{xy}(f)] = C_{xy}(f) = \frac{1}{B_e} E[x(t, f, B_e) y(t, f, B_e)],$$

$$\text{Im}[\hat{S}_{xy}(f)] = Q_{xy}(f) = \frac{1}{B_e} E[x_q(t, f, B_e) y_q(t, f, B_e)],$$

где $E[\]$ — операция усреднения; $x(t, f, B_e)$ — результат, полученный при полоснопропускающей фильтрации $x(t)$; f — центральная частота; B_e — полоса пропускания фильтра; $x_q(t, f, B_e)$ — сдвиг $x(t, f, B_e)$ по фазе на 90° (сдвигающая фильтрация). Результаты применения каждого из трех методов к одной и той же временной последовательности показаны на рис. 8.1.

Процедура, включающая преобразование Фурье, в значительной мере вытеснила метод Б—Т. Из-за некоторых существенных расхождений в обозначениях метод Б—Т вынесен в приложение В. Что же касается метода с фильтрацией, то основные вопросы, связанные с его применением, уже разбирались в гл. 5. Метод преобразования Фурье будет рассмотрен в этой главе ниже.

8.2. ПОНЯТИЕ ПЛОТНОСТИ

Всякий раз, когда речь заходит о плотности, будь то плотность мела на доске или плотность атмосферы, нужно сознавать, что добиться ее точного измерения невозможно. В самом деле, при измерении плотности сначала определяется количество вещества, приходящееся на конечную площадь или определенный объем, а затем на их величины делится полученный результат, т. е. фактически для данной поверхности или выбранного объема измеряется среднее значение. Сказанное в полной мере

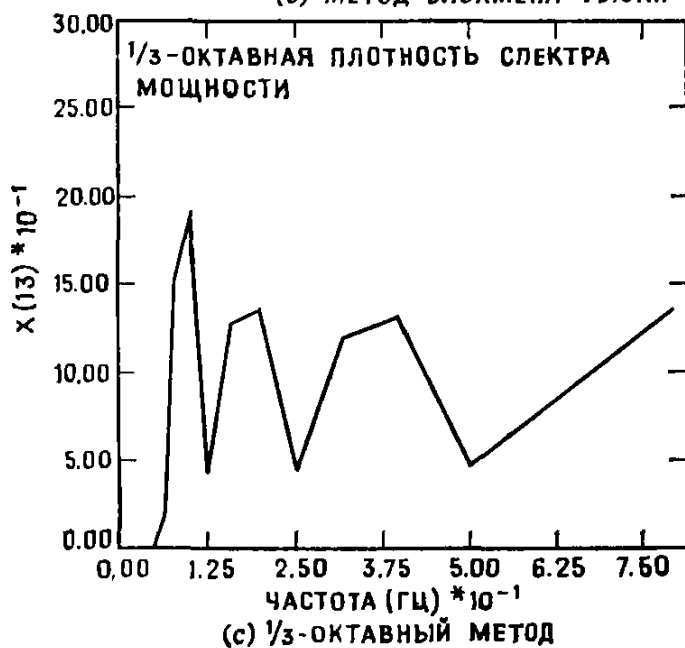
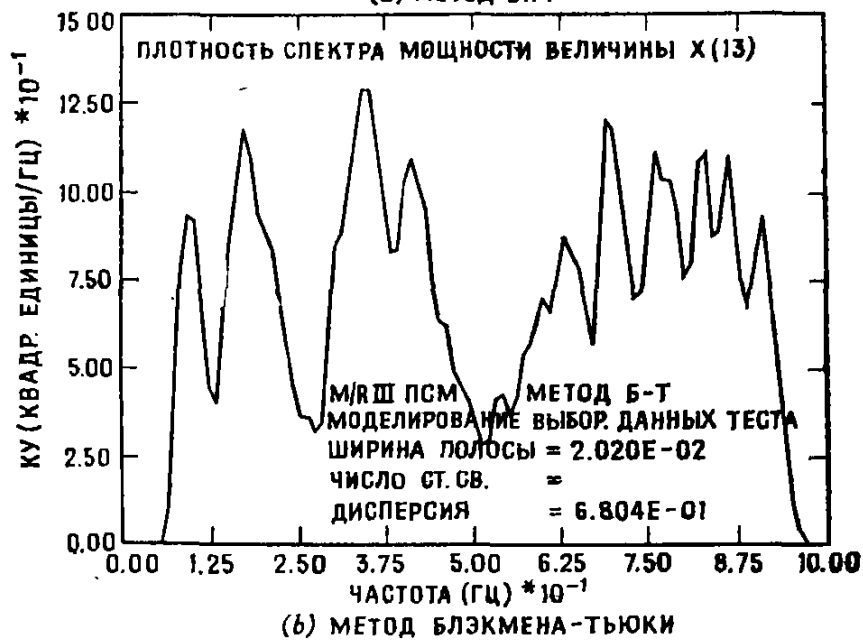
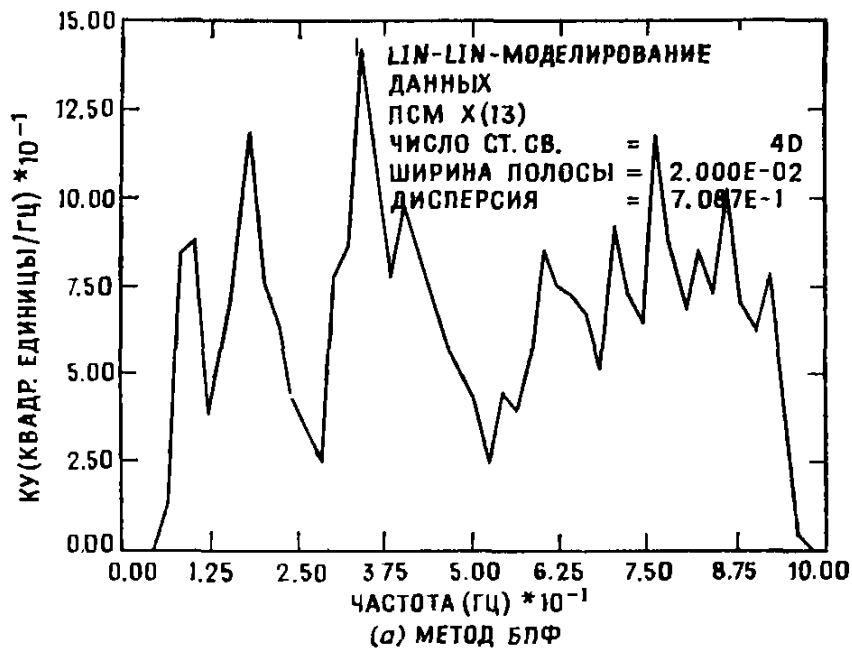


Рис. 8.1. РСМ трех типов. а — метод БПФ; б — метод Блэкмена — Тьюки; в — 1/3-октавный метод.

можно отнести к ПСМ и ПКС, представляющим собой функции плотности.

Определения (8.1), (8.2) и (8.3) дают один и тот же результат, если фигурирующие в них последовательности имеют бесконечную длину, а полосы частот — бесконечно малую ширину. Практически ни того, ни другого достичь нельзя. Семейство или совокупность временных последовательностей можно рассматривать с позиций теории случайных процессов, считая такую совокупность случайным процессом. Накопив достаточно большое число выборок этого процесса, т. е. получив набор временных последовательностей, имеющих в каждый момент времени (в зависимости от условий эксперимента ими могут быть, например, время суток или фаза луны) одинаковые статистические характеристики, можно выполнить «усреднение по совокупности», при котором для каждого момента времени определяются статистические характеристики, в том числе и ПСМ. Но такая возможность существует далеко не всегда. Из-за чрезмерной стоимости или малодоступности большого числа наблюдений чаще всего в распоряжении оказывается лишь одна временная последовательность. Если эту временную последовательность как элемент бесконечно большого пространства возможных выборок, записанных в несколько различных условиях, можно считать одной реализацией эргодического случайного процесса, то поступают иначе — статистические свойства определяют через временное усреднение. Эта возможность, имеющаяся для регулярных (т. е., в частности, стационарных) процессов, уже отмечалась в гл. 2. Временное среднее всегда определяется на конечных промежутках. Это подводит нас к важной задаче выбора полос частоты, обеспечивающих статистическую устойчивость, а также и к другим аспектам вычислений ПСМ.

Как следует из сказанного выше, понятие ПСМ тесно приемыкает к идее стационарного случайного процесса. Но помимо стационарного случая понятия преобразования Фурье и спектра мощности оказываются весьма полезными и для анализа данных, отвечающих переходным процессам, поскольку их частотные характеристики определяются подобным образом. Для случая переходных процессов потребуются, правда, некоторые изменения, касающиеся вычислений; мы остановимся на них ниже.

8.3. ШИРИНА ЭФФЕКТИВНОЙ РАЗРЕШАЮЩЕЙ ПОЛОСЫ

Основным понятием, относящимся к вычислениям, служит ширина эффективной разрешающей полосы B_e . Это понятие касается каждого из трех методов; все они дадут практически одинаковые результаты, если ширину полосы B_e и длину записи P подогнать соответствующим образом. Важно понимать, что изме-

рение средней мощности возможно только в пределах ограниченной полосы частот. Статистическая устойчивость результатов во многом зависит от ширины этой полосы (см. § 8.5).

Типичные для трех указанных методов эффективные формы фильтров изображены на рис. 8.2. Дополнительные сведения можно найти в гл. 6—8 книги Отнеса и Эноксона (1972).

Ширину полосы можно определить различными способами. Как правило, она полагается равной расстоянию между точками, в которых передаточная функция мощности (спектральное окно) принимает половину максимального значения. Напомним, что

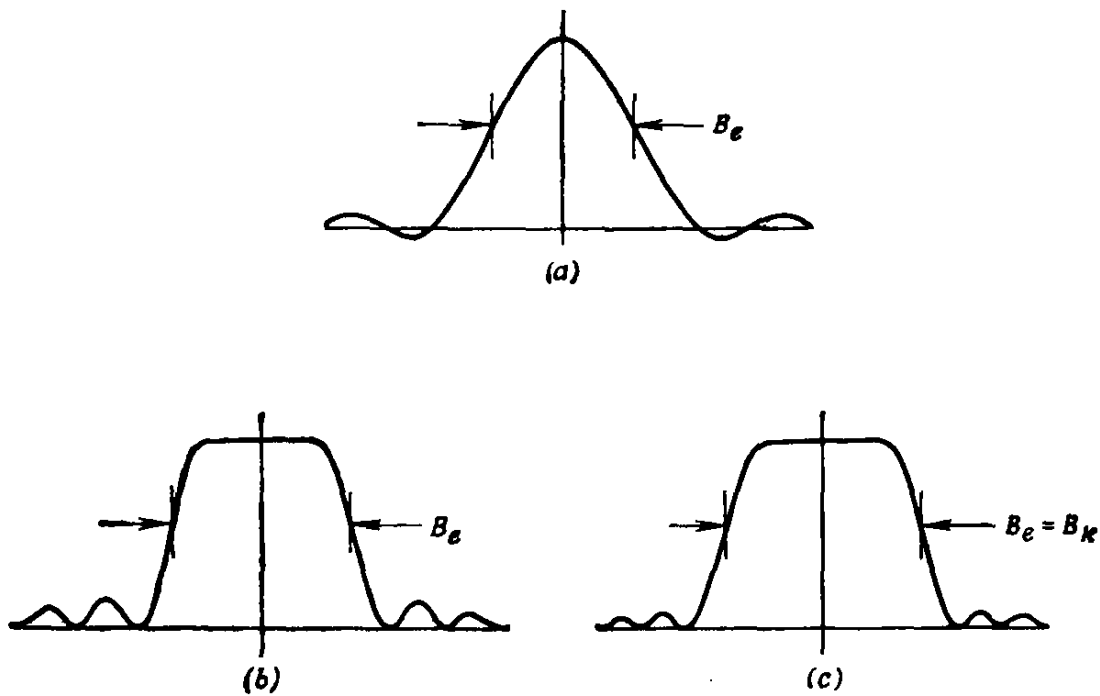


Рис. 8.2. Типичные спектральные окна. *a* — корреляционное спектральное окно; *b* — спектральное окно в методе БПФ; *c* — спектральное окно в методе фильтрации.

точка половинной мощности служит основным параметром при расчете цифровых фильтров. В методе с фильтрацией эта точка и определяет ширину полосы B_e . В методе корреляционной функции приближенное значение ширины полосы получают по формуле $B_e = 1/\tau_{\max}$, где $\tau_{\max} = mT$, m — максимальное значение запаздывания и τ_{\max} — максимальное время запаздывания, входящее в вычисления корреляции. В методе БПФ приближенное значение ширины полосы получают по формуле $B_e = m/P = m/(NT)$, где P — длина записи, N — число величин выборочных данных и m — число отдельных компонент, для которых проводится усреднение. В каждом из трех методов допускается применение различных спектральных окон (см. § 8.8). Помимо этого, в методе БПФ возможны и другие разнообразные изменения.

Различие между разрешающей и вычислительной полосами. Если значения спектральной функции вычислять в частотах, разделенных интервалами длины B_e , то спектральные величины будут (приближенно) статистически независимы (ортогональны). В полной мере это справедливо только для стационарных процессов. Значения функции можно вычислять и с меньшим шагом. Хотя дополнительные значения, строго говоря, будут нести очень мало новой информации (по крайней мере в смысле теории информации), во многих случаях значения спектральной функции вычисляют с вдвое меньшим шагом $b = B_e/2$. Такой выбор шага в области частот предохраняет от эффекта наложения при построении некоторых временных функций по обратным преобразованиям Фурье. С этой точки зрения шаг, равный $B_e/2$, оптимален. В частности, такая ширина полосы частот нужна при вычислении передаточной функции, используемой для получения свертки с помощью БПФ. Чтобы избежать кругового эффекта, нужно получить $b = B_e/2$, добавляя в каждый отрезок N_s нулей. Круговой эффект был подробно рассмотрен в гл. 1 и 7.

В пользу выбора меньшего шага говорит и то дополнительное соображение, что, вычисляя спектральную функцию с меньшим шагом в области частот, можно получить более гладкие и точные графики. По сути дела, такое вычисление эквивалентно интерполяции рядами Фурье. Последнюю можно выполнить, дополнив функцию нулями. Например, чтобы получить промежуточные значения функции, заданной таблично, интерполяцию можно провести следующим образом: сначала вычислить преобразование Фурье (для N последовательных точек) и дополнить преобразованную функцию N нулями, а затем вычислить обратное преобразование Фурье.

Как правило, ширину полосы B_e мы называем «разрешающей», а b — «вычислительной», подчеркивая тем самым их различие. В спектрах Б — Т обычно используется ширина полосы $b = B_e/2$, как это и предлагалось Блэкменом и Тьюки в их книге (1958), ставшей теперь классической. При вычислениях ПСМ методом БПФ, как правило, берется интервал $b = B_e$. Такой выбор дает минимальный шаг, при котором получаются почти неперекрывающиеся, независимые частотные величины.

8.4. ПРЕДЕЛЫ РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ

Нередко возникает вопрос о том, каким должен быть частотный диапазон для спектральных вычислений на ЭВМ. Однозначный ответ на него не имеет смысла. Причина этого кроется в том, что обычно для данных, представленных в цифровой форме, неизвестно, с каким интервалом они выбирались — им может быть и 1 мес, и 1 мкс. Возможности оборудования для анало-

го-цифрового преобразования могут быть ограничены, но, как только данные представлены в цифровой форме, потребуется частотный диапазон $0 - F_{Гц}$, где $F = 1/(2T) = S/2$ — частота Найквиста. Она меняется при каждом переходе к иному выборочному интервалу, как это происходит, например, при децимации. К этому вопросу мы вернемся во втором томе.

Если обратиться к обработке данных в истинном масштабе времени, то поставленный вопрос приобретает смысл. В этом слу-

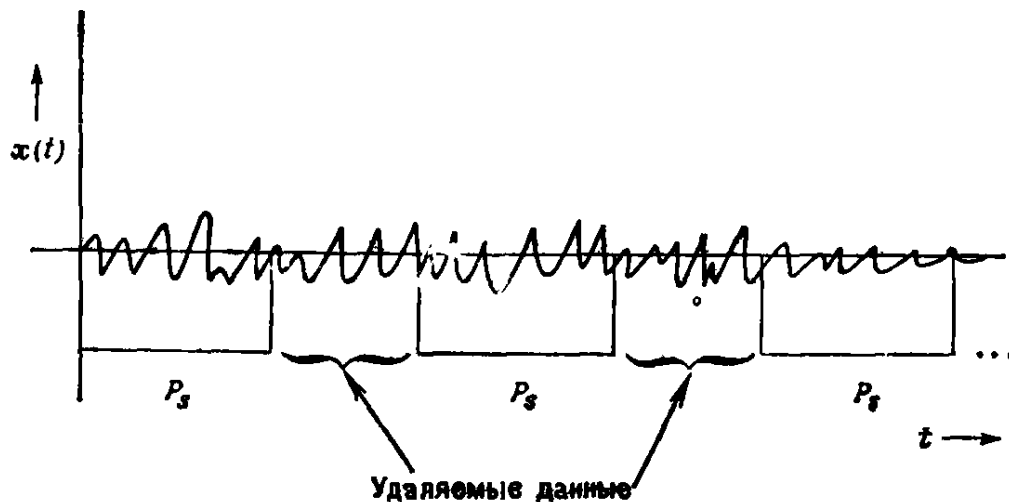


Рис. 8.3. Иллюстрация операции удаления данных при обработке в истинном масштабе времени.

чае на скорость, с которой могут обрабатываться данные, в совокупности влияют быстродействие АЦП, скорость выполнения команд ЭВМ и скорость, с которой дисплеи или периферийные устройства могут принимать данные. Кроме того, на скорость обработки влияет также установка скорости выборки, связанной с частотной полосой данных. Приведем пример. Допустим, что частота S равна 4000 Гц. Полагая, что для обработки берутся данные с полосой частот 1000 Гц, мы будем получать на период максимальной частоты по 4 значения. Для данных с полосой частот 2000 Гц будут получаться пары значений. Хотя в обоих случаях число данных для обработки одинаково, в одном случае будут обрабатываться данные с полосой частот 2000 Гц, а в другом — данные с полосой частот 1000 Гц.

Выше мы неявно предполагали, что обрабатываются все данные. Если скорости аналого-цифровых преобразователей, скорость запоминания ЭВМ и скорость работы программы обработки данных достаточны, то в машину можно вводить данные с высокой частотой выборки. Тем не менее вычисления БПФ, ПСМ и другие процедуры могут и не выполняться с такой скоростью. В этом случае вводится и обрабатывается один временной интервал, следующий за ним может выбрасываться, а затем обра-

батывается другой временной интервал, как это показано на рис. 8.3.

Мини-ЭВМ основных систем нередко снабжаются быстродействующими устройствами для БПФ. При оснащении мини-ЭВМ основных систем быстродействующими процессорами допустимая скорость обработки (ко времени написания этой книги) колеблется в интервале от нескольких сотен до примерно 50 000 выб./с. Аппаратные устройства могут достигать более высоких скоростей, во многих случаях свыше 100 000 выб./с. Быстродействующие процессоры, которые получают все большее распространение, могут обрабатывать до 1 000 000 выб./с. Несомненно, станут обычными скорости и в несколько миллионов герц.

Существуют также ограничения разрешающей способности, связанные с числом имеющихся ячеек памяти для запоминания данных. Нетрудно видеть, что эти ограничения определяются произведением $B_e P$. Запишем их в виде неравенства $B_e \geq 1/P$ или $B_e P \geq 1$, $B_e N T \geq 1$, $N B_e / S \geq 1$. Например, если $S = 4096$ выб./с, а $N = 4096$, то разрешающая полоса B_e не меньше 1 Гц. Если $S = 4.096$, а $N = 4096$, то $B_e \geq 0.001$ Гц.

8.5. СТАТИСТИЧЕСКАЯ УСТОЙЧИВОСТЬ

На статистическую устойчивость влияют два параметра — разрешающая полоса B_e и длина записи $P = NT$.

Можно показать (см., например, книги Отнеса и Эноксона (1972), гл. 5, Бриллинджера (1980), гл. 5), что в разумных предположениях оценки ПСМ приближенно имеют распределение χ_n^2 с n степенями свободы, где $n = 2B_e P$. Более того, для достаточно больших n , например $n \geq 30$, распределение χ_n^2 аппроксимируется гауссовским (нормальным) распределением. В этом случае нормированное стандартное отклонение (стандартное отклонение, связанное с оцениваемой величиной, т. е. процентная ошибка, или, в статистической терминологии, «коэффициент разброса») определяется соотношением $\epsilon_0 = 1/\sqrt{2B_e P}$. Величину ϵ_0 называют «стандартной ошибкой». Доказательства и дополнительные сведения можно найти в книге Отнеса и Эноксона (1972), гл. 5.

Формулы для чисел степеней свободы (ст. св.) и соответствующих стандартных ошибок каждого из трех методов сведены в табл. 8.1.

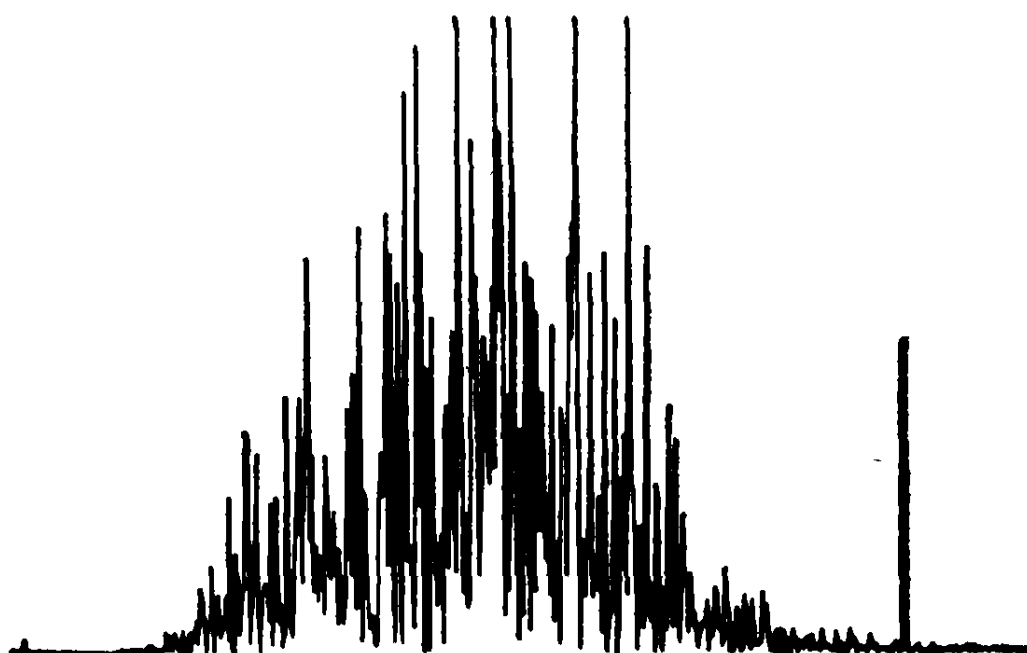
Эффект возрастания числа степеней свободы иллюстрируется на рис. 8.4. На всех пяти графиках изображены ПСМ белого шума после полоснопропускающей фильтрации и добавления синусоиды на выходе.

Для первого графика число степеней свободы $n = 2$, для второго $n = 8$, для третьего $n = 32$, для четвертого $n = 128$ и для

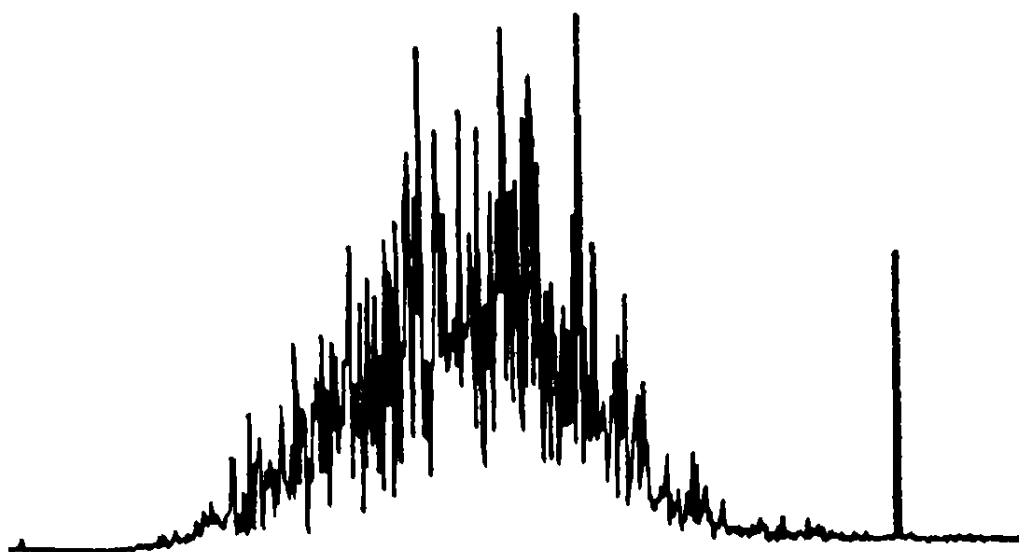
пятого $n = 512$. Стандартные ошибки имеют следующие значения:

График	Число ст. св.	$\epsilon_0 = 1/\sqrt{B_e P}$
1	2	1.0
2	8	.3
3	32	.2
4	128	.09
5	512	.05

Доверительные границы для спектра мощности можно определять (в зависимости от значения n) или по таблицам распределения χ^2 , или по таблицам гауссовского распределения. Например, полагая $n = 10$, $T = 1/2$ с, $N = 4000$, получим $P = 2000$ с,



(a) $M = 1$



(b) $M = 4$

Рис. 8.4. Пример статистической устойчивости.

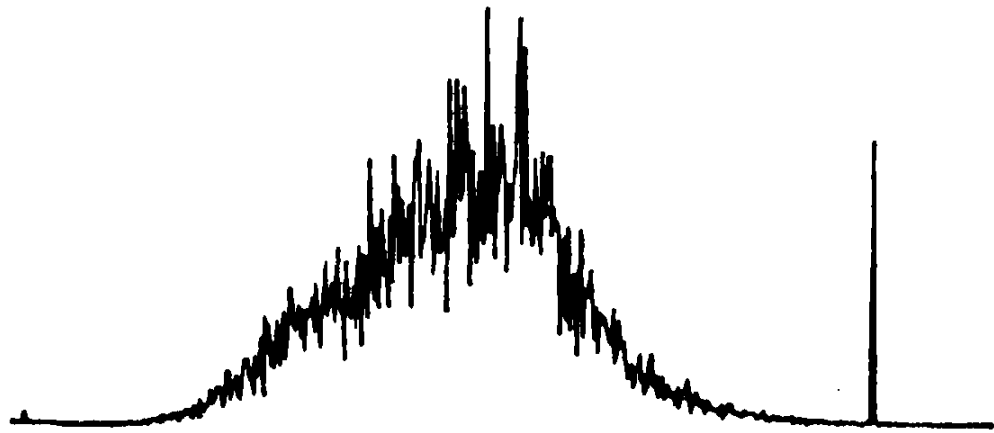
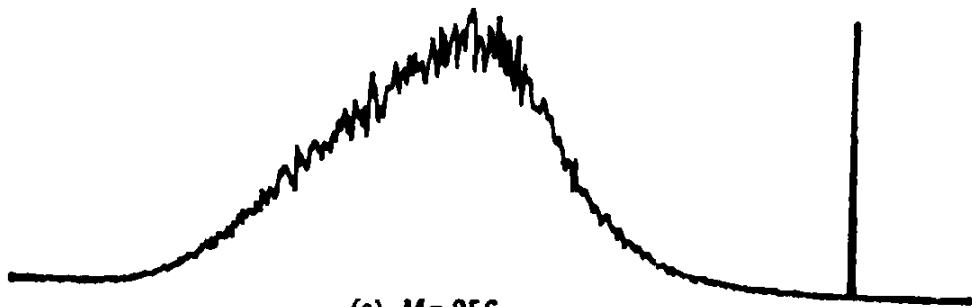
(c) $M=16$ (d) $M=64$ (e) $M=256$

Рис. 8.4. Продолжение.

Таблица 8.1

Число степеней свободы		
Метод получения ПСМ	Число ст. св. $n=2B_eP$	$\epsilon_n = 1/\sqrt{1/(B_eP)}$
Корреляции (Б—Т)	$n=2N/m$	$\sqrt{m/N}$
БПФ	$n=2M$	$\sqrt{1/M}$
Фильтрации	$n=2B_kP$	$\sqrt{1/(B_kP)}$

Здесь N — общее число значений данных, m — максимальное значение запаздывания в корреляционной функции, M — число осредняемых комплексных коэффициентов Фурье, B_k — точка половинной мощности полосы пропускания цифрового фильтра, $P=NT$ — длина записи

$F=1$ Гц, $B_e=1/400=0.0025$, так что $n=2B_eP=2\frac{2000}{400}=10$. В этом случае для 90-процентного ($=100 p$) доверительного интервала верхней доверительной границей будет служить величина $n\hat{S}(k)/\chi_{n; 1-\alpha/2}^2$, а нижней — величина $n\hat{S}(k)/\chi_{n; \alpha/2}^2$, где $\alpha=1-p$. По таблицам (см., например, таблицы Большева и Смирнова (1965) или табл. 5.4 в книге Отнеса и Эноксона (1972)) находим следующие значения: $\chi_{10; 0.975}^2=3.25$, $\chi_{10; 0.025}^2=20.48$. Следовательно, для верхней и нижней доверительных границ справедливы приближенные равенства

$$\frac{10\hat{S}(k)}{3.25} \approx 3\hat{S}(k) \quad \text{и} \quad \frac{10\hat{S}(k)}{20.48} \approx 0.5\hat{S}(k).$$

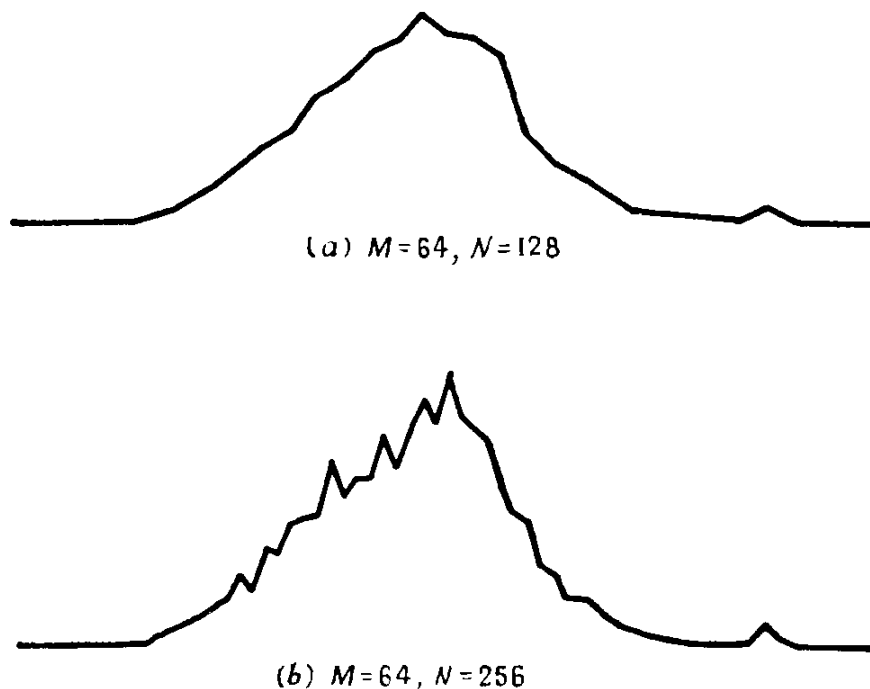
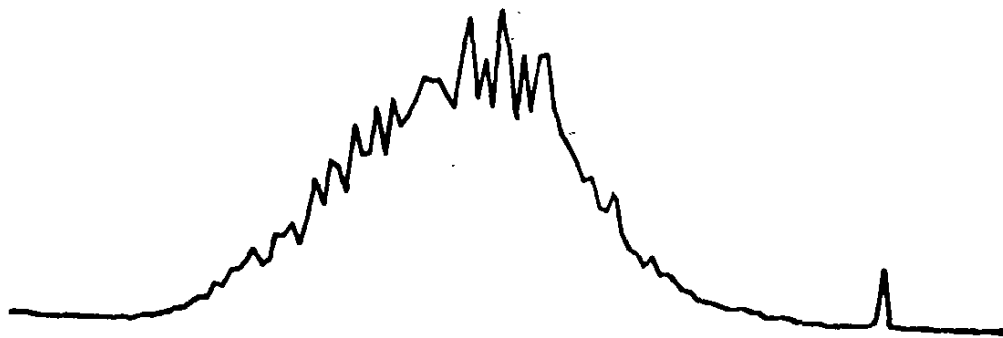
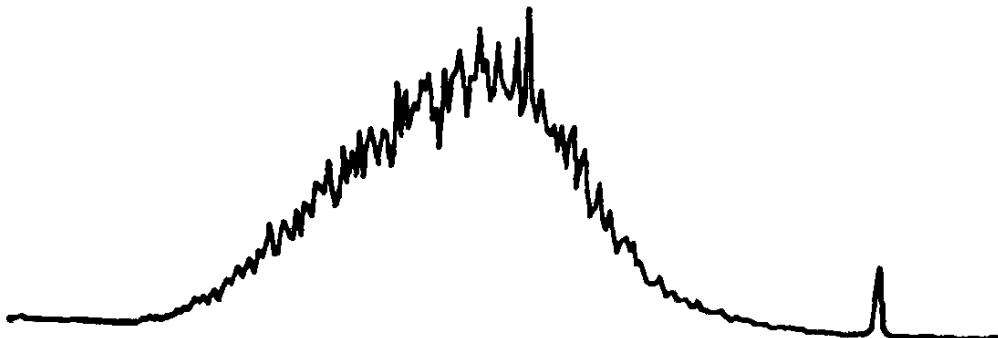


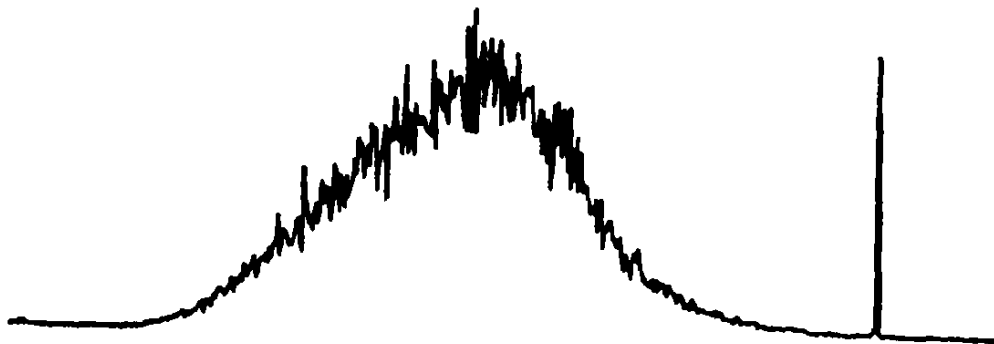
Рис. 8.5. Иллюстрация смещения (размывания), обусловленного слишком широким разрешением.



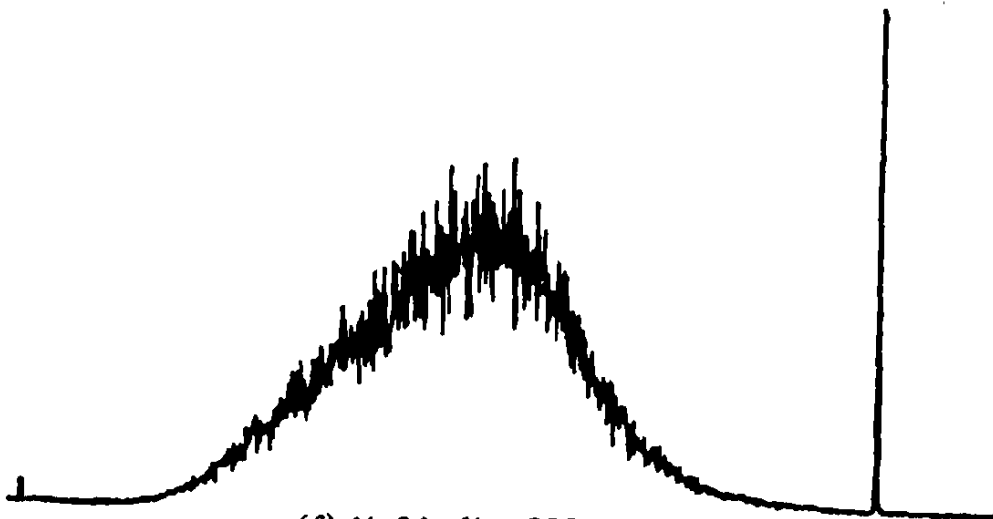
(c) $M=64, N=512$



(d) $M=64, N=1024$



(e) $M=64, N=2048$



(f) $M=64, N=4096$

Рис. 8.5. Продолжение.

Пример размывающего эффекта конечного разрешения. Если разрешающую полосу B_e , связанную с полосой частот откликов, выбрать широкой, то возникнет эффект размывания. Речь идет, выражаясь языком статистики, о смещении, относящемся к параметру в оценочной процедуре. Дать какую-либо количественную оценку этого эффекта чрезвычайно сложно. Вместе с тем появляющиеся искажения весьма наглядны и дают хорошее представление о качественной стороне явления. Обратимся к рис. 8.5. Здесь тот же полоснопропускающий фильтр, что применялся для получения графиков рис. 8.4, используется для фильтрации псевдо-

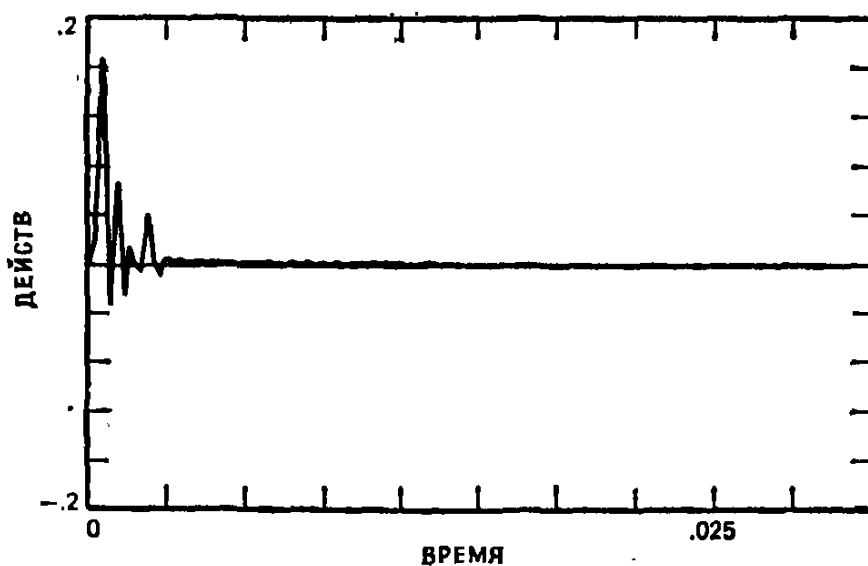


Рис. 8.6. Последовательность переходных данных с нулевым «хвостом» и ненулевым средним.

случайного белого гауссовского шума, затем на выходе добавляется синусоида. Хорошо заметно исчезновение деталей при низком разрешении. При 128-точечном разрешении синусоида исчезает почти полностью. С ростом разрешения она проявляется во все более и более сильном всплеске, достигающем максимума при $N = 4096$. Может показаться, что при большем разрешении данные имеют большую статистическую изменчивость. Но это не так. Проявление большей изменчивости объясняется не переменной статистической устойчивости — она остается постоянной, — а возрастанием числа точек, наносимых на одну и ту же площадь.

Удаление среднего. В спектральном анализе следуют полезному эмпирическому правилу — вычисляют значение выборочного среднего и вычитают его из соответствующей временной последовательности. Этот прием обоснован для анализа стационарных процессов. Однако для анализа переходных процессов он не-

приемлем. В качестве подтверждения такого вывода приведем пример (рис. 8.6). Хотя весь «хвост» процесса составляют нулевые значения, среднее оказывается положительным из-за вклада переходной части процесса. В этом случае вычитание среднего послужит причиной нежелательного искажения.

В приемах осреднения по временному промежутку, рассмотренных в этой главе ниже, имеются две вычислительные возможности — удаление среднего по всей временной последовательности и удаление среднего каждого отрезка ряда. Во втором случае величина $\hat{S}(k)$ должна быть равна нулю при $k=0$. Это связано с тем, что $\hat{S}(0) = E(\bar{x}_s - \bar{x})^2$ есть среднеквадратичное отклонение среднего по отрезку от общего среднего. Это отклонение может быть как не очень существенным, так и значительным.

8.6. УТЕЧКА

С помощью спектральных окон, на которых мы остановимся ниже, решается задача борьбы с явлением, получившим название «утечка». Утечка может происходить в двух направлениях. Во-первых, при вычислении мощности можно столкнуться с «вытеканием» в спектральной полосе, что приведет к смещению оценок мощности в сторону уменьшения. Во-вторых, спектральная оценка может сместиться в сторону увеличения в результате «втекания» (если утечке сопоставлены отрицательные веса, то смещение оценки даст заниженный результат).

Для начала продемонстрируем проявление эффекта вытекания. Рассмотрим плотность спектра мощности синусоиды, имеющей амплитуду A и частоту $f = nb$. В этом случае синусоида имеет целое число циклов. Из материала, изложенного в гл. 1 и 2, известно, что

$$S_x(k) = \begin{cases} \frac{A^2 P}{4} & \text{при } k = \pm n, \\ 0 & \text{при } k \neq \pm n. \end{cases} \quad (8.4)$$

График этой функции плотности приведен на рис. 8.7. Дополнив до вычисления преобразования исходную временную последовательность нулями, нетрудно получить более общий вид этой функции:

$$S_x(k) = \frac{PA^2}{4} \left[\frac{\sin(\pi(n-k))}{N \sin[\pi(n-k)/N]} + \frac{\sin(\pi(n+k))}{N \sin[\pi(n+k)/N]} \right]^2. \quad (8.5)$$

Заметим, что, интегрируя функцию $S_x(k)$, заданную формулой (8.4), мы получим среднеквадратичное значение синусоиды, равное $A^2/2$. Отметим также, что в этом случае все среднеквадратичное значение сосредоточено в двух точках ($k = \pm n$).

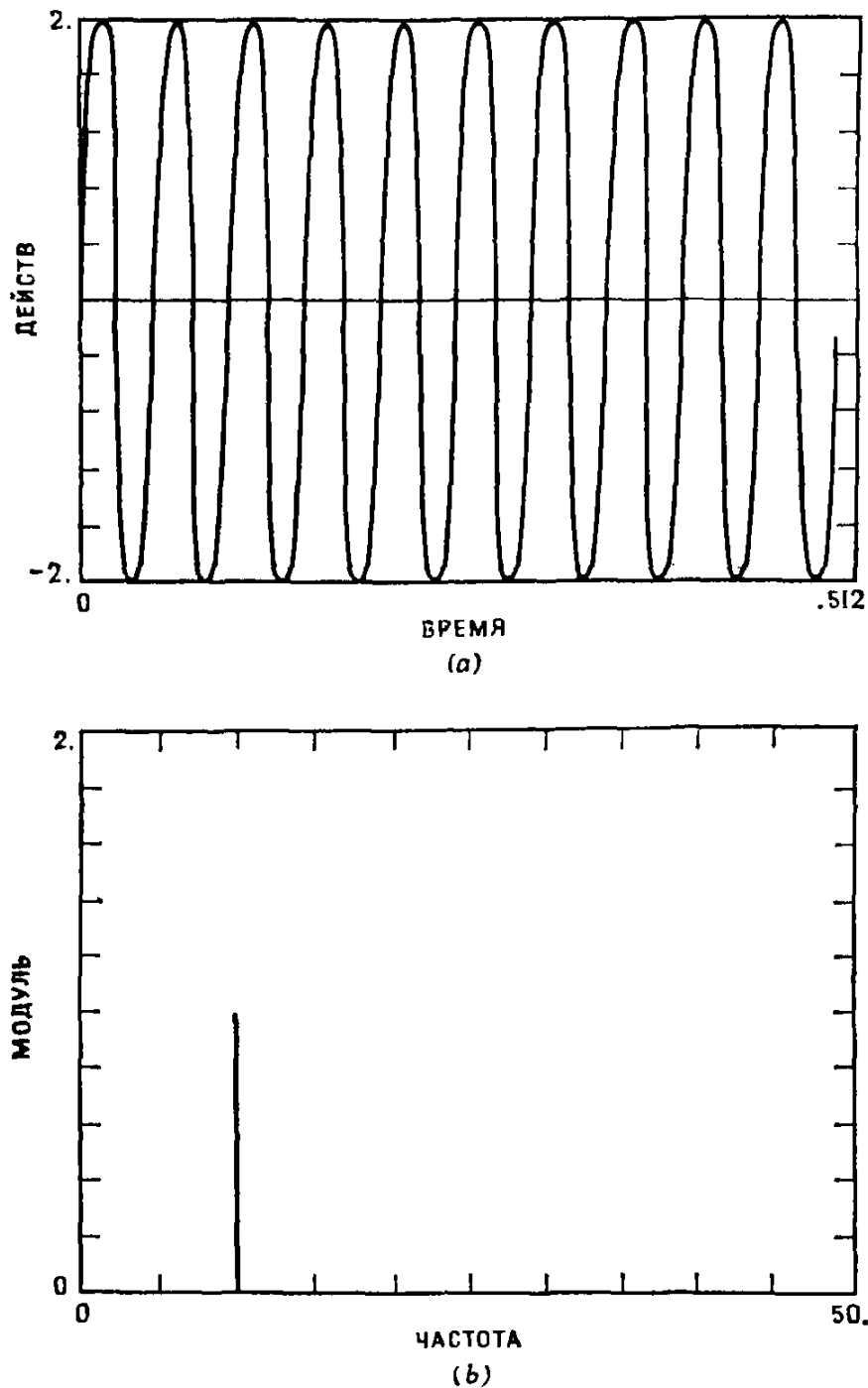


Рис. 8.7. ПСМ синусоиды с целым числом циклов (равным 10).

