

Министерство образования Российской Федерации
Ярославский государственный университет им. П.Г. Демидова
Кафедра радиофизики

А.Н. Кренёв, Т.К. Артёмова

ЦИФРОВОЙ СПЕКТРАЛЬНЫЙ АНАЛИЗ

Учебное пособие

Ярославль 2003

ББК 32.811
К 79
УДК 621.372 (075)

Цифровой спектральный анализ: Учеб. пособие / А.Н. Кренёв, Т.К. Артёмова. Ярослав. гос. ун-т. Ярославль, 2002. 114 с.

ISBN 5-8397-0238-2

В пособии дается систематическое изложение математических основ и методов цифрового спектрального анализа.

Пособие может быть полезным для студентов, аспирантов и специалистов различных специальностей, использующих методы цифрового спектрального анализа в физических исследованиях и математическом моделировании.

Ил. 53. Табл. 3 . Библиогр.: 8 назв.

Рецензенты: Кафедра радиофизики Ярославского государственного университета им. П.Г. Демидова

ISBN 5-8397-0238-2

© Ярославский государственный
университет им. П.Г. Демидова,
© А.Н. Кренёв,
Т.К. Артёмова, 2002

ОГЛАВЛЕНИЕ

	стр.
ВВЕДЕНИЕ	5
1. ОСНОВНЫЕ ПОНЯТИЯ, ЗАДАЧИ И ПРОБЛЕМЫ ЦИФРОВОГО СПЕКТРАЛЬНОГО АНАЛИЗА	6
1.1. Цифровой спектральный анализ	6
1.2. Проблемы в области спектрального оценивания	7
1.2.1. Разрешение	8
1.2.2. Обнаружимость сигналов	9
1.2.3. Оценивание параметров сигнала против спектрального оценивания	10
1.2.4. Состоятельность оценки	11
1.3. Спектральный анализ и z-преобразование, использование ДПФ в спектральном анализе	11
1.4. Взаимная связь спектров непрерывных и дискретных сигналов	14
2. СПЕКТРАЛЬНЫЙ АНАЛИЗ И БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ	18
2.1. Алгоритмы БПФ с основанием 2. Алгоритм БПФ с прореживанием по времени	18
2.1.1. Базовая операция. Алгоритм с замещением	22
2.1.2. Перестановка данных и двоичная инверсия	23
2.2. Алгоритм БПФ с прореживанием по частоте	25
2.3. Вычисление обратного ДПФ с помощью алгоритма прямого ДПФ	27
2.4. Единый подход к алгоритмам БПФ	28
2.5. Особенности цифровой реализации алгоритма БПФ	31
3. СООТНОШЕНИЕ МЕЖДУ БПФ И ЦИФРОВОЙ ФИЛЬТРАЦИЕЙ	37
3.1. Некоторые характеристики спектрального анализа на основе БПФ	37
3.2. Соотношение между "скачущим" БПФ и гребенкой фильтров	40
4. ИСПОЛЬЗОВАНИЕ ОКОН В СПЕКТРАЛЬНОМ АНАЛИЗЕ	42
4.1. Прямоугольное окно и БПФ	42
4.2. Наиболее распространенные функции окна и их характеристики	44
5. КЛАССИЧЕСКИЕ МЕТОДЫ СПЕКТРАЛЬНОГО ОЦЕНИВАНИЯ	55
5.1. Периодограммные методы спектрального анализа	56
5.1.1. Периодограмма Даньелла	57
5.1.2. Периодограмма Бартлетта	58
5.1.3. Периодограмма Уэлча	59
5.2. Корреляционный анализ и спектральная оценка	64
5.2.1. Метод получения оценки ФВК	66
5.2.2. Коррелограммный метод оценки СПМ	67
5.3. Критерии качества классической спектральной оценки	69
5.4. Комбинированные периодограммно-коррелограммные оценки	71
6. АЛГОРИТМЫ СПЕКТРАЛЬНОГО АНАЛИЗА В ОГРАНИЧЕННОМ СЕКТОРЕ Z-ПЛОСКОСТИ	73
6.1. Алгоритм анализа с использованием БПФ	73
6.2. Алгоритм Блустейна	74
6.3. Алгоритм z-преобразования с использованием ЛЧМ-фильтрации	75

7. СПЕКТРАЛЬНЫЙ АНАЛИЗ, ОСНОВАННЫЙ НА ЛИНЕЙНОМ МОДЕЛИРОВАНИИ ИССЛЕДУЕМОГО СИГНАЛА	81
7.1. Модели случайных процессов	81
7.2. Методы оценки параметров АР-моделей. Блочные алгоритмы	86
7.3. Оценка параметров АР-модели по автокорреляционной последовательности. Метод Юла – Уолкера	88
7.4. Оценка параметров АР-моделей по последовательности оценок коэффициентов отражения	90
7.4.1. Геометрический алгоритм	91
7.4.2. Гармонический алгоритм Берга	92
7.4.3. Рекурсивное оценивание по методу максимального правдоподобия .	94
7.5. Оценка параметров АР-моделей, основанная на линейном предсказании по МНК	95
7.6. Сравнительные характеристики СПМ, получаемых с помощью блочных методов АР-оценивания	97
7.7. Авторегрессионное спектральное оценивание: алгоритмы обработки последовательных данных	100
7.8. Некоторые практические вопросы спектрального оценивания, основанного на линейном моделировании	102
7.8.1. Выбор порядка модели	102
7.8.2. Авторегрессионные процессы с шумом наблюдения	104
7.8.3. Оценивание мощности синусоидальных компонент	105
7.8.4. Частотное разрешение	106
7.9. Метод Прони	107
7.9.1. Исходный метод Прони	107
7.9.2. Метод наименьших квадратов Прони	110
ЛИТЕРАТУРА	113

ВВЕДЕНИЕ

Цифровой спектральный анализ (ЦСА) является важной составной частью методов цифровой обработки сигналов.

Первоначально ЦСА использовался только как инструмент обработки данных и численного моделирования на цифровых ЭВМ. С развитием алгоритмов обработки сигналов и элементной базы цифровой электроники возникло новое направление ЦСА – аппаратный цифровой спектральный анализ. В современном ЦСА используются алгоритмы цифровой фильтрации, быстрых ортогональных преобразований, линейного моделирования исследуемого сигнала и статистической обработки. Аппаратная база ЦСА включает специализированные сигнальные процессоры и аналого-цифровые преобразователи, адаптированные к решаемым задачам.

Сегодня трудно назвать область науки и техники, где при использовании методов цифровой обработки сигналов или численного моделирования в той или иной степени не применяются методы ЦСА. Данное обстоятельство привело к выводу о необходимости разработки самостоятельного курса "Цифровой спектральный анализ". Теоретической базой "ЦСА" являются курсы "Теоретические основы радиотехники", "Теория дискретных и цифровых сигналов и цепей", "Цифровая фильтрация", "Проектирование микропроцессорных систем", а также соответствующие разделы высшей математики.

Первый раздел пособия посвящен общим вопросам ЦСА, классификации методов анализа, обсуждению задач и проблем, стоящих перед ним.

Во втором разделе рассмотрены математические основы и алгоритмы реализации быстрых преобразований Фурье (БПФ), лежащих в основе или составляющих часть большинства методов анализа, а также особенности, возникающие при квантовании сигнала и использовании при обработке чисел с конечной разрядностью.

В третьем разделе обсуждаются вопросы ЦСА на основе алгоритмов цифровой фильтрации и БПФ. Четвертый раздел пособия посвящен вопросам управления характеристиками анализатора путем использования метода окон. В пятом и шестом разделах рассмотрены основные методы ЦСА, используемые в задачах аппаратного анализа. Приведено подробное описание алгоритмов.

Основу учебного пособия составляет материал монографий С.Л. Марпла-мл. [1] и Л. Рабинера, Б. Гоулда [2], А. М. Трахтмана [4], соответствующие разделы учебного пособия Л. М. Гольденберга, Б. Д. Матюшкина, М. Н. Поляка [3], а также результаты самостоятельных исследований авторов и курс лекций, читаемый А.Н. Крениным для студентов ЯрГУ им. П.Г. Демидова.

1. ОСНОВНЫЕ ПОНЯТИЯ, ЗАДАЧИ И ПРОБЛЕМЫ ЦИФРОВОГО СПЕКТРАЛЬНОГО АНАЛИЗА

1.1. Цифровой спектральный анализ

При обработке сигналов во многих случаях приходится измерять спектры. Понятие "*спектральный анализ*" включает в себя большое число различных измерений.

Термин "*цифровой спектральный анализ*" отражает тип исследуемого колебания и используемый математический аппарат. Так, исследуемое колебание может быть дискретным или цифровым, а используемый математический аппарат должен обеспечивать разложение исследуемого колебания по дискретной или цифровой системе базисных функций. Математический аппарат, в зависимости от метода анализа, базируется, в основном, на дискретном преобразовании Фурье, z -преобразовании, дискретной свертке, разностных уравнениях, линейном моделировании.

В литературе термин "цифровой спектральный анализ" относят к обработке как дискретных сигналов, так и цифровых. Строго говоря, обработке дискретных сигналов должен соответствовать термин "дискретный спектральный анализ", где представление значений как исследуемого сигнала, так и различных констант и переменных является точным. В литературе принято пользоваться одним термином – "цифровой спектральный анализ", определяя его как "измерение, которое дает точные или приближенные значения z -преобразования дискретного или цифрового сигнала для заданных значений z ".

Спектральный анализ, и цифровой в частности, не ограничивается задачей определения спектра исследуемого сигнала – совокупности коэффициентов разложения по выбранной системе базисных функций. Так, при ЦСА решаются задачи оценки частоты спектральных линий, совместной оценки спектральной плотности случайного процесса и амплитуд спектральных линий детерминированного сигнала и ряд других. Мы говорим об *оценке*, так как в случае цифровой обработки возникают специфические искажения как исследуемого сигнала, так и результатов анализа, что не позволяет произвести точные измерения. В ряде случаев искажения (например, эффект наложения) возникают и при дискретном анализе. Кроме того, на практике все спектральные измерения проводятся на конечных временных интервалах, длина которых часто выбирается интуитивно на основе накопленного опыта. Данное обстоятельство заставляет особое внимание уделять интерпретации результатов анализа, а в некоторых случаях и их адекватности результатам соответствующего континуального анализа. Последняя проблема возникает, когда цифровой сигнал получен из континуального прототипа.

Поскольку ЦСА является численным методом анализа, то его конкретная реализация зависит от способа получения и представления исследуемого сигнала и аппаратной реализации вычислительных средств. Основными об-

ластями применения ЦСА являются математическое моделирование, обработка полученных ранее экспериментальных данных и обработка данных, получаемых непосредственно в процессе анализа. Последняя область применения – аппаратный ЦСА, функционирующий, как правило, в условиях ограничения времени анализа. В качестве цифровых вычислительных устройств, реализующих алгоритмы анализа, могут быть использованы универсальные ЭВМ, специализированные сигнальные процессоры с плавающей и фиксированной запятой. В задачах аппаратного ЦСА используются высокопроизводительные сигнальные процессоры с плавающей запятой или целочисленной арифметикой.

Классификация алгоритмов анализа, используемых в задачах ЦСА, приведена на рис. 1.1.

Алгоритмы, основанные на цифровой фильтрации, реализуются с помощью различных схем цифровых фильтров.

Алгоритмы, основанные на быстром преобразовании Фурье (БПФ), являются высокоэффективными методами.

Часть методов, объединенных под общим названием классических, использовалась уже на первых порах создания цифрового спектрального анализа. Они хорошо разработаны, относительно просты и удобны, легко реализуются программно. Дают статистически устойчивые (состоятельные) оценки. Недостатки этих методов заключаются в довольно широком круге ограничений и искажениях спектра, связанных с обработкой конечных последовательностей.

Перечисленные недостатки устраняются в методах спектрального анализа в заданном секторе z -плоскости. Они позволяют получать оценки z -преобразования в точках произвольного контура или, что то же, оценки спектральных составляющих в точках, произвольным образом расположенных на частотной оси.

Методы, основанные на моделировании исследуемого сигнала выходным сигналом линейного фильтра (авторегрессионные, АР-методы) или рядом комплексных экспонент (метод Прони), являются относительно новыми в практике ЦСА и получили признание благодаря высокой точности и, что особенно важно, высокому разрешению спектральных оценок.

Математические основы всех этих алгоритмов различны.

Перечисленные алгоритмы и их особенности, возникающие при квантовании сигнала и использовании при обработке чисел с конечной разрядностью, рассматриваются в разделах 3 – 6 настоящего пособия.