Предположим теперь, что к целому числу циклов синусоиды той же частоты добавлена еще половина цикла. При этом график ПСМ обретет вид, показанный на рис. 8.8. Это график функции, заданной формулой (8.5) и вычисленной в других точках. Полученная спектральная картина позволяет говорить о растекании мощности из частоты $(n + 1/2)b$. Если бы имелась еще одна синусоида, но с другой частотой, то мощность, которая вытекла из частоты $(n + 1/2)b$, добавилась бы к мощности второй

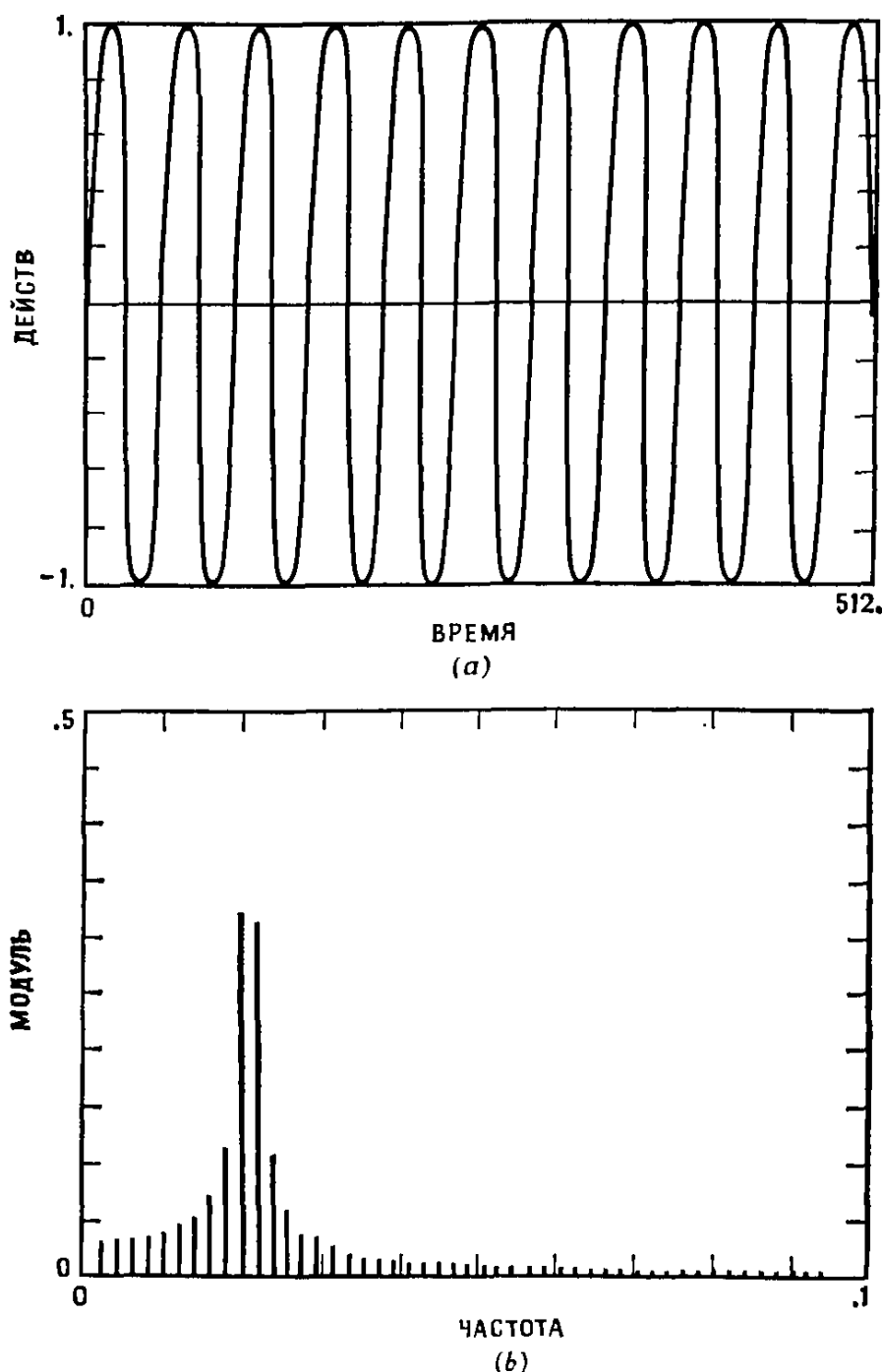


Рис. 8.8. ПСМ синусоиды, к целому числу циклов которой добавлена еще половина цикла (10.5 циклов).

синусоиды. В этом смысле можно говорить о втекании мощности. Отметим, что *при интегрировании по частотам общее средне-квадратичное значение остается неизменным.*

Применение спектральных окон позволяет изменить преобразования Фурье таким образом, чтобы эта утечка сократилась до минимума. Этой процедуре сопутствуют ухудшение разрешения (т. е. его расширение) и потеря ортогональности, что приводит к уменьшению степеней свободы.

Как только что было показано, утечки происходить не будет,

если в окно данных будут попадать только целые числа циклов. Вместе с тем не следует забывать, что очертания спектрального окна, почти повторяющие график функции $(\sin x)/x$, позволяют получить разрешение не лучше $b = 1/P$. Кроме того, следует принять во внимание, что при случайных данных частотные компоненты имеют случайные числа циклов, и, следовательно, целое число циклов будет встречаться редко, а утечка будет неизбежной.

Связь утечки с явлением Гиббса. Присутствие функции $(\sin x)/x$ в результатах обусловлено конечной длиной записи ряда. Такое ограничение длины достигается при помощи прямоугольного импульса, приводящего к разрывам. Вычисление преобразования Фурье равносильно приближению (подгонке) функции линейной комбинацией синусов и косинусов. Явление Гиббса отражает тот факт, что при такой подгонке из-за разрывов всегда возникают колебания вида $(\sin x)/x$. Если разрывы предварительно сглажены или «закруглены», то при подгонке ошибка уменьшится. Это сглаживание разрывностей достигается при помощи умножения на подходящую функцию во временной области. По теореме о свертке для преобразований Фурье такому умножению соответствует свертка (названная «спектральным окном») в области частот. Более подробно этот вопрос будет затронут ниже, при рассмотрении спектральных окон.

8.7. КАК ВЫЧИСЛЯТЬ СПЕКТРАЛЬНЫЕ ФУНКЦИИ

Среди перечисленных выше стандартных методов вычисления спектров мощности и кросс-спектров были названы методы Блэкмена и Тьюки (1958) (см. приложение В). Эти методы основаны на соотношениях Винера — Хинчина, позволяющих получать спектр как преобразование Фурье корреляционной функции. Одна из основных причин, по которой такому подходу отдавалось предпочтение, состояла в вычислительной эффективности, обусловленной возможностью получения удовлетворительного спектра по усеченной корреляционной функции: вместо N^2 операций, которые были ранее необходимы, чтобы получить спектр из преобразования Фурье исходных данных цифрового временного ряда длины N , для вычисления m значений корреляционной функции требуется лишь примерно Nm операций умножения — сложения. Из-за громоздкости с вычислительной точки зрения метод преобразования Фурье для получения спектра не находил широкого применения, поэтому связанные с ним практические вопросы никогда тщательно не изучались.

Ныне скорость алгоритмов БПФ превратила метод, прежде сбрасываемый со счетов, в самый быстрый. Хотя ряд авторов

отмечает пользу БПФ в получении корреляционной функции для последующего применения метода Блэкмена—Тьюки, чаще всего в таком подходе нет никакой необходимости. Пожалуй, нельзя найти веских причин, по которым следует отказываться от непосредственного получения спектра в области частот в пользу методов, использующих для этой цели корреляционные функции. Несмотря на известное различие результатов с чисто математической точки зрения, почти для всех практических нужд они оказываются одинаковыми. Некоторые тонкости, касающиеся этого вопроса, будут освещены во втором томе.

При использовании метода БПФ требуется около $4Nr$ операций, где $N=2^p$. Как правило, максимальное запаздывание t для корреляционной функции выбирается примерно равным $0.05N-0.1N$. Например, при $N=4096$, т. е. при $p=12$, получается выигрыш в скорости примерно в $\frac{50}{12}-\frac{100}{12}$ раз, т. е. примерно в 4—8 раз. Это главная причина, из-за которой отдается предпочтение непосредственному применению БПФ; в следующих параграфах будет рассмотрен ряд других преимуществ, связанных со сглаживанием спектра.

В гл. 6 были рассмотрены различные алгоритмы БПФ: для N , равного степени 2; для N , равного произведению степеней 2, 3, 4 и 5; для произвольно больших значений N , равных степени 2. Алгоритм для степени 2 реализован почти во всех системах для анализа Фурье, базирующихся на мини-ЭВМ, а также в аппаратных устройствах для БПФ (см. книгу Рабинера и Гоулда (1978)). В качестве примеров назовем миникомпьютерные системы Time Data и Hewlett-Packard и аппаратные системы Spectral Dynamics, и Nicolet Scientific¹⁾. Алгоритмы для произведения степеней 2, 3, 4 и 5 иногда включаются в программное обеспечение систем с более мощными ЭВМ, например University Software Systems²⁾.

Особого внимания заслуживают ограничения на длину последовательности. В принципе можно написать программу для обработки временного ряда произвольной длины N , причем скорость вычислений возрастает, если N не простое, а составное число. На практике, как правило, ограничиваются программами для случая $N=2^p$, и если длина записи ряда данных не равна целой степени двух, то ряд либо укорачивают, либо дополняют нулями. Возможен и другой подход, при котором временная последова-

¹⁾ Time Data, «TDA Time Series Analysis Systems», Santa Clara, Calif., 1973; Hewlett-Packard, «HP2451 Fourier Analysis System», Santa Clara, Calif., 1973; Spectral Dynamics, «SD501 FFT Analyzer», San Diego, Calif., 1974; Nicolet Scientific, «Omniferous FFT Analyzer», Northvale, N. J., 1974.

²⁾ University Software Systems, «MAC/RAN III, User Information Brochure», El Segundo, Calif., 1973.

тельность разбивается на более короткие временные подпоследовательности (возможно, перекрывающиеся), а получающиеся результаты осредняются. Однако во многих случаях, когда уже получен ряд сравнительно неудобной длины, отбрасывание данных может привести к потере значительной части информации. Более того, желание собрать определенное число данных может наталкиваться на дополнительные трудности.

Что касается области частот, то в ней может оказаться неудобным интервал, с которым чередуются линии спектра. Предположим, что аналого-цифровой преобразователь имеет скорость выборки S , равную 1000 выб./с, как это нередко бывает на практике. Тогда для записи ряда, длина которого равна 1024, уйдет 1.024 с, поэтому $B_e = 1/1.024 = 0.977$ Гц. Гораздо более удобной была бы в этом случае полоса 1.00 Гц. Поэтому во многих системах с мини-ЭВМ предусматриваются скорости выборок S , равные 256, 512, 1024 выб./с и т. д. (т. е. равные степени 2), что обеспечивает более удобные разрешения в области частот.

Обычно для оценивания спектров мощности, кросс-спектров и родственных им спектральных функций случайных данных выбирают один из трех способов. В первом из них проводится осреднение по области частот, во втором — осреднение по ансамблю, а в третьем способе обе эти возможности используются одновременно.

Под *осреднением по области частот* подразумевается осреднение соседних «несглаженных» спектральных оценок. Под *осреднением по ансамблю* понимается осреднение функций спектра мощности и кросс-спектра, взятых целиком; каждая из этих функций вычисляется для различных отрезков временной последовательности или пары временных последовательностей.

Основные выводы, касающиеся разрешающей полосы частот, длины записи и числа степеней свободы, без каких-либо изменений применимы к каждому из перечисленных методов. Однако некоторые особенности эффективных форм фильтров приводят к тому, что одних и тех же результатов получить нельзя, если обращаться к двум разным методам. Тем не менее при разрешении и числе степеней свободы, подогнанных друг к другу соответствующим образом, с чисто практической инженерной точки зрения результаты в целом оказываются эквивалентными.

При комбинированном подходе возможна любая последовательность чередования двух первых методов. Иными словами, можно сначала провести сглаживание каждого спектра в отдельности и потом выполнить осреднение по всему набору полученных спектров, а можно поступить и наоборот — сглаживать спектры, полученные после осреднения исходного набора спектров. Подробнее мы остановимся на этих методах в следующих параграфах.

8.8. ФУНКЦИИ ВРЕМЕННОГО СГЛАЖИВАНИЯ — ОКНА ДАННЫХ

Чтобы выделить определенные характеристики спектральных оценок, нередко прибегают к сглаживанию значений на концах случайного временного ряда. Временное сглаживание представляет собой умножение ряда на «окно данных»; эта операция аналогична умножению корреляционной функции на корреляционное окно. Поэтому временное сглаживание ряда эквивалентно применению операции свертки (см. гл. 7) к «несглаженному» преобразованию Фурье. Цель временного сглаживания, рассматри-

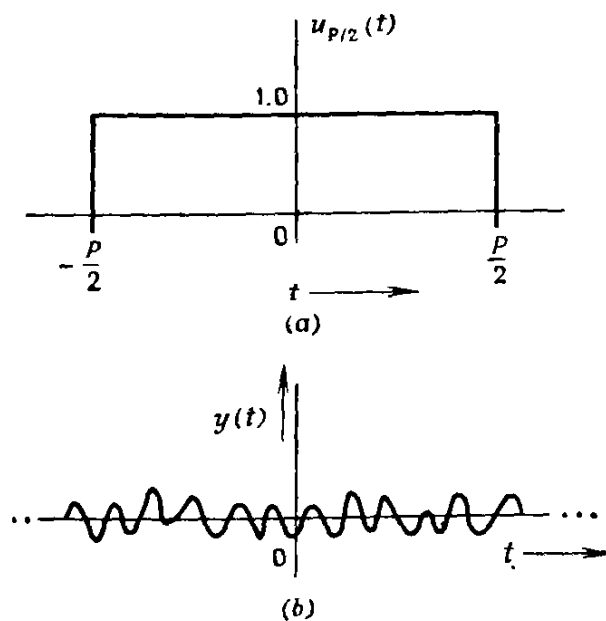


Рис. 8.9. *a* — прямоугольный импульс для БПФ; *b* — выборочная последовательность данных, имеющая длину P .

ваемая с точки зрения его проявлений в области частот, состоит в том, чтобы подавить влияние больших боковых всплесков эффективной формы фильтра, связанного с вычислением приближенного преобразования. Во временной области временное сглаживание служит для «округления» возможных разрывностей на концах конечного отрезка исследуемой временной последовательности.

Как и в случае корреляционной функции, на конечный отрезок случайного временного ряда можно смотреть как на произведение прямоугольного импульса $u_{P/2}(t)$, имеющего конечный носитель (рис. 8.9, *a*), и временной последовательности $y(t)$ бесконечной длины (рис. 8.9, *b*). Поэтому для конечного преобразо-

вания ряда $x(t)$ возможны два представления:

$$\begin{aligned} X(f) &= \int_{-P/2}^{P/2} x(t) \exp(-j2\pi ft) dt = \\ &= \int_{-\infty}^{\infty} y(t) u_{P/2}(t) \exp(-j2\pi ft) dt. \end{aligned}$$

Поскольку произведение функций преобразуется в свертку, имеем

$$X(f) = Y(f) * U_{P/2}(f),$$

где

$$U_{P/2}(f) = \int_{-P/2}^{P/2} u_{P/2}(t) \exp(-j2\pi ft) dt.$$

Без временного сглаживания эффективная форма фильтра имеет вид, показанный на рис. 8.10. Заметим, что ширина прямоугольного импульса равна P , где P — длина записи ряда.

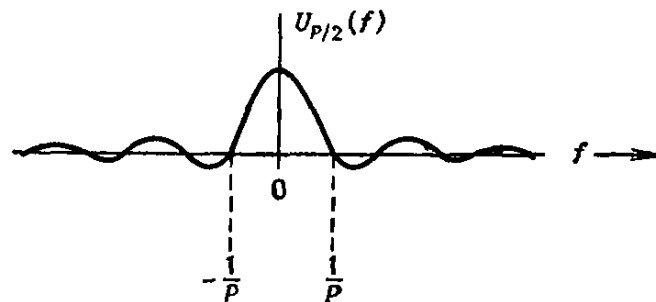


Рис 8.10. Эффективная форма фильтра без временного сглаживания.

Из-за больших боковых выступов мощность, содержащаяся в данных с частотами, некратными $1/P$, будет осредняться в величине, сосредоточенной в точке $f=0$. Несколько ухудшая разрешение, можно улучшить характеристики боковых выступов.

В работе Слоуна (1969) содержится вывод о неприемлемости временного сглаживания (названного «линейно модифицированным» оцениванием) для рядов гауссовских случайных данных. Не отрицая пользы этой операции для подавления утечки в случае данных «типа сигнала», он подчеркивает факт потери степеней свободы в случае чисто случайных данных. Эти замечания правильны, но, с нашей точки зрения, обычно шумовые данные обладают значительным сходством с данными «типа сигнала», и, следовательно, в типичных случаях целесообразно применять временные сглаживания некоторых типов. Мы считаем, что данные, встречающиеся на практике, имеют чаще всего такой же характер, как данные типа сигнала, даже если полученные данные — гауссовские. Например, искажение оценок, обусловленное

утечкой, создает серьезные препятствия при изучении вибрационных, акустических, сейсмических и большинства экономических данных, дающих неустойчивые формы спектров. Мы должны отметить, что в наших экспериментальных исследованиях искажение, вызванное утечкой, и размывание при конечном разрешении оказывались по крайней мере такой же важной проблемой, как и недостаточность имеющихся степеней свободы. В силу этого нам трудно согласиться с выводом Слоуна о том, что «линейная модификация приводит к ухудшению разрешения без компенсации в виде убывания дисперсии». Убывание дисперсии редко

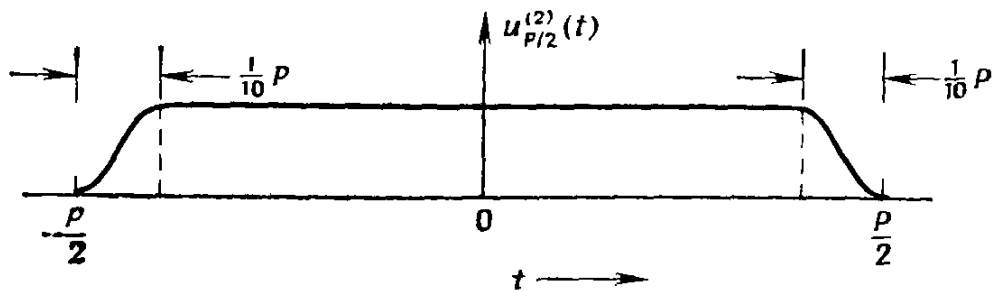


Рис. 8.11. Косинусное сглаживающее окно данных.

составляет проблему, в то время как подавление утечки служит полезным средством уменьшения ошибки.

В работе Бингхэма и др. (1967) для сглаживания на крайних отрезках ряда, составляющих примерно по 1/10 всей длины ряда, предлагается окно

$$u_{P/2}^{(2)}(t) = \begin{cases} \cos^2 \frac{5\pi t}{T}, & -\frac{P}{2} \leq t < -\frac{4P}{10}, \\ 1, & -\frac{4P}{10} \leq t < \frac{4P}{10}, \\ \cos^2 \frac{5\pi t}{T}, & \frac{4P}{10} \leq t < \frac{P}{2}, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Это окно носит название косинусного (рис. 8.11). Далее, нетрудно показать, что умножение на полный колоколообразный косинус

$$u_{P/2}^{(3)}(t) = \frac{1}{2} \left(1 + \cos \frac{2\pi t}{T} \right), \quad \text{где} \quad -\frac{P}{2} \leq t \leq \frac{P}{2},$$

эквивалентно применению дискретной свертки с весами $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$, т. е.

$$U_{P/2}^{(3)}\left(-\frac{1}{P}\right) = \frac{1}{4}, \quad U_{P/2}^{(3)}(0) = \frac{1}{2}, \quad U_{P/2}^{(3)}\left(\frac{1}{P}\right) = \frac{1}{4},$$

а во всех точках, кратных $1/P$ (тех, в которых вычисляется конечное дискретное преобразование Фурье), функция $U_{P/2}^{(3)}(f)$

равна 0. Тем не менее сокращение участков сглаживания до $(1/10)P$ на концах последовательности несколько меняет эту картину; получающаяся эффективная форма фильтра изображена на рис. 8.12.

Оба приведенных окна, согласно Слоуну, относятся к классу окон, имеющих преобразования Фурье вида

$$U_{P/2}(\omega, m) = \frac{\sin\left(\frac{2m-1}{2m}\frac{\omega P}{2}\right)}{\left(\frac{2m-1}{2m}\frac{\omega}{2}\right)} \left\{ \frac{\left(\frac{2\pi m}{T}\right)^2}{\left[\left(\frac{2\pi m}{T}\right)^2 - \omega^2\right]} \right\} \cos \frac{P}{4m}. \quad (8.6)$$

Из этой формулы при $m=5$ получаем $U_{P/2}^{(2)}(\omega)$, а при $m=1$ получаем $U_{P/2}^{(3)}(\omega)$.

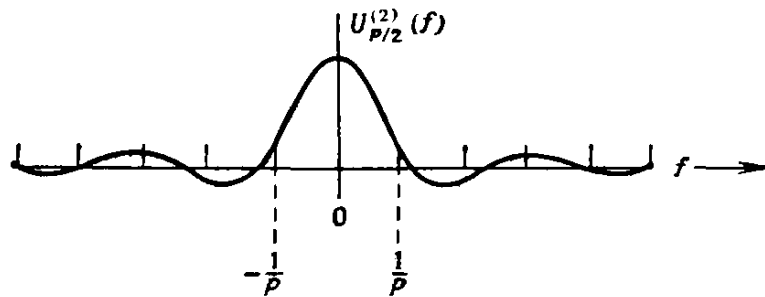


Рис. 8.12. Эффективная форма фильтра с косинусным временным сглаживанием.

Выбор функции временного сглаживания во многих отношениях аналогичен конструированию электрического фильтра в технике. Ниже будет рассмотрено построение двух специальных окон, в основе выбора которых лежит компромисс между допустимым объемом утечки, потерей разрешения и соответствующей потерей степеней свободы.

Окно Гудмэна — Эноксона — Отнеса (ГЭО). Первое окно, предложенное для ситуации, в которой необходимо сглаживание в приемлемо широкой полосе частот, носит название окна ГЭО. Оно играет роль сглаживающей функции не во временной, а в частотной области. С точки зрения вычислительных затрат сглаживания во временной и в частотной областях представляются примерно одинаковыми.

Сглаживанию во временной области (т. е. умножению) соответствует операция свертки в области частот. Для уменьшения вычислительных затрат желательно, чтобы число коэффициентов сглаживания было столь малым, насколько это возможно. По некоторым соображениям, в какой-то мере произвольным, было решено остановиться на семи весовых коэффициентах. Считается, что при этом достигается разумный компромисс между сложностью вычислений и гибкостью в выборе процедур сглаживания. Урав-

нение сглаживания имеет вид

$$\bar{X}(k) = \sum_{l=-3}^3 a(l) X(k+l) = X(k) + \sum_{l=1}^3 a(l) (X(k-l) + X(k+l)),$$

где

$$a(0) = 1, \quad a(-3) = a(3), \quad a(-2) = a(2), \quad a(-1) = a(1). \quad (8.7)$$

Чтобы избежать утечки, на эти веса накладывают условие вида

$$a(0) - 2a(1) + 2a(2) - 2a(3) = 0,$$

которое обеспечивает сокращение утечки мощности с больших расстояний пропорционально $1/f^4$.

Окно ГЭО имеет следующие коэффициенты¹⁾ (8.7):

$$a(0) = 1.0, \quad a(1) = 0.1817, \quad a(2) = -0.1707, \quad a(3) = 0.1476.$$

При таких коэффициентах спектральное окно оказывается на 10% шире, чем окно $(\sin x)/x$. Этому сопутствует потеря примерно 10% ст. св., что является приемлемым компромиссом в борьбе с боковыми всплесками.

Приведем для сравнения некоторые другие сглаживающие окна. На рис. 8.13 показаны основные спектральные окна для процедур сглаживания ГЭО, Хеннинга и 1/10-косинусной. На рис. 8.14 показано эффективное окно ПСМ при осреднении 12 оценок (что дает в спектральной оценке 24 ст. св.). На рис. 8.15 представлены соответствующие временные окна.

Критерий утечки. Для оценки относительной доли втекающей мощности, обусловленной боковыми выступами, был применен дополнительный критерий утечки, основанный на отношении общей мощности в сумме 12 спектральных окон к мощности главного выступа составного окна. Сумма 12 окон дает спектральную оценку с 24 ст. св., что в задачах, где оправдано применение спектрального анализа, считается близким к минимуму.

Отношение, на котором построен критерий, было без труда доведено до незначительных величин. При возрастании числа степеней свободы спектральных оценок это отношение становится только меньше. Для окна ГЭО его величина составляет 0.0073 (0.73%), а для окна Хеннинга — 0.0022 (0.22%).

Чтобы обеспечить, когда это требуется, свойство инвариантности дисперсии в результатах, следует принять дополнительные меры. Всякое окно данных изменяет дисперсию (или мощность) данных, так как разные участки временной последовательности

¹⁾ В этой формулировке считается, что начальная точка отсчета времени расположена посередине данных. Если эту точку перенести в начало, то у величин с нечетными индексами следует изменить знак, так как происходит поворот на 180°.

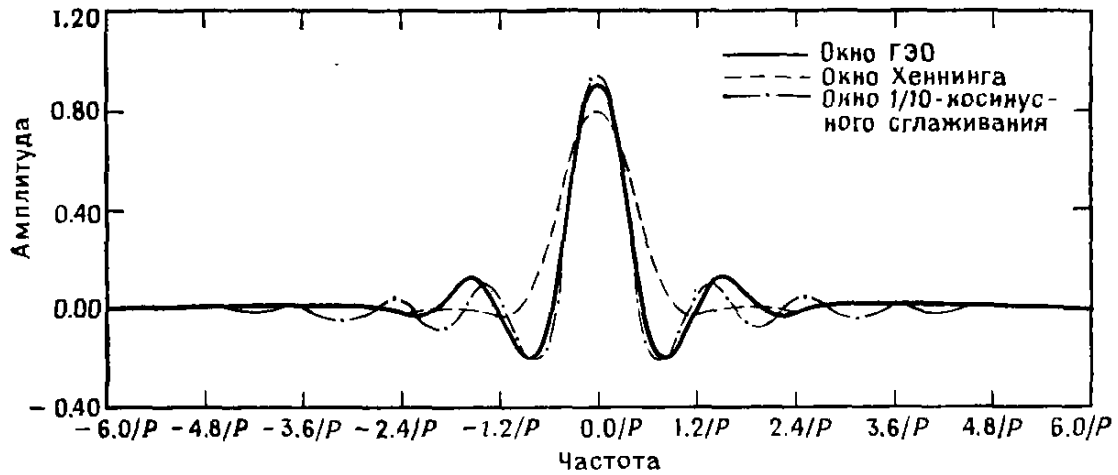


Рис. 8.13. Основные спектральные окна.

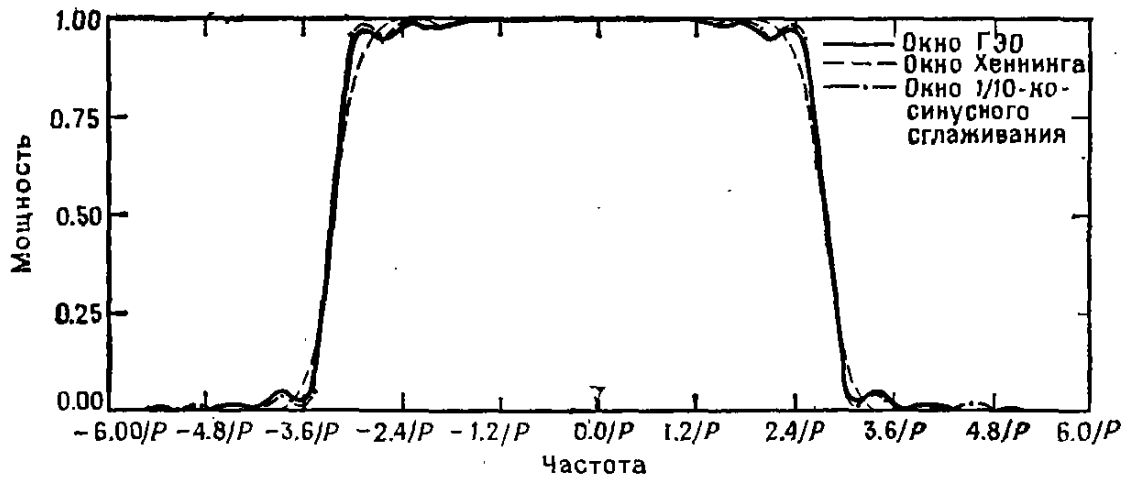


Рис. 8.14. Эффективная форма окна для ПСМ при осреднении 12 оценок основных типов.

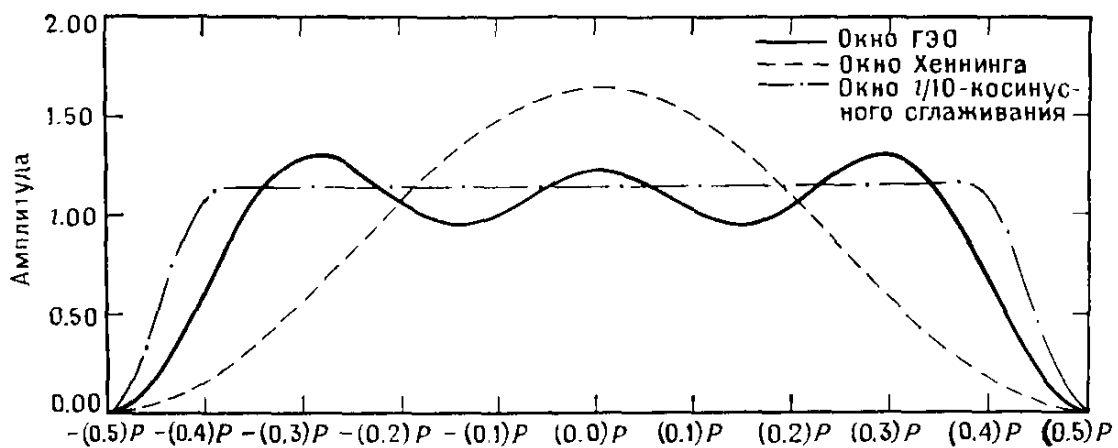


Рис. 8.15. Вид трех основных окон во временной области.

берутся с неодинаковыми весами. В результате происходит отбрасывание данных, которое приводит к потере ст. св. и к убыванию дисперсии. Величина этого сокращения характеризуется квадратом отношения площади, ограниченной функцией окна данных, к площади «прямоугольного» окна. Для окна Хеннинга, в котором первый коэффициент полагается равным 1.0, эта величина равна 0.375, для промежуточного способа Тьюки 0.875, для окна ГЭО 1.1679, Гудмэна 1.267. Чтобы площадь, ограниченная спектром мощности, оставалась инвариантной, нужно умножить этот спектр мощности на величину, обратную соответствующему множителю. С той же целью можно умножать временную последовательность или ее преобразование Фурье на квадратный корень соответствующего множителя.

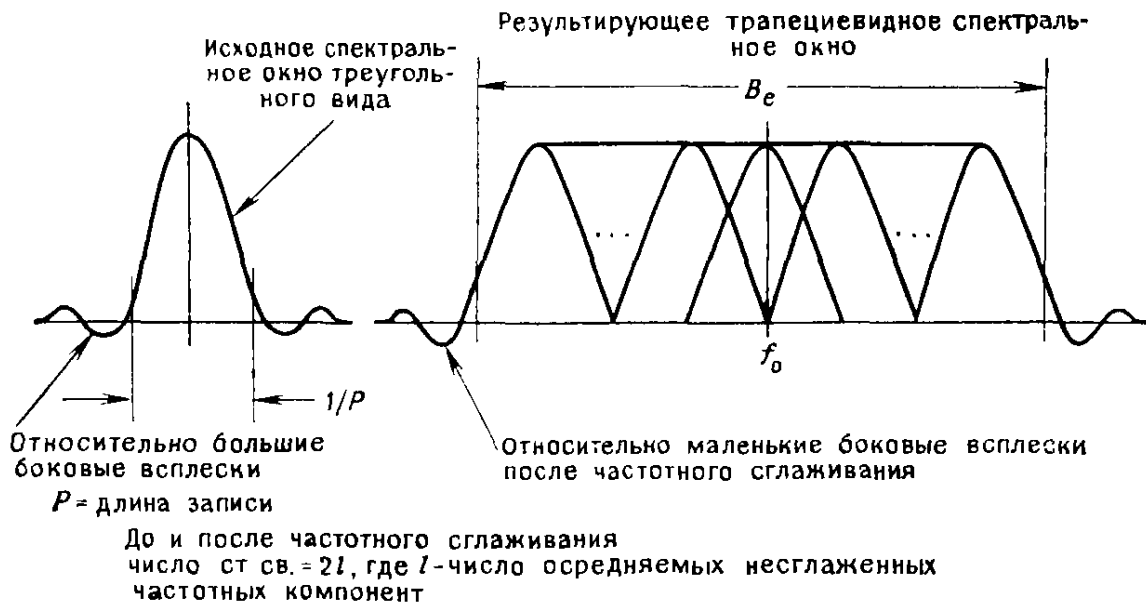
Окно для осреднения по ансамблю. При необходимости вычисления спектра мощности (или кросс-спектра) осреднением спектров, полученных по отдельным отрезкам временной последовательности, выбор окон обязан удовлетворять немного другим критериям. В этом случае более важной становится задача сокращения так называемой «соседней» утечки. При сглаживании в области частот ее решение сводится к тем мерам, которые только что рассматривались. Дело в том, что два первых относительно больших боковых всплеска составляли очень малую долю главного максимума, получаемого после сглаживания. Это объясняется тем, что главный максимум раздвигается и становится больше, в то время как боковые всплески сохраняют, по существу, один и тот же объем.

Ситуация меняется, если частотное сглаживание отсутствует. При осреднении по ансамблю главный максимум осредняется с главным, а боковой всплеск — с боковым; их соотношение остается постоянным. В этом случае нужно окно, которое в значительно большей степени подавляет влияние боковых всплесков. Чтобы достичь этого, придется пожертвовать частью ст. св.; предполагается, что потери ст. св. можно избежать за счет увеличения длины записи или обработкой не смежных, а перекрывающихся отрезков временной последовательности.

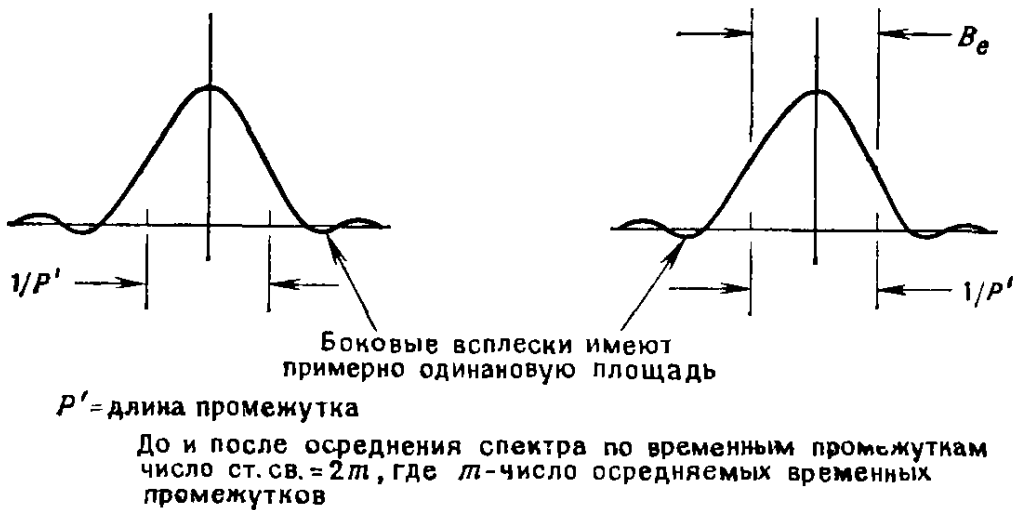
При перекрытиях часть данных используется дважды, что равносильно умножению на 2. Допустим, что хвосты данных при временном сглаживании уменьшены вдвое. Тогда множители 2 и 0.5 взаимно уничтожат друг друга, и это сведет на нет потерю ст. св. Конечно, это достигается ценой времени вычислений, что порой имеет существенное значение.

На рис. 8.16 показаны спектральные окна до и после сглаживаний.

Спектральное окно (неофициально предложенное Гудмэном) для осреднения по временным отрезкам (или осреднение по ан-



(a)



(b)

Рис. 8.16. Сглаживание плотностей спектра мощности, полученных посредством БПФ. *a* — осреднение в полосе частот; *b* — осреднение по ансамблю.

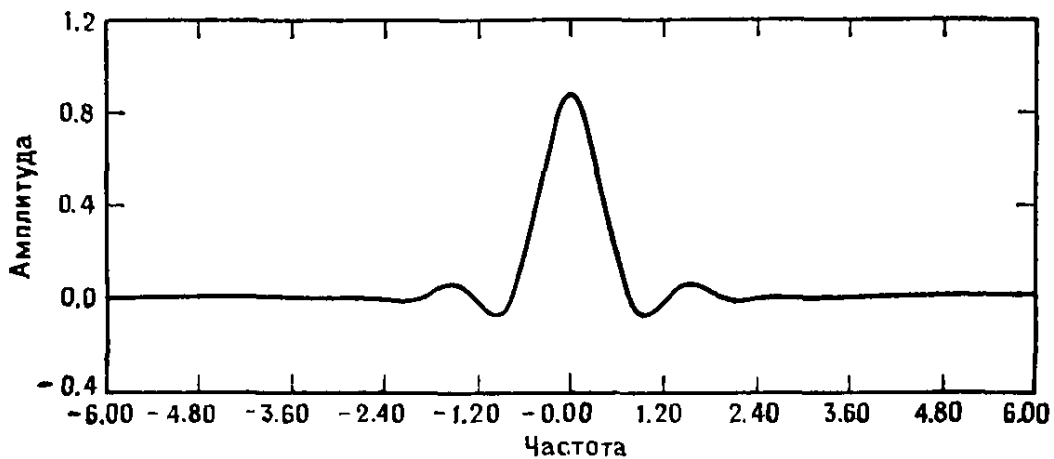


Рис. 8.17. Окно Гудмэна (в области частот).

самблю, см. ниже), построенное в отличие от окна ГЭО на менее формальной основе, имеет весовые коэффициенты

$$b(0) = 1.0, \quad b(1) = b(-1) = 0.35, \\ b(2) = b(-2) = -0.0875, \quad b(3) = b(-3) = 0.0625.$$

Для этого окна утечка составляет примерно 2%; когерентность между соседними оценками равна примерно 30%. Однако выступы, ближайšie к главному, имеют высоту, составляющую примерно

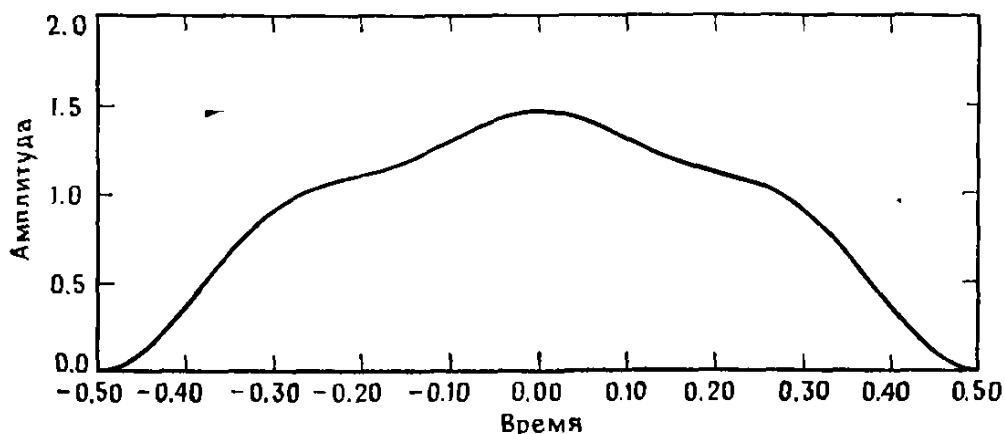


Рис. 8.18. Окно Гудмэна (во временной области).

2% высоты главного выступа. Форма этого окна представлена на рис. 8.17 (окно Гудмэна). Временной эквивалент окна показан на рис. 8.18.

Добавление нулей и интервал в области частот. Значения дискретного преобразования Фурье вычисляются с интервалом $b = 1/P = 1/(NT)$. Расстояния до первых пересечений нуля главным выступом окна $U_{P/2}$ также равны $B_e = 1/P$. Добавление нулей к последовательности никак не отражается на основной форме окна $U_{P/2}(f)$, поэтому ширина главного выступа остается неизменной. Однако, как следует из формулы для вычислений, шаг для оценок при увеличенной длине записи равен

$$b' = \frac{1}{(N + N_z)T},$$

где N_z — число добавленных нулей. Если, например, $N = N_z$, т. е. добавлено столько же нулей, сколько членов содержит исходная последовательность, то шаг уменьшится вдвое и возникнет картина, изображенная на рис. 8.19. Такое изменение шага вызывает неожиданные эффекты, поскольку свертка, применяемая к приближенному преобразованию Фурье, может и не обеспечить желаемого гашения боковых выступов. На самом деле эти выступы могут увеличиваться. Задачи модификации преобразований Фурье для тех случаев, когда необходимо добавлять про-

извольное число нулей, носят, по существу, такой же характер, что и конструирование фильтров эмпирическим способом. Эффект, к которому приводит добавление нулей, можно проследить на примере, предложенном в упр. 8.7.

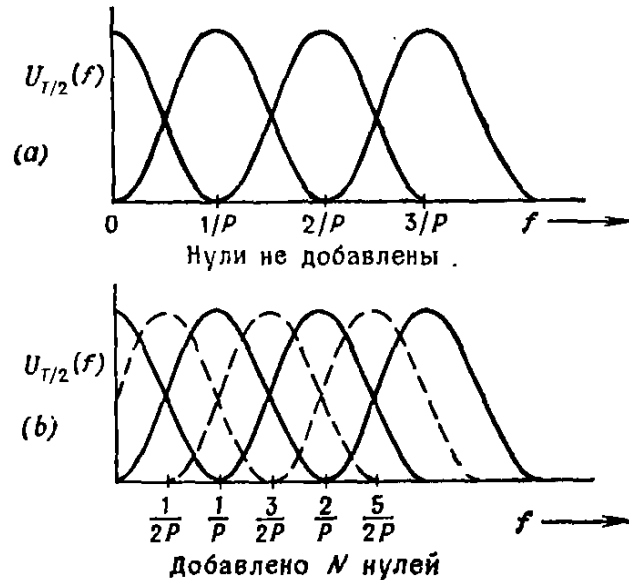


Рис. 8.19. Влияние добавления нулей на шаг чередования спектральных средних.

Спектр мощности получается из преобразования Фурье по формуле

$$\tilde{S}_x(k) = \frac{T}{N} |X(k)|^2 = \frac{T}{N} (\{\operatorname{Re}[X(k)]\}^2 + \{\operatorname{Im}[X(k)]\}^2),$$

где тильдой \sim снабжена «грубая», или «несглаженная», оценка.

Если число данных равно N , то грубые оценки разместятся так, как это показано на рис. 8.20. Можно показать, что грубые спектры мощности (см. эту главу выше) приближенно имеют распределение χ^2 с 2 ст. св. Иначе говоря, каждая точка спектра имеет такое же распределение, как сумма квадратов двух независимых случайных величин, имеющих гауссовские распределения. Стандартная ошибка оценок несглаженного спектра вычисляется по формуле из § 8.5 и равна

$$\varepsilon = \left(\frac{2n}{n^2}\right)^{1/2} = \left(\frac{2 \cdot 2}{2 \cdot 2}\right)^{1/2} = 1,$$

или 100%. Для многих целей такое значение неудовлетворительно. Оценки локально гладкого спектра, разделенные интервалом $1/P$, являются почти некоррелированными (здесь корреляция — перекрытие соседних оценок, см. рис. 8.19, а).

Способы вычислений. Для того чтобы добиться роста числа степеней свободы и сократить тем самым статистическую изменчивость, существуют три способа: осреднение по частотам, осред-

нение по ансамблю, комбинирование этих двух осреднений. Перейдем к объяснению этих методов.

Осреднение по частотам. Сглаженная оценка

$$\hat{S}(k) = \frac{1}{l} (\bar{S}(k) + \bar{S}(k+1) + \dots + \bar{S}(k+l-1)), \quad (8.8)$$

полученная осреднением l соседних оценок спектра, имеет распределение χ^2 с числом степеней свободы, равным примерно $2l$ (крышкой снабжена сглаженная оценка); это следует из теоремы

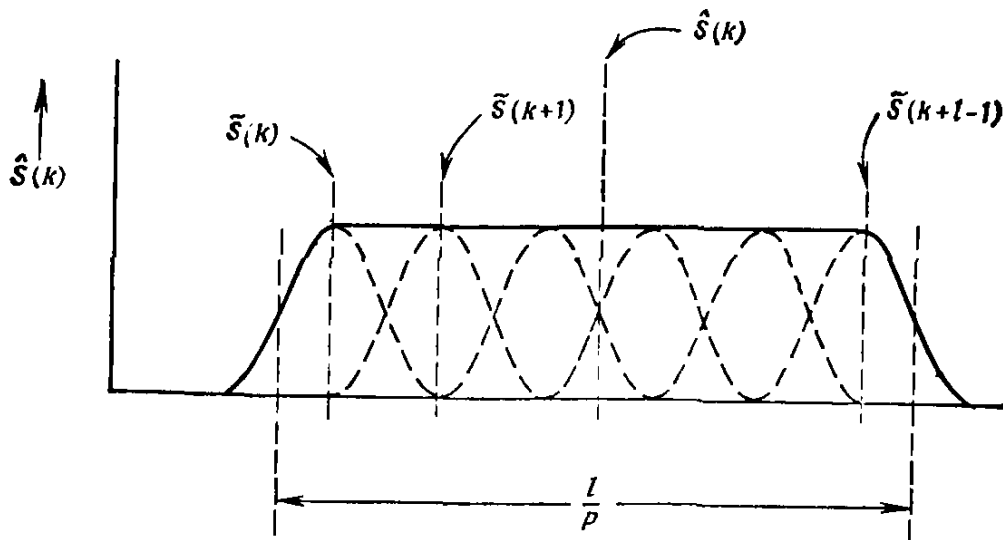


Рис. 8.20. Эффективная форма фильтра после осреднения.

о сложении величин, имеющих распределение χ^2 . Трапециевидная эффективная форма фильтра, изображенная на рис. 8.20, объясняется тем, что она складывается из треугольников, перекрывающихся в точках половинной мощности. Заметим, что эффективная полоса частот B_e теперь приблизительно равна $lb = l/P$.