1.2. Проблемы в области спектрального оценивания

Интерес к цифровым методам спектрального анализа поддерживается тем улучшением характеристик, которое они обещают, а именно: высоким частотным разрешением, повышенной способностью к обнаружению слабых

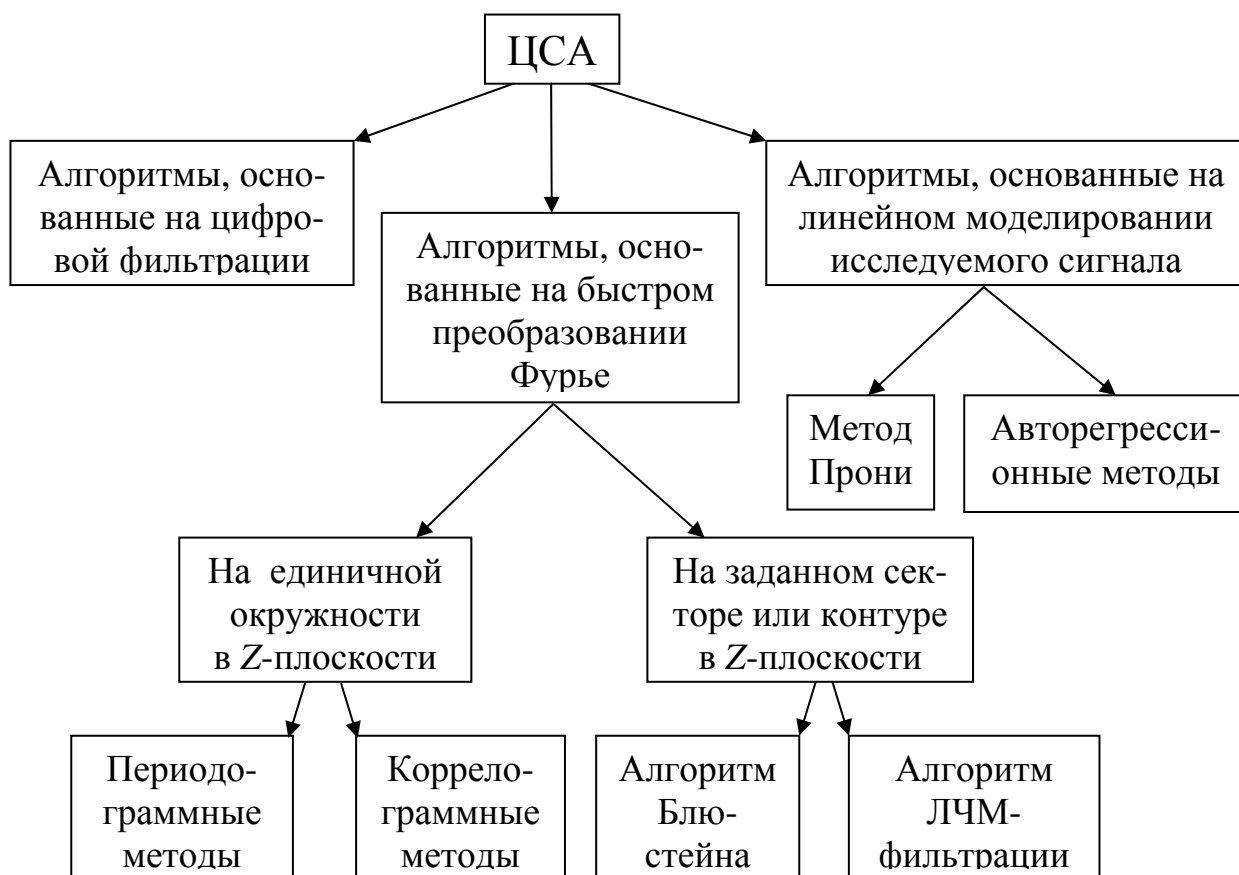


Рис. 1.1

сигналов или же сохранением "достоверности" формы спектра при меньшем числе используемых параметров. Однако при применении методов цифрового спектрального анализа к конкретным задачам встают следующие проблемы, диктующие выбор метода.

1.2.1. Разрешение

Спектральное разрешение относится к числу главных проблем современного спектрального оценивания, в особенности применительно к анализу коротких последовательностей данных. При этом то, что понимается под термином "разрешение", носит весьма субъективный характер. Одно из ранних определений принадлежит Релею, которое исходит из следующего рабочего определения для разрешения оптических телескопов с ограниченной пространственной апертурой: "Подобно тому, как оптическая сила телескопа измеряется близостью двойных звезд, которые он может разрешить, так и оптическую силу спектрографа следует измерять близостью самых близких двойных линий в спектре, которые он может разрешить".

В литературе принято характеризовать относительные величины разрешающей способности двух спектральных оценок на основе визуальных

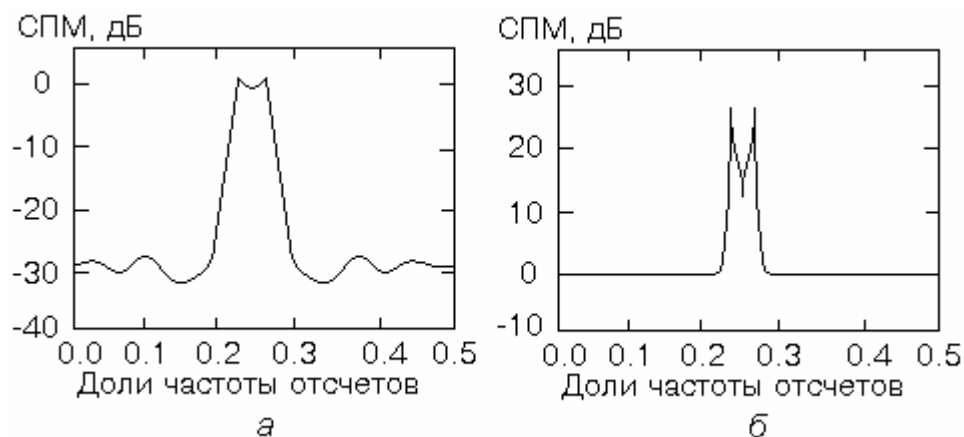


Рис. 1.2

наблюдений. Рассмотрим две спектральные оценки, показанные на рис. 1.2. Спектральная оценка на рис. 1.2, *а* имеет единичный нормированный максимум с двумя слабо различимыми пиками, что должно указывать на присутствие двух спектральных составляющих. Если теперь нелинейным масштабированием сформировать новую спектральную функцию, которая показана на рис. 1.2, *б*, то она будет выглядеть, как спектр с "более высоким" разрешением. На самом же деле оба приведенных спектра несут в себе одну и ту же информацию, а это означает, что одного лишь визуального сравнения спектральных оценок, конечно же, недостаточно для суждения о характеристиках разрешения.

1.2.2. Обнаружимость сигналов

Использование спектрального оценивания для обнаружения присутствия сигналов приводит к возникновению другой проблемы. Рассмотрим периодограммы, показанные на рис. 1.3. Обе они соответствуют одной и той же 1024-точечной реализации некоторого процесса, состоящего из аддитивной смеси одной синусоиды и белого шума при отношении сигнал/шум, равном -17 дБ. Периодограмма на рис. 1.3, *а* получена непосредственно по всем 1024 отсчетам, и в ней над СПМ шума отчетливо видна спектральная линия синусоиды с относительной частотой 0.18. На рис. 1.3, *б* представлена сглаженная периодограмма, полученная в результате разбиения исходной последовательности на 32 сегмента по 32 отсчета в каждом и последующего усреднения периодограмм этих сегментов. На этой оценке более очевидна равномерная спектральная характеристика белого шума, но отклик, обусловленный присутствием синусоиды, здесь не обнаружим. Можно показать, что обнаружимость сигнала по несглаженной периодограмме превышает обнаружимость этого сигнала по сглаженной периодограмме в число раз, равное примерно корню квадратному из числа сегментов, или на $10\lg(32^{1/2}) = 7,5$ дБ в рассматриваемом случае. С точки зрения отображения полного спектра процесса (и оценки его СПМ) лучше сглаженная спектральная оценка, показан-

ная на рис. 1.3, б, но с точки зрения обнаружимости сигнала – лучше несглаженная оценка.

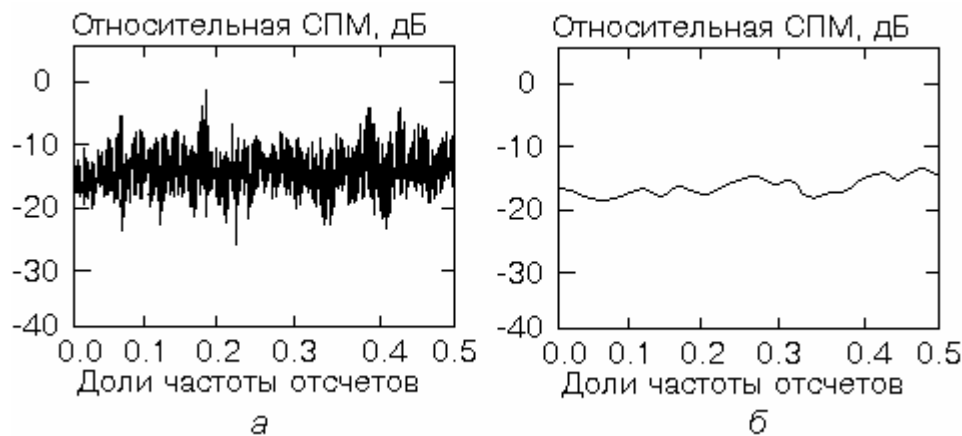


Рис. 1.3

Таким образом, аргументы в пользу выбора той или иной спектральной оценки будут зависеть от того, интересует ли нас сглаженная оценка в пределах всего диапазона анализируемых частот или же нам важна более высокая степень обнаружимости сигнала на некоторых конкретных его участках. Поэтому алгоритмы спектрального оценивания вовсе не обязательно являются также и хорошими алгоритмами обнаружения.

1.2.3. Оценивание параметров сигнала против спектрального оценивания

Задача **спектрального оценивания** подразумевает оценивание некоторой функции частоты. О характеристиках спектральной оценки судят по тому, насколько хорошо она согласуется с известным спектром тест-сигнала в некоторой непрерывной области частот.

Задача **оценивания параметров сигнала** встает, когда основной интерес представляет лишь локальное поведение спектральной оценки в некоторых заданных диапазонах частот. Примером может служить оценивание частоты синусоиды в белом шуме, результаты которого представлены на рис. 1.3.

Видно, что в качестве спектральной оценки лучше использовать оценку на рис. 1.3, б, так как она ближе к равномерному спектру белого шума. В качестве же оценки частоты синусоиды лучше использовать оценку на рис. 1.3, а, так как она позволяет более точно определить частоту этой синусоиды.

Задача оценивания параметров и задача спектрального оценивания требуют различного статистического подхода, хотя обе они могут перекрываться в некоторых частных случаях. Поэтому о характеристиках спектраль-

ной оценки нельзя полностью судить по ее характеристикам как оценки параметров. Так, например, сравнение оценок, представленных на рис. 1.3, показывает, что сглаженная периодограмма является хорошей спектральной оценкой, но плохой оценкой параметра, которым в данном случае является частота синусоиды.

1.2.4. Состоятельность оценки

Спектральная оценка, получаемая по конечной реализации данных, характеризует некоторое предположение относительно той истинной спектральной функции, которая была бы получена, если бы в нашем распоряжении имелась запись данных бесконечной длины. Именно поэтому поведение и характеристики спектральных оценок должны описываться с помощью статистических терминов.

Общепринятыми статистическими критериями качества оценки являются ее *смещение* и *дисперсия*, для получения которых нужна серия спектральных оценок. Если в некоторой области частот значение дисперсии велико, то это свидетельствует о том, что особенности, наблюдаемые в спектре отдельной реализации, не могут считаться статистически значимыми. С другой стороны, особенности совмещенных спектров в тех областях, где эта дисперсия мала, с большей достоверностью могут быть соотнесены с действительными частотными составляющими анализируемого сигнала.

Однако в случае коротких записей данных часто не удастся получить несколько спектральных оценок, да и сам статистический анализ отдельных спектральных оценок, полученный по коротким записям данных, в общем случае представляет собой весьма трудную проблему.

В предыдущих рассуждениях мы делали предположение о стационарности анализируемого процесса. В противном случае задача спектральной оценки значительно усложняется.