Можно считать, что оценка $\hat{S}(k)$ располагается в середине промежутка частот от kb до $(k+l-1)b$. Этот метод получения дополнительных ст. св. путем сглаживания мы называем «осреднением по частотам».

Окончательные оценки $\hat{S}(k)$ можно расположить любым удобным способом. Если достаточно окончательных оценок, перекрывающихся в точках половинной мощности, то после осреднения и децимации их окажется N/l штук, причем смежные пары оценок будут заметно коррелировать друг с другом.

Простое скользящее арифметическое среднее в формуле (8.8) можно заменить сверткой более общего вида

$$\hat{S}_x(k) = \sum_{i=-m}^m a(i) \bar{S}_x(k-i), \quad \text{где } k=0, 1, \dots, N/2.$$

вида (ядер). Сглаживающее ядро можно строить, имея в виду операцию во временной области (в данном случае — окно смещений). При этом выборе «хорошей» формы низкочастотного фильтра будет отвечать сглаживающее ядро с очертаниями одной из типичных импульсных функций отклика, рассмотренных в гл. 4 и 5. Напомним, что эти функции принимают как отрицательные, так и положительные значения. Поэтому можно столкнуться с классическим парадоксом метода Б — Т — отрицательными спектральными оценками. Они могут возникать для спектров неустойчивой формы. Возможные неудобства, связанные с появлением таких оценок, отступают на второй план, поскольку в самом факте их наличия содержится информация о том, что исследуемый спектр быстро меняется, или о том, что его значения очень малы и близки к отрицательным.

Метод осреднения по частотам. В качестве процедуры получения приемлемых спектральных оценок для последовательности длины N может служить следующая.

1. Если позднее нужно будет получить корреляционную функцию, для чего понадобится обратное преобразование спектра, то начинают с дополнения исходной последовательности данных нулями до длины $2N$. (Будем считать, что этого не сделано.)

2. Вычисляется конечное преобразование Фурье последовательности данных:

$$\tilde{X}(k) = \sum_{i=0}^{N-1} x(i) W_N^{ik}, \quad \text{где } k=0, 1, \dots, N-1.$$

3. К несглаженному преобразованию Фурье применяется спектральное окно ГЭО, т. е. вычисляется свертка

$$X(k) = \tilde{X}(k) + \sum_{i=1}^3 a(i) [\tilde{X}(k-i) + \tilde{X}(k+i)],$$

где $k=0, 1, \dots, N-1$, $a(1) = 0.1817$, $a(2) = -0.1707$, $a(3) = 0.1476$. Крайние значения (для индексов, меньших нуля и больших $N-1$) вычисляют, исходя из периодичности преобразования $\tilde{X}(k)$: $\tilde{X}(N+k) = \tilde{X}(k)$.

4. Вычисляются квадраты абсолютных величин, которые нормируются так, чтобы получались «несглаженные» спектральные оценки мощности:

$$\tilde{S}_x(k) = \frac{1}{P} |X(k)|^2, \quad k=0, 1, \dots, \frac{N}{2}.$$

(Если нужно получить мощности в отрицательных частотах, то здесь необходимо взять в два раза более длинную последовательность.)

5. Чтобы подправить оценки, их умножают на коэффициент, отвечающий сглаживанию ГЭО:

$$0.856\tilde{S}_x(k) \rightarrow \tilde{S}_x(k), \quad k=0, 1, \dots, \frac{N}{2}.$$

Это необходимо делать, чтобы компенсировать убывание дисперсии, вызванное временным сглаживанием, отвечающим примененному спектральному окну. Приведенное конкретное значение коэффициента определяется из того факта, что центральный весовой коэффициент был выбран (в целях экономии вычислений) равным единице.

6. После этого сглаженные оценки получаются осреднением соседних несглаженных оценок по формуле

$$\hat{S}_x(k) = \frac{1}{2l+1} \sum_{j=-l}^l \tilde{S}_x(k+j),$$

где $k=0, 2l+1, 2(2l+1), \dots, m(2l+1)$. При этом считается, что полученные значения спектра, т. е. величины $\hat{S}_x(k)$, расположены в частотах

$$f=0, (2l+1)b, 2(2l+1)b, \dots, m(2l+1)b.$$

Отметим, что это завершающее сглаживание можно заменить сверткой более общего вида, а интервал чередования частот можно сделать более мелким, чем B_p . Нередко от предложенного способа вычисления свертки заставляют отказаться в пользу подхода Б—Т соображения экономии. Действительно, свертку в области частот иногда выгодно заменить умножением функций (в данном случае — окна смещения на ковариационную функцию) во временной области. При этом сначала при помощи обратного преобразования получают корреляционную функцию, умножают ее на соответствующее окно смещения, а затем получают посредством преобразования сглаженный спектр, сопровождая это при необходимости децимацией спектральной функции.

Теперь обратимся к другому методу вычислений.

Осреднение по ансамблю. Для длинных рядов данных более подходит метод, при котором вся временная последовательность разбивается на отрезки, и по ним определяется осредненный спектр. То название метода — осреднение по ансамблю, — которое уже встречалось раньше, вообще говоря, не совсем удачно, поскольку никакого ансамбля (множества) выборочных функций случайного процесса, по которым определяются статистические средние, не имеется. Взамен мы имеем лишь набор временных последовательностей, представляющих собой отрезки одной и той же временной последовательности (выборочной функции случайного

процесса). Следует отметить, что рассматриваемый метод наиболее часто реализуется в системах для анализа Фурье, построенных на мини-ЭВМ, и в аппаратных цифровых системах.

Получая статистики осреднением по набору отрезков ряда, мы неявно предполагаем, что исследуются стационарные процессы. На практике наши методы можно приспособить для неоднородных во времени процессов, статистические характеристики которых на протяжении отрезка меняются не слишком заметно.

Наиболее естественный подход к разбиению ряда должен основываться на увязке длины отрезков с ограничениями разрешающей способности. При этом выбору $B_e = b$ отвечает длина отрезков, равная $P_s = 1/b$; эту длину наращивают за счет добавления нулей до ближайшей степени 2. В качестве спектрального окна предлагается использовать окно Гудмэна, имеющее маленькие боковые выступы (рис. 8.17). Вычислительная процедура сводится к следующему.

Метод осреднения по ансамблю.

1. По значению разрешения вычисляется пробное значение длины отрезка ряда $P'_s = 1/b$. Определяется p — наименьшая степень 2, удовлетворяющая условию

$$\frac{P'_s}{T} = N'_s \leq N_s = 2^p.$$

Если $2^{p-1} < N'_s \leq 1.25 \cdot 2^{p-1}$, то длина отрезка берется равной $N_s = 2^{p-1}$; в противном случае полагают $N_s = 2^p$. Если выбрано значение $N_s = 2^{p-1}$, то вычисляется новое разрешение $B_e = b = 1/2^{p-1}$. (Множитель 1.25 выбран произвольно; в экспериментах авторов этот выбор хорошо оправдал себя.)

2. По формуле

$$m < \left[\frac{N}{N_s} + 1 \right],$$

где скобками обозначена целая часть числа, определяется общее число отрезков. Если длина записи фиксирована, как правило, получается один «короткий» отрезок длины $N_m = N - (p-1)N_s$. При необходимости этот последний отрезок увеличивают добавлением нулевых значений, доводя его до длины N_s .

3. Для завершающего осреднения определяются весовые коэффициенты, с которыми берутся все спектры. Для первых $(m-1)$ спектров эти коэффициенты получают по формуле $\omega = N_s/(mN'_s)$; последнему отрезку соответствует весовой коэффициент $\omega_m = N_s/(mN_m)$. В случае равных отрезков все веса полагаются равными $1/m$.

4. Для первой N'_s -точечной последовательности, дополненной, если необходимо, нулями до длины $N_s = 2^p$, вычисляется конеч-

ное преобразование Фурье

$$\bar{X}_l(k) = \sum_{i=0}^{N_s-1} x_l(i) W_{N_s}^{ik}, \quad l=1,$$

где l — номер отрезка.

5. К несглаженному преобразованию Фурье применяется спектральное окно Гудмэна ($k=0, 1, \dots, N_s-1$)

$$X_l(k) = \bar{X}_l(k) + \sum_{i=1}^3 b(i) [\bar{X}_l(k-i) + \bar{X}_l(k+i)],$$

$$b(1) = 0.35, \quad b(2) = -0.0875, \quad b(3) = 0.0625.$$

6. Вычисляя квадраты абсолютных величин и соответствующим образом нормируя их, получают «грубые» спектральные оценки

$$\tilde{S}_x(k) = \frac{1}{P_s} |X(k)|^2, \quad k=0, 1, \dots, \frac{N_s}{2}.$$

7. Полученные оценки подправляют, умножая их на коэффициент, отвечающий сглаживанию Гудмэна

$$(1.267) \tilde{S}_x(k) \rightarrow \bar{S}_x(k).$$

8. Для отрезков $l=2, 3, \dots, m$ повторяют шаги 4—7.

9. Осредняя m грубых спектров, вычисляют сглаженные оценки

$$\hat{S}_x(k) = \omega \sum_{l=1}^{m-1} \bar{S}_{sl}(k) + \omega_m \bar{S}_{xm}(k), \quad k=0, 1, \dots, \frac{N_s}{2} - 1.$$

Полученные $N_s/2$ оценок спектра (возможно, коррелирующих) считают расположенными в частотах

$$f = 0, b, 2b, 3b, \dots, \left(\frac{N_s}{2}\right) b;$$

каждая из оценок имеет примерно $n = 2m$ степеней свободы.

Предложенный метод часто применяют в тех экспериментальных ситуациях, где для всех практических целей в распоряжении оказываются ряды произвольной длины. В этом случае длину отрезков ряда обычно можно подобрать так, чтобы она вместе со скоростью выборки обеспечила нужное разрешение. Кроме того, скорость выборки можно установить так, чтобы величина B_e получила удобные целые значения. Например, если $N = 256$ и $S = 256$ выб./с, то

$$P = N \frac{1}{S} = 1 \text{ с}, \quad B_e = \frac{1}{P} = 1 \text{ Гц.}$$

Чтобы получить удобное разрешение, можно выбрать $S = 2.56, 25.6, 2560$ выб./с или целые кратные этим значения.

Другим значениям степеней 2^p будут отвечать другие значения скоростей выборок. Наконец, в этом случае реального времени веса отрезков всегда одинаковы и равны $\omega = 1/m$, где m — общее число отрезков.

Комбинированный метод. Оба предыдущих метода — осреднения по частотам и осреднения по ансамблю — представляют собой линейные операции общего вида. Поэтому их последовательное применение в любом порядке также будет линейной (усредняющей) операцией. При осреднении по m отрезкам с последующим усреднением l соседних спектральных оценок в итоге получаются оценки, число степеней свободы которых равно $n \approx 2lm$; разрешение в этом случае равно $B_e \approx l/P_s$. Поскольку линейные операции осреднения перестановочны, их последовательность не имеет значения — результат оказывается одним и тем же (без учета ошибок округления).

Необходимо ответить на два вопроса.

1. Как следует определять длину отрезка?

2. Какие спектральные окна должны использоваться?

Привести совершенно точные ответы на эти вопросы нельзя. В основу выбора длины отрезка естественно положить требование минимизации времени вычислений. Довольно близко к этому требованию подходит такое условие на выбор величины отрезка, при котором число нулей, добавляемых к отрезку, минимально, а разрешение так близко к выбранному аналитиком-исследователем, насколько это возможно. Лучшего приближения к нужному разрешению можно достичь, обратившись к более общему алгоритму для произведения степеней 2, 3, 4, 5 вместо варианта для степени 2.

Рассмотрим пример. Предположим, что при скорости выборки $S = 10\,000$ Гц необходимо получить полосу частот $B_e = 15$ Гц. Минимальная возможная длина отрезка равна

$$P_s = \frac{1}{B_e} = \frac{1}{15} = 0.0667 \text{ с.}$$

Ей соответствует число точек, равное

$$N_s = \frac{P_s}{T} = P_s S = 0.0667 \times 10\,000 = 667.$$

Для 2^p -алгоритма наименьшее возможное число, до которого следует добавлять нули, равно 1024; следовательно, в этом случае нужно добавить $1024 - 667 = 357$ нулей. В качестве другой возможности выступает случай $3 \times 667 = 2001$ точки, в котором до нужного числа остается добавить 23 нуля; чтобы получить при этом разрешение, близкое к заданному, можно проводить частотное сглаживание троек соседних спектральных величин.

При этом на последнем этапе должен получиться спектр в шагом

$$3b = 3 \frac{10\,000}{2048} = 14.65.$$

На самом деле разрешение окажется несколько шире и будет определяться приближенной формулой

$$B_e = 2b + 5 \approx 14.76.$$

Это объясняется тем, что конечные компоненты спектра вносят в полную ширину несколько больший вклад. Число ст. св. на последнем этапе зависит от всей длины записи и определяется числом спектров в осреднении по ансамблю. Нужное число ст. св. можно получить, увеличивая число осредняемых спектров до необходимых (или возможных) границ. Отметим, что частотное осреднение необходимо проводить только один раз, после всех осреднений по ансамблю. Это несколько сокращает время вычислений.

Нами рассмотрен следующий алгоритм выбора длины отрезка.

1. Вычисляется величина $P_s = 1/B_e$, где B_e — заданное разрешение.

2. Определяется пробное число точек $N_s = P_s/T$.

3. Целые кратные ему числа kN_s ($k = 1, 2, \dots$) сравниваются с 2^p .

4. Фиксируется наименьшее значение p , которое дает наименьшее число нулей. Число полос частот, по которым проводится осреднение, равно соответствующему k .

Для алгоритма с произведением степеней 2, 3, 4 и 5 третий шаг заменяется сравнением со значениями из табл. 6.2. На практике произведение степеней 2, 3, 4 и 5 редко не совпадает с нужной длиной в точности.

Кросс-спектры, передаточная функция и когерентность. По существу, приемы вычисления кросс-спектров и спектров мощности одинаковы. Передаточные функции и функции когерентности будут досконально разобраны в гл. 9, но основа их вычисления проста и будет рассмотрена сейчас.

Если необходимо получить функции плотности кросс-спектра, то начинают с вычисления $Y(k)$ — второго преобразования Фурье. Вслед за этим получают оценку спектральной плотности «несглаженного» кросс-спектра ($k = 0, 1, \dots, N/2$)

$$\tilde{S}_{xy}(k) = \frac{1}{P} X^*(k) Y(k) = \tilde{C}_{xy}(k) - j\tilde{Q}_{xy}(k).$$

Будем считать, что используются, если необходимо, те модификации преобразования Фурье, которые позволяют избежать утечки.

Нередко для одновременного вычисления двух преобразований применяют формулы (6.8) и (6.9).

Заключительное сглаживание можно выполнять любым из предложенных для ПСМ способом — осреднением по частотам, осреднением по ансамблю, комбинированным способом.

Как правило, для кросс-спектра вычисляют абсолютную величину $|\hat{S}_{xy}(k)|$ или ее квадрат $|\hat{S}_{xy}(k)|^2$, а также фазу $\hat{\theta}_{xy}(k)$ (в градусах), определяемые равенствами

$$\begin{aligned} |\hat{S}_{xy}(k)|^2 &= \hat{C}_{xy}^2(k) + \hat{Q}_{xy}^2(k), \\ \hat{\theta}_{xy}(k) &= \frac{360}{2\pi} \operatorname{arctg} [\hat{Q}_{xy}(k)/\hat{C}_{xy}(k)] \text{ (в градусах)}, \\ |\hat{S}_{xy}(k)| &= \sqrt{|\hat{S}_{xy}(k)|^2}, \quad k=0, 1, \dots, \frac{N}{2}. \end{aligned} \quad (8.10)$$

Получение кросс-спектра само по себе редко служит конечной целью. Чаще всего его вычисление бывает промежуточным этапом при получении тесно связанных с ним передаточной функции и функции когерентности. Передаточная функция определяется отношением ($k=0, 1, \dots, N/2$)

$$\hat{H}_{xy}(k) = \frac{\hat{S}_{xy}(k)}{\hat{S}_x(k)}.$$

Обычно ее записывают через модуль, называемый *коэффициентом усиления*, или квадрат модуля и фазу. Коэффициент усиления вычисляется по формуле ($k=0, 1, \dots, N/2$)

$$|\hat{H}_{xy}(k)| = \frac{|\hat{S}_{xy}(k)|}{\hat{S}_x(k)}.$$

Фаза передаточной функции совпадает с фазой кросс-спектра и, следовательно, вычисляется по формуле (8.10). Передаточная функция определяет соотношение между двумя временными последовательностями, выраженное в виде функции частот. Комплексные значения этой функции обычно записывают в полярных координатах — через коэффициент усиления (модуль) и фазу. К подробному изучению передаточной функции мы вернемся в гл. 9.

К кросс-спектру тесно примыкает функция когерентности, определяемая формулой ($k=0, 1, \dots, N/2$)

$$\hat{\gamma}_{xy}^2(k) = \frac{|\hat{S}_{xy}(k)|^2}{\hat{S}_x(k) \hat{S}_y(k)}.$$

В некоторых случаях с функцией когерентности связывают также фазу. В других случаях в качестве значений функции когерентности берут положительные значения квадратного корня;

при этом функция когерентности становится обобщением понятия коэффициента корреляции, встречающегося в классической статистике. Сдвигу фазы в 0° отвечает положительная корреляция; сдвигу фазы в 180° соответствует отрицательная корреляция. Для других фаз аналогов в элементарной статистике нет. Более широко эти вопросы будут рассмотрены в гл. 9.

Линейное осреднение. Все процедуры сглаживания для ПСМ, рассмотренные выше, относятся к классу *квадратичных* сглаживающих процедур. В некоторых ситуациях более точных спектральных оценок позволяет достичь осреднение по ансамблю временных промежутков, т. е. операция

$$\bar{x}(i) = \frac{1}{m} \sum_{l=1}^m x_l(i), \quad i = 0, 1, \dots, N-1. \quad (8.11)$$

Если промежутки статистически независимы, то предложенный метод применять бессмысленно из-за быстрого стремления суммы (8.11) к среднему значению процесса, которое в случае процесса с нулевым средним будет равно 0. Покажем, каким образом можно извлечь пользу из линейного осреднения в практически важных ситуациях со статистически зависимыми промежутками.

Отметим, что операции (8.11) эквивалентно осреднение преобразований Фурье, поскольку

$$\bar{X}(k) = \mathcal{F}[\bar{x}(i)] = \mathcal{F}\left[\frac{1}{m} \sum x_l(i)\right] = \frac{1}{m} \sum \mathcal{F}[x_l(i)].$$

Иначе говоря, осреднение и преобразование Фурье как линейные операции перестановочны. Принимая это во внимание, нетрудно понять, в чем заключается польза линейного осредне-



Рис. 8.21. Импульс и переходный отклик линейной системы.

ния. Мы ограничимся здесь объяснением его преимуществ лишь на интуитивном уровне, оставляя читателю возможность более полного знакомства с вопросом по работе Норина и Слоуна (1975).

Рассмотрим отклик физической системы на импульс, поступивший на вход. Очевидно, что возникающий переходный отклик полностью затухает по прошествии некоторого промежутка времени (рис. 8.21). Отличным примером отклика такого рода может

служить звучание колокола после удара его языка. Временную последовательность можно рассматривать как последовательность импульсов, разделенных бесконечно малыми интервалами (рис. 8.22). Поэтому при вычислении спектральных оценок ограничение пары временных последовательностей некоторым конечным промежутком времени неизбежно приводит к потере той части переходного отклика на выходе, которая соответствует последней части входа. Чем больше длина ограничивающего

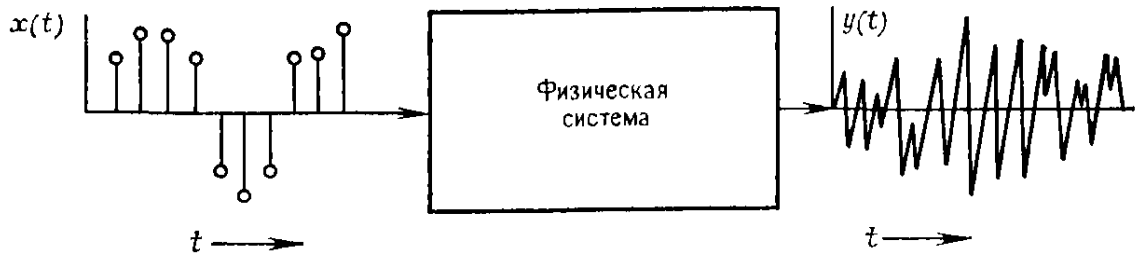


Рис. 8.22. Отклик на временную последовательность как на последовательность импульсов.

промежутка по отношению к длине переходного отклика, тем меньше ошибка, вызванная этим ограничением.

Для области частот рассмотренный эффект истолковывают как ошибку, обусловленную слишком широкой разрешающей полосой. В этой ситуации на первый план выходит тот факт, что линейное осреднение позволяет возместить ошибку, вызванную размыванием, даже при слишком широком разрешении.

Будучи связанными ограничениями, диктуемыми длиной отрезка, мы не можем получить дополнительных точек для спектральных функций и, следовательно, улучшить разрешения. В то же время можно минимизировать ошибку в тех точках, где вычисляются спектры. Помимо сокращения ошибок размывания мы получим *дополнительный выигрыш* — *уменьшим изменчивость!* Это объясняется тем, что при линейном осреднении посторонний независимый шум сглаживается.

Со статистической точки зрения оценки ПСМ и ПКС представляют собой функции второго порядка (квадратичные); их аналогами служат дисперсия и ковариация. Аналогом линейного осреднения служит оценка среднего значения. В случае некоррелирующих выборок данных оценка среднего значения просто сходилась бы к нулю. Если же выборки коррелируют, то получится более точная оценка импульсного отклика.

Напомним, что при обсуждении подмены частот подчеркивался следующий факт: если вычисляется обратное преобразование частотной функции (в частности, интересующих нас здесь в первую очередь передаточных функций), то в итоге получается искаженная (круговая) временная функция. Иными словами,

получается функция

$$h'(i) = \sum_{n=-\infty}^{\infty} h(i - nP), \quad \text{где } P = NT. \quad (8.12)$$

В этом случае

$$h'(i) = b \sum_{k=0}^{N-1} H(k) W_N^{-ik}, \quad \text{где } k=0, 1, \dots, N-1;$$

будем считать, что $H(k)$ — точные значения передаточной функции в тех точках, где она вычислена (рис. 8.23). Рассуждая в обратном порядке, нетрудно понять, что если при вычислении передаточной функции ограничиться взятым без изменений вре-

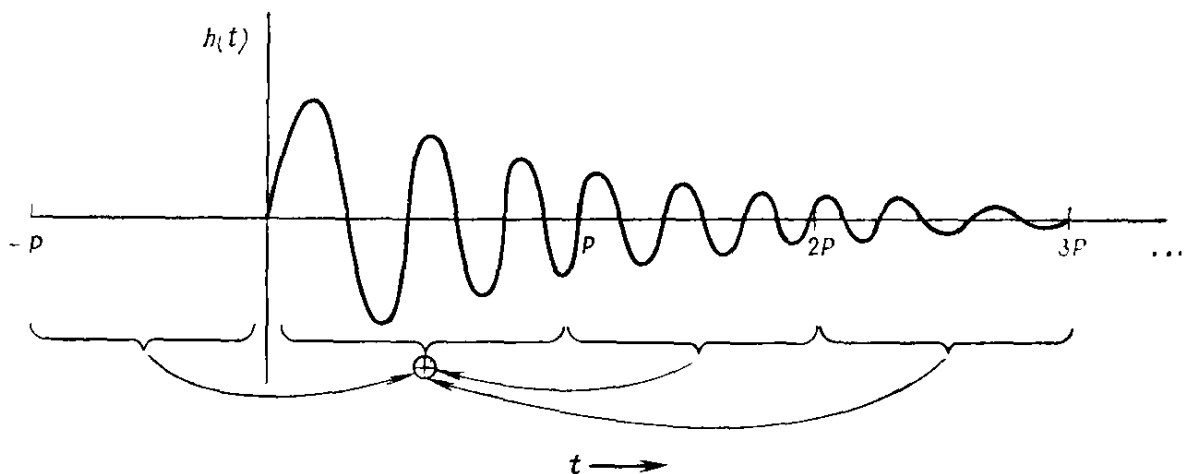


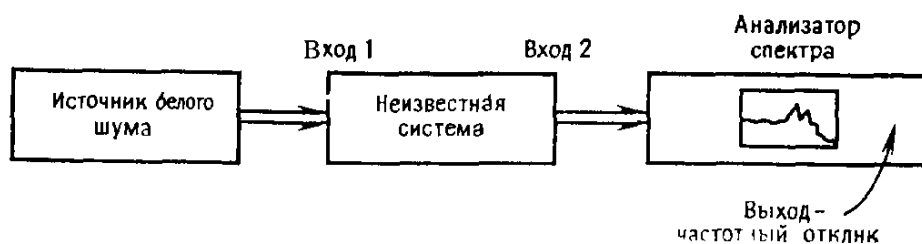
Рис. 8.23. Смещенная импульсная функция отклика. Все промежутки данных, имеющие длину P , добавляются к промежутку от 0 до P .

менным промежутком, длина которого по отношению ко времени затухания импульсного отклика мала, то возникнет ошибка. Ее можно компенсировать линейным сложением по отрезкам, определенным формулой (8.12). Это основной довод, оправдывающий использование линейного осреднения.

8.9. ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ ФУНКЦИЙ ПЛОТНОСТИ СПЕКТРА МОЩНОСТИ

Быстрые измерения частотного отклика. Поскольку у белого шума на всех частотах одна и та же интенсивность, его можно подать на один из двух входов системы. Этот шум будет выполнять роль управляющего сигнала. Второй вход можно исследовать при помощи анализатора плотности спектра мощности; в этом случае спектр выхода есть частотный отклик системы

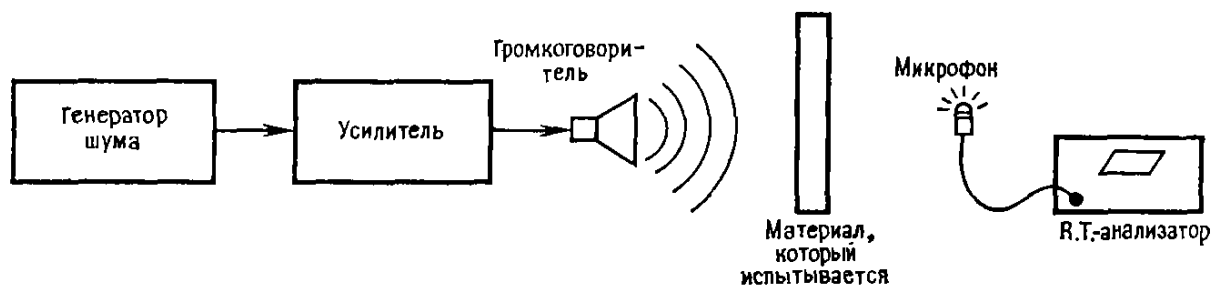
с двумя входами.



Чтобы получить для типичных генераторов развертки частотный отклик 10—20 Гц, требуется несколько минут. Анализатор спектра с истинным масштабом времени (R.T.-анализатор) может существенно ускорить процесс. К этому можно добавить, что генераторы развертки имеют ограниченные возможности и не покрывают слишком низкие частоты, а так как спектральные R.T.-анализаторы все чаще реализуются как ЦВМ, то при их использовании доступную границу низких частот определяет генератор шума.

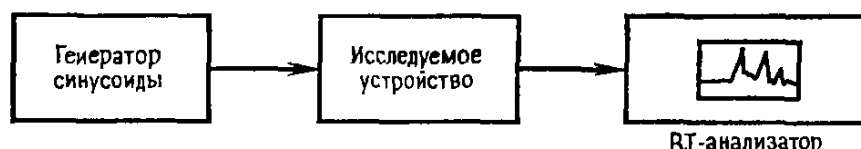
Измерение акустических свойств передачи/потери. Нередко возникает важная задача — исследовать характеристики передачи каких-то материалов. Допустим, что такого рода задачу нужно решить для некоторой металлической структуры, подобной стальной или стеклянной пластине; итогом исследования должна стать та частота, которая возбуждает данную структуру и заставляет ее вибрировать. Часто проводятся эксперименты более общего характера, в которых проверяются материалы, специально подобранные для глушения шума, подобно звукозащитным панелям. Для материалов этого типа, как правило, получают кривую затухания, которая характеризует зависимость от частоты величины потерь при прохождении звука через панель.

Измерения можно также производить со стороны громкоговорителя, вычитая их результаты из данных, полученных с другой стороны. Это позволяет исправить неточности, возникающие вследствие влияния частотных откликов усилителя и громкоговорителя.



Измерения гармонического искажения у систем со звуковыми частотами. При конструировании звуковых схем всегда пытаются подавить искажения, вызванные неидеальностью характе-

ристик электроники. Здесь важно как измерить эти искажения, так и решить вопрос о том, насколько результаты согласуются с поставленными требованиями. Как правило, гармоническое искажение вызывается третьей гармоникой основного тона. Это гармоническое искажение помогает измерить простой прием, основанный на применении Р. Т.-анализатора с истинным масштабом времени и генератора синусоидального сигнала.



Приближенную величину процента искажения ищут в виде уровня D между основной и третьей гармониками; так, уровню D , равному 40 дБ, отвечает искажение 1%, уровню D , равному 60 дБ,—искажение 0.1%. Точное гармоническое искажение можно получить, сопоставляя другие гармоники и суммируя их мощности; однако это дает небольшой выигрыш в точности.

Упражнения

8.1. Почему сделан вывод о том, что биномиальные веса сглаживания имеют в пределе очертания гауссовской плотности?

Ответ: это заключение основано на фундаментальной предельной теореме теории вероятностей о сходимости биномиального распределения к гауссовскому.

8.2. Предположим, что нужно провести анализ вибрационных данных; ожидаемое максимальное значение частоты компонент составляет $3k$ Гц. Допустим, что $S = 10\,000$ выб./с и что самая узкая ожидаемая полоса отклика предполагается равной около 40 Гц. Разрешение выбрано равным $B_e = 10$ Гц. Какие значения имеют:

- частота Найквиста F ;
- период следования промежутков P_s ;
- число данных в промежутке N_s ?

Сколько промежутков потребуется, чтобы получилось $n = 100$ ст. св.?

8.3. Когда в § 8.8 шла речь о частотном осреднении, утверждалось, что преобразование Фурье гауссовской функции есть гауссовская функция. Докажите это утверждение.

8.4. Докажите, что множитель $PA^2/4$, фигурирующий в формуле (8.5) для конечной дискретной ПСМ синусоиды частоты nb , совпадает с T/N , используемым в другом месте.

8.5. Приведите вывод формулы (8.5).

8.6. Приведите вывод формулы для временного варианта окна ГЭО с произвольным N , напишите программу для вычислений по этой формуле и получите график.

8.7. В § 8.8 при обсуждении шага частот и операции добавления нулей отмечалось, что добавление нулей может вызвать непредсказуемые эффекты, приводящие к изменениям очертаний спектральных окон. Используя результаты упр. 8.6 получите очертания окна ГЭО (в области частот). Затем получите для него преобразование Фурье после добавления $1/3N$ и $2/3N$ нулей и изобразите результаты в виде графика.

Глава 9

ПЕРЕДАТОЧНЫЕ ФУНКЦИИ И ФУНКЦИЯ КОГЕРЕНТНОСТИ

9.1. СВОЙСТВА ПЕРЕДАТОЧНЫХ ФУНКЦИЙ

Пользуясь для краткости в этой главе термином *передаточная функция*, мы будем большей частью относить его, строго говоря, к функциям частотного отклика. В частности, мы коснемся здесь вопроса о вычислении оценок передаточных функций, возникающих в анализе «линейных» моделей. В качестве такой передаточной функции в строительной механике может выступать функция, обратная функции механического сопротивления. В электротехнике она может представлять собой линейный фильтр, который в широком смысле понимается как всякая линейная операция, определяемая посредством интеграла свертки. В большинстве ситуаций передаточной функцией может быть просто какое-нибудь полезное линейное соотношение между двумя произвольными временными последовательностями, пусть и не дающее исчерпывающего описания связи. В этой главе затронуты линейные системы как с одним, так и с несколькими входами. Понятия, относящиеся к случаю многих входов, требуют знакомства читателя со статистикой многих переменных и анализом временных рядов. Поэтому большая часть материала предполагает довольно высокий уровень подготовки читателя; если он чувствует себя хотя бы в одной из названных дисциплин не очень уютно, можно при первом чтении книги этот материал пропустить. Это касается § 9.3, 9.4 и отчасти некоторых других параграфов.

Нужные понятия без труда переносятся (в теории, но обязательно на практике) на такие модели со многими переменными, как модели систем со многими входами и одним выходом. Этот вопрос рассмотрен в § 9.2.

Мерой точности и приложимости линейных соотношений служит исключительно полезный в исследованиях параметр — функция «когерентности»; она рассмотрена в § 9.4.

Рассмотрим физически реализуемую линейную систему, подобную системе, изображенной на рис 9.1 в виде схемы входа — выхода. Будем считать, что эта система не имеет никаких параметров, изменяющихся во времени. Как говорилось в гл. 1, весовая функция $h(\tau)$, связанная с этой системой, определяется как функция отклика системы (иначе говоря, ее выход) на

единичный входной импульс, представляющая собой функцию времени τ вхождения импульса. Для физически реализуемых систем необходимо, чтобы выполнялось равенство $h(\tau) = 0$ при $\tau < 0$, поскольку отклик должен следовать за входом. Полезность введения понятия весовой функции или, иначе, импульсной функции отклика объясняется тем фактом, что при произ-

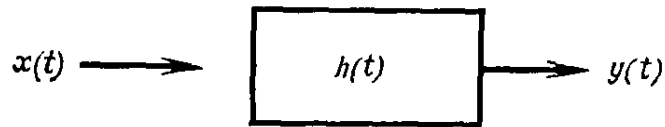


Рис. 9.1. Блок-диаграмма линейной системы с одним входом и одним выходом.

вольном входе $x(t)$ выход $y(t)$ системы можно задать посредством интеграла свертки

$$y(t) = \int_0^{\infty} h(\tau) x(t - \tau) d\tau. \quad (9.1)$$

Таким образом, значения выхода $y(t)$ для каждого момента времени t можно искать в виде линейной взвешенной (бесконечной) суммы значений всей предыдущей последовательности входа $x(t)$.

Другим способом характеризации линейных систем служит определение частотного отклика или, иначе, передаточной функции $H(f)$ этой системы. Передаточная функция определяется как преобразование Фурье функции $h(\tau)$, т. е. как преобразование

$$H(f) = \int_0^{\infty} h(\tau) e^{-j2\pi f\tau} d\tau. \quad (9.2)$$

Нижний предел интегрирования взят равным нулю вместо $-\infty$, так как $h(\tau) = 0$ при $\tau < 0$.

Идея физической реализуемости важна с точки зрения инженерного анализа реальных систем. Что же касается этого вопроса с математической, а иногда и с вычислительной точки зрения, то наиболее полезными оказываются физически нереализуемые варианты соотношений (9.1) и (9.2). В них вместо конечного нижнего предела интегрирования нужно подставить $-\infty$; в результате они принимают вид

$$y(t) = \int_{-\infty}^{\infty} h(\tau) x(t - \tau) d\tau, \quad (9.3)$$

$$H(f) = \int_{-\infty}^{\infty} h(\tau) e^{-j2\pi f\tau} d\tau. \quad (9.3')$$

Например, в указанном смысле не должны быть реализуемы нерекурсивные фильтры, используемые на цифровых вычислительных машинах. В данном случае применение симметричных весовых функций и использование фильтров с нулевым фазовым сдвигом (см. гл. 4) совершенно оправданно. Частотная функция отклика связывает переменные входа и выхода соотношением

$$Y(f) = H(f) X(f). \quad (9.4)$$

Это равенство получено применением преобразования Фурье к обеим частям соотношения (9.3) и составляет утверждение «теоремы о свертке». Частотная функция отклика представляет большой интерес, поскольку несет в себе информацию об усилении амплитуды (или ее затухании) и о сдвиге фазы. Так как функция $H(f)$ принимает комплексные значения, то для нее можно использовать тригонометрическую запись (т. е. запись в полярных координатах). Эта запись имеет вид

$$H(f) = |H(f)| e^{j\varphi(f)},$$

где

$$|H(f)| = \sqrt{(\operatorname{Re}[H(f)])^2 + (\operatorname{Im}[H(f)])^2}, \quad \operatorname{tg} \varphi(f) = \frac{\operatorname{Im} H(f)}{\operatorname{Re} H(f)}; \quad (9.5)$$

здесь абсолютная величина $H(f)$ есть коэффициент усиления, а аргумент $\varphi(f)$ — фазовый угол в радианах.

Указанное выше правило выбора знака фазового угла продиктовано желанием получить определение, согласованное с тем, которое принято в теории управления, а также в других областях, использующих идею частотной функции отклика. При таком определении отрицательные временные запаздывания (временные задержки) будут иметь фазу, представляющую собой прямую с отрицательным тангенсом угла наклона. Допустим, что функция $y(t)$ есть сдвинутая во времени функция $x(t)$, т. е. $y(t) = x(t - \tau)$. Тогда для преобразования Фурье функции $y(t)$ справедлива следующая цепочка равенств:

$$\begin{aligned} Y(f) &= \int_{-\infty}^{\infty} y(t) e^{-j2\pi ft} dt = \int_{-\infty}^{\infty} x(t - \tau) e^{-j2\pi ft} dt = \\ &= e^{-j2\pi f\tau} \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt = e^{-j2\pi f\tau} X(f). \end{aligned}$$

Поэтому для частотной функции отклика, коэффициента усиления и фазы выполняются равенства

$$H(f) = \exp(-j2\pi f\tau), \quad |H(f)| = 1, \quad \varphi(f) = -2\pi f\tau. \quad (9.6)$$

В частности, из последнего равенства следует, что фаза есть прямая с тангенсом угла наклона $-2\pi\tau$. Коэффициент усиления и фаза для рассмотренного случая изображены на рис. 9.2.

Сделаем предположение, согласно которому система, связывающая функции $x(t)$ и $y(t)$, «линейна». Когда говорят о такой системе, под линейностью подразумевают свойства, которым

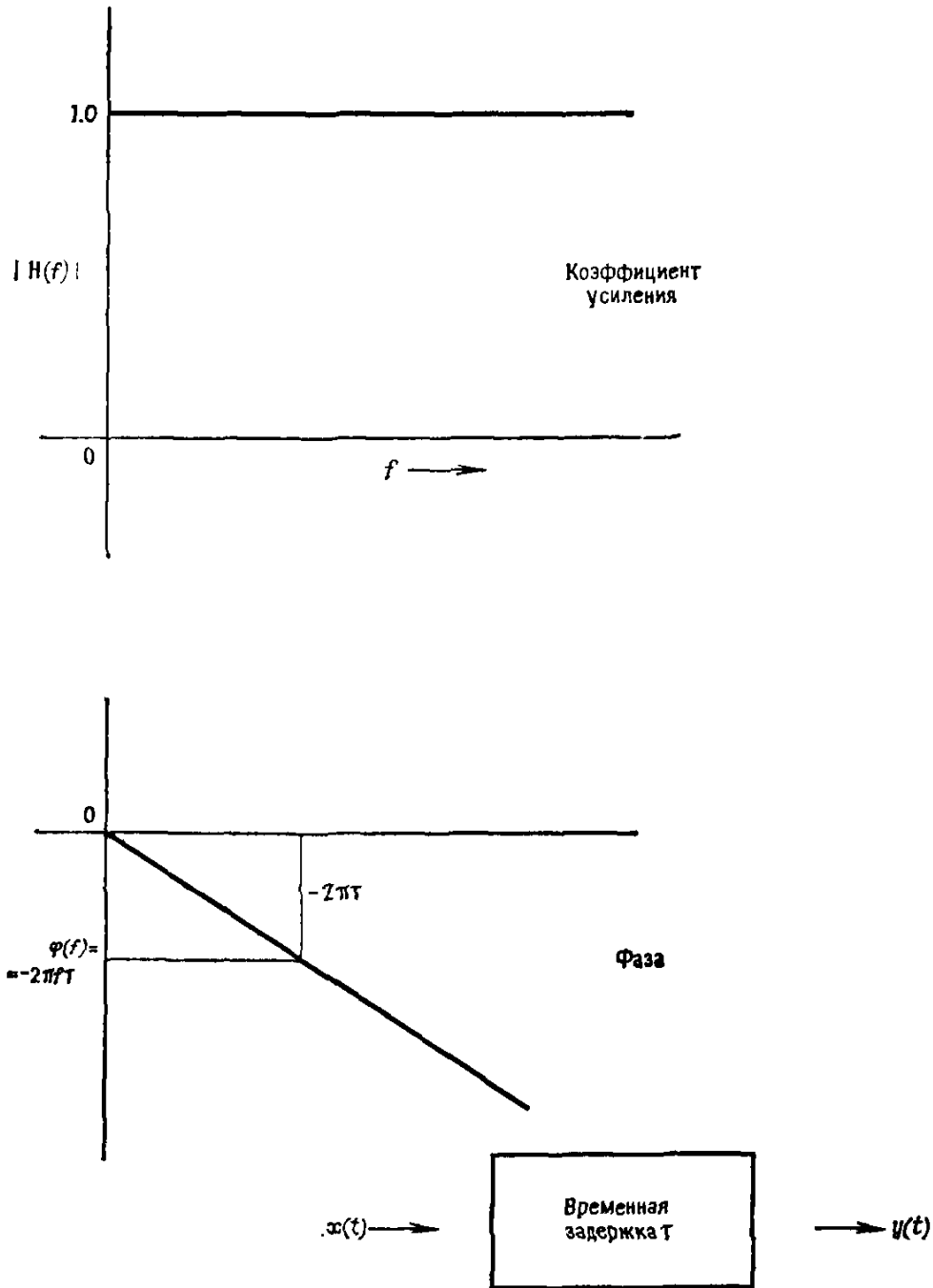


Рис. 9.2. Коэффициент усиления и фаза системы, состоящей только из временной задержки.

удовлетворяет линейный оператор L :

$$\begin{aligned} y(t) &= L[a_1 x_1(t) + a_2 x_2(t)] = \\ &= a_1 L[x_1(t)] + a_2 L[x_2(t)] = a_1 y_1(t) + a_2 y_2(t). \end{aligned}$$

Другими словами, система должна подчиняться аддитивному правилу, согласно которому сумма двух входов переходит в сумму соответствующих выходов, а также правилу однородности, в соответствии с которым входу, умноженному на константу, отвечает выход, умноженный на ту же константу.

Следует иметь в виду, что система может быть линейной, даже если ее «усиление» $|H_{xy}(f)|$ и «фаза» $\varphi_{xy}(f)$ будут нелинейными функциями частоты. Коэффициент усиления показывает

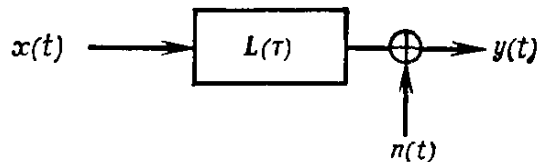


Рис. 9.3. Блок-схема линейной системы с одним входом и одним выходом; к выходу системы добавляется посторонний шум $n(t)$.

только, с каким усилением (или ослаблением) передается та или иная компонента. Что же касается фазы, то ее значения, нанесенные на график, могут показывать, каковы величины задержек компонент сигнала; для компоненты частоты f величина задержки определяется по формуле $\varphi_{xy}(f)/(2\pi f) = -\tau$, которую можно получить, решив уравнение (9.6) относительно τ и вспомнив, что положительную задержку мы рассматриваем как отрицательный сдвиг во времени. На рис. 9.3 показана блок-схема более общего вида. Она соответствует линейной системе с шумом.

На рис. 9.4 показано, как выглядят типичные графики коэффициента усиления (в линейных единицах, а не в децибелах) и фазы для (линейных) цифровых низкочастотных фильтров.

9.2. СПЕКТРАЛЬНЫЕ СООТНОШЕНИЯ ДЛЯ СИСТЕМ С ОДНИМ ВХОДОМ

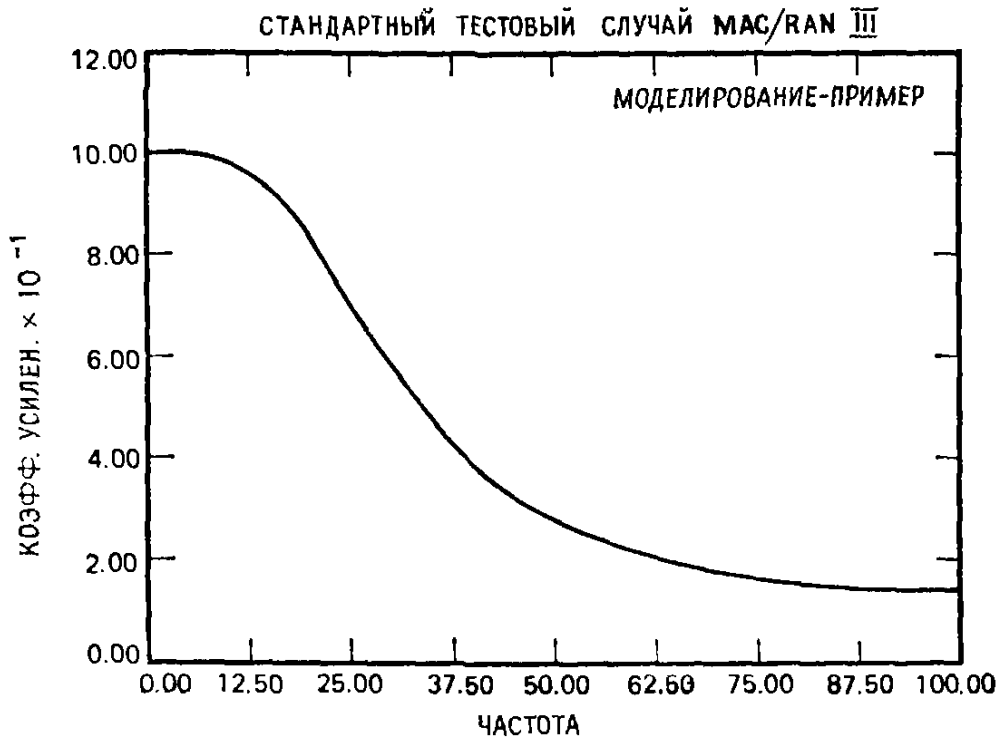
Предположим, что на вход линейной системы с одним точно определенным входом и одним выходом подается некоторый стационарный случайный процесс x с нулевым средним. Тогда, как показали Бендат и Пирсол (1971), выход y будет иметь те же свойства, что и вход. Функции плотности спектра мощности $S_x(f)$, $S_y(f)$ и функцию плотности кросс-спектра $S_{xy}(f)$ связывают соотношения

$$S_y(f) = |H(f)|^2 S_x(f) \quad (9.7)$$

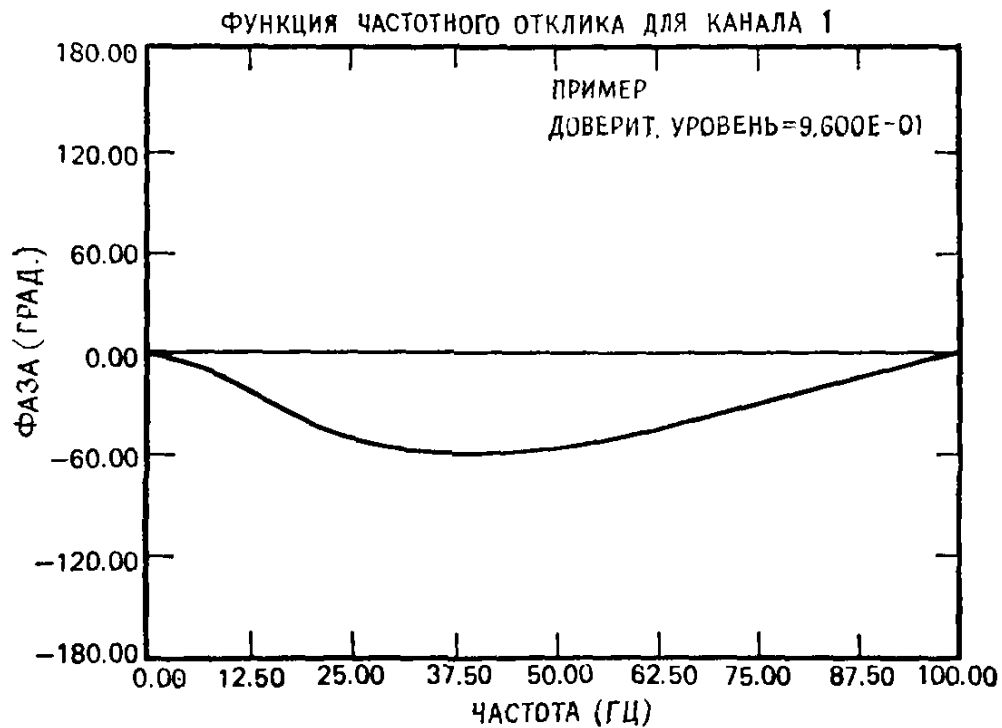
и

$$S_{xy}(f) = H(f) S_x(f). \quad (9.8)$$

Следовательно, если спектр мощности входа и кросс-спектр мощности известны, то частотная функция отклика, а значит, коэффициент усиления и фаза полностью определены.



(a) Коэффициент усиления



(b) Фаза

Рис. 9.4. Коэффициент усиления (a) и фаза (b) цифрового фильтра.

Функция когерентности $\gamma_{xy}^2(f)$ определяется формулой

$$\gamma_{xy}^2(f) = \frac{|S_{xy}(f)|^2}{S_x(f) S_y(f)} \quad (9.9)$$

и принимает только действительные значения. Некоторые авторы предпочитают называть когерентностью положительные значения $\gamma_{xy}(f)$ квадратного корня введенной величины. Можно показать, что функция плотности кросс-спектра мощности $S_{xy}(f)$ удовлетворяет неравенству (см. Дженкинс и Ваттс (1971—1972)) $|S_{xy}(f)|^2 \leq S_x(f) S_y(f)$, из которого следует, что $0 \leq \gamma_{xy}^2(f) \leq 1$. Соотношения (9.7) — (9.9) очень полезно сопоставить с формулами для регрессии и корреляции, приведенными в гл. 2. Нетрудно усмотреть, например, что передаточная функция $H(f)$ является аналогом коэффициента регрессии β и что спектр мощности и кросс-спектр служат аналогами дисперсии и ковариации соответственно. Далее, когерентность является аналогом квадрата коэффициента корреляции. Продолжает аналогию *комплексная когерентность*

$$\gamma_{xy}(f) = \sqrt{\frac{|S_{xy}(f)|}{S_x(f) S_y(f)}} e^{i\varphi_{xy}(f)},$$

представляющая собой обобщение коэффициента корреляции, причем положительная корреляция соответствует сдвигу фазы в 0° («в фазе»), а отрицательная корреляция соответствует сдвигу фазы в 180° («в противофазе»).

На протяжении этой главы мы не раз будем иметь возможность убедиться в том, что соотнесение результатов, касающихся области частот, и относительно простых понятий элементарной статистики сильно упрощает понимание многих глубоких фактов. Далее, для стационарных процессов благодаря свойству ортогональности преобразования Фурье в каждой узкой полосе частот оказывается возможным проведение независимого статистического анализа. Поэтому помимо соображений, которые касаются приложений в технике, а именно соображений о желательности работы с областью частот, что необходимо для определения таких важных параметров, как частота резонанса и т. п., есть также важный математико-статистический довод в пользу выбора области частот — упрощение анализа.

Теперь обратимся к задаче измерения плотности кросс-спектра идеальных линейных систем без шума. Для таких систем выполняется равенство $\gamma_{xy}^2 = 1$. Следовательно, функция когерентности достигает своего теоретического максимума, равного единице на всех частотах, в случае линейных систем. Если же функция когерентности не равна тождественно единице, то одной из возможных причин этого может служить отсутствие линейной зави-

симости выхода от входа рассматриваемой системы. Иными словами, эта система будет нелинейной.

Для того чтобы получить оценки передаточной функции и функции когерентности данных временных последовательностей x_i и y_i , где $i = 0, 1, \dots, N-1$, непосредственно применяют формулы (9.8) и (9.9). Эти формулы можно считать «естественными» формулами оценивания, хотя их нетрудно получить методом наименьших квадратов (см., например, книгу Дженкинса и Ватса (1971—1972)).

Приведем один метод вывода формул (9.8) и (9.9) (оказывается, что он приложим ко всем ситуациям, в которых возможен подход, использующий метод наименьших квадратов). Сначала запишем равенство, составляющее содержание теоремы о свертке:

$$Y(f) = H(f) X(f).$$

Затем, умножив обе его части на $X^*(f)$ (звездочка означает, что берется комплексно-сопряженная величина) и взяв от них математическое ожидание, получим соотношение

$$E[X^*(f) Y(f)] = E[X^*(f) H(f) X(f)]. \quad (9.10)$$

Поскольку передаточная функция для линейных систем является фиксированной характеристикой, функцию $H(f)$ можно вынести за знак математического ожидания. Тогда, согласно определениям, в левой части стоит кросс-спектр, а в правой — спектр мощности, помноженный на функцию $H(f)$. Таким образом,

$$S_{xy}(f) = H(f) S_{xx}(f).$$

Вычисления по этой формуле равносильны получению ковариации входа и выхода. Если же математическое ожидание брать от квадратов обеих частей исходного равенства и, следовательно, исходить из соотношения

$$E|Y(f)|^2 = E[|H(f) X(f)|^2],$$

то получится формула

$$S_{yy}(f) = |H(f)|^2 S_{xx}(f).$$

Все эти формулы имеют свои аналоги для дискретных спектральных функций, соответствующие случаю конечной длины записи. Никакой особенной разницы между непрерывным и дискретным случаями нет. Исключение составляют такие ошибки, которые, встретившись при вычислении преобразования Фурье (например, из-за подмены частот), повлекут за собой ошибки при вычислении функций, получаемых из этого преобразования.

9.3. СПЕКТРАЛЬНЫЕ СООТНОШЕНИЯ ДЛЯ ЛИНЕЙНЫХ СИСТЕМ СО МНОГИМИ ВХОДАМИ

Теперь обратимся к модели линейной системы, отвечающей случаю многих входов. Предположим, что число входов равно p , а число выходов — одному. Важную роль в изучении таких систем имеют функции обычной, множественной и частной когерентности. О способах их вычисления будет рассказано в § 9.4.

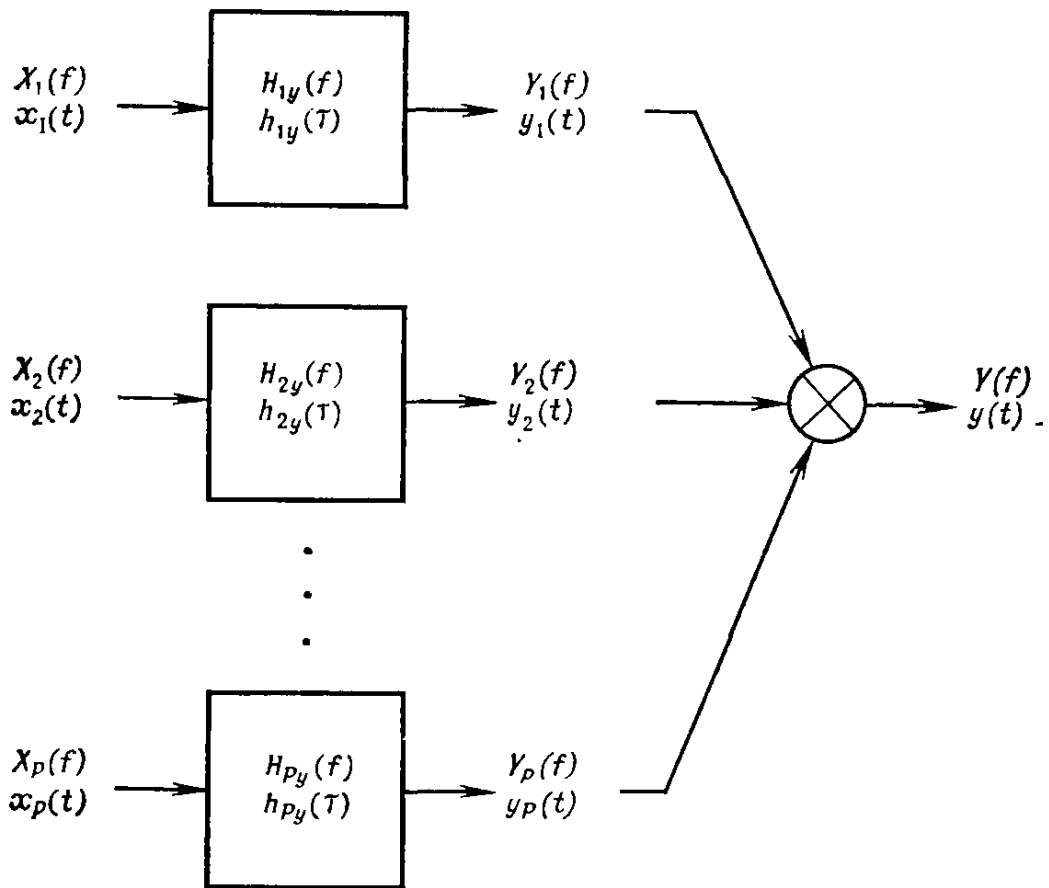


Рис. 9.5. Линейная система с несколькими входами.

Рассмотрим линейную систему, имеющую параметры, не зависящие от времени; p входов этой системы обозначим $x_l(t)$, $l=1, 2, \dots, p$. Выход $y(t)$ такой системы определяется соотношением

$$y(t) = \sum_{l=1}^p y_l(t),$$

где слагаемое $y_l(t)$ определяется как та часть выхода, которая получается при подаче на l -й вход значений $x_l(t)$, а на все остальные входы — нуля (рис. 9.5). Каждая из функций h_{ly} на рис. 9.5 определяется как весовая функция, относящаяся к ли-

нейной системе, образованной входом $x_i(t)$ и соответствующим ему частным выходом $y_i(t)$. Таким образом, функция $y_i(t)$ определяется формулой

$$y_i(t) = \int_{-\infty}^{\infty} h_{iy}(\tau) x_i(t - \tau) d\tau. \quad (9.11)$$

Применив к обеим частям формулы (9.11) преобразование Фурье, получим равенство $Y_i(f) = H_{iy}(f) X_i(f)$, в котором $Y_i(f)$ и $X_i(f)$ — преобразования Фурье функций $y_i(t)$ и $x_i(t)$ соответственно. Учитывая это равенство, для $Y(f)$ — преобразования Фурье полного выхода — получим следующее выражение:

$$Y(f) = \sum_{i=1}^p Y_i(f) = \sum_{i=1}^p H_{iy}(f) X_i(f). \quad (9.12)$$

В этом месте статистик, читающий книгу, непременно заметит, что соотношение (9.12) есть уравнение множественной регрессии для величин с нулевым средним, подобное уравнению для случая одного входа, определенному чуть выше. В качестве зависимой величины выступает преобразование Y , в качестве независимых величин выступают преобразования X_i , а в качестве коэффициентов регрессии — функции H_{iy} . Следовательно, о понятиях передаточных функций в частотной области, относящихся к областям электротехники и механики, можно говорить на языке того раздела статистики, который составляет регрессионный анализ комплексных случайных величин. Эта идея была детально разработана Акайком (1965). Суть этой идеи заключается в возможности обращаться с каждой полосой частот независимо; выполняя многомерный регрессионный анализ многих комплексных величин (по одному разу для каждой полосы частот), получают полные функции в области частот.

Предыдущим соотношениям можно придать более сжатую форму. Для этого следует, как это принято в многомерном регрессионном анализе, перейти к более удобной матричной записи. Во-первых, определим p -мерный вектор-столбец преобразований Фурье

$$\mathbf{X}(f) = \begin{bmatrix} X_1(f) \\ X_2(f) \\ \vdots \\ X_p(f) \end{bmatrix}$$

и p -мерный вектор-столбец передаточных функций

$$\mathbf{H}(f) = \begin{bmatrix} H_{1y}(f) \\ H_{2y}(f) \\ \vdots \\ H_{py}(f) \end{bmatrix}. \quad (9.13)$$

Во-вторых, введем p -мерный вектор-столбец кросс-спектров мощности выхода $y(t)$ с входами $x_l(t)$:

$$\mathbf{S}_{xy}(f) = \begin{bmatrix} S_{1y}(f) \\ S_{2y}(f) \\ \vdots \\ S_{py}(f) \end{bmatrix},$$

где $S_{ly}(f) \equiv S_{x_ly}(f)$ для $l=1, 2, \dots, p$. Наконец, определим $(p \times p)$ -матрицу спектров мощности и кросс-спектров всех входов $x_l(t)$:

$$\mathbf{S}_{xx}(f) = \begin{bmatrix} S_{11}(f) & S_{12}(f) & \cdots & S_{1p}(f) \\ S_{21}(f) & S_{22}(f) & \cdots & S_{2p}(f) \\ \vdots & \vdots & & \vdots \\ S_{p1}(f) & S_{p2} & \cdots & S_{pp}(f) \end{bmatrix} = E[\mathbf{X}^*(f)\mathbf{X}^T(f)],$$

где звездочкой помечена сопряженная величина, а индекс означает транспонирование. Для упрощения записи введем обозначение $S_{ij} = S_{x_i x_j}(f)$, где $i, j=1, 2, \dots, p$. Матрица \mathbf{S}_{xx} эрмитова, поскольку она совпадает с матрицей, полученной из нее транспонированием и заменой элементов на сопряженные им. Из этого факта, в частности, следует, что собственные значения матрицы $\mathbf{S}_{xx}(f)$ суть действительные числа. В некоторых приложениях эти параметры представляют большой интерес.

Для того чтобы получить для элементов $H_{ly}(f)$ вектора (9.13) решение по методу наименьших квадратов, необходимо решить систему линейных уравнений, которая в матричной записи имеет вид $\mathbf{S}_{xy} = \mathbf{S}_{xx} \mathbf{H}$. Это уравнение можно расписать в виде

$$\begin{bmatrix} S_{1y} \\ S_{2y} \\ \vdots \\ S_{py} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix} \begin{bmatrix} H_{1y} \\ H_{2y} \\ \vdots \\ H_{py} \end{bmatrix}. \quad (9.14)$$

Решение системы дает формула

$$\mathbf{H} = \mathbf{S}_{xx}^{-1} \mathbf{S}_{xy}. \quad (9.15)$$

На практике оценки спектров и кросс-спектров, входящих в выражение (9.14), ищут при помощи процедур, рассмотренных в гл. 8.

9.4. ФУНКЦИИ ОБЫЧНОЙ, МНОЖЕСТВЕННОЙ И ЧАСТНОЙ КОГЕРЕНТНОСТИ

Для системы с одним входом мы встречались только с одним типом когерентности, а именно с функцией обычной когерентности, определенной формулой (9.9). При этом отмечалось, что когерентность имеет аналог в элементарной статистике — квадрат коэффициента корреляции. Теперь же заметим, что понятие когерентности обобщается на случай многих входов, причем возникают два других типа когерентности — множественная и частная; их аналогами служат квадрат множественного коэффициента корреляции и квадрат частного коэффициента корреляции.

Чтобы придать этим трем разным функциям когерентности единообразное толкование, перепишем формулу (9.9) в ином виде, связав мощность входа с линейно «прогнозируемой» мощностью выхода. Вернемся к рис. 9.3, на котором выход системы $y(t)$ складывается из двух независимых (по предположению) компонент. Спектр мощности выхода такой системы имеет две компоненты, т. е. $S_y = S_n + S_u$, где $U(f)$ определяется как незашумленный выход системы $U(f) = H(f) X(f)$. Вспоминая соотношение (9.7), получим $S_u = |H|^2 S_x$. Здесь S_u — часть общей мощности входа. Доля этой мощности определяется отношением

$$\frac{S_u}{S_y} = \frac{|H|^2 S_x}{S_y} = \frac{|S_{xy}|^2}{S_x^2} \frac{S_x}{S_y} = \frac{|S_{xy}|^2}{S_x S_y} = \gamma_{xy}^2. \quad (9.16)$$

Такой вид записи показывает, что обычная когерентность дает ту долю мощности выхода, появление которой обусловлено линейной зависимостью от входа. Точно такую же интерпретацию для многомерного случая имеют множественная и частная когерентности.

Если система нелинейна, то оценка передаточной функции $\hat{H}(f)$ служит приближением «наименьших квадратов» к «истинной» нелинейной передаточной функции. В очень интересной теореме о системах утверждается, что выход $y(t)$ всегда можно расщепить на две независимые части. Первая из них $u(t)$ будет полностью когерентна с входом $x(t)$, а вторая часть $n(t)$ совершенно некогерентна с входом $x(t)$. В этом случае, как и в (9.16), функция когерентности есть отношение мощности когерентной части к общей мощности.

При $p > 1$ имеются две дополнительные формы когерентности. Первая — множественная когерентность $\gamma_{y \cdot x}^2(f)$; она определяет долю мощности выхода $y(t)$, обусловленную линейными соотношениями, которые связывают входы $x_1(t), \dots, x_p(t)$ с выходом. Вторая форма — частная когерентность $\gamma_{iy | 1, 2, \dots, i-1, i+1, \dots, p}^2(f)$; она представляет собой обычную когерентность между остатками i -го входа $\Delta x_i(t)$ и выхода $\Delta y(t)$, полученных удалением из них линейных эффектов, обусловленных остальными $p-1$ входами.

В случае многих входов дисперсия «прогноза» определяется формулой $\mathbf{S}_{yx}^T \mathbf{H}_{xy} = \mathbf{S}_{yx}^T \mathbf{S}_{xx}^{-1} \mathbf{S}_{xy}$, поэтому множественная когерентность дается равенством

$$\gamma_{y \cdot x}^2 = \frac{\mathbf{S}_{yx}^T \mathbf{S}_{xx}^{-1} \mathbf{S}_{xy}}{S_{yy}}.$$

Остаточный спектр выхода, который нередко называют «условным»¹⁾ спектром, определяется следующим образом:

$$S_{y|x} = S_{yy} - \mathbf{S}_{yx}^T \mathbf{S}_{xx}^{-1} \mathbf{S}_{xy}.$$

Множественная когерентность через условный спектр выражается просто:

$$\gamma_{y \cdot x}^2 = 1 - \frac{S_{y|x}}{S_{yy}}.$$

Функция множественной когерентности дает ту долю мощности на выходе, которую определяет система соотношений линейной фильтрации для всех входов.

Для того чтобы получить функцию частной когерентности между каждым входом (предположим, что в данный момент нас интересует вход x_1) и выходом при условии остальных ($p-1$) входов, сначала разбивают расширенную спектральную матрицу на подматрицы так, как это показано ниже:

$$\mathbf{S}_{yxx} = \left[\begin{array}{cc|ccc} S_{yy} & S_{y1} & S_{y2} & \cdots & S_{yp} \\ S_{1y} & S_{11} & S_{12} & \cdots & S_{1p} \\ \hline S_{2y} & S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ S_{py} & S_{p1} & S_{p2} & \cdots & S_{pp} \end{array} \right] = \left[\begin{array}{cc} \Sigma_{yy} & \Sigma_{y1} \\ \Sigma_{1y} & \Sigma_{11} \end{array} \right].$$

¹⁾ Появление этого слова объясняется связью с понятием условных вероятностей в теории вероятностей. Обычно оно используется в ситуациях, когда нужно знать, выражаясь иестрого, «что будет, если возникают некоторые условия».

Вслед за этим вычисляют элементы условной спектральной матрицы¹⁾:

$$S_{xy|p} = \Sigma_{yy} - \Sigma_{y1} \Sigma_{11}^{-1} \Sigma_{1y}. \quad (9.17)$$

Наконец, по формуле

$$\gamma_{1y|p}^2 = \frac{|S_{1y|p}|^2}{S_{11|p} S_{yy|p}}$$

определяют функцию частной когерентности между входом x_1 и выходом y при условии остальных $(p-1)$ входов. Аналогичные действия проводят для x_2 , переставляя x_2 и x_1 , затем для x_3 , переставляя x_3 и x_2 и т. д.

9.5. ДОВЕРИТЕЛЬНЫЕ ГРАНИЦЫ ДЛЯ КОГЕРЕНТНОСТИ

Статистическое распределение когерентности тесно связано со статистическим распределением коэффициентов корреляции. По-видимому, наиболее удобный подход к получению доверительных границ — применение нормализующего преобразования, т. е. такого преобразования, которое делает статистическое распределение оценок когерентности приближенно нормальным (гауссовским). Таким подходом пользовались Эноксон и Гудмэн (1965), его развил и усовершенствовал Бенинус (1969), в дальнейшем он рассматривался Картером (1972). Преобразование, которое используется в таком подходе, — это «z»-преобразование Фишера, названное так по имени выдающегося статистика Р. А. Фишера.

Характеристики оценок когерентности. Как следует из теории статистических выборок, оценки когерентности, помимо того что они проявляют статистическую изменчивость, еще и смещены. Приближенные значения дисперсии и смещения даются формулами

$$\sigma_z^2 = \frac{1}{n-2}, \quad b = \frac{1}{n-2},$$

где $z = \text{arth } \hat{\gamma}$. Поскольку в знаменателе имеется n — число степеней свободы, для малых значений n смещение оказывается чрезвычайно большим. В частности, для 2 ст. св., как показали Отнес и Эноксон (1972), гл. 9, справедливо при всех f равенство $\hat{\gamma}_{xy}^2 = 1.0$.

Кроме рассмотренного имеется также смещение второго типа, не затрагиваемое в теории. Его влияние можно изучать лишь на качественном уровне, как это сделано в § 9.12. Это смещение для любой спектральной оценки возникает вследствие «размы-

¹⁾ Соотношение (9.17) — комплексный аналог уравнения многомерного регрессионного анализа, приведенного, например, в книге Андерсона (1963).

вающего эффекта», если разрешающая полоса частот (т. е. B_e) широка по сравнению с действительным максимумом в спектре.

Смещение такого типа встречается не так уж редко. Кроме того, по некоторой причине такое смещение может оказывать на когерентность гораздо более сильное воздействие. Авторы в целом ряде случаев наблюдали рост оценки когерентности от $\hat{\gamma}^2 = 0.4$ до $\hat{\gamma}^2 = 0.9$ при уменьшении B_e вдвое (все остальные параметры оставались неизменными). Такого рода экспериментирование с изменениями B_e полезно проводить всякий раз, когда неясно, с какой шириной спектральных пиков придется столкнуться.

По крайней мере отчасти такое поведение когерентности можно отнести на счет временных задержек в системе. Большие задержки проявляются в слишком быстром изменении фазовой функции. Это в свою очередь влечет за собой быстрое изменение действительной и мнимой частей кросс-спектра. Поэтому во многих практических ситуациях весьма вероятно появление сильного смещения, имеющего своим источником быстрое изменение функций кросс-спектра.

Такое смещение помогает устранять процедура оценивания, которую предложили Кливленд и Парзен (1975).

Преобразование, приводящее распределение выборочной функции когерентности к точной нормальной (гауссовской) аппроксимации, имеет вид (см. Эноксон и Гудмэн (1965))

$$z = \operatorname{arth} \hat{\gamma} = \frac{1}{2} \ln \left[\frac{1 + \hat{\gamma}}{1 - \hat{\gamma}} \right]. \quad (9.18)$$

Здесь $\hat{\gamma}$ — положительное значение квадратного корня из выборочной оценки когерентности $\hat{\gamma}^2$. Приближенные значения среднего значения и дисперсии величины z дают формулы

$$\mu_z = \operatorname{arth} \gamma + \left(\frac{\rho}{n - 2\rho} \right), \quad \sigma_z^2 = \frac{1}{n - 2\rho}, \quad (9.19)$$

где $n = 2B_e P$.

Усовершенствования, предложенные Бенинусом (1969), касаются как смещения, так и изменчивости. Оценку $\hat{\gamma}^2$ когерентности можно улучшить следующей коррекцией смещения:

$$B(\hat{\gamma}^2) = \frac{1}{2n} (1 - \hat{\gamma}^2).$$

При этом в (9.18) используется подправленная оценка когерентности $\tilde{\gamma}^2 = \hat{\gamma}^2 - B(\hat{\gamma}^2)$. Вместо дисперсии, полученной по формуле (9.19), должно использоваться ее подправленное значение; поправка узнается по подгоночной кривой, которая позволяет улучшить оценку дисперсии σ_z^2 при $\hat{\gamma}^2 < 0.3$. В качестве указан-

ной кривой берется кривая $E(\sigma_z) = 1 - 0.004(1.6\hat{\gamma}^2 + 0.22)$. Вместо дисперсии из (9.19) будет в этом случае использоваться дисперсия

$$\sigma_z = \sqrt{\frac{1}{n-2p}} [E(\sigma_z)].$$

Предыдущие равенства позволяют для данных γ^2 и n построить $(1-\alpha)$ -доверительные границы для истинного значения γ^2 :

$$\text{th}(z - b - \sigma_z^Z \alpha^{1/2}) \leq \gamma \leq \text{th}(z - b + \sigma_z^Z \alpha^{1/2}).$$

Здесь Z_α есть 100α -процентная точка нормального распределения, а величина b определяется равенством

$$b = \frac{p}{n-2p}. \quad (9.20)$$

Приведенную выше формулу для доверительных границ можно использовать и для функций обычной когерентности при $p=1$, и для функций множественной когерентности при $p > 1$.

Чтобы получить доверительные границы для функции частной когерентности, требуются очень небольшие изменения. При этом число ст. св. в анализе должно, вообще говоря, уменьшиться на число условных переменных, эффекты которых должны быть устранены. Например, в случае, когда требуется устранить влияние $(p-1)$ входов, используется число n' ст. св., равное $n - (p-1)$, где $n = 2B_e T$.

9.6. ОПРЕДЕЛЕНИЕ ДОВЕРИТЕЛЬНЫХ ГРАНИЦ ДЛЯ ПЕРЕДАТОЧНЫХ ФУНКЦИЙ

Расположение и размеры доверительных областей для частотных функций отклика $H_{iy}(f)$, $i=1, 2, \dots, p$, зависят от выборочной функции когерентности между выходами и входами, выборочной функции множественной когерентности между входами, выборочных функций ПСМ входов и выборочных частотных функций отклика (см. рис. 9.5). Основой приведенных ниже рассуждений послужили книга Бендата и Пирсола (1971) и работа Гудмэна (1965).

Будем считать, что ошибки смещения всех встречающихся ниже спектральных оценок пренебрежимо малы. Допустим, что $|H_{iy}|$ — истинный коэффициент усиления, а φ_{iy} — истинная фаза; таким образом, истинная передаточная функция записывается в виде $H_{iy} = |H_{iy}| e^{j\varphi_{iy}}$. Тогда $(1-\alpha)$ -доверительные интервалы для H_{iy} и φ_{iy} при каждом i и любой фиксированной частоте f (в дискретном случае ее роль выполняет частотный индекс k)

определяются системой неравенств

$$\begin{cases} |\hat{H}_{iy}| - \hat{r}_i < |H_{iy}| \leq |\hat{H}_{iy}| + \hat{r}_i, \\ \hat{\varphi}_{iy} - \Delta\hat{\varphi}_i < \varphi_{iy} \leq \hat{\varphi}_{iy} + \Delta\hat{\varphi}_i, \end{cases} \quad i = 1, 2, \dots, p, \quad (9.21)$$

где $|\hat{H}_{iy}|$ и $\hat{\varphi}_{iy}$ — выборочные оценки. Здесь квадрат ошибки в длине радиуса, которую обозначим $\hat{r}_i \equiv \hat{r}_i(f)$, и ошибка $\Delta\hat{\varphi}_i \equiv \Delta\hat{\varphi}_i(f)$ в фазе вычисляются при каждом i по формулам

$$\hat{r}_i^2 = \frac{2p}{n-2p} (F_{n_1, n_2; \alpha}) \frac{(1 - \hat{\gamma}_{y \cdot x}^2) \hat{S}_y}{(1 - \hat{\gamma}_{i \cdot x}^2) \hat{S}_i}, \quad (9.22)$$

$$\Delta\hat{\varphi}_i = \arcsin \left(\frac{\hat{r}_i}{|\hat{H}_{iy}|} \right). \quad (9.23)$$

В формулах (9.22) и (9.23) используются следующие обозначения: p — число входов (без выхода); $n = 2B_e P$ — число ст. св. в каждой спектральной оценке; $F_{n_1, n_2; \alpha}$ есть 100α -процентная

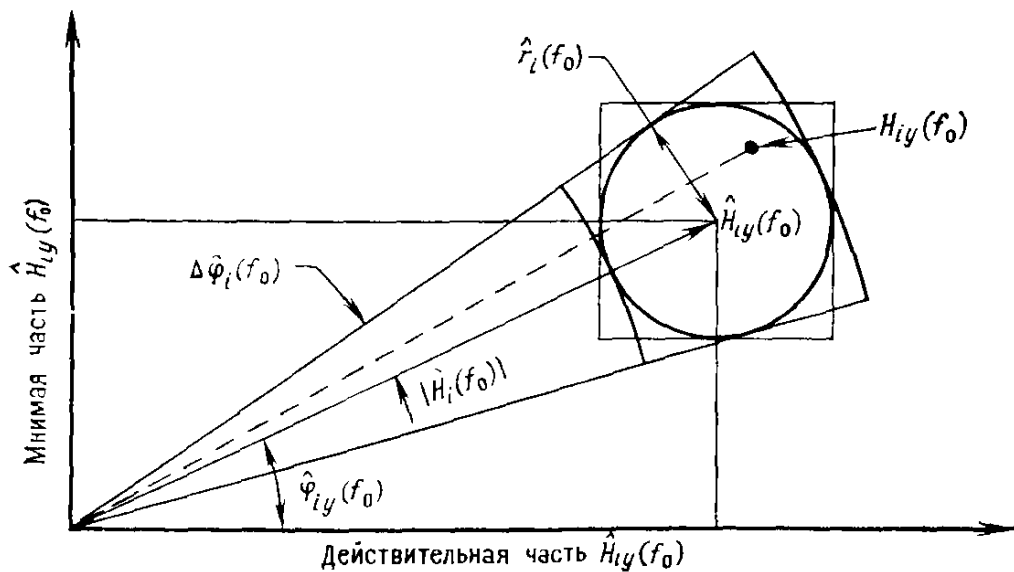


Рис. 9.6. Диаграмма, изображающая доверительную область для функций множественного частотного отклика.

точка F -распределения с $n_1 = 2p$ и $n_2 = 2n - 2p$ степенями свободы; $\hat{\gamma}_{y \cdot x}^2$ — выборочная оценка функции множественной когерентности между выходом y и всеми измеренными входами; $\hat{\gamma}_{i \cdot x}^2$ — выборочная оценка функции множественной когерентности между входом x_i и всеми остальными измеренными входами, исключая x_i ; \hat{S}_y — оценка спектра мощности для выхода y ; \hat{S}_i — оценка спектра мощности для входа x_i .

На рис. 9.6 в полярных координатах представлена доверительная область, определенная формулами (9.21), при значении

частоты f_0 . Каждой отдельной частоте и любому возможному значению $i = 1, 2, \dots, p$ отвечает своя собственная доверительная область.

Формула (9.22), дающая выражение для ошибки в радиусе, отвечает наиболее общему случаю многих когерентных входов. Если ограничиться специальным случаем p некогерентных входов или случаем только одного входа ($p = 1$), то формула для квадрата ошибки в радиусе примет вид

$$\hat{r}_i^2 = \frac{2}{n-2} (F_{n_1, n_2; \alpha}) (1 - \hat{\gamma}_{yi}^2) \frac{\hat{S}_y}{\hat{S}_i}, \quad (9.24)$$

где n , \hat{S}_y и \hat{S}_i — те же величины, что и раньше; $F_{n_1, n_2; \alpha}$ есть 100α -процентная точка F -распределения с $n_1 = 2$ и $n_2 = n - 2$ степенями свободы; $\hat{\gamma}_{yi}^2$ — выборочная оценка функции обычной когерентности между y и x_i . Ошибка в фазе, т. е. величина $\Delta\hat{\varphi}_i$, вычисляется, как и прежде, по формуле (9.23), поэтому различие с предыдущим случаем сводится к тому, что вместо формулы (9.22) для вычисления ошибки \hat{r}_i в радиусе используется формула (9.24).

Теперь коснемся различий между доверительными областями при одном и нескольких входах. В этом вопросе имеется одна тонкость, на которой мы остановимся. Допустим, что для 100 систем со многими входами построены 90% -е доверительные границы. Следует ожидать, что примерно для десяти из этих систем рассчитанные границы будут нарушены. Теперь предположим, что вместо этого доверительные границы определяются для каждого входа и выхода отдельно (как в ситуации с одним входом). В этом случае следует ожидать, что из 100 доверительных областей для любого отдельно взятого входа около десятка областей окажутся нарушенными. Однако при числе входов, равном, например, десяти (т. е. при $p = 10$), число нарушений должно оказаться больше (хотя и нельзя сказать точно, на сколько), так как разные доверительные области для моделей с единственным входом не бывают независимыми. Упомянутая тонкость заключается в том, что если наши интересы ограничиваются только одной частотной функцией отклика, то нужно пользоваться формулами, соответствующими случаю одного входа, даже если интересующий нас вход — лишь один из многих входов системы. Определенные таким образом доверительные границы дадут более узкую область, чем границы для случая многих входов; следовательно, первые более предпочтительны. Но если требуются все доверительные области, то необходимо пользоваться формулами для случая многих входов.

Обратимся к вопросу повышения точности оценок множественных передаточных функций, в основе вычисления которых

лежит формула (9.22). Нетрудно понять, что точность в экспериментах по измерению этих функций можно повысить, придерживаясь следующих правил.

1. Все входные функции следует делать одинаково мощными, чтобы величина $\hat{r}_i(f)$ уменьшилась за счет уменьшения отношения мощности выхода к мощности входа.

2. Сокращать шум на входе и делать системы линейными, чтобы уменьшить величину $\hat{r}_i(f)$ за счет большой множественной когерентности.

3. Делать входы независимыми, насколько это возможно, чтобы уменьшить величину $\hat{r}_i(f)$ за счет низкой множественной когерентности между входами.

4. Увеличивать, насколько можно, число степеней свободы, чтобы уменьшить величину $\hat{r}_i(f)$ как за счет уменьшения значения F (для данного α), так и за счет увеличения знаменателя. Следует иметь в виду, что для фиксированной длины записи данных возрастанию числа ст. св. сопутствует возрастание B_e , что в свою очередь приводит к сложностям, которые связаны со смещениями, обусловленными размыванием.

Перечисленные примеры, описанные на качественном уровне, весьма полезны и могут принести ощутимую помощь при составлении программ сбора данных. Главное, что может при этом стать предметом заботы, — это компромисс между числом степеней свободы и высокой когерентностью. Но вряд ли стоит придавать ему особое значение. Мы повторяем здесь важное положение: «Высокая когерентность позволяет получать очень точные оценки передаточных функций даже тогда, когда число ст. св. мало и спектр мощности вычислен с большими погрешностями (и наоборот)».

9.7. КАК ВЫЧИСЛЯЮТСЯ ПЕРЕДАТОЧНЫЕ ФУНКЦИИ

Методы вычислений передаточной функции системы с одним входом и передаточных функций системы с несколькими входами весьма несхожи. Здесь мы подробно остановимся на последних, а в примерах, относящихся к этому случаю, рассмотрим вопросы, общие для обеих возможностей.

Все вычислительные процедуры, необходимые для поиска решения системы (9.15), распадаются на три группы.

1. Программы вычисления функций плотности спектра мощности и плотности кросс-спектра.

2. Процедура совместной обработки $p+1$ переменных, цель которой заключается в получении эффективным способом спектральных функций плотности для всех пар переменных.

3. Комплексные арифметические и матричные операции для вычисления параметров многомерных линейных систем.

Как правило, спектральные функции плотности получают при помощи процедур БПФ — ПСМ. Главное требование состоит здесь в том, что при вычислениях должны быть получены все кросс-спектры. В силу комплексной сопряженности элементов, расположенных симметрично относительно главной диагонали матриц кросс-спектров, т. е. в силу того, что $S_{ij}(f) = S_{ji}^*(f) = S_{ij}(-f)$ (эрмитово свойство матрицы), нужно вычислять только элементы диагонали и правой верхней части над ней.

Рассмотренные здесь процедуры предназначены для вычисления параметров математической модели линейной системы с p входами $x_i(t)$, $i = 1, 2, \dots, p$, и одним выходом $y(t)$. Следует определить следующий набор параметров системы.

1. Передающую функцию между каждым входом и выходом.
2. Функции обычной когерентности между всеми парами переменных.
3. Функцию множественной когерентности между выходом и всеми входами.
4. Функции частной (условной) когерентности между каждым входом и выходом при условии на остальные входы.

На некоторое время мы откажемся от крышечки над статистическими оценками — обозначения ниже и без того достаточно сложны. Надеемся, что это не вызовет никаких недоразумений, тем более, что во всех случаях теоретические соотношения и их переформулировки для статистических оценок совпадают.

Первая операция, в которой возникает необходимость, — это сортировка. Как правило, спектральные функции плотности вычисляются как функции частоты, и программы для анализа линейных систем со многими входами должны в конечном счете оперировать со спектральными $(p+1) \times (p+1)$ -матрицами плотностей (по одной матрице на каждое значение частоты).

Таким образом, данные, с которыми оперирует программа, составляют набор матриц плотностей спектров в частотах, снабженных индексом k :

$$S_{yxx}(k) = \begin{bmatrix} S_{yy}(k) & S_{y1}(k) & S_{y2}(k) & \cdots & S_{yp}(k) \\ S_{1y}(k) & S_{11}(k) & S_{12}(k) & \cdots & S_{1p}(k) \\ \vdots & \vdots & \vdots & & \vdots \\ S_{py}(k) & S_{p1}(k) & S_{p2}(k) & \cdots & S_{pp}(k) \end{bmatrix}, \quad k = 0, 1, \dots, m.$$

Обычно частотный индекс k обозначает значения частот $f_k = kF/m$, где $k = 0, 1, \dots, m$ и F — частота Найквиста. Однако, индекс может использоваться и в более общем смысле, а именно

для обозначения частот $f_k = f_1 + k\Delta f$, где $k = 0, 1, \dots, m$, f_1 — начальная частота, Δf — приращение частоты (шаг).

Набор m отдельных спектральных $(p+1) \times (p+1)$ -матриц плотности можно представлять себе в виде трехмерного массива (рис. 9.7). Вычисление каждой из спектральных функций плотности даст одномерный столбец длины $m+1$ вдоль $(p+1) \times (p+1) \times (m+1)$ -блока с рис. 9.7. При вычислениях частотного

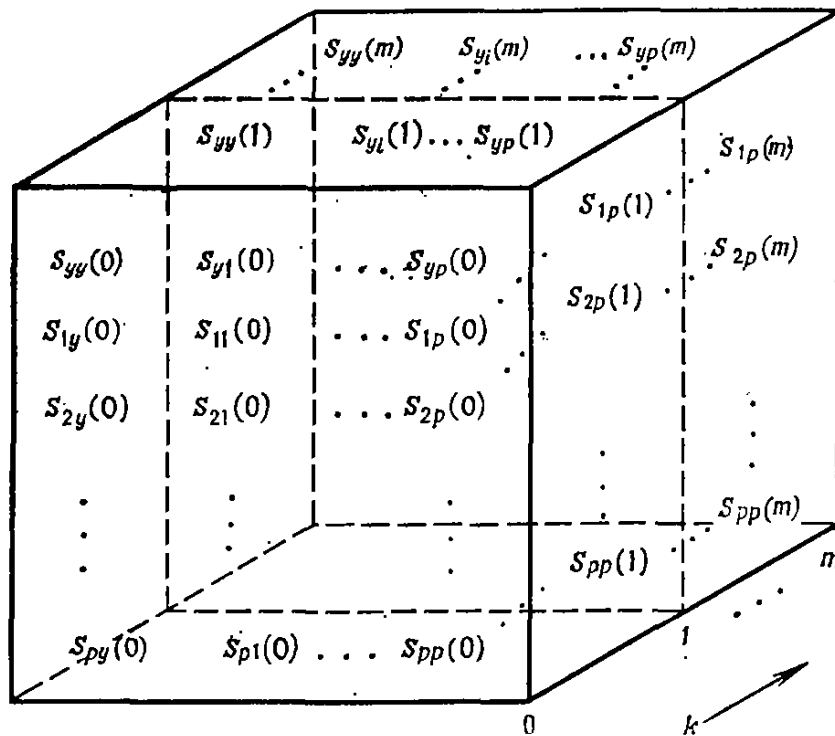


Рис. 9.7. Изображение трехмерного блока для запоминания спектральных матриц.

отклика требуется один поперечный (вертикально-горизонтальный) слой. Если для запоминания промежуточных результатов используется магнитная лента, то возникает задача разумного, расположения данных, поскольку итоговые данные должны быть упорядочены в виде ряда. Для этой цели больше подходит память файлами на дисках.

Всякий раз, когда речь будет заходить о статистическом распределении результатов, будет предполагаться, что все p входных величин, а также выход — стационарные гауссовские процессы с нулевым средним. Функции $H_{ly}(k)$, $l = 1, \dots, p$, суть характеристики частотных функций отклика (передаточных функций) линейной системы, выход $y(t)$ которой получается при прохождении через нее входных величин.

Предполагается, что величины $x_i(i)$, $i = 1, \dots, p$, и $y(i)$ составляют дискретные (преобразованные в цифровую форму) последовательности из N точек в каждой.

Функция частотного отклика определяется из матричного уравнения (9.14), в котором для упрощения записи аргумент

функции будет опускаться (т. е. вместо записи вида $S_{11}(k)$ будет использоваться запись S_{11}). Матрицы и векторы в соотношении (9.14) имеют комплексные элементы, поэтому при расчетах требуется применение комплексной арифметики. В частности, справедливо представление $S_{ii} = C_{ii} + jQ_{ii}$, где $C(k)$ и $Q(k)$ — соответствующие функции плотности коспектра и квадратурного спектра как функции индекса k .

Решение уравнения (9.14) определяется следующим соотношением:

$$\begin{bmatrix} H_{1y} \\ H_{2y} \\ \vdots \\ H_{py} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix}^{-1} \begin{bmatrix} S_{1y} \\ S_{2y} \\ \vdots \\ S_{py} \end{bmatrix}.$$

В векторных обозначениях оно принимает вид

$$\mathbf{H}_{xy} = \mathbf{S}_{xx}^{-1} \mathbf{S}_{xy}. \quad (9.25)$$

Здесь нам кажется уместным процитировать следующее место из книги Форсайта и Молера (1969): «Один из способов решения системы линейных уравнений $\mathbf{Ax} = \mathbf{b}$ состоит в том, что, вычислив матрицу, обратную \mathbf{A} , умножают \mathbf{A}^{-1} на \mathbf{b} Однако целый ряд процедур решения линейных уравнений . . . может как упростить эту задачу за счет уменьшения числа операций, так и улучшить ее решение путем увеличения точности».

Систему линейных уравнений над комплексными числами можно решать непосредственно в комплексной арифметике. Можно пойти и другим путем — разбить уравнение на мнимые и действительные части. Один из таких методов решения, предложенный Ланцошем (1961), детально разобран в книге Отнеса и Эноксона (1972), с. 347 — 349. Кроме того, можно пользоваться изоморфизмом

$$(\mathbf{C} + j\mathbf{Q}) \sim \begin{bmatrix} \mathbf{C} & -\mathbf{Q} \\ \mathbf{Q} & \mathbf{C} \end{bmatrix}, \quad (9.26)$$

который комплексным матрицам размера $p \times p$ соотносит действительные $2p \times 2p$ -матрицы в правой части (9.26). Действительную часть решений системы (9.25) будут составлять первые p элементов вектора \mathbf{H}_{xy} , а мнимую — оставшиеся p элементов этого вектора.

Систему линейных уравнений предполагается решать методом, основанным на процедурах Форсайта и Молера, приспособленных для действий с комплексными числами. Эти процедуры более других подходят для гауссовского случая, главная особенность которого состоит в наличии максимального значения, что

требует масштабирования. Другие возможные процедуры основаны либо на ортогонализации данных, либо на ортогонализации спектральной матрицы (см., например, Лонгли (1967) или Уомплер (1970)). Метод Форсайта и Молера содержит программу последовательного улучшения решения с оценкой числа точных цифр в решении. В зависимости от длины слова ЭВМ и вида матриц могут потребоваться вычисления с удвоенной точностью.

Метод решения системы (9.25) сводится к следующему:

1. Система спектральных уравнений решается методом Форсайта—Молера, работающим непосредственно в комплексной арифметике.

2. Полученное решение корректируется последовательными улучшениями. Если результаты неудовлетворительны, то используется арифметика с удвоенной точностью.

3. При помощи оператора «развертки», определенного ниже, вычисляется множественная когерентность, а также частные когерентности.

Известный интерес представляют итерационные методы, в которых число переменных увеличивается постепенно, восхождением от случая системы с одним входом. Один из таких методов, «метод эскалатора», рассмотрен в книге Отнеса и Эноксона (1972). Однако с точки зрения эффективности вычислений этот класс приемов не оптимален; более того, точность вычислений и результатов находится в зависимости от того, в каком порядке вводятся новые переменные. Поэтому такие методы, вообще говоря, применять не рекомендуется.