1.3. Спектральный анализ и z -преобразование, использование ДПФ в спектральном анализе

Сигналы, подлежащие спектральному анализу, можно свести к шести случаям, показанным на рис. 1.4. Сигналы лежат в определенной области z -плоскости. На рис. 1.4, *a*, *b* и *d* эти области представляют собой совокупности точек на единичной окружности в z -плоскости. В трех остальных случаях измерения также проводятся в отдельных точках, которые, однако, не лежат на единичной окружности. В общем случае задачу спектрального анализа можно рассматривать как задачу вычисления z -преобразования модифицированного к удобному случаю из приведенных на рис. 1.4 сигнала в некоторой области на z -плоскости.

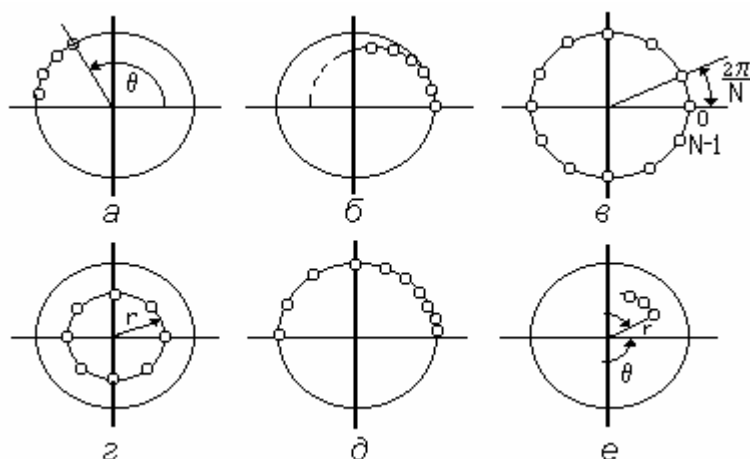


Рис. 1.4

Теоретически спектр можно измерять в любой точке z_1 на z -плоскости. Однако, практически при таких измерениях нужно учитывать время вычисления и эффекты, обусловленные квантованием. С учетом сказанного обобщенный спектр сигнала $x[n]$ можно определить как

$$S_n(z_1) = x[n] + x[n-1]z_1^{-1} + x[n-2]z_1^{-2} + \dots + x[n-N+1]z_1^{-(N-1)}, \quad (1.1)$$

или

$$S_n(z_1) = \sum_{m=n-N+1}^n x[m]z_1^{-(n-m)}, \quad (1.2)$$

где N – число отсчетов, по которым находится оценка спектра.

Когда спектр сигнала меняется во времени, возникает необходимость измерять $S_n(z_1)$ для последовательных значений n , т. е. значения $S_0(z_1)$, $S_1(z_1)$, $S_2(z_1)$ и т. д. Такой способ измерений называют **скользящим**. Он обеспечивается за счет смещения на один отсчет вперед временного окна, содержащего N отсчетов, и повторения измерения.

Анализ формул (1.1) и (1.2) показывает, что скользящее спектральное измерение в одной точке $z = z_1$ эквивалентно обработке сигнала $x[n]$ КИХ-фильтром с импульсной характеристикой вида

$$h[n] = z_1^{-n}, \quad 0 \leq n \leq N-1. \quad (1.3)$$

На рис. 1.5 приведена схема КИХ-фильтра для вычисления прямой свертки, обеспечивающая спектральные измерения в соответствии с формулой (1.2).

Проанализировав выражения для двух последовательных спектральных измерений $S_{n-1}(z_1)$ и $S_n(z_1)$, можно получить следующую рекуррентную

формулу:

$$S_n(z_1) = z_1^{-1} S_{n-1}(z_1) + x[n] - z_1^{-N} x[n-N]. \quad (1.4)$$

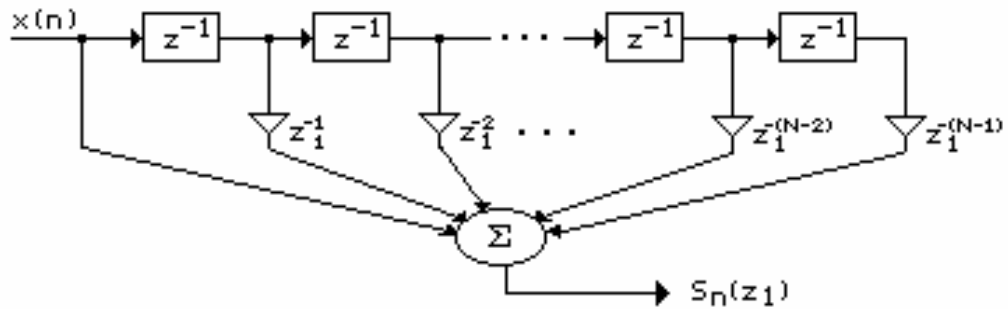


Рис. 1.5

Схема фильтра, реализующего алгоритм (1.4), приведена на рис. 1.6.

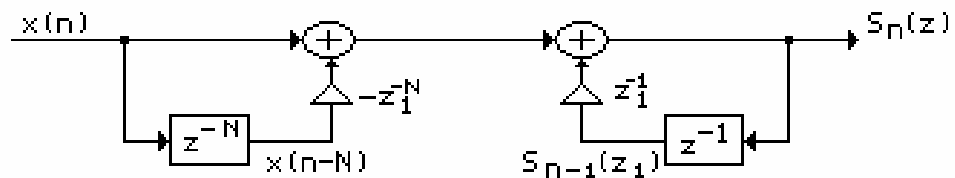


Рис. 1.6

Блоки, обозначенные на рис. 1.5 и рис. 1.6 как z^{-N} и z^{-1} , представляют собой элементы задержки, тогда как величины, равные степеням z_1 , представляют коэффициенты умножителей (в общем случае они являются комплексными). Входные сигналы и промежуточные результаты фильтра тоже могут быть комплексными.

Из (1.4) и рис. 1.6 следует, что для обеспечения скользящих спектральных измерений сигнала в одной точке $z = z_1$ достаточно выполнить всего два комплексных умножения на входной отсчет.

Чтобы найти спектр сразу во многих точках, например z_1, z_2, z_3 и т.д., необходимо повторять описанные выше вычисления для каждого нового значения z . На практике чаще всего представляет интерес поведение спектра на некоторой заданной довольно большой совокупности точек в z -плоскости. В этом случае главная задача состоит в том, чтобы выполнить измерение спектра, используя минимальное количество операций.

В подавляющем большинстве приложений задача измерения спектра сводится к нахождению значений z -преобразования конечной реализации сигнала для большого числа точек, равномерно распределенных по окружности единичного радиуса (рис. 1.4, в). Измерения такого типа соответствуют

вычислению ДПФ $S[k]$ конечной последовательности $x[n]$ и обычно наиболее эффективно выполняются с применением алгоритмов быстрого преобразования Фурье (БПФ)

$$S[k] = \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} nk} \dots$$

Точки z_k на единичной окружности расположены с интервалом $2\pi/N$.

Иногда желательно проводить измерения спектра, вычисляя значения z -преобразования последовательности $x[n]$ в равноотстоящих точках, расположенных внутри единичной окружности. Так, на рис. 1.4, z представлен случай, когда все точки равномерно распределены на окружности радиуса $r < 1$. Такое преобразование также можно получить с помощью ДПФ

$$S[re^{-j \frac{2\pi}{N} nk}] = \sum_{n=0}^{N-1} [x[n] r^{-n}] e^{-j \frac{2\pi}{N} nk}.$$

Это преобразование соответствует ДПФ последовательности $x[n] r^{-n}$, поэтому в данном случае спектральный анализ сводится к предварительному умножению массива сигнала $x[n]$ на r^{-n} и последующему выполнению ДПФ.

1.4. Взаимная связь спектров непрерывных и дискретных сигналов

Допустим, что имеется непрерывный сигнал $S(t)$ (рис. 1.7) в виде последовательности равноотстоящих импульсов, и этот сигнал задан на интервале $[0, T)$, на котором уместается ровно N импульсов, например, $N=4$. Эти импульсы можно пронумеровать. Приняв номер импульса за новую дискретную переменную n , мы можем утверждать, что тот же рис. 1.7 изображает дискретный сигнал $S(n)$, определенный на интервале $[0, N)$. Оба эти сигнала, непрерывный $S(t)$ и дискретный $S(n)$, будут иметь дискретные спектры $C_H(k)$ и $C_D(k)$ (рис. 1.8), отличающиеся друг от друга.

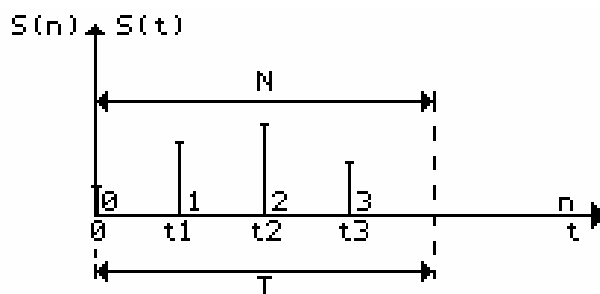


Рис. 1.7

Установим разницу между ними. Дискретная переменная на интервале $[0, N)$ принимает только фиксированные значения $n = 0, 1, 2, \dots$, а непрерывная переменная t на интервале $[0, T)$ принимает всевозможные значения. Таким образом, функция $S(n)$ задает сигнал только в четы-

рех точках, а функция $S(t)$ – на бесконечном множестве точек, как совпадающих с точками отсчетов, так и в паузах между ними.

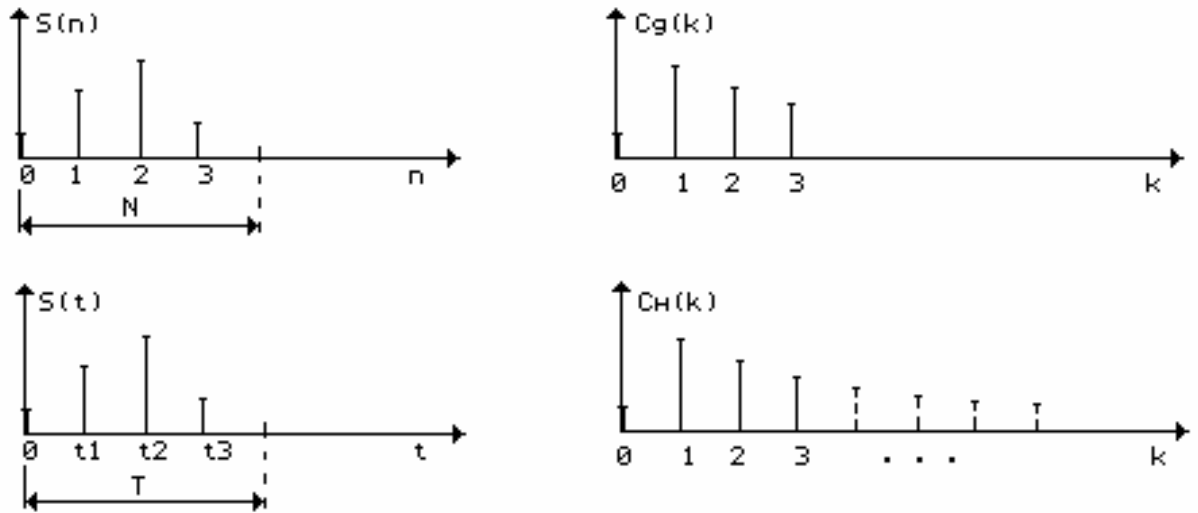


Рис. 1.8

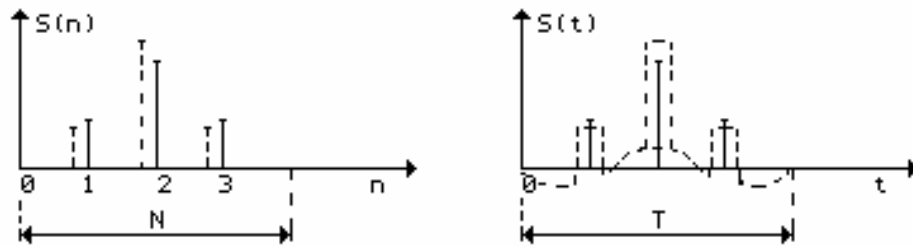


Рис. 1.9

Из этого различия вытекает важная особенность: разложение сигнала $S(n)$ можно проводить только по системе дискретных решетчатых функций $\{\eta(k,n)\}$, а разложение сигнала $S(t)$ – только по системе непрерывных функций $\{\eta(k,t)\}$. Спектры $C_D(k)$ и $C_H(k)$, полученные при таких разложениях (рис. 1.8), будут отличаться друг от друга тем, что спектр $C_D(k)$ будет содержать только 4 (по числу степеней свободы дискретного сигнала), а спектр $C_H(k)$ – бесконечное множество компонент, так как непрерывный сигнал $S(t)$ имеет бесконечную размерность. Однако, поскольку сигнал $S(t)$ задан на конечном интервале $[0,T)$, его спектр будет дискретным. В отличие от спектра $C_D(k)$, коэффициенты разложения $C_H(k)$ будут бесконечно малыми, так как сигнал $S(t)$ имеет бесконечно малую среднюю мощность (вследствие этого изображение спектра $C_H(k)$ на рис. 1.8 является условным).