9.8. ОПЕРАТОР РАЗВЕРТКИ

Допустим, что решается система уравнений, которая имеет следующую матрицу:

$$\Sigma = \begin{array}{c} \begin{array}{cc} p & 1 \\ \hline \begin{array}{c} S_{xx}^{-1} \\ \hline -(S_{xx}^{-1}S_{xy})^* \end{array} & \begin{array}{c} (S_{xx}^{-1}S_{xy})^T \\ \hline S_{yy} - S_{yx}S_{xx}^{-1}S_{xy} \end{array} \\ \hline p & 1 \end{array} \end{array} \begin{array}{c} p \\ 1 \end{array}.$$

Теперь напомним, что множественная когерентность определяется по формуле

$$\gamma_{y \cdot x}^2 = 1 - \frac{S_{yy} - S_{yx}S_{xx}^{-1}S_{xy}}{S_{yy}} = \frac{S_{yx}S_{xx}^{-1}S_{xy}}{S_{yy}},$$

а множественные когерентности между входами — равенством

$$\gamma_{x \cdot ij}^2 = 1 - \frac{1}{S_{ii}S_{jj}},$$

где S^{ii} есть i -й диагональный элемент матрицы S_{xx}^{-1} . Эти величины необходимы при вычислениях доверительных границ.

Определим *развертку* (см. Шатцов и др. (1968)) r -х строки и столбца следующим образом. Пусть $A = (a_{ij})$ и $B = (b_{ij})$ суть квадратные матрицы. Матрица A преобразуется разверткой в матрицу B по такому правилу:

$$b_{rr} = \frac{1}{a_r} ; \quad b_{ir} = -\frac{a_{ir}}{a_{rr}}, \quad i \neq r;$$

$$b_{rj} = \frac{a_{rj}}{a_{rr}}, \quad j \neq r; \quad b_{ij} = a_{ij} - \frac{a_{ir}}{a_{rr}} a_{rj}, \quad i, j \neq r.$$

Оператор развертки обратим. Следовательно, применив к матрице Σ развертку, например, по p -м строке и столбцу, получим

$$\Sigma' = \begin{array}{c} p-1 \\ 2 \end{array} \left[\begin{array}{c|c} \Sigma_{11}^{-1} & (\Sigma_{11}^{-1} \Sigma_{1y})^T \\ \hline -(\Sigma_{11}^{-1} \Sigma_{1y})^* & \Sigma_{yy} - \Sigma_{y1} \Sigma_{11}^{-1} \Sigma_{1y} \end{array} \right] \begin{array}{c} p-1 \\ 2 \end{array}.$$

Как видно, здесь получена и условная спектральная матрица $S_{xy|p}$ из соотношения (9.17). Поэтому, вычислив по этой 2×2 -матрице обычную когерентность, мы получим частную когерентность, между p -м входом и выходом y .

Для того чтобы получить остальные частные когерентности, нужно снова применить оператор развертки p -х строки и столбца, восстановив тем самым матрицу Σ из Σ' .

Вслед за этим, применяя развертку по другим столбцам и строкам, получают остальные частные когерентности.

Если требуются частные и множественные когерентности первого порядка, действия при вычислениях можно значительно упростить. Здесь в матрице Σ' необходимо знать только элементы $S_{pp|p}$, $S_{yy|p}$ и $S_{py|p}$, соответствующие b_{rr} , $b_{p+1, p+1}$ и $b_{r, p+1}$. Поэтому для того чтобы получить условную спектральную матрицу, вычисляются величины

$$S_{pp|p} = b_{rr} = \frac{1}{a_{rr}}, \quad (9.27)$$

$$S_{py|p} = b_{r, p+1} = \frac{a_{r, p+1}}{a_{rr}}, \quad (9.28)$$

$$S_{yy|p} = b_{p+1, p+1} = a_{p+1, p+1} - \frac{a_{p+1, r}}{a_{rr}} a_{r, p+1} =$$

$$= a_{p+1, p+1} - \frac{|a_{r, p+1}|^2}{a_{rr}}. \quad (9.29)$$

С точки зрения качества вычислений приведенный метод получения частной когерентности и условной спектральной матрицы, разумеется, наиболее эффективен и точен.

9.9. ОПРЕДЕЛЕНИЕ ПЕРЕДАТОЧНОЙ ФУНКЦИИ ПРИ ПОМОЩИ СИНУСОИД

В лабораторной практике прошлого метод вычисления передаточной функции при помощи синусоид находил самое широкое применение. Цифровой эквивалент этого метода уже был рассмотрен в § 5.2. При этом мы пользовались тем фактом, что изменение амплитуды и сдвиг фазы определяют характеристики передаточной функции, которую с учетом формул (9.4) и (9.5) задает соотношение

$$Y(f) = |H(f)| e^{j\varphi(f)} X(f).$$

Если на вход подается косинус частоты f_c , который можно рассматривать как действительную часть комплексной экспоненты, т. е.

$$x(t) = \operatorname{Re} [Ae^{j2\pi f_c t}] = A \cos(2\pi f_c t),$$

то

$$X(f) = \frac{A}{2} [\delta(f - f_c) + \delta(f + f_c)],$$

и, следовательно, на выходе будет получена функция, имеющая преобразование Фурье вида

$$\begin{aligned} Y(f) &= |H(f)| e^{j\varphi(f)} \frac{A}{2} [\delta(f - f_c) + \delta(f + f_c)] = \\ &= \frac{A}{2} [|H(f_c)| e^{j\varphi(f_c)} \delta(f - f_c) + |H(f_c)| e^{j\varphi(-f_c)} \delta(f + f_c)], \end{aligned}$$

т. е. будет получена функция

$$y(t) = A |H(f_c)| \operatorname{Re} (e^{j[2\pi f_c t + \varphi(f_c)]}). \quad (9.30)$$

Таким образом, действительная часть выхода есть косинус с амплитудой $A |H(f_c)|$, фаза которого сдвинута относительно входа на $\varphi(f_c)$. Покажем это. На рис. 9.8,а изображена синусоида на входе; 9.8,б дает изображение выхода. Легко заметить, что амплитуда на выходе, как, впрочем, и фаза, изменилась. Поэтому, располагая методом получения отношения амплитуд и способом измерения сдвига фазы, мы можем полностью определить характеристики передаточной функции.

Следует отметить, что синусоида — функция «установившегося» поведения. Даже если соотношение (9.4) и остается в силе для произвольных функций, то в частном случае синусоиды следует до проведения измерений подождать, пока система не начнет

давать установившийся отклик. Поскольку в физических системах промежуток времени, который занимает переходный отклик, примерно равен величине, обратной полосе частот, длительность этого промежутка можно взять за основу при определении ин-

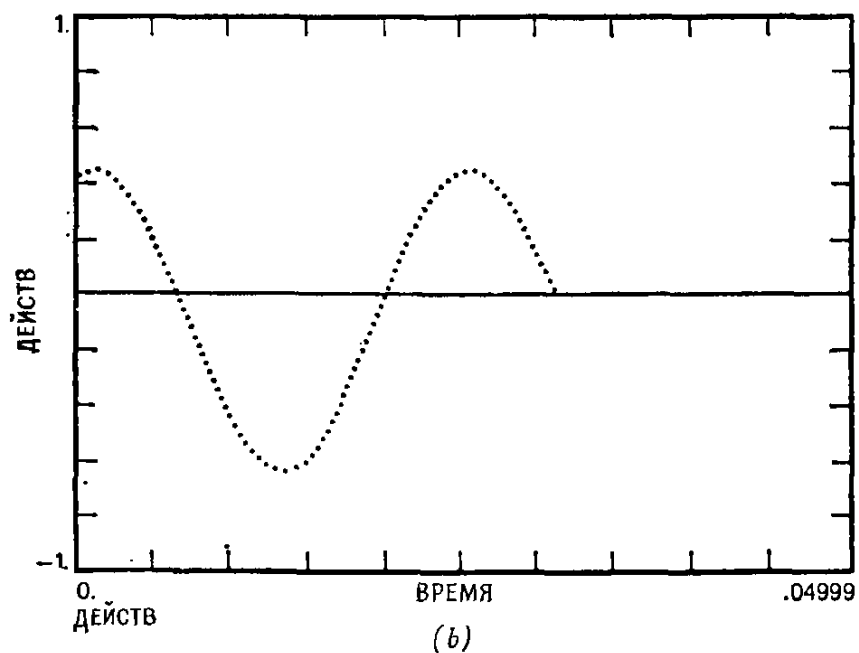
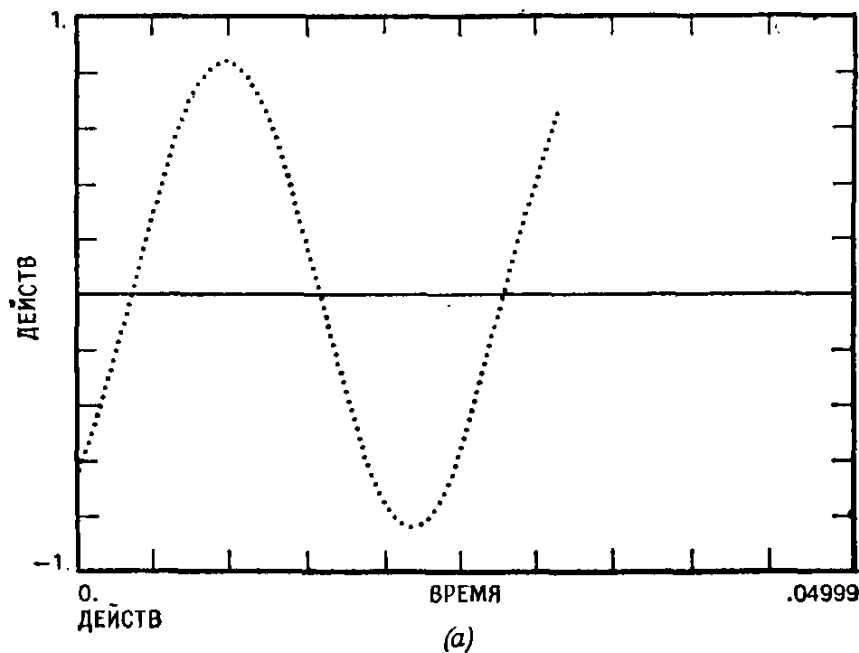


Рис. 9.8. *a* — синусоида на входе системы; *b* — синусоида на выходе линейной системы; амплитуда уменьшилась и фаза сдвинулась.

тервала ожидания. Более точно, если P_t — длина переходного отклика, а B_e — полоса частот отклика, то справедливо приближенное равенство $P_t \approx 1/B_e$; длину интервала времени ожидания перед измерениями амплитуды и фазы при помощи синусоиды нужно выбрать равной $2P_t$.

Рассматриваемый метод применяется в такой последовательности.

1. Выбирается осциллятор или генератор синусоиды нужной частоты.

2. Назначается тракт передачи данных на вход фильтра, смесителя или иного исследуемого устройства, который дает возможность получить на входе этих систем синусоиду.

3. Выжидается промежуток времени, необходимый для того, чтобы получился установившийся отклик.

4. Измеряется отношение амплитуды выхода к амплитуде входа и фазовый сдвиг выхода относительно входа.

5. Шаги 1—4 повторяются для всех частот, составляющих полосу частот отклика.

Преимущества метода исследования передаточных функций при помощи синусоид состоят в следующем. Во-первых, вся энергия сигнала концентрируется в одной частоте, поэтому отношение сигнала к шуму велико. Во-вторых, этот метод прост, применяется к системам непосредственно и допускает относительно простое оборудование. К недостаткам метода следует отнести в первую очередь исключительно долгое время, которое может потребоваться для полного исследования передаточной функции. Это неудобство заставляет искать иные подходы. Здесь мы остановимся на методе использования развернутой синусоиды, а затем — на методе случайных входов.

Метод получения передаточных функций при помощи развернутой синусоиды. Для того чтобы полностью определить передаточную функцию системы, необходимо иметь такую функцию на входе, которая несет достаточную мощность на всех частотах, имеющих заметные отклики. Как уже говорилось, метод одной синусоиды требует неприемлемо большого времени. Сократить время исследования можно при помощи развернутой синусоиды. *Линейно развернутая синусоида* определяется формулой

$$x(t) = \sin(\omega t),$$

где $\omega = at + b$; следовательно,

$$x(t) = \sin(at^2 + bt).$$

Мгновенная частота есть производная аргумента синусоиды, имеющая вид $\omega(t) = 2at + b$. При $t = 0$ по этой формуле получаем начальную частоту $\omega(0) = 2\pi f_0 = b$. Конечная частота, которая достигается на P -й секунде, получается из равенств $\omega(P) = 2\pi f_1 = 2aP + b$. Таким образом, скорость развертки равна $2a$, где

$$2a = \frac{2\pi(f_1 - f_0)}{P} \quad \text{или} \quad a = \frac{\pi(f_1 - f_0)}{P} \text{ рад/с.}$$

Чтобы пройти частотную область от f_1 до f_0 за P с, требуется скорость развертки $a/(2\pi) = (f_1 - f_0)/(2P)$.

Типичное поведение плотности спектра мощности развернутой синусоиды показано на рис. 9.9. Нередко приходится сталкиваться с заблуждением, что развернутая синусоида с постоянной амплитудой имеет «плоский» спектр. Как можно видеть из

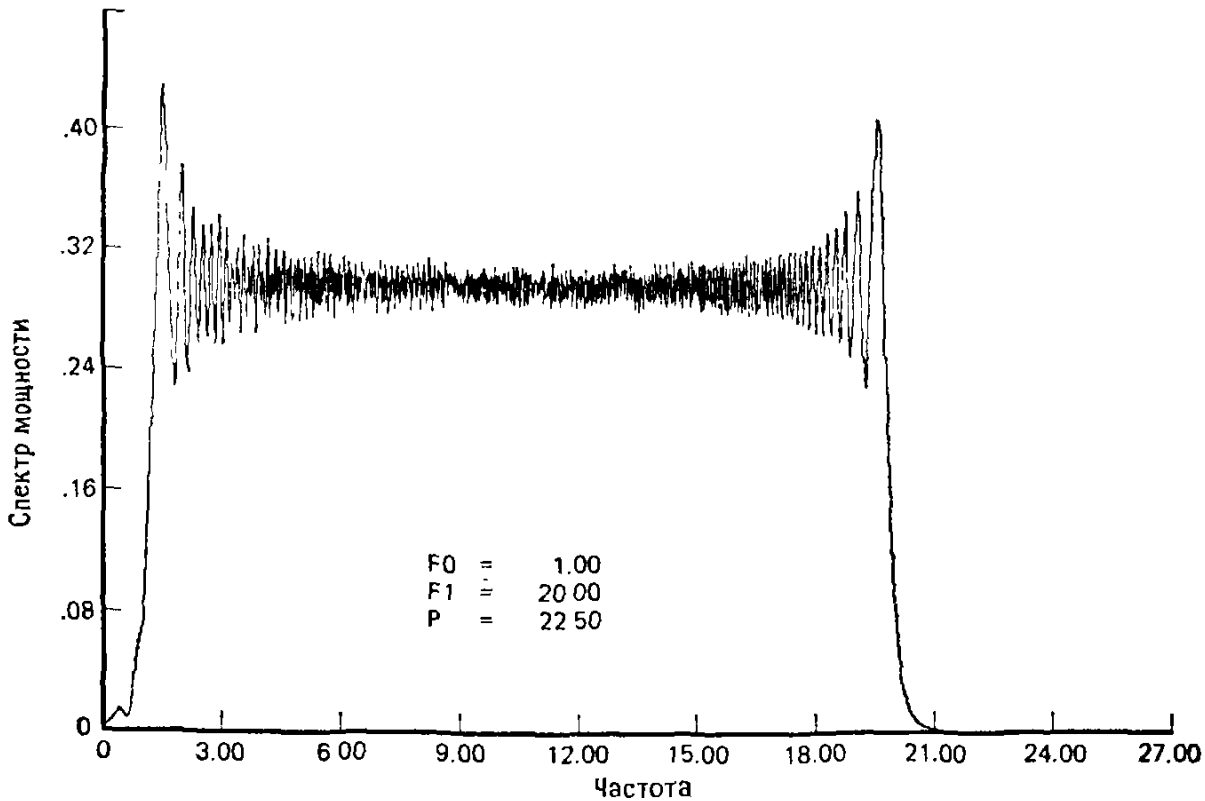


Рис. 9.9. Спектр развернутой синусоиды.

рисунка, это не так. В спектре выделяется «испорченный диапазон», величина которого не зависит от скорости развертки. По приближенной формуле

$$S_{\text{макс}} = \frac{1.38\pi}{4a} = 1.38 S_{\text{средн}}$$

можно получить крайние значения отклонений от среднего уровня спектра. Если бы спектральная плотность входа была функцией, постоянной вдоль всей области частот, то вычисления преобразования Фурье входа и не требовалось бы: достаточно было бы взглянуть на значения квадратного корня из плотности спектра мощности — по ним и определялся бы коэффициент усиления передаточной функции.

При использовании методов, основанных на применении кросс-спектров или отношений преобразований Фурье, дело обстоит не так просто, поскольку исследования осложняются

ограничениями на разрешающую способность и число ст. св. (при наличии существенного уровня шума). Упомянем еще об одной тонкости, связанной с методом развернутой синусоиды. Дело в том, что мощность, которую развернутая синусоида несет на взятом диапазоне частот, подается лишь в очень короткие промежутки времени. «При слишком скорой» развертке подчас не удается получить точного отклика систем, и немало примеров тому имеется по крайней мере в прикладных структурных исследованиях, где малых промежутков времени может не

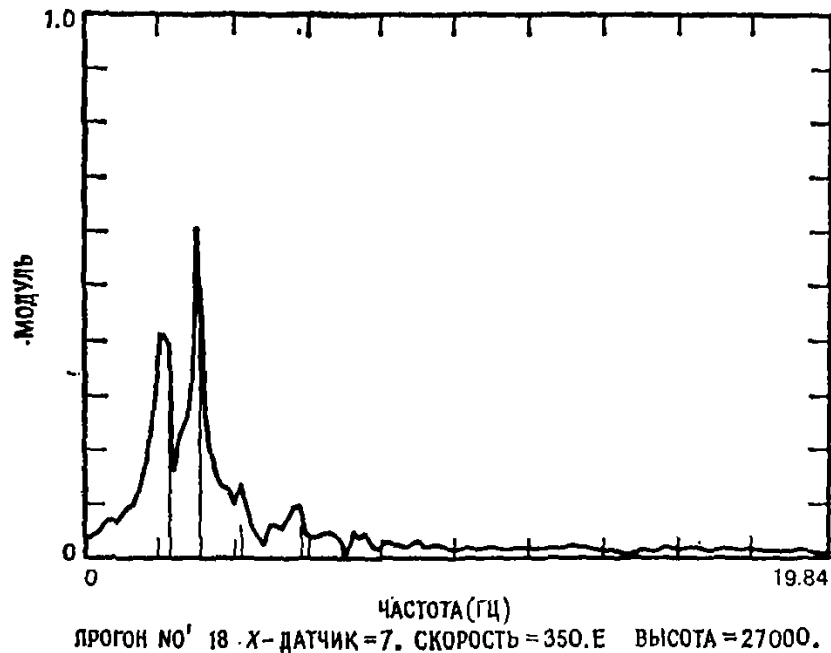


Рис. 9.10. Результаты теста по определению флаттера, проводившегося в AFFTC (один из научно-исследовательских центров военно-воздушных сил США.— Прим. перев.): модуль передаточной функции (27 000 футов = 8230 м).

хватать для того, чтобы все типы колебаний структуры при возбуждении полностью выходили на устойчивые режимы. Поэтому в каждом отдельном приложении, прежде чем сделать вывод о том, что такую возможность позволительно игнорировать, нужно поэкспериментировать с несколькими разными скоростями развертки.

На рис. 9.10 показана передаточная функция, полученная для одной из частей конструкции бомбардировщика B-1 (Рокуэлл) в ходе тестовых испытаний на полетные вибрации. В этих тестах использовалось возбуждение развернутой синусоидой.

9.10. ОПРЕДЕЛЕНИЕ ПЕРЕДАТОЧНОЙ ФУНКЦИИ ПРИ ПОМОЩИ СЛУЧАЙНЫХ ВХОДОВ

Соотношение (9.4) связывает выход с произвольным входом. При этом, по существу, никаких ограничений на вход не нало-

жено. В частности, в его качестве может выступать выборочная функция случайного процесса. В своей основе возможность использования таких входов восходит к тому факту, что для точной линейной системы без постороннего шума *одного измерения* достаточно для *точной оценки* передаточной функции системы.

В частности, представляется очень подходящим кандидатом для использования в качестве входной функции белый шум, несущий мощность на всех частотах. Разумеется, в реальных ситуациях посторонний шум всегда присутствует, поэтому приходится проводить какие-то операции усреднения, которые позволяют избавляться от влияния шума. Тем не менее тот факт, что имеется стационарная (в среднем) функция, несущая равные мощности на всех частотах (опять-таки в среднем), придает выбору белого шума в качестве входной функции исключительное значение по сравнению со всеми другими возможностями, существующими для измерений передаточной функции.

Перечислим преимущества, которые дает использование в качестве входа случайной функции — белого шума.

1. Белый шум несет в среднем равные мощности на всех частотах.

2. Для данного уровня постороннего шума очень удобно проводить осреднения и по отрезкам ряда соответствующей длины, и по частотам в достаточно широких диапазонах, что дает возможность увеличивать число ст. св. и повышать тем самым точность оценок.

3. Чтобы получить необходимую точность при данном разрешении, которое требуется в исследовании, и постороннем шуме, необходимо лишь доводить время теста до соответствующей границы.

Недостатки метода случайных входов обусловлены только традицией. В прошлом из-за отсутствия полного понимания преимуществ этого подхода выбор исследовательского оборудования, предназначенного для генерации белого шума, был крайне ограничен. Однако применение ЭВМ для оценивания передаточных функций в приложениях (например, в тестах вибрационного контроля) заложило прочную практическую основу под методики, использующие случайные входы.

Если используется случайный шум, то применяется метод кросс-спектра, основанный на соотношении (9.8). При этом кросс-спектры и спектры мощности должны оцениваться правильно. Если спектральные оценки окажутся размытыми из-за недостаточного разрешения, то ошибки перейдут и в оценку передаточной функции.

На рис. 9.11 показан контрольный спектр, полученный в цифровой схеме контроля вибрации. Белый шум был получен при

помощи цифрового датчика случайных чисел. Вслед за оцениванием передаточной функции системы спектру входа придавалась такая форма, чтобы выходной спектр укладывался в заранее заданную полосу.

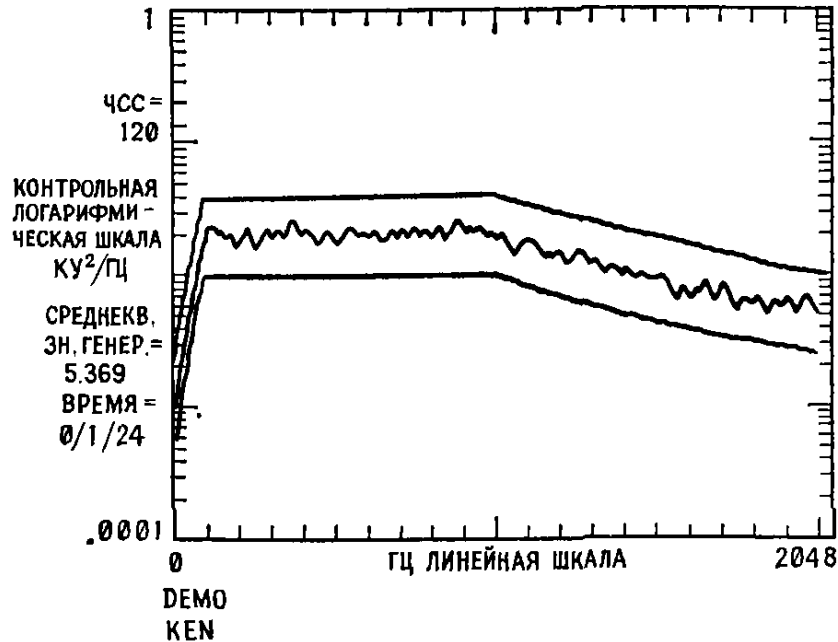


Рис. 9.11. Поведение согласованного спектра (КУ — коэффициент усиления).

9.11. ФУНКЦИЯ КОГЕРЕНТНОСТИ ПРИ $B_e = 1/P$

Функция когерентности служит очень полезным параметром, описывающим системы типа вход—выход. Вместе с тем когерентность обладает рядом особых статистических свойств. Если внимательно посмотреть на формулы для ошибок, приведенные в § 9.5, в частности на равенство (9.20), то обнаружится необычное поведение смещения, которое при $n \rightarrow 2$ стремится к бесконечности. Это объясняется тем, что при $n = 2$ функция когерентности равна в точности единице. Поясним это следующим соотношением:

$$\gamma_{xy}^2(k) = \frac{|S_{xy}(k)|^2}{S_x(k) S_y(k)} = \frac{(1/P)^2 X(k) Y^*(k) X^*(k) Y(k)}{(1/P) X^*(k) X(k) (1/P) Y^*(k) Y(k)} \equiv 1.$$

Оно справедливо в силу того, что при $n = 2$ (или $B_e = 1/P$) никакого осреднения нет. Вообще говоря, среднее отношения, разумеется, не равно отношению средних, и, следовательно, в общем случае когерентность не должна быть единицей.

Рассмотренное явление помогает уяснить более наглядное представление о когерентности как о линейной связи. Предположим, что имеется одна точка на комплексной плоскости, представляющая собой преобразование Фурье $X(f)$ в некоторой частоте. Далее, допустим, что на этой же плоскости находится точ-

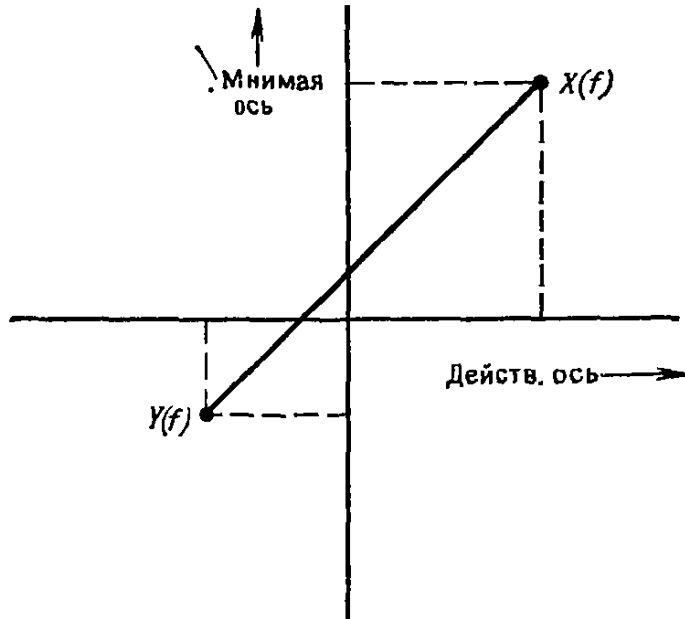
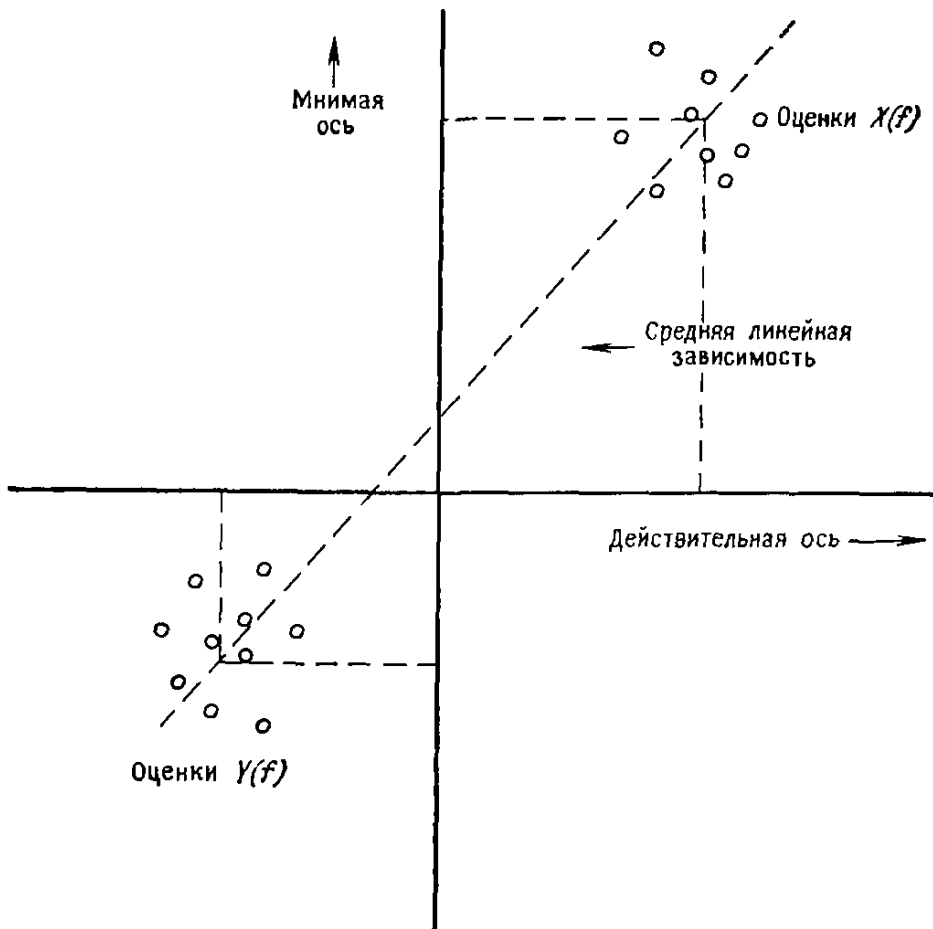


Рис. 9.12. Иллюстрация полиой когерентности в одной частоте.

Рис. 9.13. Осредненное линейное соотношение между группами оценок $X(f)$ и оценок $Y(f)$.

ка второго преобразования Фурье, на этот раз $Y(f)$. Как показано на рис. 9.12, эти две точки определили единственную прямую; следовательно, они связаны лишь одним линейным соотношением, и, значит, когерентность должна быть равна 1.

Если же имеется не две, а несколько оценок преобразования Фурье, каждое со своей ошибкой, то все они сгруппируются вокруг некоторых средних значений $X(f)$ и $Y(f)$. В этом случае прямую можно провести по средним, определенным методом наименьших квадратов. Здесь когерентность будет служить мерой точности подгонки линии (рис. 9.13).

Кроме того, как показал простым способом Купмэнс (1974), с. 309, при слишком широком разрешении может возникать смещение когерентности, вызванное размытием. Это касается, в частности тех областей исследования, где встречаются быстро меняющиеся фазы. Такие фазы возникают, например, при резонансах механических систем.

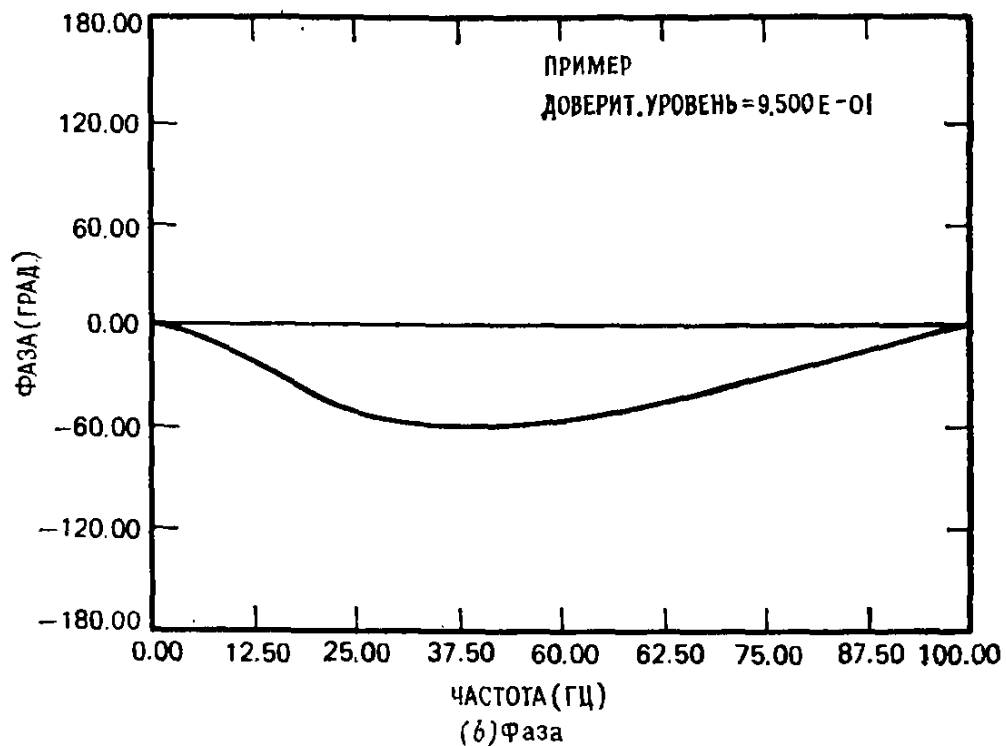
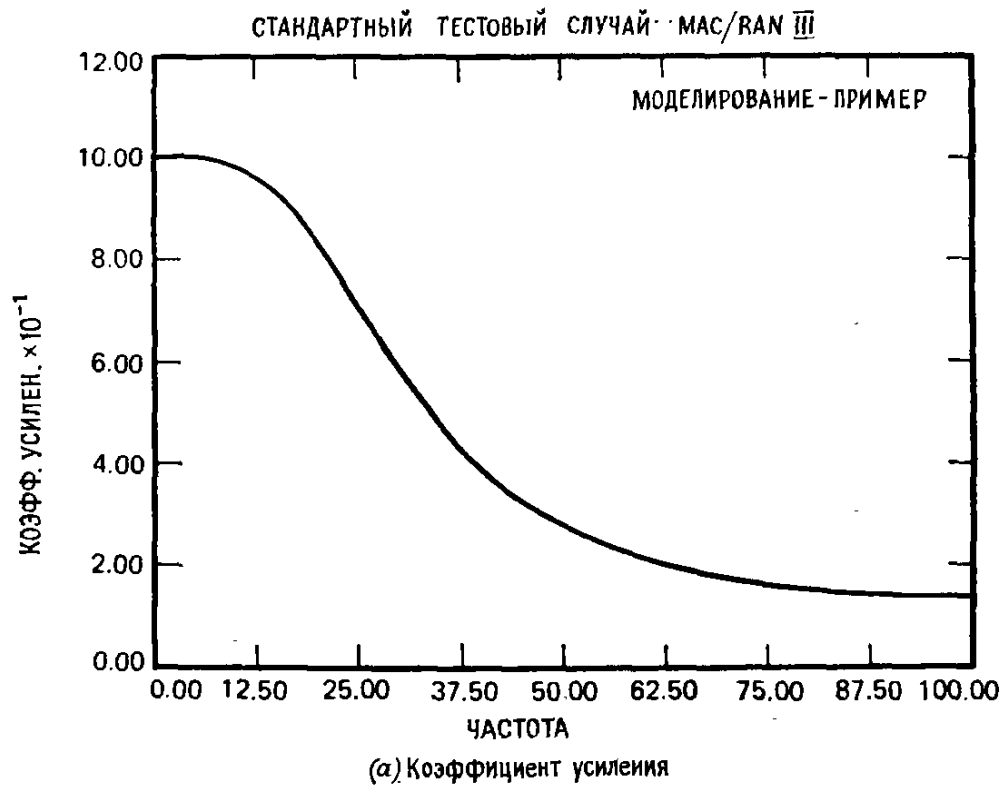
9.12. ПРИМЕРЫ ВЫЧИСЛЕНИЙ ПЕРЕДАТОЧНЫХ ФУНКЦИЙ

Цифровой фильтр. Если конструируется цифровой фильтр, то мы сталкиваемся с примером совершенно когерентной линейной системы, для которой получить хорошие оценки передаточной функции не составляет труда. Они показаны на рис. 9.14. Здесь изображена передаточная функция низкочастотного рекурсивного фильтра Баттеруорта, полученная при помощи вычислений кросс-спектра.

Как видно из рисунка, коэффициент усиления и фаза — совершенно гладкие функции. Те статистические ошибки, которые здесь имеются, очень малы.

Эффект добавления шума. Продемонстрируем влияние добавления шума на примере простого электрического фильтра, к выходу которого добавлен независимый белый шум. На рис. 9.15—9.17 рассмотрены ситуации с 0, 10 и 50%-м уровнем постороннего шума. В каждом из случаев число ст.св. в анализе одинаково и равно 128. Следует заметить, что белый шум, добавляемый к выходу, имеет по определению постоянную мощность вдоль всей полосы частот. Тем не менее на выходе мощность распределена по частотам неравномерно. В полосе пропускания полоснопропускающий фильтр оставляет большую мощность, чем на всех остальных частотах. Поэтому отношение сигнала к постороннему шуму ниже вне полосы пропускания. Это находит свое отражение в гораздо большей статистической изменчивости вне полосы пропускания.

Функция когерентности неразрывно связана с передаточной функцией. Это объясняется исключительно важной ролью

Рис. 9.14. *a* — коэффициент усиления; *b* — фаза цифрового фильтра.

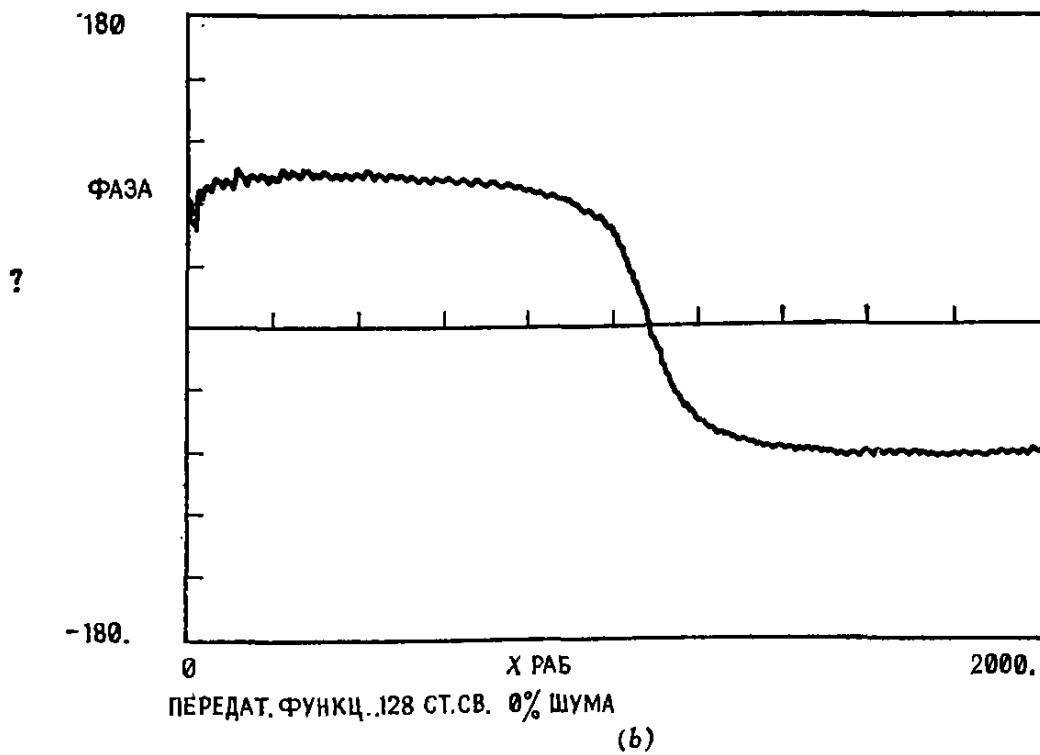
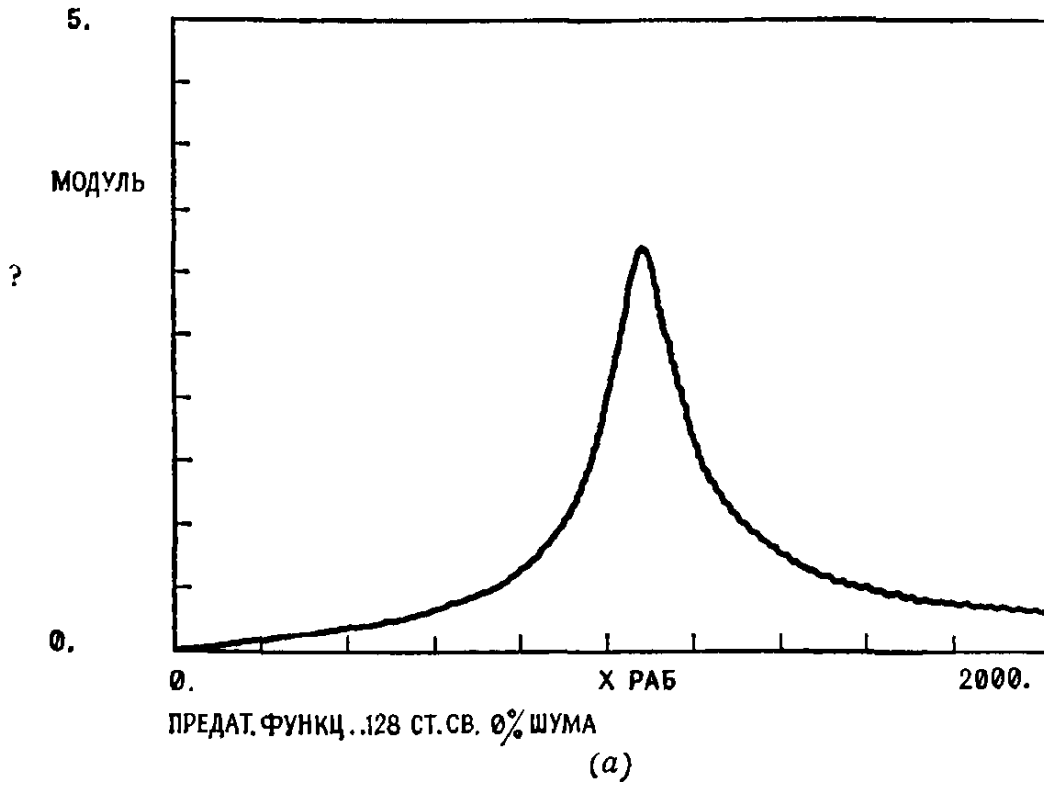


Рис. 9.15. Оценка передаточной функции цифрового фильтра при 0%-м уровне шума.

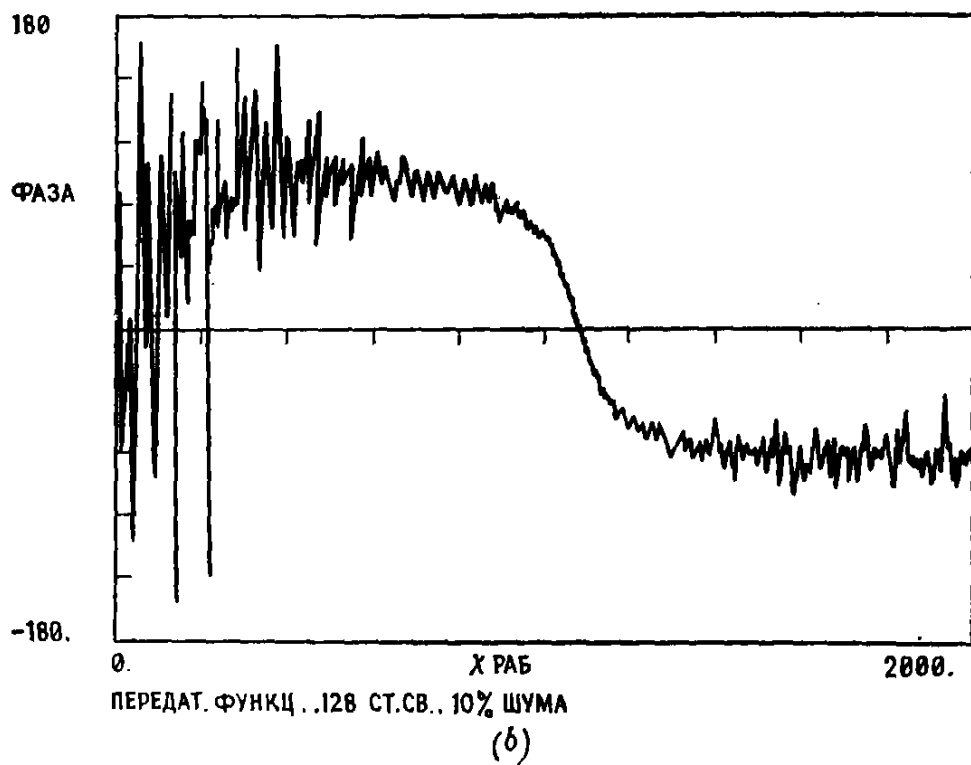
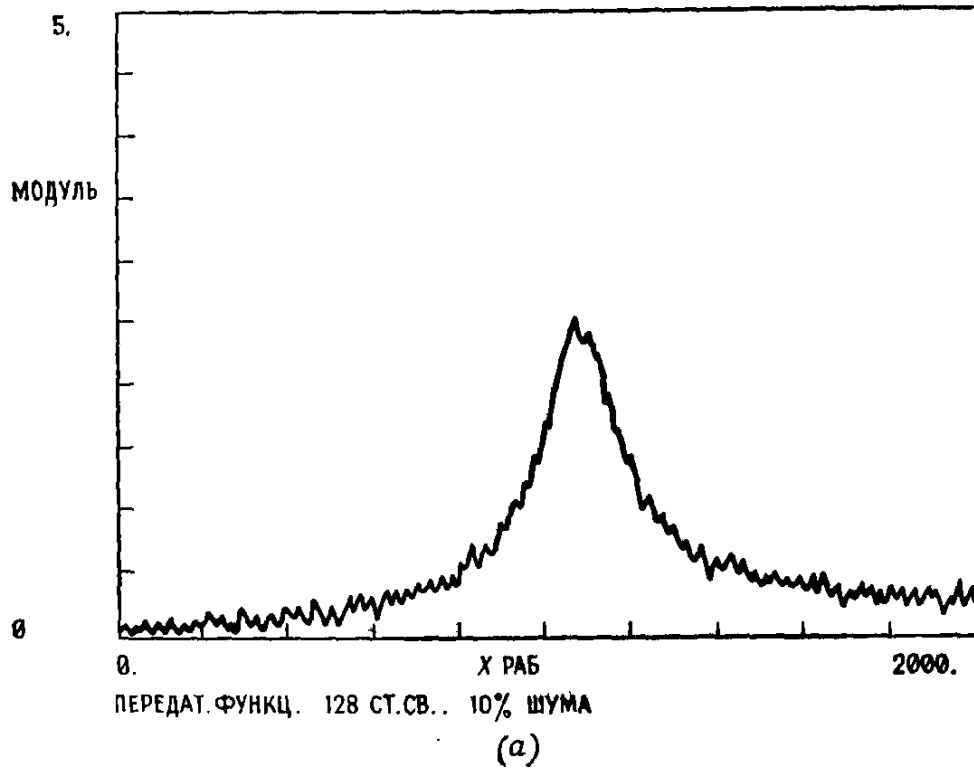


Рис. 9.16. Оценка передаточной функции цифрового фильтра при 10%-м уровне шума.

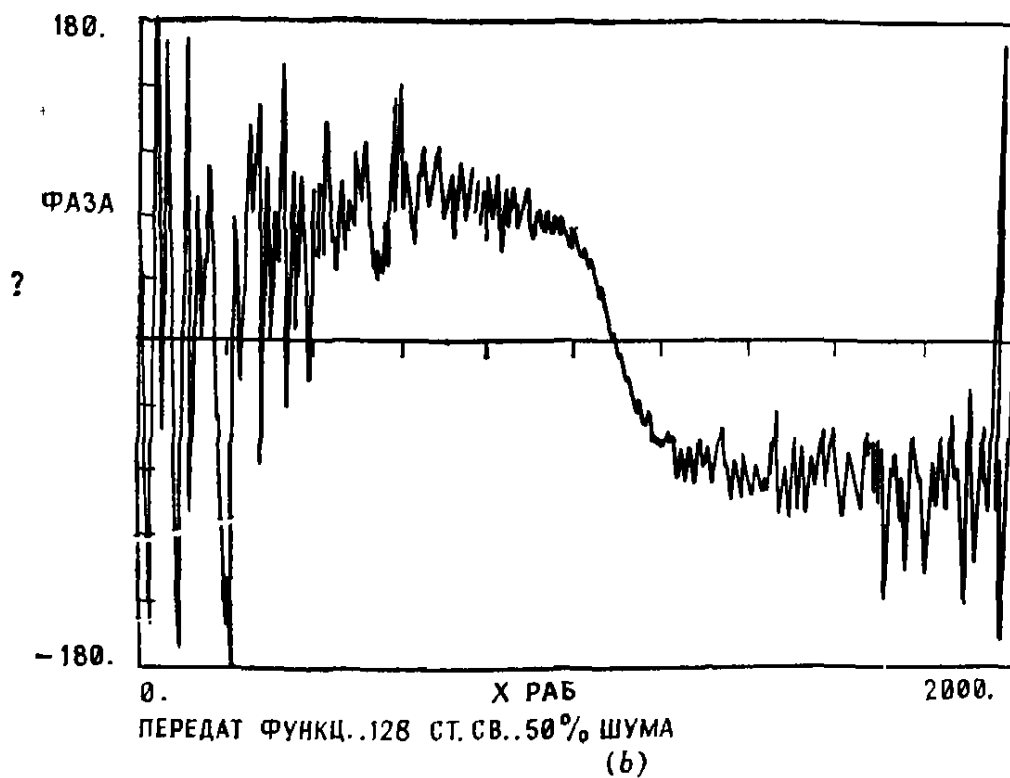
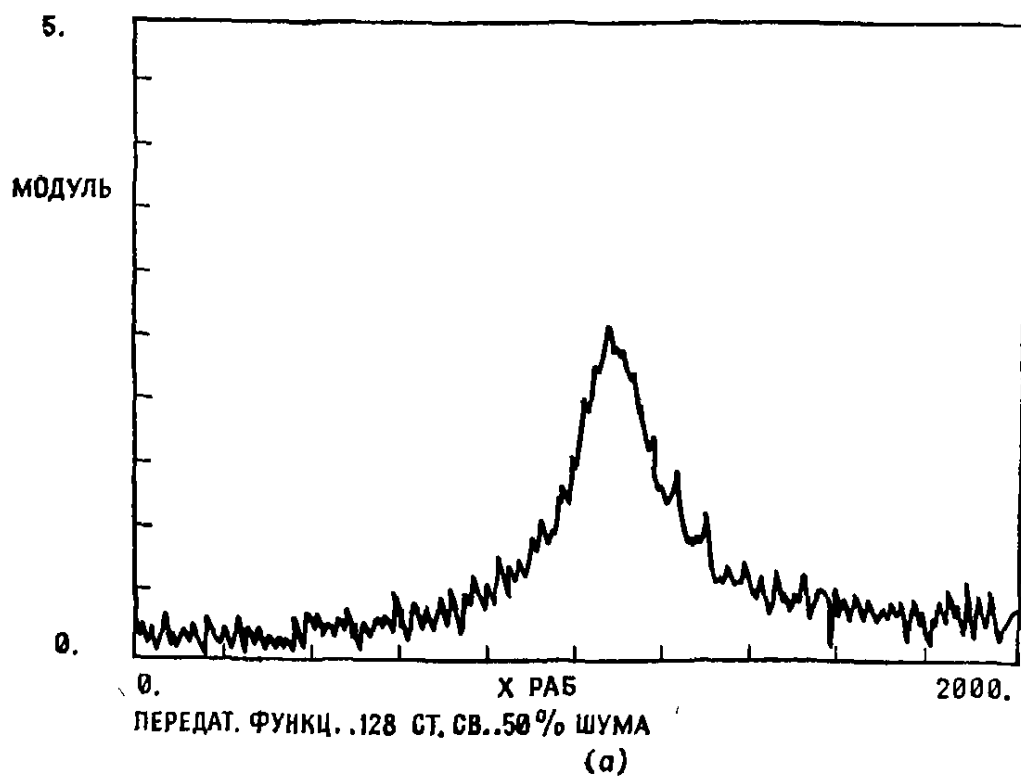
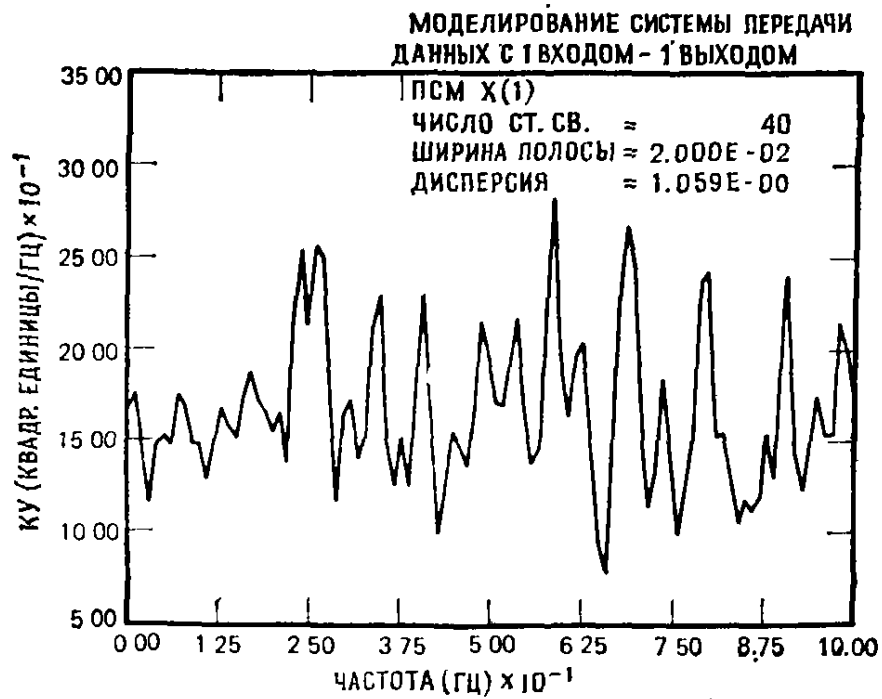
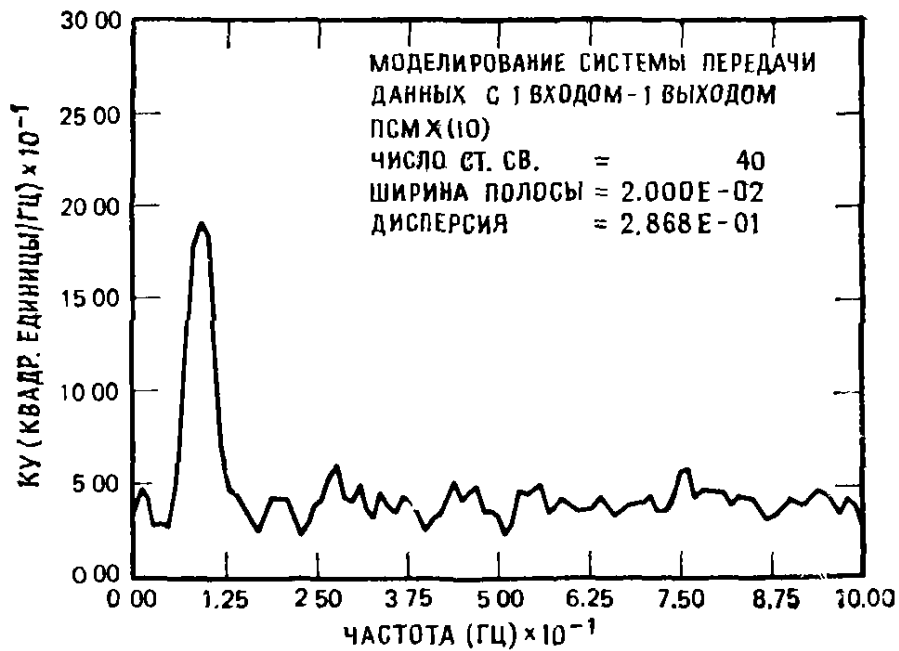
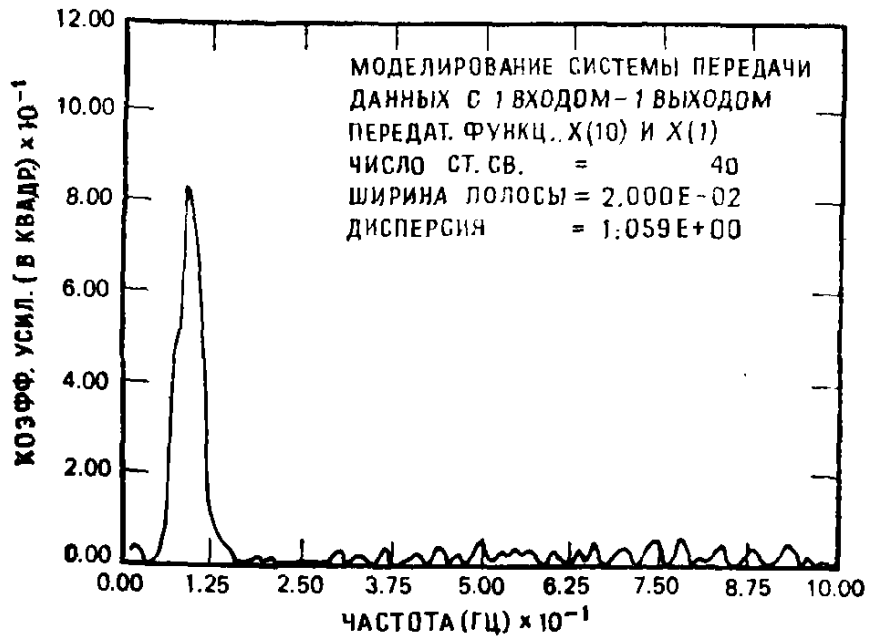
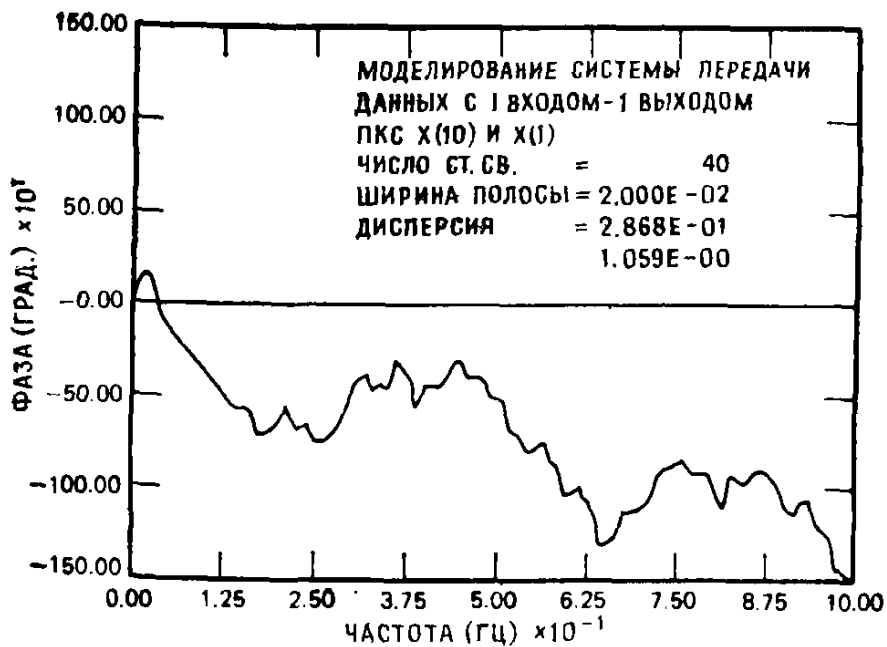


Рис. 9.17. Оценка передаточной функции цифрового фильтра при 50%-м уровне шума.

Рис. 9.18. ПСМ входа $\hat{G}_{xy}(f)$.Рис. 9.19. ПСМ выхода $\hat{G}_{yy}(f)$.

Рис. 9.20. Коэффициент усиления фильтра $|\hat{H}_{xy}(f)|$.Рис. 9.21. Фаза фильтра $\hat{\varphi}_{xy}(f)$.

функции когерентности в оценке статистической точности результатов — количественную характеристику этой точности дают вычисления доверительных границ. Что же касается качественной характеристики, то, учитывая ее немалую пользу в исследованиях, мы коснемся сейчас связанных с ней вопросов подробнее. На рис. 9.18—9.22 приведены $\hat{G}_{xx}(f)$, $\hat{G}_{yy}(f)$, $|\hat{H}_{xy}(f)|$, $\hat{\varphi}_{xy}(f)$ и $\hat{\gamma}_{xy}^2(f)$ соответственно; здесь $x(t)$ — псевдослучайный белый шум,

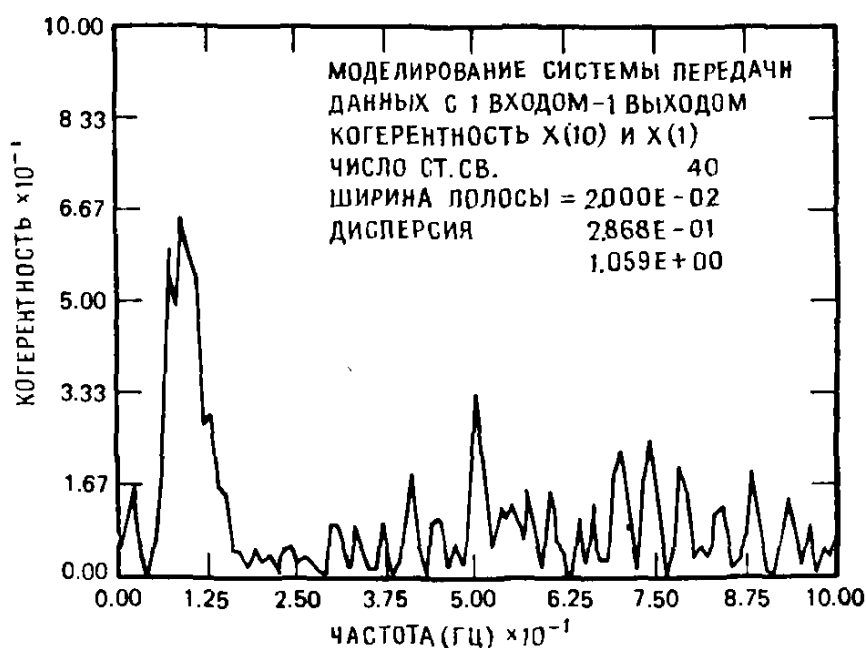


Рис. 9.22. Когерентность $\hat{\gamma}_{xy}^2(f)$.

функция $H_{xy}(f)$ соответствует полоснопропускающему цифровому фильтру. График когерентности показывает, что при высокой истинной когерентности в полосе пропускания фильтра (около 0.8 в этом случае, поскольку добавлен посторонний шум) оценка когерентности становится ближе к истинному значению. С другой стороны, низкая когерентность (почти равная нулю вне полосы пропускания) приводит к тому, что оценка проявляет высокую изменчивость. Отметим, что важно принимать в расчет и общий вид функции. Очень отчетливо выделяется диапазон высокой когерентности, расположенный в низких частотах. Это заставляет «полагать», что истинная когерентность на самом деле высока (начиная примерно с 0.2 Гц наблюдается очень нерегулярное поведение графика). Поэтому, даже встречая отдельные довольно высокие пики (например, $\hat{\gamma} = 0.4$ около 0.5 Гц), не следует придавать им большого значения ввиду «общей» нерегулярности функции. В итоге наибольшее доверие должны вызывать части графиков передаточной функции и фазы, заключенные в диапазоне от 0.05 до 0.15 Гц, и меньшее — вне этого диапазона. К сожалению, нам неизвестен более точный коли-

чественный способ описания поведения такого типа. Однако подобного рода ситуации на практике не редкость, поэтому такие рассуждения весьма полезны.

Пример системы со многими входами. Рассмотрим искусственную систему с двумя входами (рис. 9.23). Передаточные функции системы $H_1(f)$ и $H_2(f)$ отвечают двум восьмиполосным полоснопропускающим фильтрам с полосами пропускания от 0.07 до 0.12 Гц и от 0.14 до 0.24 Гц соответственно. Корреляция входов моделируется добавлением к первому входу $x_1(t)$ зна-

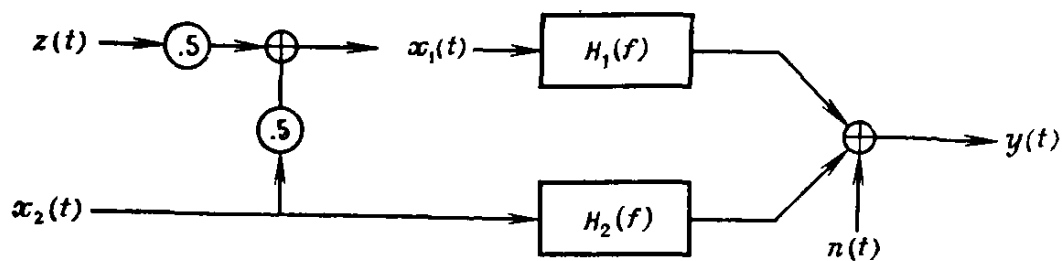


Рис. 9.23. Система с двумя коррелирующими входами.

чений $x_2(t)$ второго входа, уменьшенных вдвое. Если бы оба входа были полностью независимыми, то достаточно было бы решений для $H_1(f)$ и $H_2(f)$, полученных так же, как в ситуации с одним входом. Наличие корреляции между входами заставляет для вычисления точных результатов пользоваться подходом, предназначенным для систем со многими входами и включающим в себя операцию обращения матриц.

Изучение графиков на рис. 9.24—9.32 помогает сделать следующие выводы.

1. Характеристики полоснопропускающих фильтров по графикам коэффициентов усиления можно установить лишь довольно грубо.

2. График множественной когерентности показывает, что данные на выходе проявляют высокую степень линейной связанности со входами на диапазонах частот, определяемых коэффициентами усиления полоснопропускающих фильтров. Как следует из формул доверительных границ, наиболее надежные результаты для передаточных функций получаются там, где множественная когерентность имеет большие значения. Экспериментально определяем, что отдельным пикам графика когерентности в диапазоне высоких частот можно не придавать значения ввиду «общего» нерегулярного поведения там графика.

3. Сопоставление обычной и частной когерентностей для двух каналов подтверждает тот факт, что линейная связь каналов ограничена диапазонами примерно 0.03—0.10 Гц и 0.13—0.26 Гц для каналов 1 и 2 соответственно. Эти диапазоны выявились



Рис. 9.24. Пример со многими входами. Частная когерентность для второго фильтра. График показывает, что эффекты, сохраняющиеся после первой полосы пропускания, почти полностью исчезают, а эффекты после прохождения второй полосы пропускания остаются.

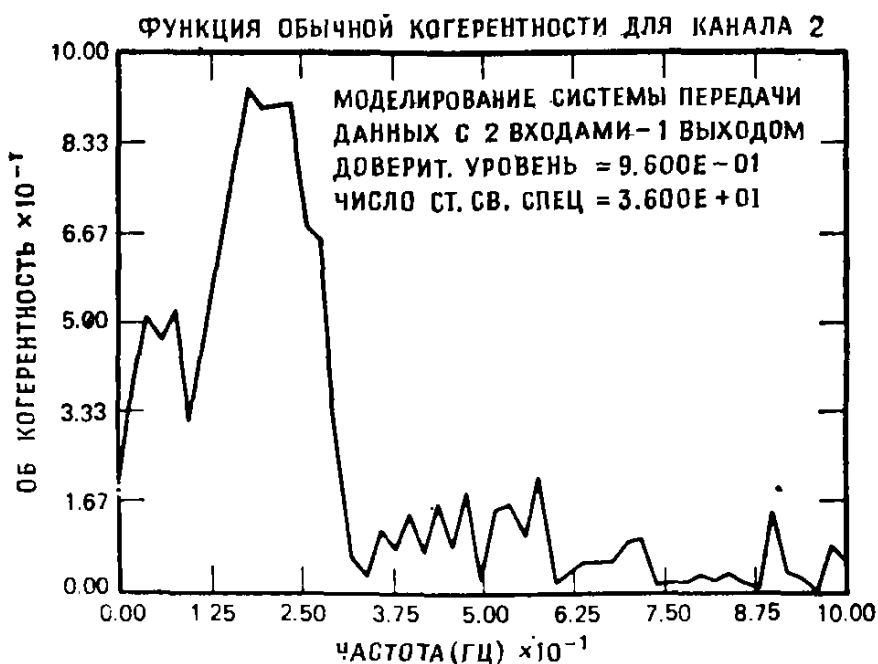


Рис. 9.25. Пример со многими входами. Обычная когерентность для второго входа оставляет полосу пропускания первого фильтра примерно на половинном уровне мощности (когерентность равна примерно 0.5).

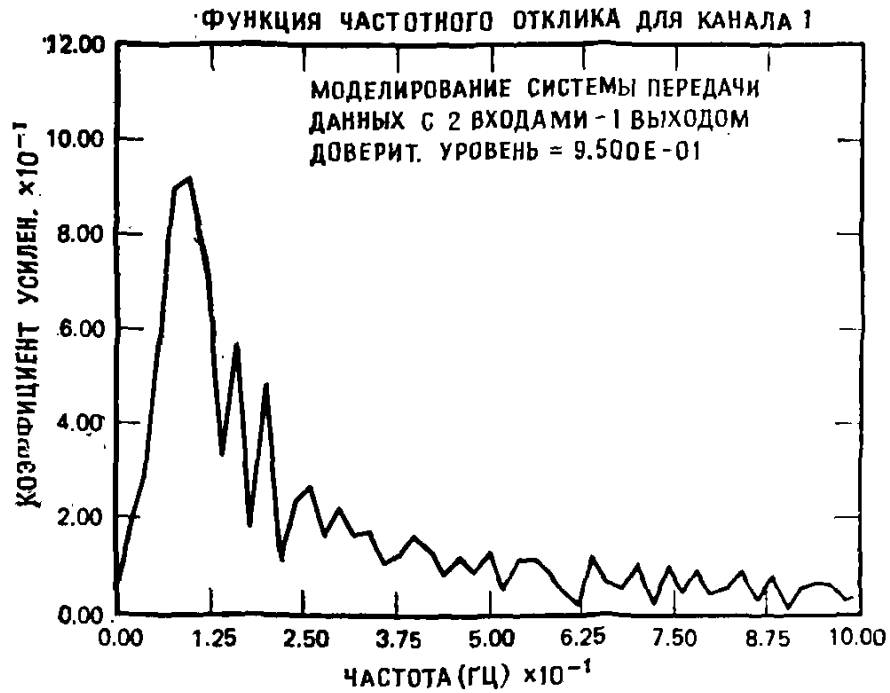


Рис. 9.26. Пример со многими входами. Коэффициент усиления в полосе пропускания (примерно 0.07—0.12 Гц) первого фильтра.

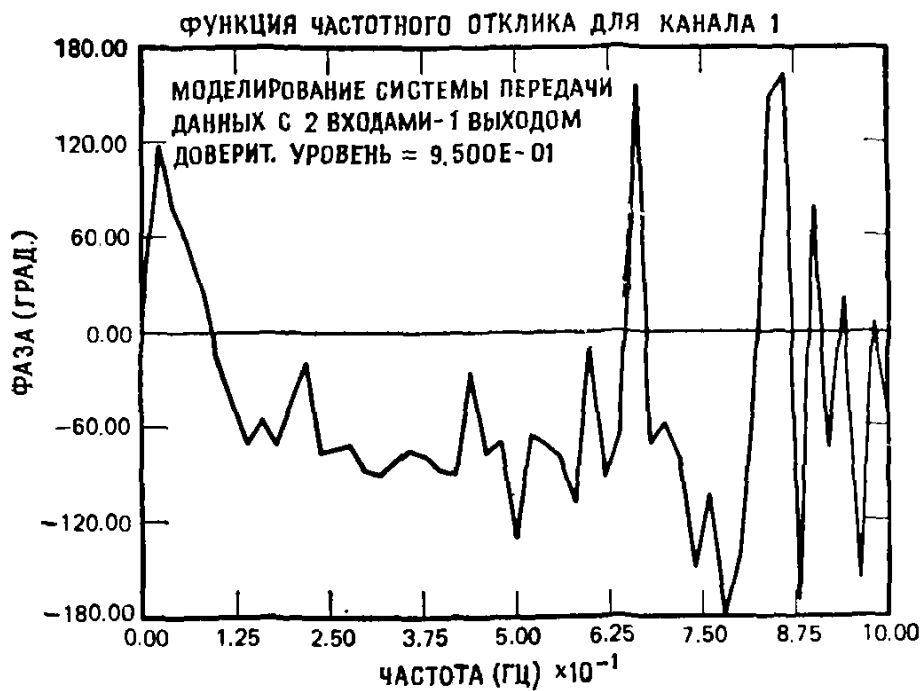


Рис. 9.27. Пример со многими входами. Выборочная фаза первого фильтра.



Рис. 9.28. Пример со многими входами. Коэффициент усиления в полосе пропускания (примерно 0.14—0.24 Гц) второго фильтра.

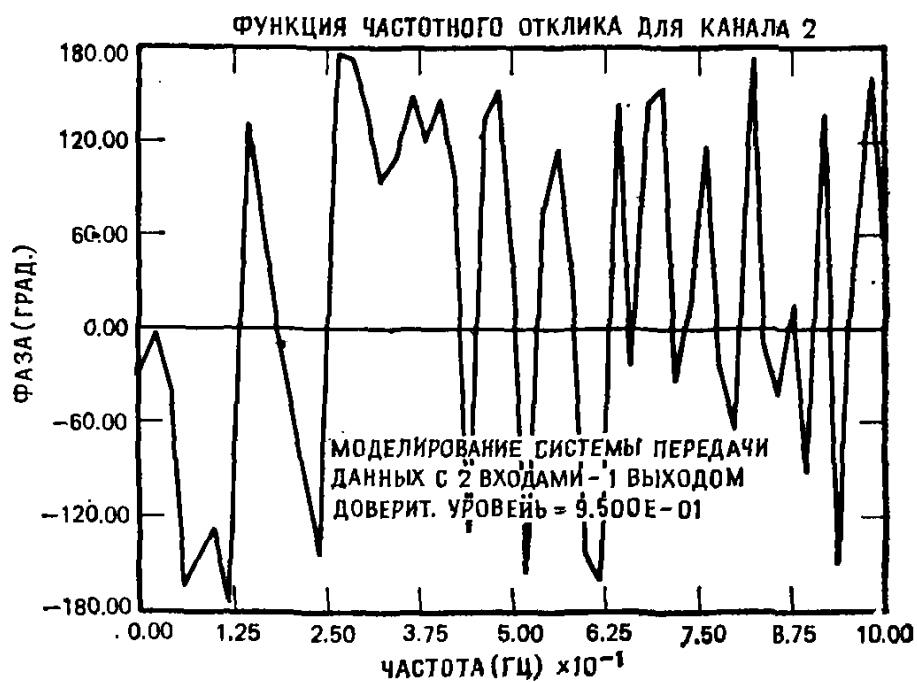


Рис. 9.29. Пример со многими входами. Выборочная фаза второго фильтра.

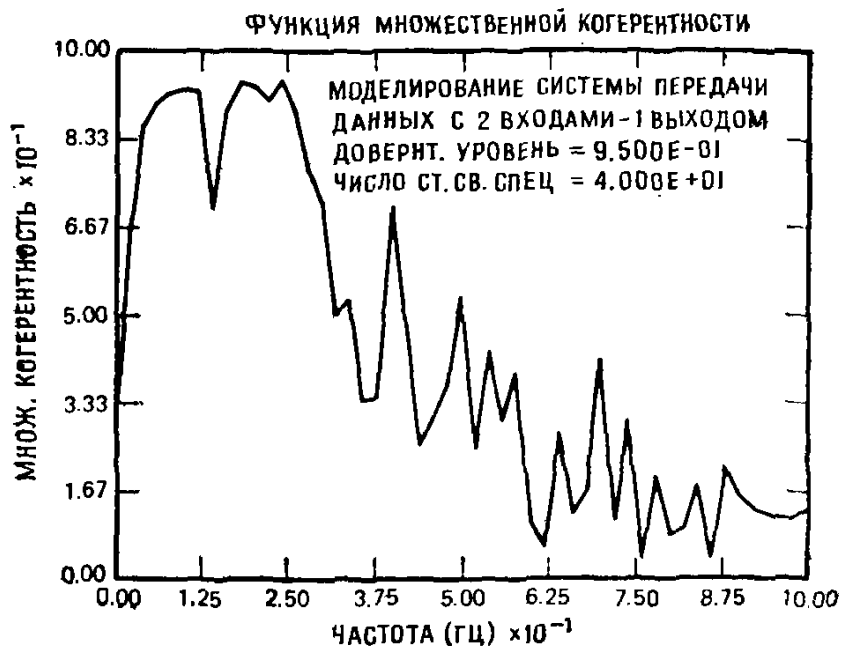


Рис. 9.30. Пример со многими входами. Функция множественной когерентности. Обратите внимание на высокую когерентность в полосах пропускания фильтров.

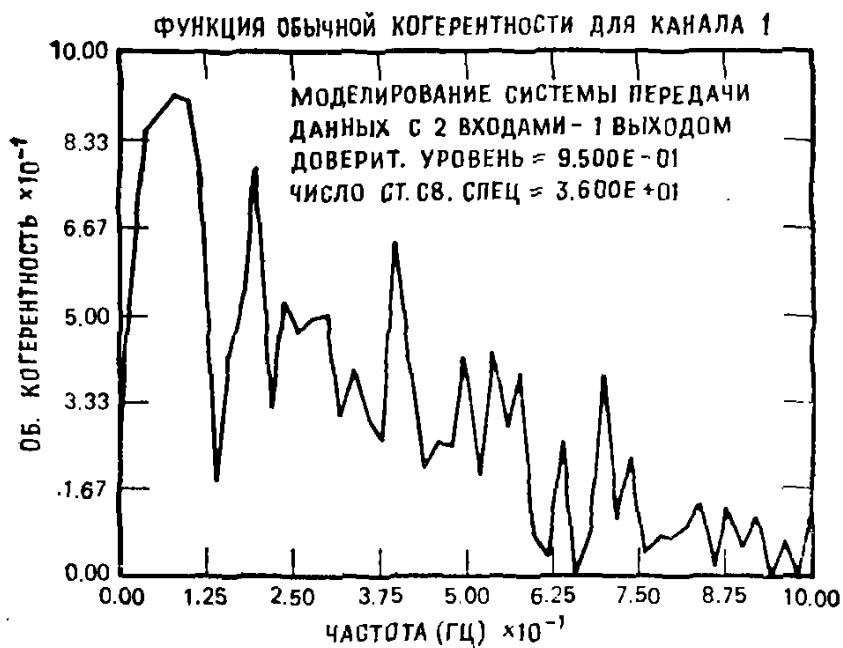


Рис. 9.31. Пример со многими входами. Обычная когерентность для первого фильтра. Обратите внимание на высокую когерентность в полосе пропускания первого фильтра и несколько неустойчивую высокую когерентность в полосе второго фильтра.

в результате того, что видимая мощность в высоких частотах в первом случае и в низких частотах во втором случае пропала при вычислении частной когерентности.

Рассмотренные примеры были призваны объяснить, как делаются различные полезные выводы на основе вычислений (в случае многих входов) передаточных функций и различных функций когерентности. Выводы становятся более точными, если учитываются физические условия. Например, если можно выдвинуть

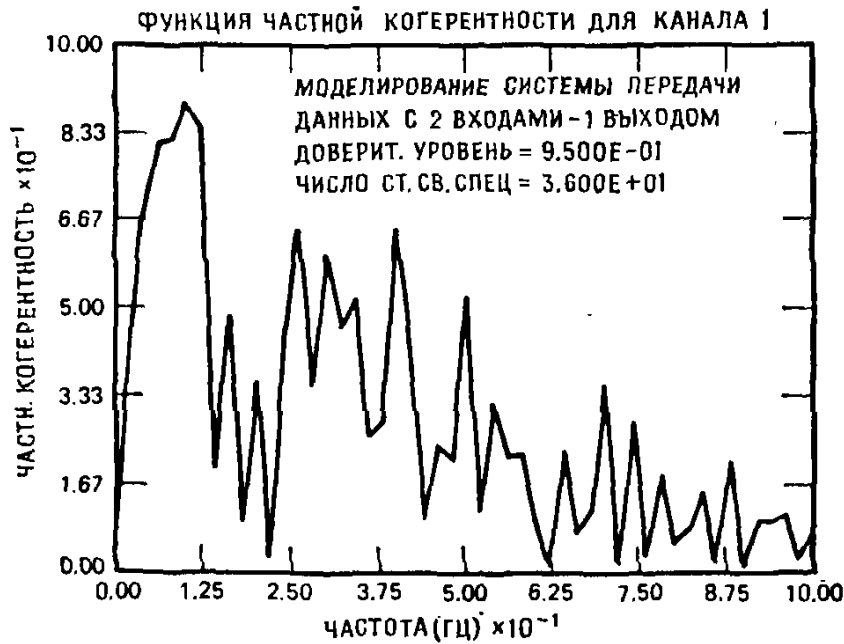


Рис. 9.32. Пример со многими входами. Частная когерентность для первого фильтра. Обратите внимание на то, что большая часть значений когерентности на протяжении полосы пропускания — очень малые величины. Исключение составляет первый всплеск.

гипотезу о том, что на двух каналах с передачей стоят полоснопропускающие фильтры, то, получив результаты, легко сделать вывод о том, противоречат они гипотезе или нет.

Некоторые замечания к графикам. В области теории управления передаточные функции по традиции переводят на графиках в дБ, используя логарифмическую шкалу. Мы предпочитаем графики типа линейная шкала — линейная шкала, хотя для выбора каждого представления есть свои аргументы, говорящие в его пользу. Обозначение дБ пришло из акустики, где оно отвечает децибелам, определяемым формулой $1\text{дБ} = 10 \log(x/\text{ед})$, где «ед» — единица измерения переменной x . Обычно в качестве такой единицы в теории управления выступает напряжение $1\text{дБ} = 20 \log(v_1/v_2)$ или, через отношение мощностей, $1\text{дБ} = 10 \log(v_1^2/v_2^2) = 20 \log(v_1/v_2)$, поскольку мощность пропорциональна квадрату напряжения.

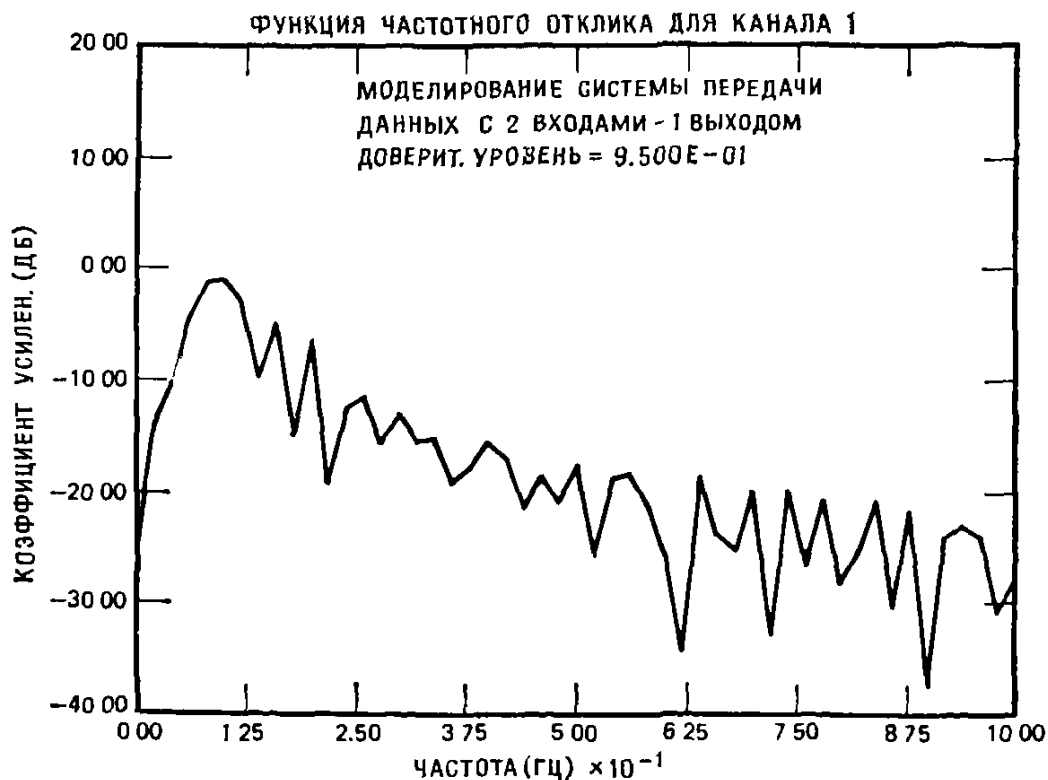


Рис. 9.33. График коэффициента усиления фильтра для канала 1 (шкала децибел). Сравните с рис. 9.26.



Рис. 9.34. График коэффициента усиления фильтра для канала 2 (шкала децибел). Сравните с рис. 9.26.

Для коэффициентов усиления передаточных функций роль «ед» выполняет коэффициент $1: |H| \text{ дБ} = 10 \log \frac{|H(f)|^2}{1} = 20 \log |H(f)|$.

Единица выбирается потому, что выход физической системы нельзя усилить (если не используется внешняя мощность за счет усилителя), и, значит, коэффициент, равный единице, будет давать фильтр, пропускающий все данные без изменения (на всех частотах). Значению 1 соответствует в дБ значение 0, поэтому затухание характеризуется отрицательными значениями.

Логарифмическое преобразование расширяет диапазон представимости данных на графиках. Это положение иллюстрируют графики на рис. 9.33 и 9.34, которые получены из графиков на рис. 9.26 и 9.28 переходом к дБ. Заметим, что логарифмическая шкала после перехода к дБ дана в линейных единицах. Следует иметь в виду, что 10 дБ отражают изменение, характеризующееся множителем 10. Почти все понятия, затронутые сейчас, пришли из акустики — человеческое ухо воспринимает силу звука логарифмическим образом. Иначе говоря, при возрастании звука в 100 раз его громкость будет казаться лишь вдвое больше.

Логарифмическое преобразование частоты растягивает нижние и сжимает верхние частоты. Полосы пропускания многих физических систем шире на высоких частотах. В этом случае логарифмическое преобразование сделает полосы низких и высоких частот постоянными.

Упражнения

9.1. Приведите опущенные при выводе формулы (9.10) выкладки.

9.2. Докажите, что формулы (9.16) и (9.9) эквивалентны.

9.3. Докажите формулу (9.29).

9.4. Перепишите соотношения (9.27) — (9.29) в их простейшей форме через элементы спектральной матрицы.

9.5. Получите (9.4) из (9.3), применив преобразование Фурье к (9.3). Предполагается, что все необходимые математические условия существования преобразований и т. п. выполнены.

Почти все формулы для линейных систем можно вывести при помощи комплексных аналогов величин, к которым применяется преобразование Фурье (в каждой частоте f), и понятиями классической статистики. Определим дисперсию формулой $E[X^*X]$, а ковариацию — формулой $E[X^*Y]$. Пусть ПСМ служит аналогом дисперсии, а кросс-спектр — ковариации. Эти соглашения используйте в упр. 9.6—9.8

9.6. Выведите соотношение (9.7) как аналог статистической формулы

$$\sigma_y^2 = \beta^2 \sigma_x^2 \quad y = \beta x + \epsilon,$$

где ϵ — независимый шум с нулевым средним.

9.7. Выведите соотношение (9.8) как аналог статистической формулы $\sigma_{xy} = \beta_{xy} \sigma_x^2$, где $\sigma_{xy} = \text{cov}[x, y]$.

9.8. Что служит аналогом обычной когерентности (9.9) в классической статистике?

9.9. Получите необходимые соотношения для частного случая системы с двумя входами и одним выходом.

9.10. Получите обратным преобразованием Фурье формулу (9.30).

Приложение А

ПРОГРАММЫ ДЛЯ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ

А.1. ВВЕДЕНИЕ

Это приложение должно вооружить читателя, изучающего методы анализа временных рядов, набором тех программ, которые понадобятся при работе над упражнениями.

Все включенные в программы команды записаны на упрощенном Фортране, что дает возможность прогонять эти программы на большинстве ЭВМ, снабженных трансляторами с Фортрана.

Характер этих программ во многом определен целями книги. Главное, на что обращалось внимание при их составлении, — это прозрачность смысла и возможно более широкий спектр применимости. Разумеется, порой это достигалось ценой некоторого снижения эффективности и увеличения необходимой для вычислений памяти. Поэтому читателю не стоит удивляться, если ему удастся подобрать программу, которая будет отвечать его собственным требованиям в большей степени.

Хотелось бы сразу предостеречь и еще от одной ошибки. Хотя приведенные программы в том виде, как они здесь представлены, полезны при решении задач, затронутых в этой книге, ими не исчерпываются все приемы, которые нужны при работе с временными рядами. Рассмотрение предложенных программ преследует более скромную цель — привить читателю элементарные навыки работы с задачами, возникающими при исследованиях реальных временных рядов. Сложность программ, при помощи которых они обрабатываются, оказывается почти на два порядка выше.

В приложение включены следующие подпрограммы: TDRAND, реализующая генератор случайных чисел с равномерным распределением; PRPLOT, предназначенная для получения графического изображения результатов; FFTRAN — для вычисления быстрого преобразования Фурье; LPSB — для вычисления весовых коэффициентов низкочастотных фильтров Баттеруорта; TTRAN — для определения передаточных функций цифровых фильтров.

Помимо этого, приложение содержит небольшую программу, которая позволяет частично проверять приведенные программы. Прежде чем начать работать над любой задачей, рекомендуется прогнать эту тестовую программу и снять возникшие вопросы.

Ведь достаточно неправильно ввести символ, чтобы затем часами из-за ошибки в программе безуспешно пытаться решить задачу. Если же сначала прогнать тест, то большая часть ошибок такого рода будет сразу выявлена. Но, пожалуйста, не забывайте, что проверка при помощи тестовой программы не всеобъемлюща: эта программа может работать правильно и тогда, когда имеются ошибки перфорации или ошибки в программе.

Другое обстоятельство, на которое следует обратить внимание, состоит в том, что все приведенные подпрограммы *не* выполняют проверки диапазона входных параметров для определения точности. Если для некоторых параметров берутся плохие значения, то арифметические ошибки, по своей природе подобные тем, что служат причиной преждевременного уничтожения программы на некоторых операционных системах, станут почти неизбежными.

А.2. ГЕНЕРАТОР СЛУЧАЙНЫХ ЧИСЕЛ

При помощи подпрограммы TDRAND (рис. А.1) при каждом обращении вырабатывается одно равномерно распределенное

```
SUBROUTINE TDRAND (X)
DATA I/783637/
I=125*I
I=I-(I/2796203)*2796203
X=FLOAT(I)/2796202.
RETURN
END
```

Рис. А.1. Генератор равномерно распределенных случайных чисел.

псевдослучайное число. Последовательность операторов вызова сводится к предложению

CALL TDRAND(X),

где X — получающееся случайное число. Оно имеет равномерное распределение на интервале (0, 1), математическое ожидание $1/2$ и дисперсию $1/12$.

Здесь рассматривается, по существу, версия Л. Хэнсона реализации алгоритма 266 [G5], предложенного Пайком и Хиллом (Пайк, Хилл (1965)).

Вернуться в исходное положение подпрограмма TDRAND практически не позволяет.

На рис. А.2 и А.3 приведены гистограмма и плотность спектра мощности 1000 псевдослучайных чисел, полученных при помощи подпрограммы TDRAND, из которых вычиталась $1/2$ (чтобы среднее было равно 0).