Чтобы продемонстрировать различие спектров $C_D(k)$ и $C_H(k)$, проведем следующий мысленный эксперимент. Опустим из этих спектров составляющую с одним и тем же номером. Это вызовет искажения сигнала (рис. 1.9), но в случае дискретного сигнала эти искажения будут выражаться только в

изменении величин отсчетов (пунктирные линии), а в случае непрерывного сигнала помимо таких изменений в паузах между отсчетами еще появляются колебания.

Несмотря на указанное различие, спектры $C_D(k)$ и $C_H(k)$ принадлежат родственным сигналам и имеют некоторую общность, а именно: первые N коэффициентов разложения в этих спектрах совпадают с точностью до постоянного множителя.

Покажем, что это действительно так. Непрерывный сигнал $S(t)$ может быть записан в виде

$$S(t) = \sum_{n=0}^{N-1} S(t) S_0(t - t_n).$$

Перейдем к нормированному непрерывному времени $\tau = t / \Delta t$ и нормированному дискретному времени $n = [t / \Delta t]$, где Δt – интервал дискретизации, а $[b]$ обозначает целую часть числа b , и введем дельта-функцию:

$$S(\tau) = \sum_{n=0}^{N-1} S(\tau) S_0(\tau - n) = \lim_{\Delta\tau \rightarrow 0} S(\tau) \Delta\tau \sum_{n=0}^{N-1} \delta(\tau - n). \quad (1.5)$$

Спектр непрерывного сигнала $S(t)$ будет определяться выражением

$$C_H(k) = \frac{1}{P_{H,\eta}} \frac{\Delta t}{T} \int_0^{T/\Delta t} S(\tau) \eta(k, \tau) d\tau = \frac{1}{P_{H,\eta}} \frac{1}{N} \int_0^{T/\Delta t} S(\tau) \eta(k, \tau) d\tau.$$

Подставим сюда выражение (1.5) для сигнала $S(t)$ в виде ряда Фурье:

$$C_H(k) = \frac{1}{P_{H,\eta}} \frac{1}{N} \lim_{\Delta\tau \rightarrow 0} \sum_{n=0}^{N-1} [\Delta\tau \int_0^{T/\Delta t} S(\tau) \eta(k, \tau) \delta(\tau - n) d\tau].$$

Учитывая фильтрующее свойство дельта-функции, получим

$$C_H(k) = \lim_{\Delta\tau \rightarrow 0} \frac{\Delta\tau}{P_{H,\eta}} \frac{1}{N} \sum_{n=0}^{N-1} S(n) \eta(k, n). \quad (1.6)$$

Ранее было показано, что разложение дискретного сигнала имеет вид:

$$C_D(k) = \frac{1}{P_{D,\eta}} \frac{1}{N} \sum_{n=0}^{N-1} S(n) \eta(k, n). \quad (1.7)$$

Сравнив (1.6) и (1.7), замечаем, что при $k \in [0, N-1]$ спектры $C_H(k)$ и $C_D(k)$ совпадают с точностью до масштабного коэффициента. Из этого результата следует также, что у всех непрерывных огибающих данного дискретного сигнала с N отсчетами первые $N/2$ составляющих спектра совпадают с точностью до постоянного множителя.

Таким образом, спектр дискретного сигнала одновременно определяет ансамбль непрерывных сигналов, порождающих данный дискретный сигнал. Чтобы выбрать из этого ансамбля конкретный непрерывный сигнал, необходимо знать дополнительно составляющие спектра при $k \geq n$, которые отличаются спектр непрерывной огибающей от спектра самого дискретного сигнала.

При разложении дискретного сигнала в ряд Фурье по системе комплексных экспоненциальных функций $\{e^{jkn2\pi/N}\}$ спектр является периодическим с периодом N . Другими словами, периодический дискретный сигнал $S(n)$ и его спектр $C(k)$ обладают свойством

$$S(n) = S(n + aN); \quad C(k) = C(k + aN),$$

где N – период, а b – произвольное целое число.

Действительно, спектр периодического дискретного сигнала при разложении по системе функций $\{e^{jkn2\pi/N}\}$ может быть записан в виде

$$C(k) = \frac{1}{P_\eta} \frac{1}{N} \sum_{n=0}^{N-1} S(n) e^{-jkn2\pi/N}.$$

Отсюда имеем

$$C(k + aN) = \frac{1}{P_\eta} \frac{1}{N} \sum_{n=0}^{N-1} S(n) e^{-j(k+aN)n \frac{2\pi}{N}} = \frac{1}{P_\eta} \frac{1}{N} \sum_{n=0}^{N-1} S(n) e^{-jkn \frac{2\pi}{N}} e^{-jan2\pi} = C(k).$$

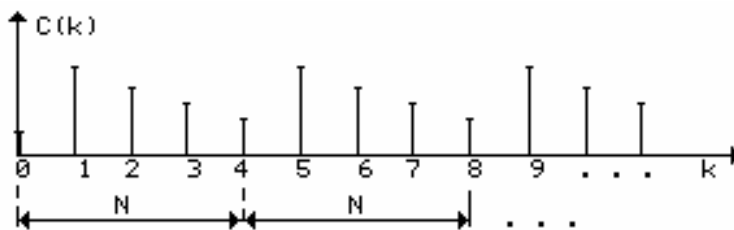


Рис. 1.10

При переходе от дискретного сигнала к непрерывному число отсчетов в одном периоде $N \rightarrow \infty$, поэтому по мере приближения дискретного сигнала к непрерывному период его спектра будет неограниченно возрастать.

В пределе, при $N \rightarrow \infty$ получим известную картину: периодический непрерывный сигнал имеет линейчатый, но непериодический спектр.

2. СПЕКТРАЛЬНЫЙ АНАЛИЗ И БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Набор алгоритмов, называемых *алгоритмами быстрого преобразования Фурье* (БПФ), включает разнообразные методы уменьшения времени вычисления дискретного преобразования Фурье (ДПФ). Поскольку вычисление ДПФ является основной операцией в большинстве задач спектрального анализа, то использование БПФ, позволяющее многократно ускорить вычисление ДПФ по сравнению с методом прямого вычисления ДПФ, имеет чрезвычайно важное значение и должно рассматриваться как неотъемлемая часть применения методов цифровой обработки сигналов для спектрального анализа.

2.1. Алгоритмы БПФ с основанием 2. Алгоритм БПФ с прореживанием по времени

Рассмотрим преобразование дискретного сигнала, определенного на $N = 2^i = 0, 2, 4, 8, 16, \dots$ точках. ДПФ конечной последовательности $\{x(n)\}$, $n \in [0, N-1]$ определяется выражением

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j \frac{2\pi}{N} nk}, \quad k \in [0, N-1], \quad (2.1)$$

или в более удобной форме

$$X(k) = \sum_{n=0}^{N-1} x(n) W^{nk}, \quad (2.2)$$

где $W = e^{-j \frac{2\pi}{N}}$. W^{nk} является периодической последовательностью с периодом N , т. е.

$$W^{(n+mN)(k+lN)} = W^{nk}, \quad m, l = 0, 1, \dots \quad (2.3)$$

Периодичность W^{nk} является одним из ключевых моментов БПФ. Часто периодичность W^{nk} подчеркивают тем, что вместо W записывают W_N .

Из соотношения (2.1) следует, что в случае, когда последовательность $x(n)$ является комплексной, при прямом вычислении N -точечного ДПФ нужно выполнить $(N-1)^2$ комплексных умножений и $N(N-1)$ комплексных сложений. Основная идея БПФ состоит в том, чтобы разбить исходную N -точечную последовательность на две более короткие последовательности, ДПФ которых могут быть скомбинированы таким образом, чтобы получился ДПФ исходной N -точечной последовательности.

Так, например, если N четное (что всегда имеет место при основании два), а исходная N -точечная последовательность разбита на две $(N/2)$ -точечные последовательности, то для вычисления искомого N -точечного ДПФ потребуется порядка $2(N/2)^2 = N^2/2$ комплексных умножений, т. е. вдвое меньше по сравнению с прямым вычислением. Здесь множитель $(N/2)^2$ дает число умножений, необходимое для прямого вычисления $(N/2)$ -точечного ДПФ, а множитель 2 соответствует двум ДПФ, которые должны быть вычислены. Эту операцию можно повторить, вычисляя вместо $(N/2)$ -точечного ДПФ два $(N/4)$ -точечных ДПФ и сокращая тем самым объем вычислений еще в два раза. Вы-

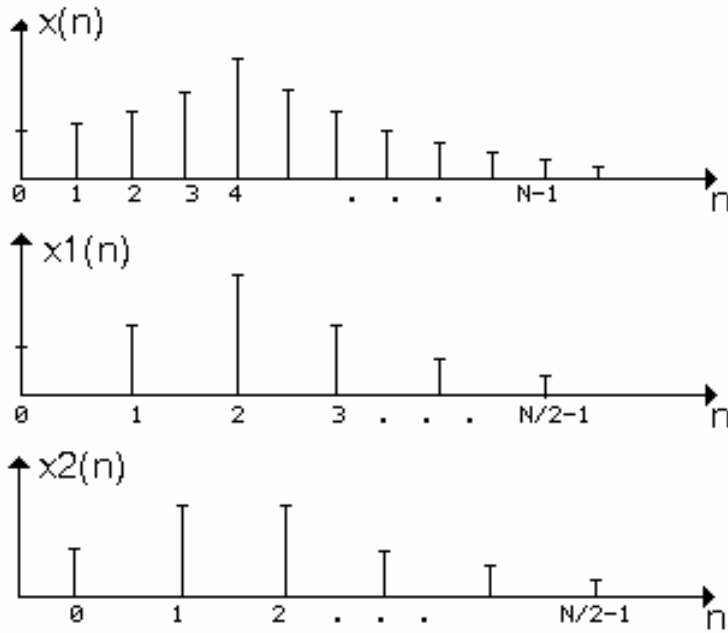


Рис. 2.1

игрыш в два раза является приближенным, поскольку не учитывается, каким образом из ДПФ меньшего размера образуется искомое N -точечное ДПФ.

Проиллюстрируем описанную методику для N -точечной последовательности $\{x(n)\}$, считая, что N равно степени 2. Введем, как показано на рис. 2.1, две $(N/2)$ -точечные последовательности $\{x_1(n)\}$ и $\{x_2(n)\}$ из четных и нечетных членов $x(n)$ соответственно, т. е.

$$x_1(n) = x(2n), \quad x_2(n) = x(2n+1), \quad n = 0, 1, \dots, \frac{N}{2} - 1. \quad (2.4)$$

N -точечное ДПФ последовательности $\{x(n)\}$ можно записать как

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n) W_N^{nk} = \sum_{n=0}^{N/2-1} x_1(n) W_N^{nk} + \sum_{n=0}^{N/2-1} x_2(n) W_N^{nk} = \\ &= \sum_{n=0}^{N/2-1} x(2n) W_N^{2nk} + \sum_{n=0}^{N/2-1} x(2n+1) W_N^{(2n+1)k}. \end{aligned} \quad (2.5)$$

С учетом того, что

$$W_N^2 = [e^{-j(2\pi/N)}]^2 = e^{-j[2\pi/(N/2)]} = W_{N/2}, \quad (2.6)$$

перепишем выражение (2.5) в виде

$$X(k) = \sum_{n=0}^{N/2-1} x_1(n) W_{N/2}^{nk} + W_N^k \sum_{n=0}^{N/2-1} x_2(n) W_{N/2}^{nk} \quad (2.7)$$

или

$$X(k) = X_1(k) + W_N^k X_2(k), \quad (2.8)$$

где $X_1(k)$ и $X_2(k)$ – $(N/2)$ -точечные ДПФ последовательностей $x_1(n)$ и $x_2(n)$ соответственно.