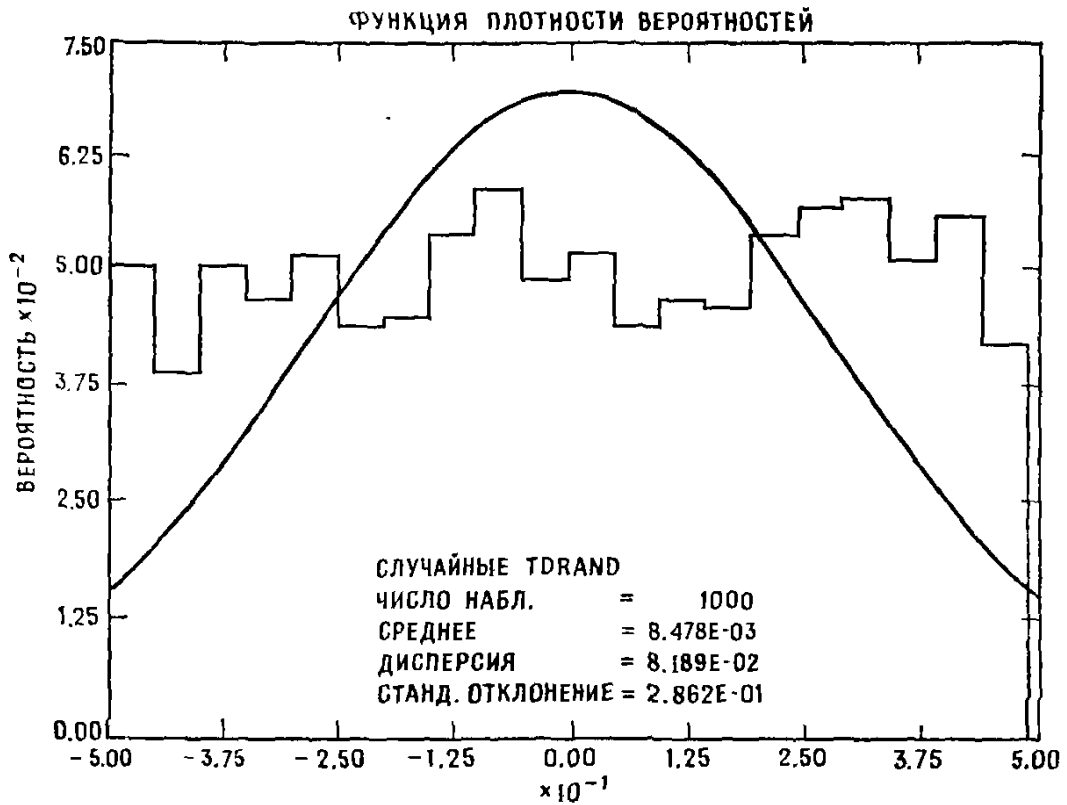


Рис. А.2. Гистограмма 1000 значений, полученных процедурой TDRAND. Чтобы гистограмма была симметричной относительно нуля, из каждого значения вычиталась 1/2.

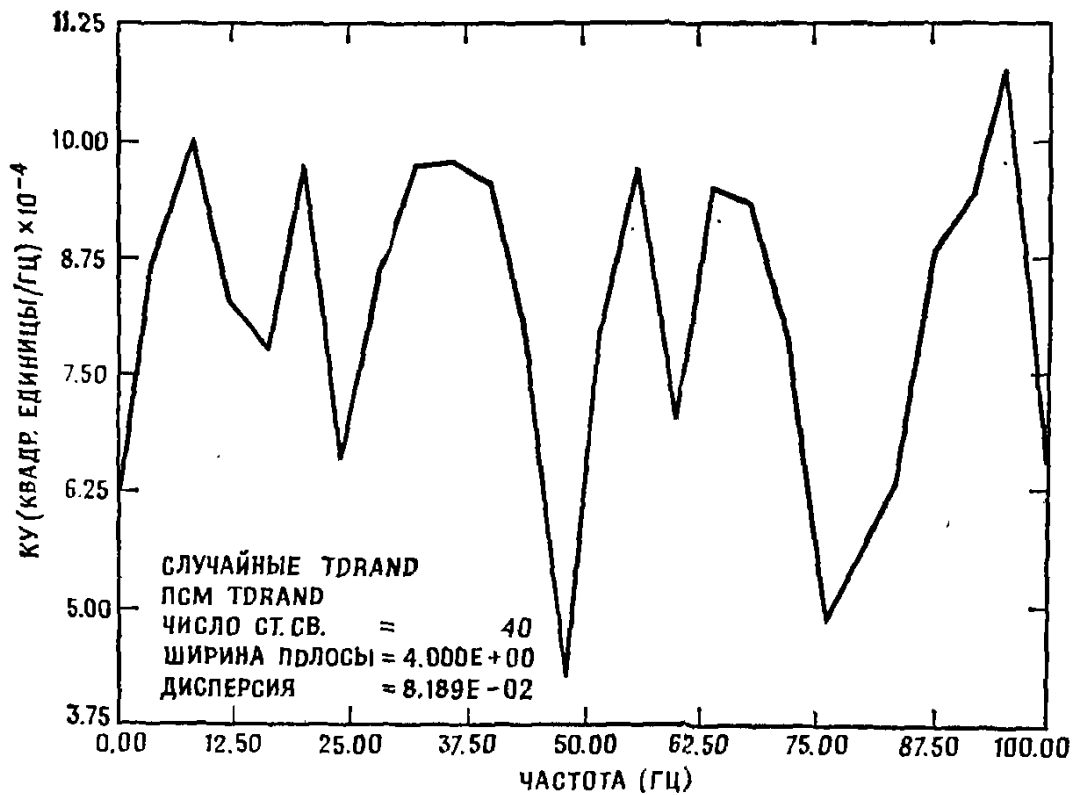


Рис. А.3. Плотность спектра мощности 1000 значений. Предполагается, что $T = 0,005$.

Гистограмма была построена с 20 одинаковыми карманами. Вероятность попадания в каждый из них равна 0.05. Ясно, что положение, при котором в каждый из карманов попадает ровно по 50 значений, слишком маловероятно. Тем не менее, как можно видеть, полученные результаты довольно близки к этим величинам.

Теоретическое значение плотности спектра мощности должно в рассматриваемом случае быть постоянным и равным $8.3333 \cdot 10^{-4}$. Действительно, полученные результаты располагаются около этого значения.

А.3. ПОЛУЧЕНИЕ ГРАФИКОВ НА ПЕЧАТАЮЩЕМ УСТРОЙСТВЕ

Подпрограмма PRPLOT (рис. А.4) представляет собой простую подпрограмму, позволяющую получать на печатающем устройстве графики данных с равным шагом. Последовательность операторов вызова имеет вид

```
CALL SUBROUTINE PRPLOT (Y,N,FIRST,DX,ISKIP)
```

Перечисленные параметры обозначают следующие величины: Y — данные (зависимая величина), значения которых располагаются в равноотстоящих точках, разделенных интервалом DX ; N — общее число точек, наносимых на график; $FIRST = X_0$ — первое значение независимого переменного (как правило, равное 0); DX — шаг, с которым чередуются последовательные значения величины X ; $ISKIP \leq 1$ — график всех значений Y ; $ISKIP > 1$ — график N значений Y , начиная с $Y(1)$, с пропуском $ISKIP - 1$ значений перед каждым новым значением.

Эта программа не выдает страницу до начала печати графика. Пользователь должен до начала работы PRPLOT пропустить несколько строк (чтобы отделить график). Кроме того, в информации выше нужно для выдачи страницы предусмотреть управление кареткой.

Последний параметр для выбора пропуска позволяет получать на графике действительные и мнимые части комплексных чисел (данных) отдельно и обходиться при этом без нарушения того естественного порядка, в котором эти данные хранились. Обратите внимание на то, что в качестве печатающего устройства выбран блок б. Такой выбор требуется на многих системах.

А.4. БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Подпрограмма FFTRAN (рис. А.5) позволяет вычислять быстрое преобразование Фурье равноотстоящих комплексных данных при помощи метода Кули — Тьюки. Предложение вызова имеет вид

CALL FFTRAN (SIGN, T, X, NPOW)

где SIGN = -1. для преобразования из времени в частоты,
= 1. — из частот во время; T — выборочный интервал во времен-

```

SUBROUTINE PRPLOT (Y,N,FIRST,DX,JSKIP)                                100010
DIMENSION Y(1), A(26), SYM(4)                                       100020
DATA BLNK/4H /, SYM(1)/4H+ /, SYM(2)/4H + /,                          100030
1SYM(3)/4H + /, SYM(4)/4H + /, IPRINT/6/                             100040
ISKIP=JSKIP                                                            100050
IF (ISKIP.LE.0) ISKIP=1                                              100060
DO 5 I=1,26                                                            100070
5 A(I)=BLNK                                                            100080
ZERO=FIRST                                                            100090
C=Y(1)                                                                100100
D=C                                                                    100110
DO 10 I=2,N                                                           100120
LOC=1+ISKIP*(I-1)                                                    100130
E=Y(LOC)                                                             100140
IF (E.GT.0) D=E                                                       100150
IF (E.LT.C) C=E                                                       100160
10 CONTINUE                                                           100170
IGT=1                                                                 100180
IF (D.GT.9999.) IGT=2                                                 100190
IF (C.LT.-9999.) IGT=2                                               100200
IF (D.LT.0.1.AND.C.GT.-0.1) IGT=2                                    100210
DMC=D-C                                                               100220
AA=(DMC)/100.                                                         100230
BB=(D+C)/2.                                                           100240
WRITE (IPRINT,45) C,BB,D                                             100250
IF (ABS(DMC).GT.1.E-20) GO TO 15                                       100260
WRITE (IPRINT,50)                                                    100270
GO TO 40                                                              100280
15 DO 35 I=1,N                                                         100290
LOC=1+ISKIP*(I-1)                                                    100300
XP=Y(LOC)                                                            100310
J=(XP-C)/AA+0.5                                                       100320
J4=J/4                                                                100330
JU=J-J4*4+1                                                           100340
J4=J4+1                                                              100350
A(J4)=SYM(JU)                                                         100360
GO TO (20,25),IGT                                                    100370
20 WRITE (IPRINT,55) ZERO,XP,A                                       100380
GO TO 30                                                              100390
25 WRITE ((PRINT,60) ZERO,XP,A                                       100400
30 ZERO=ZERO+DX                                                       100410
A(J4)=BLNK                                                            100420
35 CONTINUE                                                           100430
40 RETURN                                                             100440
C                                                                      100450
C                                                                      100460
C                                                                      100470
45 FORMAT (20X,F10.2,35X,F10.2,36X,F10.2)                             100480
50 FORMAT (26HODATA ALL EQUAL...NO PLOT.,//)                         100490
55 FORMAT (1X,F8.4,F10.4,1X,26A4)                                     100500
60 FORMAT (1X,F7.4,E11.4,1X,26A4)                                     100510
END                                                                    100520

```

Рис. А.4. Программа для графического изображения данных. Написана Р. К. Отнесом.

ной области; X — данные; NPOW — степень 2, определяющая длину N записи данных, т. е. $N = 2^{NPOW}$.

Предполагается, что в массиве переменного X запоминается $N = 2^{NPOW}$ комплексных значений данных. Таким образом, раз-

мерность X для действительных значений должна быть равной по крайней мере $2 \cdot (2^{**} NPOW)$.

Величина T всегда обозначает интервал, с которым производится выборка. Если необходимо, то по ней вычисляется величина b — шаг в области частот: $b = 1/(NT)$.

Вычисления проводятся по формулам

$$X(k) = T \sum_{i=0}^{N-1} x(i) \exp\left(-\frac{j2\pi ik}{N}\right),$$

где $N = 2^{NPOW}$.

Величины $x(i)$, так же как и $X(k)$, имеют комплексные значения. Они хранятся следующим образом:

Временная функция	Хранение в программе	Функция частоты	Хранение в программе
$x(0)$	$X(1)$	$X(0)$	$X(1), X(2)$
$x(1)$	$X(3)$	$X(1)$	$X(3), X(4)$
⋮	⋮	⋮	⋮
$x(i)$	$X(2i+1)$	$X(k)$	$X(2k+1), X(2k+2)$
⋮	⋮	⋮	⋮
$x(N-1)$	$X(2N-1)$	$X(N-1)$	$X(2N-1), X(2N)$

Решение многих задач упрощается, если 1) каждый раз преобразуется только одна функция, 2) для переменного x используется двумерный массив, определяемый описанием

DIMENSION X (2,1024)

Здесь $N = 1024$; функция, для которой ищется преобразование, запоминается следующим образом:

$$\begin{aligned} x(i) &\rightarrow X(1, I \text{ или } i+1), \quad I = 1, \dots, N, \\ 0 &\rightarrow X(2, I \text{ или } i+1), \quad i = 0, \dots, N-1. \end{aligned}$$

После преобразования результаты запоминаются так:

	Частота
Re $[X(0)]$ в $X(1, 1)$ Im $[X(0)]$ в $X(2, 1)$	0
⋮	⋮
Re $[X(k)]$ в $X(1, K \text{ или } k+1)$ Im $[X(k)]$ в $X(2, K \text{ или } k+1)$	kb
⋮	⋮
Re $[X(N/2)]$ в $X(1, N/2+1)$ Im $[X(N/2)]$ в $X(2, N/2+1)$	F

```

SUBROUTINE FFTRAN (SIGN,T,X,NROW)                                90010
C                                                                 90020
C   МЕТОД КУЛИ-ТЬЮКИ ВЫЧИСЛЕНИЯ ПРЕОБРАЗОВАНИЯ ФУРЬЕ          90030
C   ВКЛЮЧАЕТ ВЫЧИСЛЕНИЯ СИНУСОВ, КОСИНУСОВ И ПЕРЕУПОРЯДОЧЕНИЕ 90040
C   ДАННЫХ ЧЕРЕЗ ОБРАТНУЮ ПЕРЕСТАНОВКУ БИТОВ ИНДЕКСОВ          90050
C                                                                 90060
C   SIGN = УКАЗАТЕЛЬ НАПРАВЛЕНИЯ ПРЕОБРАЗОВАНИЯ ФУРЬЕ         90070
C   = -1. ДЛЯ ПРЯМОГО ПРЕОБРАЗОВАНИЯ                           90080
C   = 1. ДЛЯ ОБРАТНОГО ПРЕОБРАЗОВАНИЯ                          90090
C   T     = ШАГ ПО ВРЕМЕНИ                                     90100
C   X     = МАССИВ ДЛЯ ПРЕОБРАЗОВАНИЯ ФУРЬЕ                   90110
C   NROW  = СТЕПЕНЬ 2 (ОБЪЕМ БЛОКА = 2**NROW)                 90120
C                                                                 90130
DIMENSION X(1), CS(2), MSK(13)                                  90140
COMPLEX X,CXCS,HOLD,HA                                          90150
EQUIVALENCE (CXCS,CS)                                          90160
NMAX=2**NROW                                                    90170
ZZ=6.283185306*SIGN/FLOAT(NMAX)                                90180
DELTA=T                                                         90190
IF (SIGN) 10,10,5                                              90200
5 DELTA=1./(T*FLOAT(NMAX))                                     90210
10 MSK(1)=NMAX/2                                               90220
DO 15 I=2,NROW.                                               90230
15 MSK(I)=MSK(I-1)/2.                                         90240
NN=NMAX                                                         90250
MM=2                                                            90260
C                                                                 90270
C   ВНЕШНИЙ ЦИКЛ ДЛЯ СЛОВ NROW                                 90280
C                                                                 90290
DO 45 LAYER=1,NROW                                           90300
NN=NN/2                                                         90310
NW=0                                                            90320
DO 40 I=1,MM,2                                               90330
I=NN*I                                                         90340
C                                                                 90350
C   CXCS = CEXP(2*PI*NW*SIGN/NMAX)                             90360
C                                                                 90370
W=FLOAT(NW)*ZZ                                                 90380
CS(1)=COS(W)                                                   90390
CS(2)=SIN(W)                                                  90400
C                                                                 90410
C   ВЫЧИСЛЕНИЕ ЭЛЕМЕНТОВ ДЛЯ ОБЕИХ ПОЛОВИН КАЖДОГО БЛОКА    90420
C                                                                 90430
DO 20 J=1,NN                                                  90440

```

Рис. А.5. Программа быстрого преобразования Фурье. Написана

Не забудьте, что соотношение между k и K имеет следующий вид:

$$k: 0 \quad k \quad \frac{N}{2}$$

$$K: 1 \quad k+1 \quad \frac{N}{2} + 1$$

А.5. ПОЛУЧЕНИЕ ВЕСОВЫХ КОЭФФИЦИЕНТОВ НИЗКОПОЛОСНОГО ФИЛЬТРА

Подпрограмма LPSB (рис. А.6) позволяет получать весовые коэффициенты, необходимые при цифровой низкочастотной филь-

	II=II+1	90450
	IJ=II-NN	90460
	XA=CXCS*X(II)	90470
	X(II)=X(IJ)-XA	90480
20	X(IJ)=X(IJ)+XA	90490
C		90500
C		90510
C		90520
C	ВЫЧИСЛЕНИЕ ОБРАТНЫХ АДРЕСОВ	90530
C		90540
	DO 25 LOC=2,NPOW	90550
	LL=NW-MSK(LOC)	90560
	IF (LL) 30,35,25	90570
25	NW=LL	90580
30	NW=MSK(LOC)+NW	90590
	GO TO 40	90600
35	NW=MSK(LOC+1)	90610
40	CONTINUE	90620
C		90630
C	ПРОВОДИТСЯ ЗАКЛЮЧИТЕЛЬНОЕ ПЕРЕУПОРЯДОЧЕНИЕ	90640
C	А ТАКЖЕ УМНОЖЕНИЕ НА DELTA	90650
C		90660
45	MM=MM*2	90670
	NW=0	90680
	DO 80 I=1,NMAX	90690
	NW1=NW+1	90700
	HOLD=X(NW1)	90710
	IF (NW1-I) 60,55,50	90720
50	X(NW1)=X(I)*DELTA	90730
55	X(I)=HOLD*DELTA	90740
C		90750
C		90760
C	ВЫЧИСЛЕНИЕ ОБРАТНОГО АДРЕСА	90770
C		90780
60	DO 65 LOC=1,NPOW	90790
	LL=NW-MSK(LOC)	90800
	IF (LL) 70,75,65	90810
65	NW=LL	90820
70	NW=MSK(LOC)+NW	90830
	GO TO 80	90840
75	NW=MSK(LOC+1)	90850
80	CONTINUE	90860
	RETURN	90870
	ENO	90880

Л. Д. Эноксоном, Ховардом Натансом и Р. К. Отнесом.

трации. В частности, она дает возможность определять весовые коэффициенты для синусного варианта фильтра Баттеруорта. Предложение обращения к этой подпрограмме имеет вид

CALL LPSB (M, T, BW, A1, A2, BZERO)

Входящие в него символы имеют следующий смысл.

На входе: M — порядок фильтра, который должен быть ограничен промежутком $1 \leq M \leq 20$ (обычно $M=6$); T — интервал, с которым производится выборка, в секундах; BW — частота отсечки фильтра (точка половинной мощности), Гц, которая должна быть ограничена промежутком $0 < BW < 1/(2T)$.

```

SUBROUTINE LPSB (MM,T,BW,A1,A2,BZERO)
C
C   ЭТА ПРОГРАММА ВЫЧИСЛЯЕТ ВЕСА НИЗКОПОЛОСНЫХ СИНУСНЫХ ФИЛЬТРОВ
C   БАТТЕРУОРТА
C   ИЗ ... ПРИКЛАДНОГО АНАЛИЗА ВРЕМЕННЫХ РЯДОВ ...
C
C   РОБЕРТА К. ОТНЕСА И ЛОРЕНА ЭНОКСОНА
C
C   COPYRIGHT 1977, WILEY INTERSCIENCE, NEW YORK
C
C   СЧИТАЕМ, ЧТО ...
C
Y(I)=-A1*Y(I-1)-A2*Y(I-2)+BZERO*X(I)
C
DIMENSION A1(1), A2(1)
DOUBLE PRECISION A,B,C,D,E,F,G,H, FN, FACT, WEDGE, SECTOR, ANG
FACT=3.14159265*T*BW
FACT=DSIN(FACT)
F=1.00
M=MM
M1=M/2
M3=M1
IF (M1.EQ.0) GO TO 10
M2=M
A=M2
SECTOR=3.14159265D0/A
WEDGE=SECTOR/2.00
DO 5 I=1,M1
FN=I-1
B=FACT*DSIN(FN*SECTOR+WEDGE)
C=1.00-FACT*FACT
D=0.500*(-C+DSQRT((C*C+4.00*B**2)))
E=DSQRT(D+1.00)+DSQRT(D)
G=2.00*((2.00*B*B/D)-1.00)/(E**2)
H=-1.00/(E**4)
F=F*(1.00-G-H)
A1(I)=-G
A2(I)=-H
5 CONTINUE
10 IT=M-2*M1
IF (IT.EQ.0) GO TO 15
M3=M1+1
A=FACT*FACT
G=2.00*A+1.00-2.00*FACT*DSQRT(A+1.00)
A1(M3)=-G
A2(M3)=0.
F=F*(1.00-G)
15 A=M3
BZERO=F**2*(1.00/A)
RETURN
END

```

Рис. А.6. Программа получения весовых коэффициентов для низкочастотной фильтрации Баттеруорта.

На выходе: $A1$ — коэффициенты фильтра a_{1p} , $p = 1, \dots, P$, где P определяется соотношениями

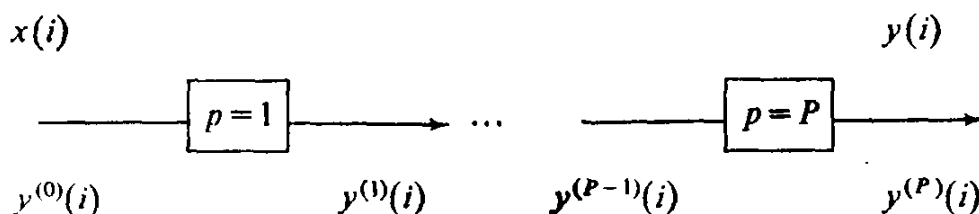
$$P = \frac{M}{2} \quad (M \text{ — четное}), \quad P = \frac{M+1}{2} \quad (M \text{ — нечетное});$$

$A2$ — коэффициенты a_{2p} фильтра, значение P здесь то же, если M — четное, если же M — нечетное, то последний P -й коэффициент равен нулю; $BZERO$ — постоянный коэффициент, одинаковый для всех составляющих фильтра.

Предполагается, что на p -м шаге фильтрации выполняется операция

$$y^{(p)}(i) = b_0 y^{(p-1)}(i) - a_{1p} y^{(p)}(i-1) - a_{2p} y^{(p)}(i-2),$$

которую схематически можно изобразить так:



Отметим, что $y^{(p-1)}(i)$ есть вход на p -м шаге фильтрации, а $y^{(p)}(i)$ — выход. Далее, $y^{(0)}(i) = x(i)$ есть начальное значение при фильтрации, $y^{(P)}(i) = y(i)$ есть последнее значение, полученное на выходе.

А.6. ПЕРЕДАТОЧНАЯ ФУНКЦИЯ ЦИФРОВОГО ФИЛЬТРА

Подпрограмма TTRAN, приведенная на рис. А.7, позволяет определить передаточную функцию цифрового фильтра на выбранном множестве частот при заданных весовых коэффициентах фильтра. Считается, что фильтр реализуется в каскадной форме.

Предложение обращения имеет вид

CALL TTRAN (A1,A2,B0,B1,B2,M,NPO,T,FREQ,ABZ,PHS)

Входящие в него величины имеют следующий смысл.

На входе программы: $A1$, $A2$, $B0$, $B1$ и $B2$ — коэффициенты a_{1p} , a_{2p} , b_{0p} , b_{1p} и b_{2p} соответственно (все они образуют массивы размерности P или больше); M — число рекурсивных членов фильтра; NPO — число точек на выходе, нужное для исследуемой передаточной функции; T — интервал выборки в секундах; $FREQ$ — таблица частот, в которых необходимо вычислить передаточную функцию (все частоты измеряются в Гц, массив $FREQ$ может иметь размерность NPO или больше).

На выходе программы: ABZ — квадрат абсолютной величины передаточной функции в дБ (массив должен иметь размерность

NPO или больше); PHS—фаза в градусах (размерность NPO или больше).

```

SUBROUTINE TTRAN (A1,A2,B0,B1,B2,M,NPO,T,FREQ,ABZ,PHS)
DIMENSION A1(1),A2(1),B0(1),B1(1),B2(1),FREQ(1),ABZ(1),
1PHS(1)
FACT=6.2831853*T
IP=M-M/2
ADD=0.
PREV=0.
OO 10 I=1,NPD
FD=FREQ(I)*FACT
S1=SIN(FD)
C1=COS(FD)
A=2.*FD
S2=SIN(A)
C2=COS(A)
ABSA=1.
PHSA=0.
DO 5 J=1,IP
AR=B0(J)+B1(J)*C1+B2(J)*C2
AI=-B1(J)*S1-B2(J)*S2
ANM=AR**2+AI**2
PND=0.
IF (AI.NE.0..OR.AR.NE.0.) PND=ATAN2(AI,AR)
AR=1.+A1(J)*C1+A2(J)*C2
AI=-A1(J)*S1-A2(J)*S2
ABSA=ABSA*ANM/(AR**2+AI**2)
DUM=0.
IF (AI.NE.0..DR.AR.NE.0.) DUM=ATAN2(AI,AR)
5 PHSA=PHSA+PND-DUM
ABZ(I)=10.*ALOG10(ABSA+1.E-30)
ANG=PHSA*180./3.14159265
CUR=ANG+ADD
TEST=CUR-PREV
IF (ABS(TEST).LT.180.) GO TO 9
IF (TEST.LT.0.) GO TO 7
ADD=ADD-360.
CUR=CUR+360.
GO TO 9
7 ADD=ADD+360.
CUR=CUR-360.
9 PREV=CUR
PHS(I)=CUR
10 CONTINUE
RETURN
END

```

Рис. А.7. Программа вычисления передаточной функции цифрового фильтра в терминах коэффициента усиления в децибелах и фазы в градусах,

Предполагается, что фильтрация проводится за P шагов, где

$$P = \begin{cases} M/2, & \text{если } M \text{ — четное,} \\ (M+1)/2, & \text{если } M \text{ — нечетное.} \end{cases}$$

На p -м шаге фильтрации выполняется операция

$$y^{(p)}(i) = b_{0p}y^{(p-1)}(i) + b_{1p}y^{(p-1)}(i-1) + b_{2p}y^{(p-1)}(i-2) - a_{1p}y^{(p)}(i-1) - a_{2p}y^{(p)}(i-2), \quad p = 1, \dots, P.$$

Так же, как и в предыдущей программе, исходное значение на входе фильтра $x(i)$ обозначается $y^{(0)}(i)$ и $y(i)$. Последнее значение, получаемое на выходе фильтра, также обозначается $y^{(P)}(i)$.

Рассмотренная подпрограмма предназначена для нерекурсивных фильтров более общего вида, чем те, что фигурировали в описании программы LPSB. Чтобы найти передаточную функцию фильтра, коэффициенты которого получены при помощи LPSB, нужно использовать те же значения A1, A2, M и T, которые в ней использовались. Для величин B0, B1 и B2 проводятся следующие присваивания: для всех $I = 1, \dots, P$

$$\text{BZERO} \rightarrow \text{B0}(I) \quad 0. \rightarrow \text{B1}(I) \quad 0. \rightarrow \text{B2}(I)$$

Отметим, что в LPSB величина BZERO — скаляр, а в программе TTRAN величины BZERO образуют массив.

А.7. ТЕСТ-ПРОГРАММА И РЕЗУЛЬТАТЫ ПРОВЕРОК

Для проверки работы подпрограмм можно воспользоваться простой тест-программой, приведенной на рис. А.8. Она состоит из четырех частей.

Номер теста	Строки	Операции
1	4—21	Получение 10 000 псевдослучайных чисел при помощи программы TDRAND и вычисление по ним среднего и дисперсии
2	22—42	Получение симметричного импульса и вычисление значений его преобразования Фурье с последующим нанесением их на график
3	43—58	Вычисление и графическое изображение значений весовых коэффициентов низкочастотного пятиполосного фильтра Баттеруорта
4	59—69	Определение передаточной функции фильтра, веса которого получены на предыдущем шаге

Эта программа и подпрограммы записывались и опробовались на различных ЭВМ. При этом добавлялся язык управления заданиями, который здесь не приводится из-за его особенностей на опробованных ЭВМ. Как правило, этот язык на ЭВМ разных марок сильно различается, поэтому мы сочли нецелесообразным приводить соответствующие операторы. Заметим, что для распечаток используется блок IPRINT, который определен в программе как блок 5. Этот выбор на некоторых машинах следует подправить.

Результаты прогона тест-программы показаны на рис. А.9—А.11.

Из-за разного объема слов результаты на разных ЭВМ могут оказаться различными. Тем не менее они должны быть примерно такими же, как приведенные ниже.

```

DIMENSION A1(3),A2(3),B0(3),B1(3),B2(3),FREQ(101),X(101),Y(101)
DOUBLE PRECISION D,SUM,SUMS
DATA IPRINT /6/
WRITE (IPRINT,1000)
1000 FORMAT (13H1TEST CASE 1.,/,
*54HOCHECK OF STATISTICS OF THE FIRST 10,000 POINTS OUT OF,/,
*29H THE RANDOM NUMBER GENERATOR.)
SUM=0.00
SUMS=0.00
DO 10 I=1,10000
CALL TDRAND (A)
D=A-0.5
SUM=SUM+D
10 SUMS=SUMS+D*D
SUM=SUM/10000.00
SUMS=(SUMS-10000.00*SUM*SUM)/9999.00
A=1.00/SUMS
WRITE (IPRINT,1001) SUM,SUMS,A
1001 FORMAT (25HMEAN =,F12.6,/,
*25HOVARIANCE =,F12.6,/,
*25HORECIPROCAL OF VARIANCE =,F12.6)
DO 20 I=1,64
20 X(I)=0.
X(I)=1.
X(3)=1.
X(5)=1.
X(61)=1.
X(63)=1.
CALL FFTRAN (-1.,1.,X,5)
WRITE (IPRINT,1002)
1002 FORMAT (13H1TEST CASE 2.,/,
*47HOREAL PART OF THE FOURIER TRANSFORM OF A PULSE.)
CALL PRPLOT (X,17,0.,1.,2)
WRITE (IPRINT,1003)
1003 FORMAT (130H1IMAGINARY PART OF THE FOURIER,/,
*24H TRANSFORM OF THE PULSE.)
CALL PRPLOT (X(2),17,0.,1.,2)
CALL FFTRAN (1.,1.,X,5)
WRITE (IPRINT,1004)
1004 FORMAT (29H1INVERSE FOURIER TRANSFORM OF,/,
*23H THE TRANSFORMED PULSE.)
CALL PRPLOT (X,32,0.,1.,2)
WRITE (IPRINT,1005)
1005 FDMAT (13H1TEST CASE 3.,/,
*26HOFIVE POLE LOWPASS FILTER.,/,
*18HOFILTER WEIGHTS...,/,
*37HO I A1 A2,
*15H B0,/)
T=0.005
BW=5.
M=5
CALL LPSB (M,T,BW,A1,A2,BZERO)
DO 30 I=1,3
B0(I)=BZERO
B1(I)=0.
B2(I)=0.
30 WRITE (IPRINT,1006) I,A1(I),A2(I)
1006 FORMAT (1H ,I6,3F15.8)
DO 40 I=1,101
FREQ(I)=I-1
CALL TTRAN (A1,A2,B0,B1,B2,M,101,T,FREQ,X,Y)
WRITE (IPRINT,1007)
1007 FORMAT (30H1GAIN OF THE TRANSFER FUNCTION,/,
*21H OF THE FILTER IN DB.)
CALL PRPLOT (X,101,0.,1.,1)
WRITE (IPRINT,1008)
1008 FORMAT (131H1PHASE OF THE TRANSFER FUNCTION,/,
*26H OF THE FILTER IN DEGREES.)
CALL PRPLOT (Y,101,0.,1.,1)
STOP
END

```

Рис. А.8. Программа с четырьмя простыми тестовыми случаями для программ, часть 1. Внимание: члены в квадратных скобках форматов 1001 и 1005 перфорировать не надо. Они означают числа пробелов в длинных пустых промежутках.

TEST CASE 1.

CHECK OF STATISTICS OF THE FIRST 10,000 POINTS OUT OF THE RANDOM NUMBER GENERATOR.

MEAN = 0.000974
 VARIANCE = 0.083891
 RECIPROCAL OF VARIANCE = 11.920185

Рис. А.9. Результаты тестового случая для генератора псевдослучайных чисел.

TEST CASE 2.

REAL PART OF THE FOURIER TRANSFORM OF A PULSE.

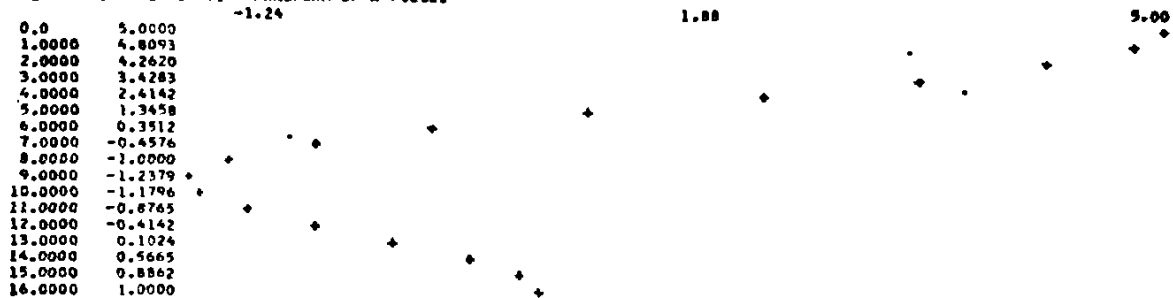


Рис. А.10а. Действительная часть преобразования импульса.

IMAGINARY PART OF THE FOURIER TRANSFORM OF THE PULSE.



Рис. А.10б. Эта мнимая часть должна быть примерно нулем, как это и получилось. Приведенный график на разных машинах имеет неодинаковый вид.

```

INVERSE FOURIER TRANSFORM OF
THE TRANSFORMED PULSE.
      0      -0.00      0.50      1.00
0.0      1.0000
1.0000      1.0000
2.0000      1.0000
3.0000      0.0000 *
4.0000      0.0000 *
5.0000      0.0000 *
6.0000      0.0000 *
7.0000      0.0000 *
8.0000      0.0 *
9.0000      0.0000 *
10.0000      0.0000 *
11.0000      0.0000 *
12.0000      -0.0000 *
13.0000      0.0000 *
14.0000      0.0000 *
15.0000      0.0000 *
16.0000      0.0000 *
17.0000      0.0000 *
18.0000      0.0000 *
19.0000      -0.0000 *
20.0000      -0.0000 *
21.0000      -0.0000 *
22.0000      0.0000 *
23.0000      0.0000 *
24.0000      0.0 *
25.0000      0.0000 *
26.0000      0.0000 *
27.0000      0.0000 *
28.0000      0.0000 *
29.0000      0.0000 *
30.0000      1.0000
31.0000      1.0000

```

Рис. А.10с. Вариант PRPLOT с пропусками, дающий только действительные части.

TEST CASE 3.

FIVE POLE LOWPASS FILTER.

FILTER WEIGHTS...

I	A1	A2	BD
1	-1.88388157	0.90733629	0.04194543
2	-1.75400639	0.77569389	0.04194543
3	-0.85491133	0.0	0.04194543

Рис. А.11а. Здесь параметры имеют следующие значения: $T=0.005$, $B=5$ Гц. Поэтому частоты получаются сразу в процентах частоты Найквиста.

GAIN OF THE TRANSFER FUNCTION
OF THE FILTER IN DB.

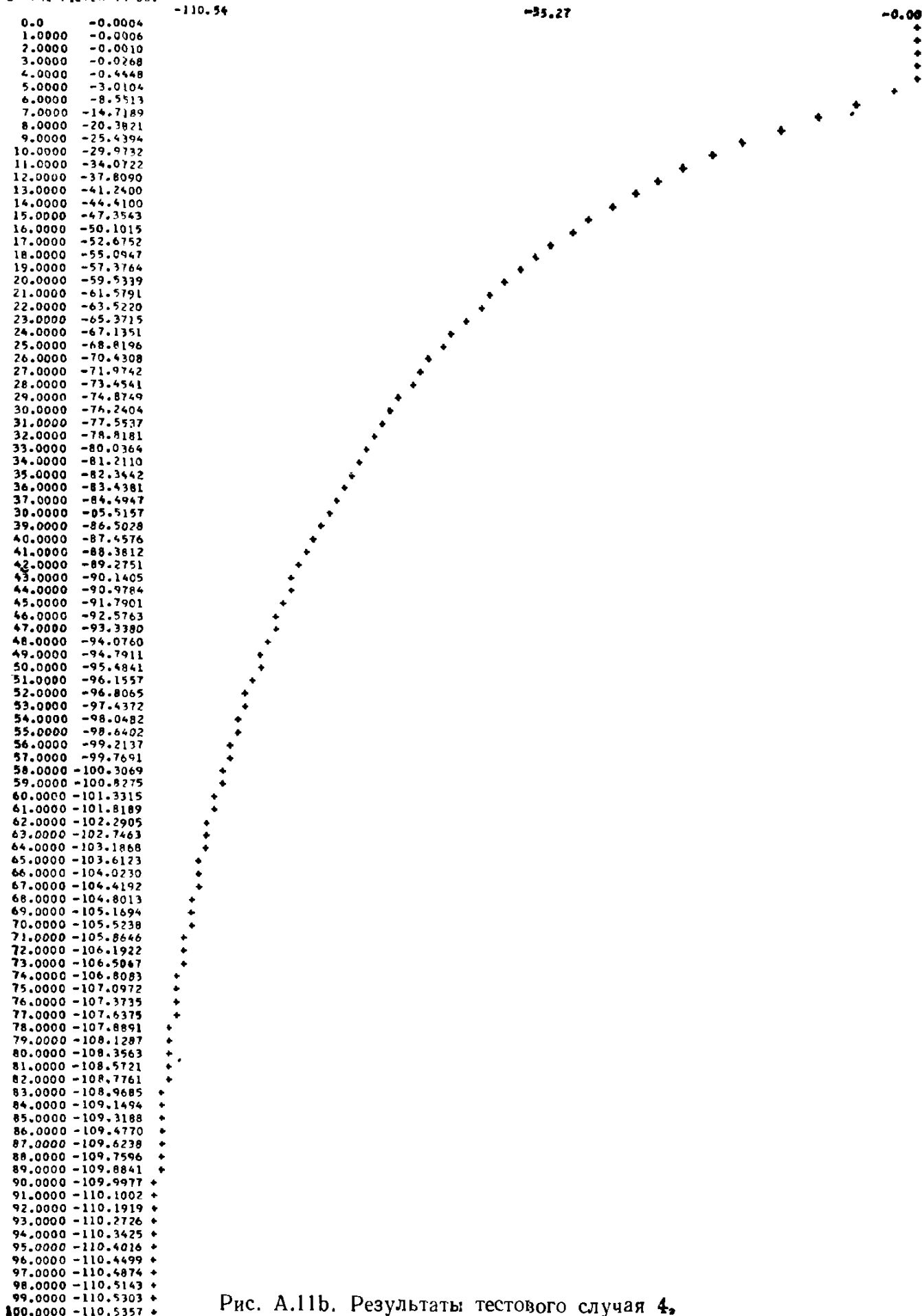


Рис. А.11в. Результаты тестового случая 4.

На рис. А.9 приведен результат прогона 1-го теста программы. Значения параметров, которые должны быть получены, равны 0., $1/12$ (0.83333...) и 12 соответственно. Как видно, результаты прогона хорошо согласуются с этим; отличия от ожидаемых величин легко объяснить статистической изменчивостью.

На рис. А.10а, А.10б, А.10с приведены результаты прогона теста 2. Цель этого теста — вычисление преобразования Фурье прямоугольного импульса. Общее число комплексных точек, использованных при этом, равнялось 32, причем

$$x(i) = \begin{cases} 1, & i = -2, -1, 0, 1, 2, \\ 0 & \text{в противном случае.} \end{cases}$$

В запоминающем устройстве происходит следующая расстановка данных:

$$\begin{array}{cccccc} i & -2 & -1 & 0 & 1 & 2 \\ I & 61 & 63 & 1 & 3 & 5 \end{array}$$

Величина T выбирается равной 1. Поэтому, пользуясь симметрией и формулами из приложения В, получаем

$$X(k) = \sum_{i=-2}^2 e^{-j \frac{2\pi ik}{32}} = \sum_{i=-2}^2 \cos\left(\frac{\pi ik}{16}\right) = \sin\left(\frac{5\pi k}{32}\right) / \sin\left(\frac{\pi k}{32}\right).$$

Как и в тесте, возьмем $k=8$:

$$X(8) = \sin\left(\frac{5\pi}{4}\right) / \sin\left(\frac{\pi}{4}\right) = -1.$$

На рис. А.10а приведены значения $\text{Re}[X(k)]$; величина $X(8)$ равна -1 . Можно проверить и другие значения.

В силу того что для преобразования выбиралась симметричная функция, значения $\text{Im}[X(f)]$ должны быть равными нулю. Однако из-за ошибок округления, возникающих при выполнении операций в арифметике с плавающей точкой, полученные значения в точности нулю не равны. Это видно из рис. А.10б, на котором приводится график величин $\text{Im}[X(f)]$. Часть значений на нем равна нулю, а остальные имеют ничтожно малые по сравнению с 5, максимальным значением $\text{Re}[X(f)]$, величины. Итак, на рис. А.10б изображен график ошибки округления. На других ЭВМ он должен иметь другой вид, поскольку совпадение таких графиков — очень маловероятное событие.

Результаты вычисления обратного преобразования Фурье и график полученных значений показаны на рис. А.10с. Поскольку выбран вариант PRPLOT с пропусками, изображаются только действительные части временной последовательности.

На рис. А.11а приведены результаты теста 3, в котором требуется получить весовые коэффициенты пятиполюсного низко-

полосного фильтра с точкой отсечки 5% частоты Найквиста (в данном случае 5 Гц). Заметим, что коэффициент a_{23} равен нулю, что и следовало ожидать для пятиполюсного фильтра.

На рис. А11в приведен график результатов для передаточной функции, полученных применением подпрограммы TTRAN к фильтру, определенному выше. Здесь изображен только коэффициент усиления этой передаточной функции в дБ. Отметим, что в точке 5 Гц передаточная функция равна -3.0103 дБ, как и следовало ожидать.

Приложение В

ПРОЦЕДУРА БЛЭКМЕНА — ТЬЮКИ ВЫЧИСЛЕНИЯ ПЛОТНОСТЕЙ СПЕКТРА МОЩНОСТИ

В.1. ВВЕДЕНИЕ

На возможность вычисления плотностей спектра мощности через ковариационные функции впервые указали Винер и Хинчин. В виде практической цифровой процедуры эта возможность была реализована позднее Блэкменом и Тьюки (1958). Их книга внесла весомый вклад в развитие методов анализа временных рядов.

Хотя теперь этот метод в значительной мере вытеснен методами БПФ, представляется разумным рассмотреть соответствующий алгоритм, чтобы понять связанные с ним проблемы.

Сначала определим автоковариационную функцию для функции $x(t)$ формулой

$$s_{xx}(\tau) = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P x(t) x(t + \tau) dt.$$

Предполагается, что функция $x(t)$ имеет нулевое среднее.

В этом случае плотность спектра мощности $S_{xx}(f)$ имеет вид

$$\begin{aligned} S_{xx}(f) &= \lim_{P \rightarrow \infty} \int_{-P}^P s_{xx}(\tau) \exp(-j2\pi f\tau) d\tau = \\ &= \lim_{P \rightarrow \infty} \int_{-P}^P s_{xx}(\tau) \cos(2\pi f\tau) d\tau = \\ &= \lim_{P \rightarrow \infty} 2 \int_0^P s_{xx}(\tau) \cos(2\pi f\tau) d\tau. \end{aligned}$$

Два последних перехода в цепочке равенств возможны в силу симметричности функции $s(\tau)$.

В реальных ситуациях пределы интегрирования конечны, так как данные не могут собираться бесконечно. Таким образом, справедливы формулы

$$\hat{s}_{xx}(\tau) = \frac{1}{2P'} \int_{-P'}^{P'} x(t) x(t + \tau) dt, \quad -P' \leq \tau \leq P',$$

$$\hat{S}_{xx}(f) = \int_{-P}^P s_x(\tau) \cos(2\pi f\tau) d\tau,$$

где крышечка указывает на ограниченность времени выборки; длина $2P'$ исходного ряда данных должна быть больше, чем длина $2P$ автоковариации, используемая в вычислениях.

Если величина P' достаточна для того, чтобы функция $\hat{s}_{xx}(\tau)$ служила хорошей оценкой величины $s_{xx}(\tau)$, то выборочный спектр можно записать в следующем виде:

$$\hat{S}_{xx}(f) = \int_{-\infty}^{\infty} v(\tau) s_{xx}(\tau) \exp(-j2\pi f\tau) d\tau,$$

где

$$v(\tau) = \begin{cases} 1, & -P \leq \tau < P, \\ 0 & \text{в противном случае.} \end{cases}$$

Последняя функция $v(\tau)$ имеет такой же вид, что и прямоугольный импульс $u(t)$, введенный в гл. 1. Имеется, правда, одно существенное различие: ширина носителя функции $v(\tau)$ вдвое больше, чем у функции $u(t)$. Поэтому преобразование Фурье $V(f)$ функции $v(\tau)$ имеет вид

$$V(f) = \frac{2 \sin(2\pi fP)}{2\pi f}.$$

Следовательно, при том же значении P функция $V(f)$ будет вдвое выше, чем функция $U(f)$, а ширина главного всплеска $V(f)$ — вдвое уже.

Наблюдаемые значения ПСМ представляют собой значения свертки функций $S_x(f)$ и $V(f)$:

$$\hat{S}_{xx}(f) = \int_{-\infty}^{\infty} S_{xx}(\eta) V(f - \eta) d\eta.$$

Заметим, что в этом выражении функция $V(f)$ не возведена в квадрат. Поскольку эта функция может принимать и отрицательные значения при некоторых значениях f , в ряде случаев могут получиться отрицательные мощности.

В качестве примера рассмотрим синусоиду

$$x(t) = A \sin(2\pi f_0 t + \varphi).$$

Как было установлено ранее, автоковариационная функция в этом случае имеет вид

$$s_{xx}(\tau) = \frac{A^2}{2} \cos(2\pi f_0 \tau),$$

а спектр мощности определяется формулой

$$S_{xx}(f) = \frac{A^2}{4} [\delta(f - f_0) + \delta(f + f_0)].$$

Чтобы смоделировать вычисление величин $S_{xx}(f)$, предположим, что вместо значений $\hat{s}_{xx}(\tau)$ используются значения $s_{xx}(\tau)$, причем функция $\hat{S}_{xx}(f)$ вычисляется на конечном интервале $(-P, P)$:

$$\begin{aligned}\hat{S}_{xx}(f) &= \int_{-\infty}^{\infty} S_{xx}(\eta) V(f-\eta) d\eta = \\ &= \int_{-\infty}^{\infty} \frac{A^2}{4} [\delta(\eta-f_0) + \delta(\eta+f_0)] \frac{2 \sin(2\pi(f-\eta)P)}{2\pi(f-\eta)P} d\eta = \\ &= \frac{A^2}{4} \left\{ \frac{2 \sin[2\pi(f-f_0)P]}{2\pi(f-f_0)P} + \frac{2 \sin[2\pi(f+f_0)P]}{2\pi(f+f_0)P} \right\}.\end{aligned}$$

Исходная ПСМ $S_{xx}(f)$ состояла из двух дельта-функций, расположенных в частотах $\pm f_0$ Гц. Наблюдаемая ПСМ представляет собой сумму двух функций вида $(\sin x)/x$, максимальные значения которых располагаются в частотах $\pm f_0$ Гц. В пределе при $P \rightarrow \infty$ эти функции дают две дельта-функции.