Поскольку $X(k)$ определено при $k \in [0, N-1]$, а $X_1(k)$ и $X_2(k)$ определены при $k \in [0, N/2-1]$, необходимо доопределить формулу (2.8) для $k \geq N/2$. Принимая во внимание свойство периодичности ДПФ, можем записать

$$X(k) = \begin{cases} X_1(k) + W_N^k X_2(k), & k \in [0, N/2-1], \\ X_1(k - N/2) + W_N^k X_2(k - N/2), & k \in [N/2, N-1]. \end{cases} \quad (2.9)$$

Заметим также, что $W_N^{k+N/2} = -W_N^k$, поэтому формулу (2.9) можно переписать в виде

$$X(k) = \begin{cases} X_1(k) + W_N^k X_2(k), & k \in [0, \frac{N}{2}-1], \\ X_1(k - \frac{N}{2}) - W_N^{k-N/2} X_2(k - \frac{N}{2}), & k \in [\frac{N}{2}, N-1]. \end{cases} \quad (2.10)$$

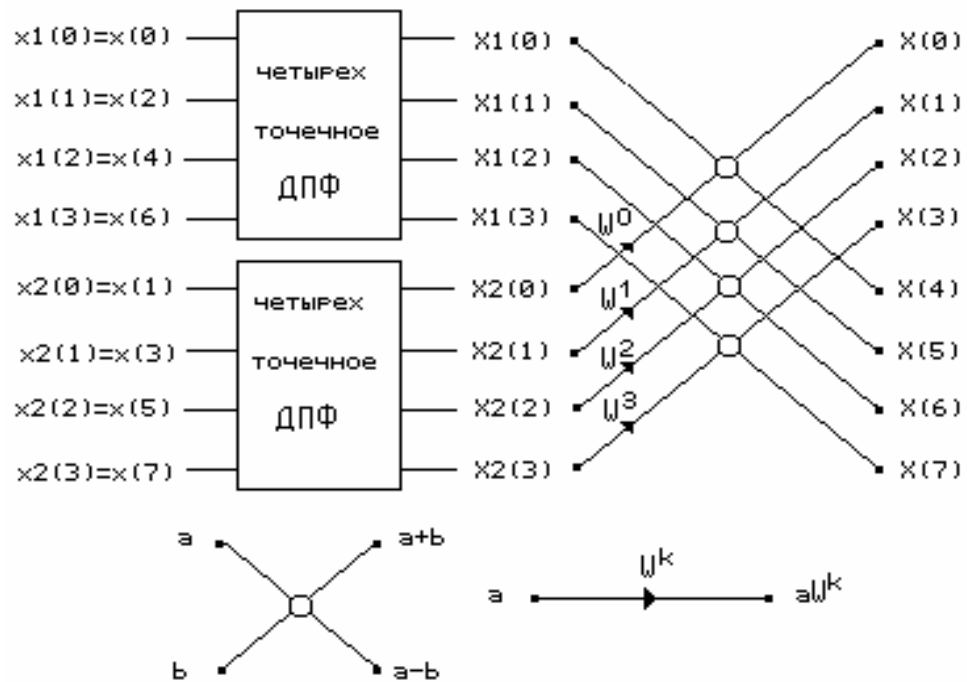


Рис. 2.2

На рис. 2.2 с помощью направленного графа представлена последовательность операций при выполнении восьмиточечного ДПФ с использованием двух четырехточечных преобразований. Входная последовательность $x(n)$ сначала разбивается на две последовательности $x_1(n)$ и $x_2(n)$ из четных и нечетных членов $x(n)$, после чего вычисляются их преобразования $X_1(k)$ и $X_2(k)$. Затем в соответствии с (2.10) получают $X(k)$.

Рассмотренная схема вычислений может быть использована для расчета $(N/2)$ -точечных ДПФ в соответствии с формулами (2.8) и (2.9). Каждая из последовательностей $x_1(n)$ и $x_2(n)$ разбивается на две последовательности, состоящие из четных и нечетных членов. Аналогично $(N/2)$ -точечные ДПФ могут быть записаны как комбинации двух $(N/4)$ -точечных ДПФ, т. е.

$$X_1(k) = A(k) + W_{N/2}^k B(k) \quad (2.11)$$

или

$$X_1(k) = A(k) + W_N^{2k} B(k), \quad k \in [0, N/2-], \quad (2.12)$$

где $A(k)$ и $B(k)$ – $N/2$ -точечные ДПФ соответственно четных и нечетных членов $x_1(n)$. На рис. 2.3 показан результирующий направленный граф, в котором четырехточечные ДПФ из рис. 2.2 рассчитываются согласно (2.12).

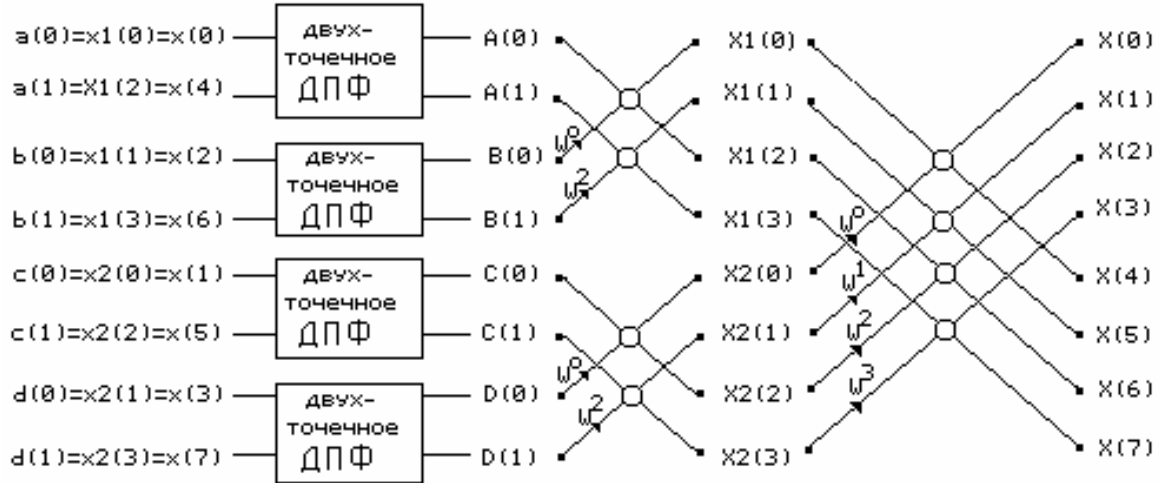


Рис. 2.3

Процесс уменьшения размера ДПФ от L до $L/2$, где L равно степени 2, может быть продолжен до тех пор, пока не останутся только двухточечные ДПФ. Двухточечное ДПФ $F(k)$, $k = 0, 1$ может быть рассчитано без использования умножений по формулам

$$\begin{aligned} F(0) &= f(0) + f(1)W_8^0, \\ F(1) &= f(0) + f(1)W_8^4. \end{aligned} \quad (2.13)$$

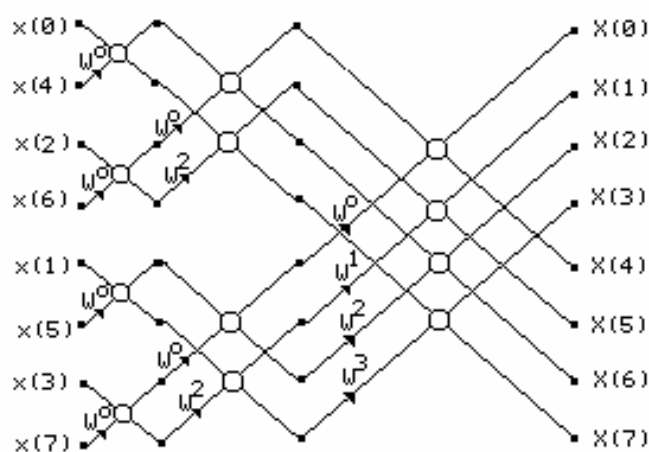


Рис. 2.4

Здесь $f(n)$, $n = 0, 1$ – преобразуемая двухточечная последовательность. Поскольку $W_8^0 = 1$ и $W_8^4 = -1$, для вычислений по формулам (2.13) действительно не нужны операции умножения. Таким образом, восьмиточечное ДПФ (рис. 2.1 и рис. 2.3) в итоге сводится к алгоритму, описываемому направленным графом, представленным на рис. 2.4.

2.1.1. Базовая операция. Алгоритм с замещением

Анализ графа на рис. 2.4 и процедуры последовательного сокращения вдвое размера преобразований показывает, что на каждом этапе БПФ (т. е. при каждом сокращении размера ДПФ) необходимо выполнить $N/2$ комплексных умножений. Поскольку общее количество этапов равно $\log_2 N$, то число комплексных умножений, необходимое для нахождения N -точечного ДПФ, приблизительно равно $(N/2)\log_2 N$. Слово "приблизительно" использовано по той причине, что умножения на W_N^0 , $W_N^{N/2}$, $W_N^{N/4}$ и $W_N^{3N/4}$ в действительности сводятся просто к сложениям и вычитаниям комплексных чисел. Как следует из направленного графа на рис. 2.4, вместо ожидаемых 12 (т. е. $4\log_2 8$) достаточно выполнить всего два нетривиальных умножения. Однако для большинства значений N фактическое число нетривиальных умножений хорошо аппроксимируется выражением $(N/2)\log_2 N$.

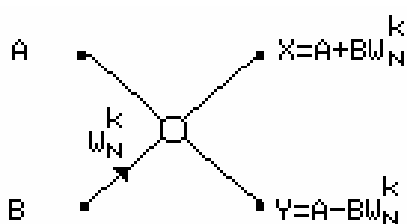


Рис. 2.5

Базовая операция алгоритма с прореживанием по времени (так называемая "бабочка") состоит в том, что два входных числа A и B объединяются для получения двух выходных чисел X и Y следующим образом:

$$X = A + W_N^k B, \quad Y = A - W_N^k B$$

На рис. 2.5 изображен направленный граф базовой операции.

Внимательное рассмотрение графа на рис. 2.4 показывает, что каждый из этапов содержит $N/2$ базовых операций. В случае, когда множитель W_N^k – нетривиальный, для каждой базовой операции необходимо выполнить только одно умножение, поскольку величину BW_N^k можно вычислить и запомнить.

Таким образом, структура базовых операций такова, что для выполнения БПФ N -точечной последовательности, размещенной в памяти, достаточно иметь лишь одну дополнительную ячейку памяти. Результаты всех промежуточных этапов БПФ можно размещать в те же ячейки памяти, где находятся исходные данные. Поэтому для хранения и входной и выходной последовательностей можно использовать один и тот же массив ячеек памяти. Алгоритм, в котором для размещения входной и выходной последовательностей используются одни и те же ячейки памяти, называется **алгоритмом БПФ с замещением**.

2.1.2. Перестановка данных и двоичная инверсия

Еще одной особенностью алгоритма с прореживанием по времени (как, впрочем, и большинства других алгоритмов БПФ) является необходимость такой перестановки элементов входной последовательности, чтобы выходная последовательность $X(k)$ имела естественный (прямой) порядок расположения, т.е. $k = 0, 1, \dots, N-1$. В примере на рис. 2.4 для этого требовался следующий порядок размещения входной последовательности: $x(0)$, $x(4)$, $x(2)$, $x(6)$, $x(1)$, $x(5)$, $x(3)$, $x(7)$. Характер перестановки элементов входной последовательности может быть описан сравнительно просто. В случае, когда N является степенью 2, входная последовательность должна быть расположена в памяти в двоично-инверсном порядке для того, чтобы выходная последовательность получилась в прямом порядке. Двоично-инверсный порядок определяется следующим образом. Если записать порядковые номера элементов входной последовательности в двоичном коде, используя L двоичных разрядов, причем $N = 2^L$, а затем инвертировать порядок следования разрядов, то получаемые при этом числа и будут номерами элементов входной последовательности после их перестановки. Так, для случая $N = 8 = 2^3$ прямой порядок номеров приведен в таблице 2.1 слева, а двоично-инверсный – справа. Таким образом, для двоичной инверсии входной последовательности необходим соответствующий алгоритм.

Таблица 2.1

Номер	Двоичное представление	Двоичная инверсия	Двоично-инверсный номер
0	000	000	0
1	001	100	4
2	010	010	2
3	011	110	6
4	100	001	1
5	101	101	5
6	110	011	3
7	111	111	7

Перестановку входной последовательности можно произвести с замещением, меняя в парах местами числа с прямыми и двоично-инверсными номерами и используя для этого лишь одну вспомогательную ячейку памяти.