При конечных значениях P наблюдается заметное размывание или *размывание* ПСМ. Наблюдаемый уход мощности с того места, где ей полагалось быть (в данном случае в частотах $\pm f_0$ Гц), называется *утечкой*.

Для уменьшения этой утечки можно подобрать окно $v(\tau)$ или применить его частотный аналог. Существует целый ряд окон, которые вместе с уменьшением утечки решают и ряд других проблем, связанных с ней.

В.2. ФОРМУЛЫ ДЛЯ ВЫЧИСЛЕНИЙ НА ЭВМ

В этом параграфе собраны все формулы, используемые при вычислении ПСМ.

1. Допустим, что последовательность $x(i)$ состоит из N величин, имеющих нулевое среднее. Если среднее не равно нулю, то сначала его вычисляют и вычитают из всех значений данных. Этот шаг может включать в себя удаление субгармонических членов.

2 Для $(m+1)$ значений индекса i вычисляются значения автоковариационной функции

$$s_{xx}(i) = \frac{1}{N-i-1} \sum_{p=0}^{N-i-1} x(p) x(p+i), \quad i=0, \dots, m.$$

3. Выбирается окно. Возможными кандидатами здесь служат:

а) окно Хеннинга

$$v_m^{(1)}(i) = \frac{1}{2} \left(1 + \cos \frac{\pi i}{m} \right),$$

b) окно Хемминга

$$v_m^{(2)}(i) = 0.54 + 0.46 \cos \frac{\pi i}{m},$$

с) окно Парзена

$$v_m^{(3)}(i) = \begin{cases} \left(1 - 6 \frac{i}{m}\right)^2 \left(1 - \frac{i}{m}\right) & \text{при } i < \frac{m}{2}, \\ 2 \left(1 - \frac{i}{m}\right)^3 & \text{при } i > \frac{m}{2}. \end{cases}$$

После применения одного из них к корреляции получают новую корреляцию $\tilde{s}_x(i)$:

$$\tilde{s}_x(i) = s_x(i) v_m^{(l)}(i).$$

4. Значения ПСМ для различных частот вычисляются посредством интегрирования методом трапеции:

$$\hat{S}_{xx}(f) = T \left(\tilde{s}_x(0) + 2 \sum_{i=1}^{m-1} \hat{s}_x(i) \cos(2\pi f iT) + \tilde{s}_{xx}(i) \cos(2\pi f m T) \right).$$

В качестве «стандартного» набора частот можно рассматривать

$$f_k = \frac{k}{2mT} \text{ для } k = 0, 1, \dots, m.$$

Для него получаются $(m+1)$ равноотстоящих перекрывающихся оценок ПСМ. Предыдущую формулу можно переписать в виде ($k = 0, 1, \dots, m$)

$$\hat{S}_{xy}(k) = T \left(\tilde{S}_x(0) + 2 \sum_{i=1}^{m-1} \tilde{S}_x(i) \cos \frac{\pi i k}{m} + \tilde{S}_x(m) \cos(\pi i) \right).$$

До шага 2 можно провести любое приведение к белому шуму, а после шага 4 — соответствующее раскрашивание.

Если ПСМ вычисляется во всех частотах $k/(2mT)$, $k = 0, 1, \dots, m$, то порядок вычислений можно изменить. Шаг 3 приведенной последовательности действий можно выбросить, а в конце добавить один новый шаг, на котором для ПСМ вычисляются взвешенные суммы вида

$$S_x(k) = \begin{cases} D_0 \tilde{G}_{x(k-1)} + D_1 \tilde{G}_{xk} + D_0 \tilde{G}_{x(k+1)} & \text{при } k \neq 0, m, \\ D_1 \tilde{G}_{x0} + 2D_0 \tilde{G}_{x1} & \text{при } k = 0, \\ D_1 \tilde{G}_{xm} + 2D_0 \tilde{G}_{x(m-1)} & \text{при } k = m. \end{cases}$$

Здесь коэффициенты D_0 и D_1 определяются формулами

$$D_0 = \begin{cases} 1/2 & \text{(Хемминг)}, \\ 0.54 & \text{(Хемминг)}, \end{cases} \quad D_1 = \begin{cases} 1/4 & \text{(Хемминг)}, \\ 0.23 & \text{(Хемминг)}. \end{cases}$$

Как выяснилось, никакой реальной выгоды такое использование окна запаздывания не приносит. Время, необходимое для вычислений автокорреляции и преобразования Фурье, оказывается по сравнению с временем применения окна смещения значительно большим. Этот вариант включен сюда только потому, что читатель, вероятно, уже сталкивался с ним в обзорах методов вычислений, составляемых некоторыми организациями.

На шаге 4, который, по существу, представляет собой вычисление косинус-преобразования Фурье, можно воспользоваться приемами вычисления БПФ, рассмотренными в гл. 6. Однако при малом числе запаздываний обращение к сложному методу БПФ может и не принести выгоды. Кроме того, при использовании БПФ может оказаться, что нужный шаг частот недостижим. Если метод БПФ использовать нельзя, ему на смену выбирают один из нескольких рекурсивных методов \sin/\cos -вычислений, снижающих стоимость вычислений.

Здесь следует воспользоваться рекурсивным методом, описанным в § 4.3. Он основан на соотношении $c_i = hc_{i-1} - c_{i-2}$, где $h = 2 \cos(2\pi T f_0)$. С его помощью можно получить и синус, и косинус в зависимости от того, что служит начальными значениями. В частности, при $c_0 = 1$ и $c_{-1} = \cos(2\pi T f_0)$ вырабатывается косинус, при $c_0 = 0$ и $c_{-1} = -\sin(2\pi T f_0)$ вырабатывается синус.

В.3. ПРОЦЕДУРА ПОЛУЧЕНИЯ КРОСС-СПЕКТРА

Последнее, на чем мы остановимся в этом приложении, — это вычисление плотности кросс-спектра (ПКС). Что касается подхода, основанного на применении окна запаздывания, то для него остаются в силе все соображения, высказанные для ПСМ, поэтому будут рассмотрены только определения и схема процедуры вычисления. В непрерывном случае кросс-ковариационная функция определяется формулой

$$s_{xy}(\tau) = E[x(t)y(t+\tau)].$$

Временное осреднение можно записать в виде

$$s_{xy}(\tau) = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P x(t)y(t+\tau) dt.$$

Плотность кросс-спектра есть преобразование Фурье последнего выражения:

$$S_{xy}(f) = \int_{-\infty}^{\infty} s_{xy}(\tau) \exp(-j2\pi f\tau) d\tau.$$

Значения ПКС можно записать в виде суммы действительной и мнимой частей: $S_{xy}(f) = C_{xy}(f) + jQ_{xy}(f)$. Здесь C_{xy} — функция, которая называется *коспектральной плотностью* или *коспектром*, а Q_{xy} называется *плотностью квадратурного спектра* или *квадратурным спектром*. Эти функции определяются формулами

$$C_{xy}(f) = \int_0^{\infty} [s_{xy}(\tau) + s_{xy}(-\tau)] \cos(2\pi f\tau) d\tau,$$

$$Q_{xy}(f) = \int_0^{\infty} [s_{xy}(\tau) - s_{xy}(-\tau)] \sin(2\pi f\tau) d\tau.$$

Процедура вычислений выглядит в основном так же, как и для ПСМ.

1. Предполагается, что последовательности $x(i)$ и $y(i)$ имеют нулевые средние и состоят из N точек каждая. Если средние не равны нулю, то вычисляются средние этих последовательностей, которые затем вычитаются из соответствующих рядов данных.

2. Для $(m+1)$ значений индекса i вычисляются значения выборочных ковариационных функций ($i = 0, 1, \dots, m$)

$$\hat{s}_{xy}(i) = \frac{1}{N-i-1} \sum_{p=0}^{N-i-1} x(p) y(p+i),$$

$$\hat{s}_{yx}(i) = \frac{1}{N-i-1} \sum_{p=0}^{N-i-1} x(p+i) y(p).$$

3. Как и выше, выбирается подходящее окно запаздываний и вычисляются новые ковариационные функции ($i = 0, 1, \dots, m$)

$$\tilde{s}_{xy}(i) = v_m^{(l)}(i) \hat{s}_{xy}(i), \quad \tilde{s}_{yx}(i) = v_m^{(l)}(i) \hat{s}_{yx}(i).$$

4. По ним вычисляются промежуточные значения $A_{xy}(i)$ и $B_{xy}(i)$ ($i = 0, 1, \dots, m$):

$$A_{xy}(i) = \tilde{s}_{xy}(i) + \tilde{s}_{yx}(i), \quad B_{xy}(i) = \tilde{s}_{xy}(i) - \tilde{s}_{yx}(i).$$

5. Интегрированием по методу трапеции вычисляются для различных частот коспектр и квадратурный спектр:

$$\hat{C}_{xy}(f) = \frac{T}{2} \left[A_{xy}(0) + 2 \sum_{i=1}^{m-1} A_{xy}(i) \cos(2\pi fiT) + A_{xy}(m) \cos(2\pi fmT) \right],$$

$$\hat{Q}_{xy}(f) = \frac{T}{2} \left[B_{xy}(0) + 2 \sum_{i=1}^{m-1} B_{xy}(i) \sin(2\pi fiT) + B_{xy}(m) \sin(2\pi fmT) \right].$$

Как и в случае ПСМ, может использоваться стандартный набор значений частот

$$f_k = \frac{k}{2mT}, \text{ где } k = 0, 1, \dots, m$$

6. По полученным значениям $\hat{C}_{xy}(f)$ и $\hat{Q}_{xy}(f)$ можно осуществить вывод целого ряда различных характеристик. Обычно вычисляются следующие дополнительные параметры.

а) Абсолютное значение ПКС

$$|\hat{S}_{xy}(f)| = \sqrt{\hat{C}_{xy}^2(f) + \hat{Q}_{xy}^2(f)}.$$

б) Фазовый угол ПКС (в градусах)

$$\hat{\phi}(f) = \frac{360}{2\pi} \operatorname{arctg} \left[\frac{\hat{Q}_{xy}(f)}{\hat{C}_{xy}(f)} \right].$$

Заметим, что нужный квадрант всегда известен, поэтому диапазон значений $\hat{\phi}$ составляет 360° . Чаще всего используется диапазон от -180° до $+180^\circ$. Чтобы правильно определять нужный квадрант, используется проверка по знакам \hat{Q}_{xy} и \hat{C}_{xy} . Во многих программах, предусматривающих вычисление арктангенсов, этому вопросу специально уделяется внимание.

в) Передаточная функция между x и y

$$\hat{H}(f) = \frac{\hat{C}_{xy}(f) + j\hat{Q}_{xy}(f)}{\hat{G}_{xx}(f)}.$$

Она обычно вычисляется в терминах модуля и фазового угла. Модуль определяется формулой

$$|\hat{H}(f)| = \sqrt{\frac{\hat{C}_{xy}^2(f) + \hat{Q}_{xy}^2(f)}{\hat{G}_{xx}^2(f)}}.$$

Фазовый угол вычисляется по той же формуле, что и для ПКС.

В.4. ЗАМЕЧАНИЯ

При вычислении ПСМ и последующем их изучении необходимо иметь в виду следующее.

1. Вместо $S_x(k)$ обычно получают величину $G_x(k)$, которая носит название «односторонняя ПСМ» и определяется формулой

$$G_x(k) = \begin{cases} 2S_x(k) & \text{при } k \geq 0, \\ 0 & \text{при } k < 0. \end{cases}$$

2. Спектральные оценки, разделенные интервалом $f_k = k/(2mT)$, накладываются одна на другую. Поэтому необходимо учесть следующее:

а) если для определения дисперсии интегрируется ПСМ, то результат должен быть поделен на 2, иначе результат окажется вдвое большим, чем он должен быть,

б) кругового эффекта в отличие от метода БПФ, применяемого без добавления нулей, не происходит; обратное преобразование, примененное к $S_x(k)$, даст точно $s_x(t)$ (если не учитывать ошибок округления)

3 Число степеней свободы для каждой спектральной оценки в случае белого шума равно $d = 2N/m$. Через bP эту величину можно записать так: $d = 2bP$ (здесь $b = 1/(mT)$ и $P = NT$)

4. Рассмотренный метод получения ПСМ «сохраняет энергию», в то время как метод преобразования Фурье таким свойством не обладает. Полную энергию несет член $s_x(0)$, который есть не что иное, как дисперсия. Все сглаживающие окна равны в точке $\tau = 0$ единице, поэтому применение окон не меняет энергии, содержащейся в ПСМ. Это не всегда имеет место при вычислении плотностей спектра мощности посредством процедуры быстрого преобразования Фурье, в которых также применяется сглаживание.

ЛИТЕРАТУРА ¹⁾

- Абрамовиц, Стиган (Abramowitz M., Stegun I. A.)
Справочник по специальным функциям. Пер. с англ. — М.: Наука, 1979.
- Акайк (Akaike H.)
Undamped Oscillation of the Sample Autocovariance Function and the Effect of Prewhitening Operation, *Annals of the Institute of Statistical Mathematics*, Vol. 13, pp. 127—144, 1962.
On the Statistical Estimation of the Frequency Response Function of a System Having Multiple Input, *Annals of the Institute Statistical Mathematics*, Vol. 17, No. 2, 1965.
- * Андерсон (Anderson T.)
Введение в многомерный статистический анализ. Пер. с англ. — М.: Физматгиз. 1963.
- * Ахо, Хопкрофт, Ульман (Aho A. V., Hopcroft J. E., Ullman J. D.)
Построение и анализ вычислительных алгоритмов. Пер. с англ. — М.: Мир, 1979.
- Балакршнан (Balakrishnan A. V.)
On the Problem of Time Jitter in Sampling, *IRE Transactions on Information Theory*, Vol. IT-8, No. 3, pp. 226—236, April 1962.
- Баттеруорт (Butterworth S.)
On the Theory of Filter Amplifiers, *Experimental Wireless*, Vol. 7, pp. 536—541, Oct. 1930.
- Бендат, Пирсол (Bendat J. S., Piersol A. G.)
Измерение и анализ случайных процессов. Пер. с англ. — М.: Мир, 1971.
- Бенинус (Benignus V. A.)
Estimation of the Coherence Spectrum and Its Confidence Interval Using the Fast Fourier Transform, *IEEE Transactions on Audio and Electroacoustics*, Vol. AU-17, No. 2, 1969.
- Бингхэм Годфри, Тьюки (Bingham C., Godfrey M. D., Tukey J. W.)
Modern Techniques of Power Spectrum Estimation, *IEEE Transaction on Audio and Electroacoustics*, Vol. AU-15, No. 2, pp. 56—66, 1967.
- Блэкмен, Тьюки (Blackman R. B., Tukey J. W.)
The Measurement of Power Spectra from the Point of View of Communications Engineering, Dover, New York, 1958.

¹⁾ Знаком * отмечены работы, добавленные при переводе. — *Прим. перев.*

- * Большев Л. Н., Смирнов Н. В.
Таблицы математической статистики. — М.: Наука, 1965.
- * Бриллинджер (Brillinger D. R.)
Временные ряды. Обработка данных и теория. Пер. с англ. — М.: Мир, 1980.
- Брэйсуэлл (Bracewell R. M.)
The Fourier Transform and Its Applications, McGraw-Hill, New York, 1965.
- Винер (Wiener N.)
Extrapolation, Interpolation, and Smoothing of Stationary Time Series, MIT Press, Cambridge, Mass., 1949.
* Интеграл Фурье и некоторые его приложения. Пер. с англ. — М.: Физматгиз, 1963.
- Габор (Gabor D.)
Theory of Communication, Journal IEEF (London), Vol. 93, Part 3, No. 26, 1946.
- Гийеман (Guilleman E. A.)
Synthesis of Passive Networks, Wiley, New York, 1957.
- Грэй, Мэркел (Gray A. H., Jr., Markel J. D.)
Digital Lattice and Ladder Filter Synthesis, IEEE Transaction on Audio and Electroacoustics, Dec. 1973.
A Spectral-Flatness Measure for Studying the Autocorrelation Method of Linear Speech Analysis, IEEE Transactions on Speech Acoustics and Signal Processing, pp. 207—271, 1974.
- Гудмэн (Goodman N. R.)
Measurement of Matrix Frequency Response Functions and Multiple Coherence Functions, AFFDL TR-65-56, Research and Technology Division, AFSC, Wright Patterson Air Force Base, Ohio, Feb. 1965.
- Джексон (Jackson L. B.)
Roundoff Noise Analysis for Fixed Point Digital Filters Realized in Cascade or Parallel Form, IEEE Transactions on Audio and Electroacoustics, June 1970.
- Дженкинс, Ваттс (Jenkins G. M., Watts D. G.)
Спектральный анализ и его приложения, вып. 1, 2. Пер. с англ. — М.: Мир, 1971—1972.
- Джентльмен, Сэнд (Gentlemen W. M., Sande G.)
Fast Fourier Transforms for Fun and Profit, AFIPS Conference Proceedings, Vol. 29, 563—578, 1966.
- Джури (Jury E. I.)
A Stability Test for Linear Discrete Systems Using a Simple Division, Proceedings of the IRE, Dec. 1961.
- Джури, Блэнчард (Jury E. I., Blanchard J.)
A Stability Test for Linear Discrete Systems in Table Form, Proceedings of the IRE, Dec. 1961.
- Диксон, Мэсси (Dixon W. G., Massey F. J.)
Introduction to Statistical Analysis, 3rd ed., McGraw-Hill, New York, 1969.
- Дим, Маккин (Dym H., McKean H. P.)
Fourier Series and Integrals, Academic Press, New York, 1972.
- Жюльен, Сид-Ахмед (Jullien G. A., Sid-Ahmed M. A.)
A Computer Program for Filter Design Having Arbitrary Magnitude Speci-

- fications in the Frequency Domain, International Journal of Numerical Methods (Great Britain), Vol. 6, No. 2, pp. 275—285, 1973.
- Кайзер (Kaiser J. F.)
Design Methods for Sampled Data Filters (Z-transform), Proceedings of the First Annual Allerton Conference on Circuit and System Theory, Nov. 1963.
- Картер (Carter G. C.)
Estimation of the Magnitude-Squared Coherence Function, Naval Undersea Systems Center Report 4343, May 19, 1972.
- Келли, Эноксон, Рондинелли (Kelly R. D., Enochson L. D., Rondinelli L. A.)
Techniques and Errors in Measuring Cross-Correlation and Cross-Spectral Density Functions, NASA CR-74505, Feb. 1966.
- Кендалл, Стьюарт (Kendall M. G., Stuart A. G.)
Статистические выводы и связи. Пер. с англ. — М.: Наука, 1973.
- Кеннеди, Сэффорд (Kennedy J. E., Safford F. B.)
The Use of Vibration/Impedance Measurements to Predict Blast Induced Structural Vibrations. Fourth International Symposium on Military Applications of Blast Simulation, Atomic Weapon Research Establishment, Foulness, England, Sept. 1974.
- Кливленд, Парзен (Cleveland W. S., Parzen E.)
Estimation of Coherence, Frequency Response, and Envelope Delay, Technometrics, May 1975.
- * Колмогоров А. Н., Фомин С. В.
Элементы теории функций и функционального анализа. 4-е изд. перераб. — М.: Наука, 1976.
- Крамер (Cramer H.)
Математические методы статистики. Пер. с англ. — М.: Мир, 1975.
- Кули, Тьюки (Cooley J. W., Tukey J. W.)
An Algorithm for the Machine Calculation of Complex Fourier Series, Mathematics of Computation, Vol. 19, p. 297, 1965.
- Купманс (Koopmans L. H.)
The Spectral Analysis of Time Series, Academic Press, New York, 1974.
- Ланцош (Lanczos C.)
Практические методы прикладного анализа. Пер. с англ. — М.: Физматгиз, 1961.
- Лонгли (Longley J. W.)
An Appraisal of Least Squares Computer Programs for the Electronic Computer from the Point of View of the User, JASA, Vol. 62, pp. 819—841, 1967.
- Лью (под ред.) (Liu B., Ed.)
Digital Filters and the Fast Fourier Transform, Dowden, Hutchinson, and Ross, Stroudsburg, PA., 1975.
- Мак-Клеллан, Паркс, Рабинер (McClellan J. H., Parks T. W., Rabiner L. R.)
A Computer Program for Designing Optimum FIR Linear Filters, IEEE Transactions on Audio and Electroacoustics, Dec. 1973.
- Мак-Кован (McCowan D. W.)
Finite Fourier Transform Theory and Its Application to the Computation of Convolutions, Correlations, and Spectra, Research Department Technical Memorandum No. 8-66, Earth Sciences Division, Teledyne, Inc., Dec. 1966.

- Масри, Сэффорд (Masri S. F., Safford F. B.)
Dynamic Environment Simulation by Pulse Techniques, Journal of the Engineering Mechanics Division, American Society of Civil Engineers, EM1, 11923, Feb. 1976.
- Митра, Шервуд (Mitra S. K., Sherwood R. J.)
Digital Ladder Networks, IEEE Transactions on Audio and Electroacoustics, Feb. 1973.
- Мэгреб, Бломквист (Magrab E. D., Blomquist D. S.)
The Measurement of Time-Varying Phenomena, Wiley, New York, 1971.
- Мэйнерд (Maunard H. W.)
An Evaluation of Ten Fast Fourier Transfer Programs, Army Electronics Command, White Sands Missile Range, March 1973.
- Найквист (Nyquist H.)
Certain Factors Affecting Telegraph Speed, Bell Systems Journal, Vol. 3, April 1924.
- Норин, Слоун (Norin R., Sloane E.)
A New Algorithm for the Improving Digital Random Control System Speed and Accuracy, IES Proceedings, pp. 46—52, 1975.
- Оливер, Пирс, Шеннон (Oliver B. M., Pierce J. R., Shannon C. E.)
The Philosophy of PCM, Proceedings of IRE, Nov. 1948.
- Оппенгейм и др. (под ред.) (Oppenheim A. V. et al., Eds.)
Selected Papers in Digital Signal Processing, II, IEEE Press, New York, 1976.
- Ормсби (Ormsby J. F.)
Design of Numerical Filters with Applications to Missile Data Processing, Journal of the Association for Computing Machinery, Vol. 8, No. 3, July 1961.
- Отнес, Мак-Нейми (Otnes R. K., McNamee L. P.)
Instability Thresholds in Digital Filters Due to Coefficient Rounding, IEEE Transactions on Audio and Electroacoustics, Vol. AU-18, pp. 456—463, 1970.
- Отнес, Нэтэнс, Эноксон (Otnes R. K., Nathans H. A., Enochson L.)
A Procedure for Computing Power Spectral Density of Gust Data, AFFDL-TR-69-11, Wright Patterson Air Force Base, Ohio, March 1969.
- Отнес, Эноксон (Otnes R. K., Enochson L.)
Digital Time Series Analysis, Wiley-Interscience, New York, 1972.
- Пайк, Хилл (Pike M. C., Hill J. D.)
Algorithm 266, Pseudo Random Numbers (65), Collected Algorithms from SACM, July 1965.
- Парзен (Parzen E.)
Mathematical Considerations in the Estimation of Spectra, Technometrics, Vol. 3, pp. 167—190, 1961.
- Паркс, Мак-Клеллан (Parks T. W., McClellan J. H.)
A Program for the Design of Linear Phase Finite Impulse Response Digital Filters, IEEE Transactions on Audio and Electroacoustics, Aug. 1972.
- Поттер (Potter R. W.)
Compilation of Time Windows and Time Shapes for Fourier Analysis, 02-5952-0705, Hewlett-Packard, 1971.
- Рабинер, Гоулд (Rabiner L. R., Gold B.)
Теория и применение цифровой обработки сигналов. Пер. с англ.—М.: Мир, 1978.

- Рабинер, Рейдер (под ред.) (Rabiner L. R., Rader C. M., Eds.)
Digital Signal Processing, IEEE Press, New York, 1972.
- Синглтон (Singleton R. C.)
A Method for Computing the Fast Fourier Transform with Auxiliary Memory and Limited High-Speed Storage, IEEE Transactions on Audio and Electroacoustics, Vol. AU-15, No. 2, June 1967.
An Algorithm for Computing the Mixed Radix Fast Fourier Transform, IEEE Transactions on Audio and Electroacoustics, Vol. AU-17, No. 2, June 1969.
- Слоун (Sloane E. A.)
Comparison of Linearly and Quadratically Modified Spectral Estimates of Gaussian Signals, IEEE Transactions on Audio and Electroacoustics, Vol. AU-17, No. 2, 1969.
- Стейглиц (Steiglitz K.)
An Introduction to Discrete Systems, Wiley, New York, 1974.
- Стокхэм (Stockham T. G.)
High Speed Convolution and Correlation, AFIPS Conference Proceedings, Vol. 28, pp. 229—233, 1966.
- Стоурер (Storer J. E.)
Passive Network Synthesis, McGraw-Hill, New York, 1957.
- Су (Hsu H. P.)
Fourier Analysis, rev. ed., Simon and Schuster, New York, 1970.
- Сэнд (Sande G.)
On an Alternative Method for Calculating Covariance Functions, Princeton Computer Memorandum, Princeton, N. J., 1965.
- Тейлхеймер (Theilheimer F.)
A Matrix Version of the Fast Fourier Transform, IEEE Transactions on Audio and Electroacoustics, Vol. AU-17, No. 2, 1969.
- Уомплер (Wampler R. N.)
On the Accuracy of Least Squares Computer Programs, JASA, Vol. 65, p. 549, June 1970.
- Уэйнреб (Weinreb S.)
A Digital Spectral Analysis Technique and Its Application to Radio Astronomy, MIT Research Laboratory of Electronics, Technical Report 412, Aug. 1963.
- Флетчер, Пауэлл (Fletcher R., Powell M. J. D.)
A Rapidly Convergent Descent Method for Minimization, Computer Journal, Vol. 6, No. 2, 1963.
- Форсайт, Молер (Forsythe G., Moler C.)
Численное решение систем линейных алгебраических уравнений. Пер. с англ. — М.: Мир, 1969.
- Фрэйзер (Fraser D. A. S.)
Statistics: An Introduction, Wiley, New York, 1958.
- Хастингс (Hastings C.)
Approximations for Digital Computers, Princeton University Press, Princeton, N. J., 1955.
- Хинич (Hinich M.)
Estimation of Spectra after Hard Clipping of Gaussian Processes, Technometrics, Vol. 9, p. 391, 1967.

- Хольц (Holtz H.)
The Synthesis of Linear Recursive Digital Filters Optimal in a Minimax Sense, Seventh Asilomar Conference on Circuits, Systems, and Computers, Pacific Grove, Calif., pp. 513—517, Nov. 1973.
- * Шабат Б. В.
Введение в комплексный анализ. Ч. 1, 2. 2-е изд. перераб. и доп.—М.: Наука, 1976.
- Шатцов, Файнберг, Цзаэ (Schatzoff M., Fienberg M., Tsao S.)
Efficient Calculation of All Possible Regressions, Technometrics, Vol. 9, pp. 531—540, 1968.
- Шмид (Schmid L. P.)
Efficient Autocorrelation, Communications of the ACM, Vol. 8, p. 115, 1965.
- Шеннон (Shannon C. E.)
Связь в присутствии шума.—В кн.. Работы по теории информации и кибернетике.—М.: ИЛ, 1963, с. 433—460.
- Шэфер Рабинер (Schafer R. W., Rabiner L. R.)
A Digital Signal Processing Approach to Interpolation, Proceedings of the IEEE Vol. 61, No. 6, June 1973.
- Эноксон, Гудмэн (Enochson L. D., Goodman N. R.)
Gaussian Approximations to the Distribution of Sample Coherence, AFFDL TR-65-57, Research and Technology Division, AFSC, Wright—Patterson Air Force Base Ohio, Feb. 1965.

СПИСОК ОБОЗНАЧЕНИЙ

- a_{1p} — первый рекурсивный весовой коэффициент фильтра в каскадной реализации
- a_{2p} — второй рекурсивный весовой коэффициент фильтра в каскадной реализации
- b — шаг в области частот ($= 1/(NT)$)
- b_0 — общий нерекурсивный постоянный множитель
- b_{0p} — первый нерекурсивный весовой коэффициент фильтра в каскадной реализации
- b_{1p} — второй нерекурсивный весовой коэффициент фильтра в каскадной реализации
- b_{2p} — третий нерекурсивный весовой коэффициент фильтра в каскадной реализации
- B — ширина полосы сигнала
- B_e — эффективная разрешающая полоса при вычислениях ПСМ
- $c(i)$ — свертка
- e — 2.7182818 ...
- $\exp(x) = e^x$
- f — частота
- F — частота Найквиста
- $h(i)$ — импульсная функция отклика
- $H(k)$ — преобразование Фурье функции $h(i)$ (передаточная функция)
- i — временной индекс; величина запаздывания
- $\text{Im} []$ — мнимая часть числа, заключенного в скобки
- $j = \sqrt{-1}$

- k — индекс для области частот
 N — число данных в выборке
 n — число степеней свободы
 p — индекс
 P — полное время выборки; длина ряда ($=NT$)
 q — индекс
 $r(i)$ — корреляция
 $\operatorname{Re} [\]$ — действительная часть числа, заключенного в скобки
 $s_{xx}(i)$ — ковариационная функция или сигнал
 S — скорость выборки (квантования) ($=1/T$)
 $S_{xx}(k)$ — спектр мощности ряда $x(i)$ в частоте $k/(NT)$
 t — время (непрерывное)
 T — интервал, с которым производится выборка, т. е. интервал квантования
 $u(t)$ — прямоугольный импульс длины P
 $U(f)$ — преобразование Фурье функции $u(t)$, равное $(\sin(\pi f P))/(\pi f)$
 $W_N = \exp(-j2\pi/N)$
 $x(i)$ — временной ряд (обычно вход системы)
 $X(k)$ — преобразование Фурье ряда $x(i)$
 $y(i)$ — временной ряд (обычно выход системы)
 $Y(k)$ — преобразование Фурье ряда $y(i)$
 $z(i)$ — временной ряд (иногда комплексный)
 $Z(k)$ — преобразование Фурье ряда $z(i)$
 α — весовой коэффициент фильтра первого порядка
 β — коэффициент регрессии
 τ — временная задержка (запаздывание)
 η — переменная интегрирования
 φ — фазовый угол; вероятностная функция плотности
 Φ — функция распределения
 σ — стандартное отклонение
 ζ — коэффициент затухания цифрового фильтра

Σ — знак суммы

$\pi = 3.14159265\dots$

Π — знак произведения

ω — частота в радианах

ω_n — собственная частота в радианах

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Автоковариация круговая (circular) 267
Акселерометр, переходные данные 244
Аналогии между понятиями математической статистики и частотными функциями 293, 331, 342
Антипереполнение (underflow) 57
Аргумент 205
Асимметрия (skewness) 43
- Байт (byte) 82
Бит (bit) 81
Бортовой наклон 220, 221
- Величина абсолютная 10
— выборочная 45
— гауссовская, независимость 51
— полученная линейным преобразованием из гауссовской 47
— случайная 42
— хи-квадрат 244
— экстремальная 64
— — тип пика 66
Волна развернутой синусоиды 362
— ударная 289
Восьмеричная (octal) запись 83
Временное запаздывание и фаза 338
Выборка и фиксация 78
- Гауссовская модель, пример 48
Гауссовское распределение см. Распределение гауссовское
Гиббса (Gibbs) явление 309
Гистерезис 77
Гистограмма вероятностная 58
— данных с неправдоподобными значениями 73
- Гистограмма срезаемых данных 74
Граф потока сигналов, децимация по времени 228
— — — — — частотам 228
- Данные (data) выборочные 28
— — периодичность рядов 35
— кадр (frame) 80
Датчик (transducer) 77
дБ 383
Двоично-десятичное кодирование (ДДК) 81
Двоичное целое число 81
Деление комплексных чисел 9
Дельта-функция 12
Демодуляция комплексная 194
Децимация (decimation) 185
— каскадная 188
— обратная 191
— по времени 219
— — частотам 219
— частных сумм 190
Динамический диапазон 89
Дисперсия 43
— выборочная несмещенная 45
— оценки 46
— прогноза для систем с несколькими входами 348
— смещение 45
Длина записи (record length) 215
— — для корреляции 280
— — конечная 25
Доверительная область для системы с одним входом 352
Дрожание (jitter) 91

- Единица, корни 220
- Задержка (delay) максимальная 256
— по времени 256, 285, 338
- Запаздывание (lag) 256
— максимальное для корреляции 281
- Звук, скорость 286
- Значение среднеквадратичное 44, 53
— преобразования в частоте Найквиста 214
- Импульс прямоугольный (boxcar) 18, 403
- Интервалы группировки (class intervals) 58
— — число 63
— доверительные 50
- Интерполяция гармоническая 215
— на ограниченной полосе частот 245
— теорема о выборке 31
- Искажение (distortion) 163, 165
- Калибровка (calibration) шаговая и синусоидальная 93
- Квантование однобитовое 264
- Ковариация (covariance) 51
— большая длина записи 263
— БПФ-метод вычисления 267
— нормировка 284
— ошибка при нормировке 285
— различие со сверткой 262
— смещенной во времени функции 258
— удаление среднего 284
- Когерентность (coherence), влияние на точность оценок передаточной функции *см. также* Функция когерентности 354
— доверительные границы 351
— изменчивость 349
— как доля мощности 347
— комплексная 342
— множественная (multiple) 348
— связь с передаточной функцией 375
— смещение (bias) 349, 350
— толкование для нелинейных систем 348
- Когерентность, устранение смещения 350
— частная (partial) 348, 349
- Кодирование линейно упреждающее 293
- Комплексные числа 9, 10
- Контрольная сумма 82
- Координаты, поворот осей 220
— полярные 10, 205, 338
- Корреляция (correlation) *см. также* Функция корреляции 51, 52
— выборочная 257
— доверительные границы 281
— максимальное значение запаздывания 281
— пример подмены 289
— срезание сигнала, допущения 265
— — — дисперсия 264
— — — коэффициенты 264
- Косинус-преобразование 16, 203
- Кросс-ковариация (cross covariance) 255, 256
- Кросс-корреляция, оценка (cross correlation) 257
- Кросс-спектр (cross spectrum) 257
- Кросс-спектральная плотность 52, 329
- Линейная система 20
— — кросс-спектры 343
— — моделирование 336
— — со многими входами 344
— — — — решение по методу наименьших квадратов 346
— — — — — как произведение матриц 357
— — спектральные соотношения 340
— — спектры мощности (power spectra) 343
— — физическая осуществимость 25, 337
- Линейность системы 339
- Линия регрессии 52
- Максимум энтропии 293
- Мантисса 83
- Математическое ожидание максимального значения 64

- Матрица, изоморфная комплексной 357
 — перестановки 222
 — спектральная 346
 — — в памяти 356
 — — расширенная 348
 — — условная 349
 Метод эскалатора (escalator) 358
 Модуль 205
 Моменты центральные и гауссовского
 распределения 49
 Мультиплексор 80
- Найквист см.* Частота Найквиста
 Наложение (w/ground) 37, 273, 298
 Независимость 50
 Неправдоподобные (wild) данные *см.*
 Редактирование данных с неправдо-
 подобными значениями
 Неустойчивость (instability) фильтра
 164, 168
 Нули 10, 105
 — добавление 267, 268
- Область импульсного отклика фильтра
 282
 Обратная (reversal) перестановка битов
 226
 Окно (window) Гудмэна 318, 320
 — ГЭО 315, 317
 — косинусное 314, 315
 — Парзена 405
 — прямоугольное 312, 313
 — спектральное 309, 404
 — Хемминга 405
 — Хеннинга 405
 Округление (rounding) 85
 Операция развертки (sweep) 358, 359
 — сглаживания (smoothing) 261
 Ортогонализация 358
 Осреднение ПСМ 311, 321
 — — комбинированное 322, 328,
 — — линейное 331
 — — по ансамблю 311, 322, 325
 — — — частотам 311, 321, 322
 Осциллятор 117
 Отклонение стандартное 43
- Оценка 45
 — несостоятельная (inconsistent) 54
 — смещенная 46
 — частоты 28
 Ошибка округления 57
 — при наложении 289
 — стандартная оценки 258
 — щелевая (aperture) 90
- Перестановка циклическая для кова-
 риации 270, 271
 — — — в области частот 270
 — — — во временной области 274
 Периодограмма (periodogram) 46
 ПКС (плотность кросс-спектра) 52, 329
 Плавающая точка 82
 Плотность вероятностная пиков 64
 — выборочная вероятностная 59
 Поворот в координатных плоскостях 220
 — умножение на W_N 219
 Подмена (aliasing) 29
 — во временной области 273, 289
 — диаграмма 30
 — пример 29
 Полоса (bandwidth) вычислительная 298
 — разрешающая (resolution) 298, 299
 — элементарная 201
 — эффективная разрешающая (effective
 resolution) 296
 Полюсы 10, 105
 — действительные 111
 — комплексные 114
 Почти всюду 12
 Преобразование аналого-цифровое
 (analog to digital conversion) 85
 — косинуса 14, 27
 — синуса 15
 — Фишера z 349
 — Фурье 11, 200
 — — быстрое (БПФ), алгоритм Сэнда —
 Тьюки 219, 227
 — — — 2, 3, 4, 5-алгоритм Синглтона
 234
 — — — алгоритм Синглтона для внеш-
 них запоминающих устройств (out of
 core) 235

- Скорость развертки синусоиды 363
 Слово 82
 Смещение гистограммы 60
 Сопротивление механическое 336
 Сопряженность комплексная 9
 Состоятельность (consistency) 46
 Спектральная плотность *см.* ПКС и ПСМ
 Спектр условный (conditioned) 348
 Среднее (mean), вычисление 56
 — истинное, множественное 43
 — удаление 305
 Срезание (truncation) 85
 Стандартизация 48
 Степени свободы (ст. св.) 62
- Теорема Винера — Хинчина** 54, 257
 — о выборке 31
 — — свертке 291, 338
 — основная алгебры 10
 — центральная предельная 49, 72
 Тест хи-квадрат 60
 Точка 3-дБ 135
 Точность двойная 85
- Удаление (removal) треида** 98
Умножитель на базе «четверти квадратов» 266
 — — «суммы квадратов» 266
Уравнение интегральное 105
 — линейное дифференциальное 20, 105
Усиления коэффициент (gain) 205, 338
 — фильтра системы с одной степенью свободы 21
Усреднение (average), скользящее 261
 — — авторегрессионное 293
Устойчивость (stability) 106
 — ПСМ статистическая 300
Уточнение (validation) данных 72
- Фаза (phase)** 205, 338
 — для системы с одной степенью свободы 21
 — линейная 292
 — ошибка 352
- Фаза, продолжение (straightening)** 218
Фиксированная точка 82
Фильтр Баттеруорта (Butterworth) 79, 117
 — — синусный 129
 — — — высоких частот 185
 — — — полоснопропускающий 138
 — второго порядка 110
 — — — низкочастотный 117
 — дифференциатор 107
 — заграждающий (band reject) 142, 150
 — импульсный с бесконечной областью отклика (ИБО) 104
 — — — конечной областью отклика (ИКО) 152
 — Ормсби (Ormsby) 153
 — параллельный 120, 123
 — Паркса — Мак-Клеллана (Parks, McClellan) 153
 — первого порядка 106
 — полоснопропускающий (bandpass) 104, 136, 150
 — последовательный (series) 122
 — противоположенный 79
 — прямой (direct) 120
 — системы с одной степенью свободы 20
 — тангенсный 145
- Фильтрация безфазовая** 188
 — полоснопропускающая 282
 — посредством свертки 282
Фильтры идеальные базисные 126, 127
Фишера (Fisher) Z-преобразование 281
Фортран 384
Функция автоковариации 54
 — вероятностная плотности 59, 64, 66
 — дробно-рациональная 10
 — импульсного отклика 20, 337
 — — — системы с одной степенью свободы 22
 — ковариации 255, 256
 — когерентности 330, 336, 342
 — корреляции 255, 284
 — многомерная плотности 66
 — нечетная 15
 — передаточная (transfer) 21, 104, 329, 336

- Функция автоковариации в векторной записи 346
 — — доверительная область 352
 — — доверительные интервалы 352
 — — коэффициент усиления 108
 — — методы измерений 360
 — — программа вычислений 393
 — — фаза 108
 — плотности (density) 294
 — свертки 255, 258
 — четная 15
 — $(\sin x)/x$ 19
 Фурье преобразование, см. Преобразование Фурье
- Характеристика 83
 Хеннинг 323
- Частота мгновенная 362
 — Найквиста 30
 — пиковая системы с одной степенью свободы 21
- Часть действительная 9
 — мнимая 9
 Число комплексное 9
- Шаг (spacing, increment) в области частот 208, 213
 Шестнадцатеричная (hexadecimal) запись 83
 Шум (noise) аддитивный белый 55
 — — — как вход системы 365
 — — влияние на передаточную функцию 368
 — — при округлении 85
 — — равномерный, среднее и дисперсия 68
 — — узкополосный 248
 — — цифрового преобразователя 87
- Щелевая (aperture) функция 79
- Эйлера соотношение 10, 253
 Эксцесс (kurtosis) 49
 Эллипс постоянной вероятности 52
 Эргодичность 43

ОГЛАВЛЕНИЕ

Предисловие редактора перевода	5
Предисловие	6
Глава 1. Предварительные понятия	9
1.1. Введение	9
1.2. Операции над комплексными числами	9
1.3. Преобразование Фурье	11
1.4. Прямоугольный импульс и его преобразование Фурье	18
1.5. Линейные системы	20
1.6. Свертка и физическая осуществимость	23
1.7. Эффекты конечной длины выборки	25
1.8. Выборочные данные и частота Найквиста	28
1.9. Дискретное преобразование Фурье длины N	32
1.10 z -преобразование	39
Упражнения	40
Глава 2. Элементы теории вероятностей и статистики	42
2.1. Введение	42
2.2. Выборочные величины и оценки	45
2.3. Нормальное распределение	47
2.4. Корреляция и регрессия	51
2.5. Функция плотности спектра мощности	53
2.6. Как вычислять среднее и дисперсию	56
2.7. Вероятностные гистограммы	58
2.8. Вероятностные функции плотности пиков	64
2.9. Многомерные функции плотности	66
2.10. Примеры и применения вероятностных функций плотности	67
Упражнения	75
Глава 3. Сбор и предварительная обработка данных	77
3.1. Введение	77
3.2. Сбор данных	78

3.3.	Цифровое представление информации	81
3.4.	Аналого-цифровое преобразование	85
3.5.	Другие погрешности	89
3.6.	Переход к физическим единицам	93
3.7.	Редактирование данных с неправдоподобными значениями	95
3.8.	Удаление тренда	98
	Упражнения	102
Глава 4.	Конструкции цифровых фильтров	104
4.1.	Основные понятия	104
4.1.	Фильтры первого порядка	106
4.3.	Фильтры второго порядка	110
4.4.	Фильтры более высоких порядков	119
4.5.	Идеальные базисные фильтры	126
4.6.	Синусный фильтр низких частот Баттеруорта	129
4.7.	Синусный фильтр высоких частот Баттеруорта	135
4.8.	Полоснопропускающие фильтры	136
4.9.	Заграждающие фильтры	142
4.10.	Тангенсные фильтры	145
4.11.	Другие рекурсивные фильтры	151
4.12.	Нерекурсивные (ИКО) фильтры	152
4.13.	Приемы аппроксимации фильтров	157
	Упражнения	160
Глава 5.	Практические аспекты цифровой фильтрации	163
5.1.	Введение	163
5.2.	Шум и искажение	163
5.3.	Ухудшение фильтра	168
5.4.	Реализация фильтра	179
5.5.	Децимация	185
5.6.	Обратная децимация	191
5.7.	Сведение к общей скорости выборки	193
5.8.	Комплексная демодуляция	194
	Упражнения	196
Глава 6.	Преобразования Фурье	200
6.1.	Элементы теории	200
6.2.	Алгоритм быстрого преобразования Фурье	219
6.3.	Примеры	240
	Упражнения	250
Глава 7.	Ковариация и свертка	255
7.1.	Элементы теории	255
7.2.	Различия между ковариацией и сверткой	262

7.3.	Длинные ряды и основная вычислительная процедура для ковариации	263
7.4.	Вычисление ковариации и свертки при помощи БПФ	266
7.5.	Эффекты подмены и наложения	273
7.6.	Как вычислять функции ковариации и свертки	280
7.7.	Область импульсного отклика и полоса частот при фильтрации посредством свертки	282
7.8.	Нормировка и удаление среднего при вычислении ковариаций	284
7.9.	Примеры использования ковариации и свертки	285
	Упражнения	291
Глава 8.	Плотности кросс-спектра и спектра мощности	293
8.1.	Обзор основных методов вычисления спектров	293
8.2.	Понятие плотности	294
8.3.	Ширина эффективной разрешающей полосы	296
8.4.	Пределы разрешающей способности	298
8.5.	Статистическая устойчивость	300
8.6.	Утечка	306
8.7.	Как вычислять спектральные функции	309
8.8.	Функции временного сглаживания — окна данных	312
8.9.	Примеры использования функций плотности спектра мощности	333
	Упражнения	335
Глава 9.	Передаточные функции и функция когерентности	336
9.1.	Свойства передаточных функций	336
9.2.	Спектральные соотношения для систем с одним входом	340
9.3.	Спектральные соотношения для линейных систем со многими входами	344
9.4.	Функции обычной, множественной и частичной когерентности	347
9.5.	Доверительные границы для когерентности	349
9.6.	Определение доверительных границ для передаточных функций	351
9.7.	Как вычисляются передаточные функции	354
9.8.	Оператор развертки	358
9.9.	Определение передаточной функции при помощи синусоид	360
9.10.	Определение передаточной функции при помощи случайных входов	364
9.11.	Функция когерентности при $B_e = 1/P$	366
9.12.	Примеры вычисления передаточных функций	368
	Упражнения	383

Приложение А. Программы для анализа временных рядов	384
А.1. Введение	384
А.2. Генератор случайных чисел	385
А.3. Получение графиков на печатающем устройстве	387
А.4. Быстрое преобразование Фурье	387
А.5. Получение весовых коэффициентов низкочастотного фильтра	390
А.6. Передаточная функция цифрового фильтра	393
А.7. Тест-программа и результаты проверок	395
Приложение В. Процедура Блэкмена — Тьюки вычисления плотностей	
 спектра мощности	402
В.1. Введение	402
В.2. Формулы для вычислений на ЭВМ	404
В.3. Процедура получения кросс-спектра	406
В.4. Замечания	408
Литература	410
Список обозначений	416
Предметный указатель	419

УВАЖАЕМЫЙ ЧИТАТЕЛЬ!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу 129820 Москва, И-110, ГСП, 1-й Рижский пер., д. 2, издательство «Мир».