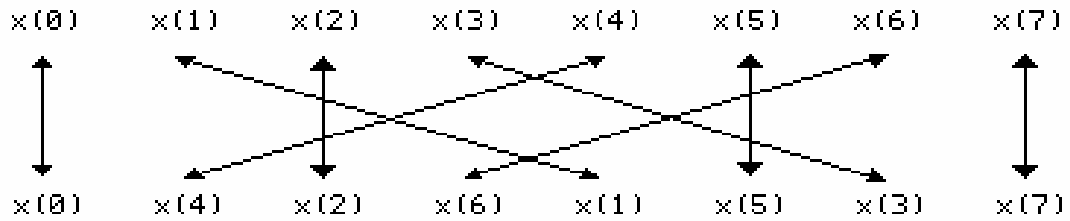


Рис. 2.6

На рис. 2.6 показана схема перестановки данных, представленных в табл. 2.1.

Отметим еще одну особенность алгоритма БПФ, заключающуюся в том, что на всех этапах преобразования используются коэффициенты W_N^k , $k = 0, 1, \dots, N-1$. Существует несколько способов получения этих коэффициентов. Простейший способ – составление таблицы, к которой можно обращаться в процессе счета. Единственный недостаток этого способа заключается в том, что для хранения этих коэффициентов необходима дополнительная память примерно из N ячеек.

Второй способ заключается в непосредственном вычислении коэффициентов $W_N^k = \cos[2\pi k / N] - j \sin[2\pi k / N]$ с использованием каждый раз стандартных подпрограмм расчета синуса и косинуса. Этот способ связан с большими затратами времени, поскольку вычисление этих функций, как правило, достаточно продолжительно.

Третий способ основан на применении простой рекуррентной формулы

$$W_N^k = (W_N^{k-L})W_N^L \quad (2.14)$$

с начальным условием $W_N^0 = 1$, так как степени W на каждом этапе БПФ меняются с постоянным шагом. Так, в примере на рис. 2.4 на первом этапе используются коэффициенты W^0 и W^4 , на втором – W^0 , W^2 , W^4 , W^6 , а на третьем – W^k , $k = 0, 1, 2, \dots, 7$. Поэтому, чтобы иметь возможность на каждом из этапов использовать формулу (2.14), достаточно запомнить или вычислить только множители W^4 , W^2 и W^0 .

2.2. Алгоритм БПФ с прореживанием по частоте

Другая распространенная форма алгоритма БПФ (при условии, что N равно степени 2) – так называемый алгоритм БПФ с прореживанием по частоте. В этом варианте алгоритма БПФ последовательность $\{x(n)\}$ разбивается на две последовательности, содержащие по $N/2$ отсчетов каждая, следующим образом: первая последовательность $\{x_1(n)\}$ состоит из первых $(N/2)$ отсчетов $\{x(n)\}$, а вторая $\{x_2(n)\}$ – из остальных $(N/2)$ отсчетов $\{x(n)\}$, т. е.

$$x_1(n) = x(n), \quad x_2(n) = x(n + N/2), \quad n = 0, 1, \dots, N/2 - 1.$$

Тогда N -точечное ДПФ последовательности $x(n)$ можно записать в виде

$$X(k) = \sum_{n=0}^{N/2-1} x(n)W_N^{nk} + \sum_{n=N/2}^{N-1} x(n)W_N^{nk} = \sum_{n=0}^{N/2-1} x_1(n)W_N^{nk} + \sum_{n=N/2}^{N-1} x_2(n)W_N^{(n+N/2)k}.$$

Учитывая, что $W^{kN/2} = e^{-j\pi k}$, получим

$$X(k) = \sum_{n=0}^{N/2-1} [x_1(n) + e^{-j\pi k} x_2(n)]W_N^{nk}.$$

Запишем выражения отдельно для четных и нечетных отсчетов ДПФ

$$\begin{aligned} X(2k) &= \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)](W_N^2)^{nk} = \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)]W_{N/2}^{nk}, \\ X(2k+1) &= \sum_{n=0}^{N/2-1} [x_1(n) - x_2(n)]W_N^{n(2k+1)} = \sum_{n=0}^{N/2-1} \{[x_1(n) + x_2(n)]W_N^n\}W_{N/2}^{nk}. \end{aligned}$$

Отсюда видно, что четные и нечетные отсчеты ДПФ можно получить из $(N/2)$ -точечных ДПФ последовательностей $f(n)$ и $g(n)$, равных

$$f(n) = x_1(n) + x_2(n), \quad g(n) = [x_1(n) - x_2(n)]W_N^n, \quad n = 0, 1, \dots, N/2 - 1.$$

Таким образом, снова вычисление N -точечного ДПФ удалось свести к вычислению двух $(N/2)$ -точечных ДПФ. На рис. 2.7 эта методика иллюстрируется для случая $N = 8$. Ее можно применить повторно и представить каждое из $(N/2)$ -точечных ДПФ в виде комбинаций двух $(N/4)$ -точечных ДПФ. На рис. 2.8 и 2.9 показан переход от четырехточечных ДПФ (рис. 2.7) к двухточечным ДПФ с последующим прямым вычислением двухточечных ДПФ.

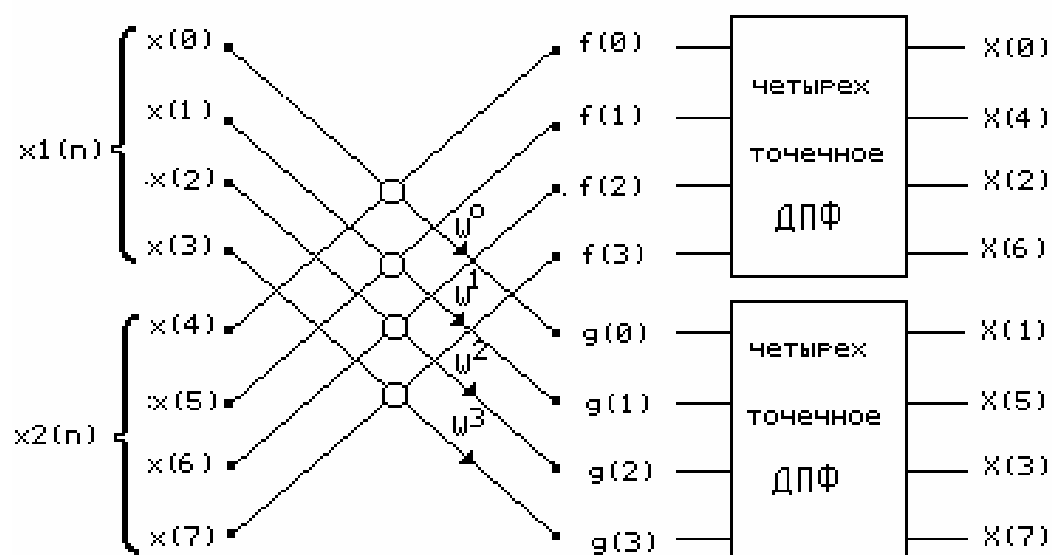


Рис. 2.7

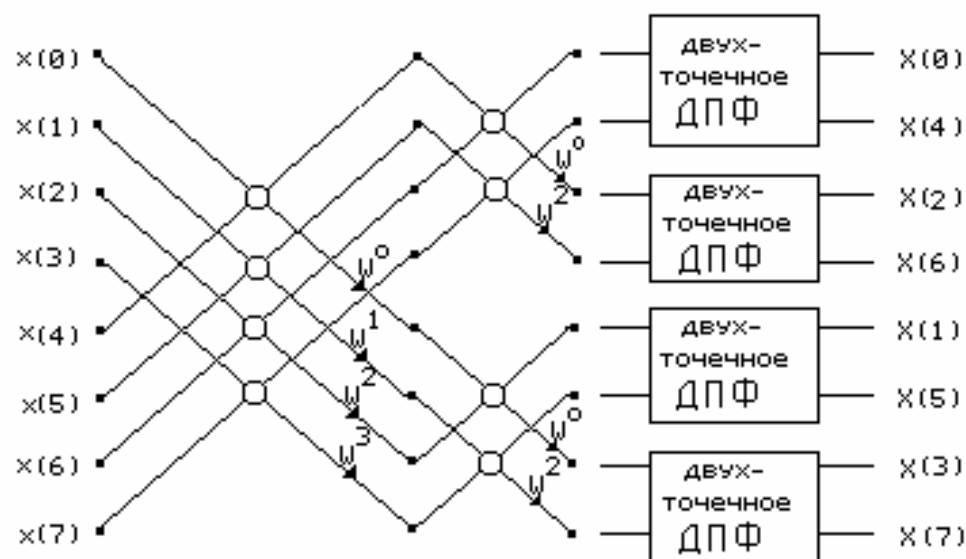


Рис. 2.8

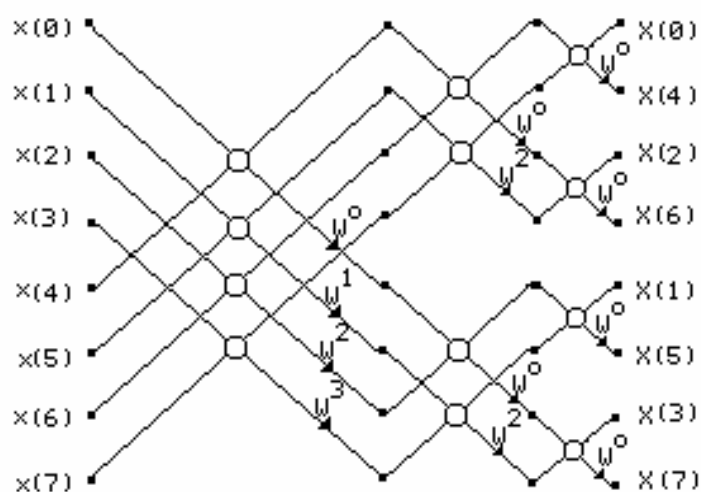


Рис. 2.9

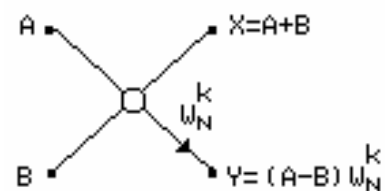


Рис. 2.10

Сравнение алгоритмов, иллюстрированных на рис. 2.4 и 2.9, позволяет выявить два очевидных различия между ними.

Во-первых, при прореживании по времени порядок следования входных отсчетов двоично-инверсный, а выходных – прямой, а при прореживании по частоте – наоборот (рис. 2.9). Однако это отличие кажущееся, поскольку в обоих алгоритмах порядок следования входных отсчетов может быть прямым, а выходных – двоично-инверсным, и наоборот. Второе отличие заключается в несколько ином выполнении базовой операции (см. рис. 2.5 и 2.10): при прореживании по частоте комплексное умножение выполняется после сложения (вычитания).

Легко заметить и сходство между алгоритмами с прореживанием по времени и по частоте. В обоих случаях при вычислении БПФ требуется около $N \log_2 N$ операций, вычисления могут быть проведены с замещением и должно быть предусмотрено выполнение двоичной инверсии.

2.3. Вычисление обратного ДПФ с помощью алгоритма прямого ДПФ

Для вычисления обратного ДПФ можно без каких-либо изменений использовать алгоритм БПФ. Обратное ДПФ N -точечной последовательности $\{X(k)\}$, $k = 0, 1, \dots, N-1$, определяется следующим образом:

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W^{-nk}. \quad (2.15)$$

Взяв выражение, комплексно-сопряженное с (2.15), и умножив его на N , получим

$$Nx^*(n) = \sum_{k=0}^{N-1} X^*(k) W^{nk}. \quad (2.16)$$

Правая часть формулы (2.16) представляет собой ДПФ последовательности $\{X^*(k)\}$ и может быть вычислена с использованием одного из описанных алгоритмов БПФ. Искомую последовательность $\{x(n)\}$ можно получить, взяв комплексно-сопряженное с (2.16) выражение и разделив его на N , т. е.

$$x(n) = \frac{1}{N} \left(\sum_{k=0}^{N-1} X^*(k) W^{nk} \right)^*.$$

Таким образом, алгоритм БПФ обеспечивает вычисление и прямого и обратного ДПФ.

2.4. Единый подход к алгоритмам БПФ

Существует много различных алгоритмов БПФ. Однако оказывается, что все они могут быть получены с помощью последовательного применения единственной операции, а именно представления одномерного массива чисел двумерным.

При вычислении N -точечного ДПФ N -точечной последовательности целое N может быть либо простым, либо составным числом (до сих пор считалось, что N состоит из большого числа сомножителей и равно степени 2). Если N простое, его нельзя разложить на произведения меньших целых чисел. В этом случае одномерный сигнал $x(0), x(1), \dots, x(N-1)$ невозможно представить в виде двумерного массива, поэтому для такого сигнала не существует алгоритма БПФ.

В большинстве практических задач вполне допустимо искусственное удлинение обрабатываемой последовательности путем добавления нулей, приводящее к тому, что результирующий спектр представляет собой некоторую интерполяцию спектра неудлиненной последовательности. Пусть, например, N равно 60. Это число можно представить как произведение меньших чисел различным образом: $60 = 3 \cdot 4 \cdot 5 = 4 \cdot 3 \cdot 5 = 5 \cdot 4 \cdot 3 = 12 \cdot 5 = 2 \cdot 2 \cdot 3 \cdot 5$, и т. д. В зависимости от порядка следования сомножителей и их общего количества могут быть получены различные формы алгоритма БПФ. Для характеристики разложения обычно используют понятие "основание". Понятие "смешанное основание" означает, что не все сомножители N одинаковы. Для $N = 60$ все формы алгоритма БПФ имеют смешанные основания. Если N можно представить в виде произведения одинаковых сомножителей r , то соответствующий алгоритм называют алгоритмом БПФ с основанием r . Например, если $N = 64 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2$, то получаются рассмотренные ранее алгоритмы БПФ с основанием 2. Если же N записать как $64 = 8 \cdot 8$, то получится алгоритм БПФ с основанием 8.

Очень важно, что разложение числа на множители можно выполнить различными способами. Следовательно, для выявления общих закономерностей следует провести тщательный анализ свойств разложения очередного числа на два сомножителя. Возьмем в качестве примера снова $N = 60$ и запишем одно из возможных разложений обрабатываемого массива в виде матрицы из (5×12) номеров отсчетов сигнала:

0	1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31	32	33	34	35
36	37	38	39	40	41	42	43	44	45	46	47
48	49	50	51	52	53	54	55	56	57	58	59

Далее, поскольку столбцы содержат по пять (т. е. простое число) отсчетов, они больше не могут быть разложены. Однако строки, состоящие из 12

отсчетов, можно представить в виде матриц размером (3×4) . Например, первая строка будет иметь вид:

0	1	2	3
4	5	6	7
8	9	10	11

Остальные строки можно представить аналогично. Итак, теперь нужно установить, каким образом, оперируя с двумерным массивом, можно получить ДПФ исходного одномерного массива.

Для получения основного результата будем считать, что входные отсчеты пронумерованы по строкам и столбцам, поэтому их номера могут быть представлены следующими парами чисел:

0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,10	0,11
1,0	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	1,10	1,11
2,0	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9	2,10	2,11
3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	3,9	3,10	3,11
4,0	4,1	4,2	4,3	4,4	4,5	4,6	4,7	4,8	4,9	4,10	4,11

Далее, пусть текущий номер столбца равен m ($m = 0, 1, \dots, 11$), а l ($l = 0, 1, \dots, 4$) – текущий номер строки. Если исходный номер отсчета обозначить через n ($n = 0, 1, \dots, 59$), то

$$n = Ml + m, \quad (2.17)$$

где M – число столбцов, а L – число строк (в данном примере $M = 12$, $L = 5$).

Допустим, что мы можем найти ДПФ двумерного массива с двойными номерами, тогда результат должен иметь вид двумерного массива с двойными номерами. Пусть m и l – переменные исходного сигнала, а r и s – переменные двумерного ДПФ по столбцам и строкам соответственно, которые преобразуются в одну переменную следующим образом

$$k = Lr + s. \quad (2.18)$$

Теперь коэффициенты одномерного ДПФ $X(k) = X(s, r)$ можно выразить через преобразование массива $x(n) = x(l, m)$, используя простую подстановку формул (2.17) и (2.18) в выражение для ДПФ (2.2), что дает

$$X(k) = X(s, r) = \sum_{m=0}^{M-1} \sum_{l=0}^{L-1} x(l, m) W^{(Ml+m)(Lr+s)}. \quad (2.19)$$

Разлагая $W^{(Ml+m)(Lr+s)}$ с учетом того, что $W^{MLrl} = W^{NrL} = 1$, и располагая со-

ответствующие переменные под знаком суммирования, преобразуем формулу (2.19) к следующему виду:

$$X(s, r) = \sum_{m=0}^{M-1} W^{Lmr} W^{ms} \underbrace{\sum_{l=0}^{L-1} x(l, m) W^{Msl}}_{g(s, m)}. \quad (2.20)$$

Эта формула содержит все необходимые сведения, позволяющие связать преобразование одномерного массива с преобразованием того же массива, представленного в виде двумерной матрицы.

Заметим прежде всего, что внутренняя сумма представляет собой ДПФ m -го столбца исходного массива с ядром преобразования W^M . Таким образом, можно сформулировать первый шаг последовательности расчета $X(k)$:

1) Вычислить L -точечные ДПФ всех столбцов. Результат является функцией s и m , причем m меняется от 0 до $M-1$. Обозначим его через $g(s, m)$ и перепишем формулу (2.20):

$$X(s, r) = \sum_{m=0}^{M-1} W^{Lmr} W^{ms} g(s, m). \quad (2.21)$$

Отсюда следует, что второй шаг вычисления $X(k)$ сводится к следующему.

2) Найти новый массив $h(s, m)$, умножая каждый элемент $g(s, m)$ на поворачивающий множитель W^{ms} . Теперь формула (2.21) примет вид

$$X(s, r) = \sum_{m=0}^{M-1} h(s, m) W^{Lmr}.$$

Она представляет M -точечное ДПФ каждой из строк с номерами s . Поэтому последний шаг алгоритма заключается в следующем.

3) Вычислить M -точечные ДПФ всех строк матрицы $h(s, m)$ с ядром преобразования W^L .

Описанная методика напоминает вычисление двумерного ДПФ, когда сначала вычисляются ДПФ строк, а затем столбцов, но шаг 2 отсутствует. Разделимость ядра преобразования с более высокой размерностью и является причиной того, что при расчете ДПФ с более высокой размерностью требуется меньше операций, чем при расчете одномерного ДПФ при одинаковом общем числе отсчетов.

Важно отметить, что после введения поворачивающих множителей, т.е. после второго шага описанной выше методики, способы расчета двумерного ДПФ и одномерного ДПФ массива, представленного в виде двумерной матрицы, становятся эквивалентными, причем при каждом таком представлении

для выполнения шага 2 требуется дополнительно до N умножений. Изменение порядка суммирования в формуле (2.19) дает

$$X(s, r) = \underbrace{\sum_{l=0}^{L-1} W^{Msl}}_{\text{шаг 3}} \underbrace{\sum_{m=0}^{M-1} W^{Lmr}}_{\text{шаг 2}} \underbrace{x(l, m) W^{ms}}_{\text{шаг 1}}.$$

Так что порядок вычисления $X(k)$ становится следующим.

- 1) Умножить отсчеты сигнала $x(l, m)$ на поворачивающие множители W^{ms} .
- 2) Вычислить M -точечные ДПФ всех строк.
- 3) Вычислить L -точечные ДПФ всех столбцов.

В этом случае умножение на поворачивающие множители предшествует вычислению ДПФ строк, тогда как раньше они выполнялись после вычисления ДПФ столбцов. Эти отличия от рассмотренного ранее порядка вычисления $X(k)$ не только напоминают, но и действительно связаны с обсуждавшимся выше различием между алгоритмами ДПФ с основанием 2 при прореживании по времени и по частоте.

Отметим еще одно важное свойство методики преобразования, вытекающее из формулы (2.19), в которой переменные m и r являются номерами столбцов, а l и s – номерами строк. При увеличении m на единицу номер отсчета исходного массива ($Ml + m$) также возрастает на единицу, тогда как при увеличении на единицу номера столбца преобразованного массива r аргумент $X(s, r)$ возрастает на L . Это означает, что в результате преобразования номера строк и столбцов меняются местами.

2.5. Особенности цифровой реализации алгоритма БПФ

Вычисление быстрого преобразования Фурье является одним из ключевых моментов спектрального анализа, поэтому эффектам конечной разрядности данного алгоритма должно быть уделено особое внимание. К таким эффектам относятся шумы округления при усечении результатов умножений, погрешности масштабирования промежуточных результатов с целью устранения возможных переполнений, а также погрешности представления коэффициентов W^k (поворотных множителей). Уэлч и ряд других авторов [2] подробно проанализировали влияние шума округления и масштабирования для алгоритмов БПФ по основанию 2 с прореживанием по времени и по частоте при условии, что все арифметические операции выполняются с фиксированной запятой. Вопрос, связанный с точностью представления коэффициентов, не рассматривался столь подробно, поэтому будут приведены лишь качественные результаты.

Рассмотрим реализацию алгоритма БПФ с основанием 2 при выполнении арифметических операций с фиксированной запятой.

Если последовательность $\{x(n)\}$ из N отсчетов имеет коэффициенты ДПФ $\{X(k)\}$, то, согласно теореме Парсеваля,

$$\sum_{n=0}^{N-1} x^2(n) = \frac{1}{N} \sum_{k=0}^{N-1} X^2(k), \quad (2.22)$$

т. е. средняя мощность выходных гармоник в N раз превышает среднюю мощность исходной последовательности. Следовательно, значения ДПФ последовательности будут, вообще говоря, существенно превышать значения самой последовательности. При теоретическом рассмотрении этот факт не играет особой роли, но как только дело доходит до практической реализации на устройстве, имеющем целочисленную арифметику, то тут уже приходится заботиться о том, чтобы числа не выходили за пределы разрядной сетки, т. е. чтобы не было переполнения. Проиллюстрируем возникновение переполнений на примере базовой операции ("бабочки") на m -ом этапе БПФ с прореживанием по времени:

$$\begin{aligned} f_{m+1}(i) &= f_m(i) + W^P f_m(j), \\ f_{m+1}(j) &= f_m(i) - W^P f_m(j). \end{aligned} \quad (2.23)$$

Здесь $f_m(i)$ и $f_m(j)$ – входные числа базовой операции, $f_{m+1}(i)$ и $f_{m+1}(j)$ – выходные числа, W^P – поворотный множитель. Очевидно, что при переходе от этапа к этапу модули чисел, вообще говоря, увеличиваются и удовлетворяют неравенству

$$\max\{f_m(i), f_m(j)\} \leq \max\{f_{m+1}(i), f_{m+1}(j)\} \leq 2 \max\{f_m(i), f_m(j)\}. \quad (2.24)$$

Следовательно, уровень сигнала на каждом этапе увеличивается не более чем в 2 раза, т. е. не быстрее, чем на один двоичный разряд. Единственный способ бороться с этим – масштабировать числа (т. е. умножать их на какое-либо число, меньшее единицы).

Учитывая вышеизложенное, можно предложить три метода масштабирования:

1) Сдвиг вправо на один разряд на каждом этапе (это эквивалентно делению чисел на два). Если модуль $f_0(i)$ не превышает 0.5 для любых i и числа сдвигаются вправо на один разряд после каждого этапа, кроме последнего, то переполнения не будет.

2) Контроль последовательности, гарантирующий выполнение условия $f_m(i) \leq 0,5$ для всех i . На каждом этапе БПФ вычисляется массив $f_m(i)$ и, если хотя бы один отсчет массива превысит по модулю 0.5, весь массив мас-

штабируется сдвигом вправо на один разряд.

3) Проверка на переполнение. Исходная последовательность масштабируется так, чтобы модули действительной и мнимой частей $f_0(i)$ были меньше 1 (а не 0.5, как в предыдущих методах). Если в ходе выполнения базовой операции фиксируется переполнение, то все числа последовательно, включая результаты уже выполненных на данном этапе (без переполнения) базовых операций, сдвигаются на один разряд вправо, после чего итерация продолжается с той базовой операции, где произошло переполнение. На каждом этапе может произойти более одного переполнения, но не более двух.

Первый метод связан с наименьшими затратами времени и наиболее прост для практической реализации. Однако он дает наименьшую точность, так как масштабирование на каждом этапе приводит к ненужному ухудшению точности в случаях, когда его можно не производить. Второй метод требует больших затрат времени (поскольку на каждом этапе приходится вычислять модули всех чисел массива), и в то же время он не слишком точный, так как модули всех элементов массива не превышают 0.5 и, как следствие, один разряд памяти данных не используется. Третий метод является самым точным, но и самым медленным, так как приходится повторно обрабатывать всю последовательность всякий раз, когда обнаруживается переполнение.

Уэлч разработал довольно простую методику расчета граничных значений уровня шума округления, возникающего при БПФ. Наличие шума округления обусловлено двумя факторами:

1) появлением погрешности при округлении произведения $W^P f_m(j)$. Если W^P и $f_m(j)$ представлены b -разрядными числами (т.е. действительная и мнимая части W^P и $f_m(j)$ содержат по b разрядов, то после округления каждого из четырех действительных произведений, входящих в $W^P f_m(j)$, до b -разрядного числа получается ошибка, равномерно распределенная на интервале $[-2^{-b/2}, 2^{-b/2}]$ и имеющая нулевое среднее и дисперсию $2^{-2b}/12$;

2) сдвигом суммы на один разряд вправо, если при сложении происходит переполнение. Если младший разряд, при сдвиге выходящий за пределы регистра, равен нулю, то ошибки не будет. Если же он равен единице, то в зависимости от знака числа возникает ошибка величиной $\pm 2^{-b}$. Дисперсия этой ошибки равна $2^{-2b}/2$.

Верхнюю границу отношения дисперсии шума округления к дисперсии выходных гармоник можно получить, предположив, что переполнение происходит на каждом этапе, так что числа каждый раз приходится масштабировать. Если $f_k(j)$ – отсчеты, получаемые на k -м этапе, то $D(f_k)$ – их дисперсия, определяемая как

$$D(f_k) = \frac{1}{N} \sum_{j=0}^{N-1} D\{f_k(j)\}, \quad (2.25)$$

причем

$$D(f_0) = \frac{2^{-2b}}{2} = 6\Delta^2, \quad (2.26)$$

где $\Delta^2 = 2^{-2b} / 12$. Равенство (2.26) получается в предположении, что на первом же этапе может произойти переполнение, поэтому в числах нулевого этапа (т. е. в исходных отсчетах) один младший разряд отбрасывается. Базовая операция первого этапа содержит умножения только на ± 1 , поэтому шум округления произведений отсутствует.

На втором этапе, согласно предположению, также может произойти переполнение, поэтому массив $f_1(j)$ нужно предварительно промасштабировать сдвигом вправо каждого элемента на один разряд. Тогда дисперсия $D(f_1)$ будет равна

$$D(f_1) = 2D(f_0) + (4 \times 6\Delta^2) = 36\Delta^2, \quad (2.27)$$

причем множитель 4 отражает тот факт, что ошибка при усечении выходных чисел на первом этапе будет вдвое больше ошибки, обусловленной усечением выходных чисел на нулевом этапе, поэтому дисперсия возрастает в четыре раза.

На втором этапе БПФ с прореживанием по времени выполняются умножения только на ± 1 и $\pm j$, поэтому ошибка, связанная с округлением произведений, опять будет равна нулю. Итак, в предположении, что числа $f_2(j)$ также масштабируются, получим следующую формулу для дисперсии ошибки на втором этапе БПФ:

$$D(f_2) = 2D(f_1) + (4^2 \times 6\Delta^2) = 168\Delta^2, \quad (2.28)$$

причем множитель 4^2 указывает на удвоение ошибки усечения при переходе от первого этапа ко второму.

На последующих этапах в большинстве базовых операций выполняются нетривиальные умножения. Так, на третьем этапе половина базовых операций содержит такие умножения, на четвертом этапе – три четверти, и т. д. Согласно формулам (2.23), общие выражения для действительной и мнимой частей результата базовой операции на m -ом этапе имеют вид:

$$\begin{aligned} \operatorname{Re}(f_{m+1}(i)) &= \operatorname{Re}(f_m(i)) + \operatorname{Re}(f_m(j))\operatorname{Re}(W^P) - \operatorname{Im}(f_m(j))\operatorname{Im}(W^P), \\ \operatorname{Im}(f_{m+1}(i)) &= \operatorname{Im}(f_m(i)) + \operatorname{Re}(f_m(j))\operatorname{Im}(W^P) + \operatorname{Im}(f_m(j))\operatorname{Re}(W^P). \end{aligned} \quad (2.29)$$

Для тех базовых операций, где умножение на W^P не является тривиальным, дисперсия результата f_{m+1} (включая шум округления произведений и шум

усечения при масштабировании) имеет вид:

$$D(f_{m+1}) = D(f_m) + D(W^P) \left\{ \text{Re}^2(f_m(j)) + \text{Im}^2(f_m(j)) \right\} + \\ + D(f_m) \left\{ \text{Re}^2(W^P) + \text{Im}^2(W^P) \right\} + D(e_{m+1}^{ml}) + D(e_{m+1}^s), \quad (2.32)$$

где $D(e_{m+1}^{ml})$ – дисперсия шума округления умножения на m -ом этапе, а $D(e_{m+1}^s)$ – дисперсия шума масштабирования на том же этапе.

$$D(e_{m+1}^{ml}) = 4^m \Delta^2, \quad D(e_{m+1}^s) = 4^m \times 6\Delta^2. \quad (2.31)$$

Третье слагаемое формулы (2.30) содержит произведение квадрата модуля W^P и $D(f_m)$. Поскольку $W^P = 1$, это слагаемое совпадает с первым членом суммы. Второе слагаемое равно произведению $D(W^P)$ на среднее значение квадрата модуля переменных, преобразуемых на m -ом этапе. Дисперсия W^P равна Δ^2 (поскольку W^P представляется b -разрядным числом с погрешностью). Чтобы определить среднее квадрата модуля переменных, обрабатываемых на m -м этапе, нужно знать величину K , представляющую собой среднее квадрата модуля исходного массива чисел, равное

$$K = f_0^2(j) = \frac{1}{N} \sum_{j=0}^{N-1} f_0^2(j). \quad (2.32)$$

Учитывая, что среднее значение квадрата модуля увеличивается от этапа к этапу вдвое, получим, что на выходе m -го этапа оно будет равно $2^m K$. Тогда $D(f_{m+1})$ становится равной

$$D(f_{m+1}) = 2D(f_m) + 2^m K \Delta^2 + 4^{m+1} \Delta^2 + (4^{m+1} \times 6\Delta^2), \quad (2.33)$$

если во всех базовых операциях m -го этапа выполняются нетривиальные умножения. Если же ввести a – долю базовых операций с нетривиальными умножениями ($0 \leq a \leq 1$), то второе и третье слагаемые суммы (2.33) следует умножить на a . Так, если $a = 0$ (как это имеет место на первом и втором этапах), то равенство (2.33) переходит в (2.27) (при $m = 0$) или в (2.28) (при $m = 1$). Если принять, что $a = 0.5$ на третьем этапе и $a = 1$ на всех последую-

щих этапах, причем номер последнего этапа равен M , то получим

$$\begin{aligned}
 D(f_m) &= 2^M (6\Delta^2) + 2^{M-1} (4 \times 6\Delta^2) + \dots + 2 (4^{M+1} \times 6\Delta^2) + 2^{M-2} K\Delta^2 + \\
 &+ (M-3)2^{M-1} K\Delta^2 + 2^{M-4} (4^2 \Delta^2) + \dots + (4^M \Delta^2) = \\
 &= (1.5)2^{M+2} \Delta^2 (1+2+\dots+2^{M-1}) + (M-2.5)2^{M-1} K\Delta^2 + 2^{M+4} (1+2+\dots+2^{M-4}) \approx \\
 &\approx 2^{2M+3} \Delta^2 + (M-2.5)2^{M-1} K\Delta^2 + 2^{M+2} \Delta^2.
 \end{aligned} \tag{2.34}$$

Поскольку среднее значение квадратов модулей элементов выходного массива $\{f_m(j)\}$ равно $2^M K$, то среднее квадрата их действительных (или мнимых) частей равно $2^M K/2$. Таким образом, оценкой отношения дисперсии шума к дисперсии сигнала на выходе (в пределе при больших M) является величина

$$\frac{\sigma_n}{\sigma_s} \approx \frac{2^{(M+3)/2} \Delta}{\sqrt{K/2}} = \frac{2^{(M+3)/2} 2^{-b} (0.3)}{\sigma_s} = \frac{0.3N2^{-b}8}{\sigma_s}. \tag{2.35}$$

Итак, верхняя граница отношения ошибки к выходному сигналу возрастает как \sqrt{N} , т. е. по $1/2$ разряда на этап.

Формулу, аналогичную (2.35), можно получить и для нижней границы отношения ошибок шума и сигнала, если предположить, что переполнений вообще не происходит, т. е. ошибки, связанные с масштабированием, отсутствуют. Следовательно, вплоть до третьего этапа эта ошибка вообще не возникает. Дисперсия результата f_M будет иметь вид:

$$\begin{aligned}
 D(f_M) &= 2^{M-2} K\Delta^2 + (M-3)2^{M-1} K\Delta^2 + 2^{M-3} \Delta^2 + 2^{M-5} \Delta^2 + 2^{M-6} \Delta^2 + \dots + \Delta^2 \approx \\
 &\approx (M-2.5)2^{M-1} K\Delta^2 + 2^{M-3} \Delta^2 + 2^{M-4} \Delta^2.
 \end{aligned} \tag{2.36}$$

Таким образом, при большом числе этапов M нижняя граница отношения дисперсий ошибки и сигнала на выходе имеет вид

$$\frac{\sigma_n}{\sigma_s} \approx 0.3\sqrt{M-2.5}2^{-b}. \tag{2.37}$$

Рассмотренная модель возникновения ошибок была экспериментально подтверждена Уэлчем на многочисленных примерах. Результаты его экспериментов приведены в [2].

3. СООТНОШЕНИЯ МЕЖДУ БПФ И ЦИФРОВОЙ ФИЛЬТРАЦИЕЙ

В подавляющем большинстве приложений задача измерения спектра сводится к нахождению значений z -преобразования конечной реализации сигнала для большого числа точек, равномерно распределенных по окружности единичного радиуса (рис. 1.4, в).

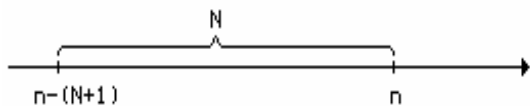


Рис. 3.1

Пусть анализу подвергается последовательность $x(n)$, положение аргумента которой на временной оси показано на рис. 3.1. Тогда спектральное измерение сводится к вычислению ДПФ последовательности $x(n)$ и обычно наиболее эффективно выполняется с применением алгоритмов БПФ. С другой стороны, как было показано в главе 1, задача может быть решена с помощью цифровой фильтрации. Там же были приведены схемы, реализующие процедуру скользящего спектрального измерения (спектрального измерения в одной точке z -плоскости). В нашем случае придется повторять ее для каждой из точек z_1, z_2 и т. д. на единичной окружности, в которых оценивается спектр сигнала.

Иногда желательно проводить измерения спектра, вычисляя значения z -преобразования последовательности в равноотстоящих точках, расположенных на окружности радиуса $r < 1$. Задача сводится к предыдущей с помощью предварительного умножения массива сигнала на r^{-n} .

3.1. Некоторые характеристики спектрального анализа на основе БПФ

Двумя наиболее важными характеристиками спектрального анализа являются:

- количество частот, на которых желательно измерить спектр;
- "разрешающая способность" измерения спектра.

Для анализа обеих характеристик лучше всего использовать отмеченную выше эквивалентность между спектральным измерением и фильтрацией.

Чтобы показать, какие факторы влияют на параметры спектрального анализа, рассмотрим пример, представленный на рис. 1.4, з, когда необходимо найти спектр сигнала в 16 точках, равномерно распределенных по единичной окружности. Пусть число L отсчетов сигнала, используемых при измерении спектра, равно 16. Спектральный анализ, удовлетворяющий этому условию, может быть выполнен двумя эквивалентными способами: либо с помощью 16-точечного БПФ, либо с помощью гребенки из 16 фильтров. Импульсная характеристика k -го фильтра, обеспечивающего измерения спектра в точке $z_k = e^{j(2\pi/N)k}$, равна

$$h(n) = e^{-j(2\pi/N)kn}, \quad 0 \leq n \leq N-1, \quad (3.1)$$

так что ее z -преобразование имеет вид

$$H(z) = \sum_{n=0}^{N-1} e^{-j(2\pi/N)kn} z^{-n} = \frac{1 - z^{-N}}{1 - z^{-1} e^{-j(2\pi/N)k}}. \quad (3.2)$$

Вычислив значения $H(z)$ на единичной окружности, получим

$$H(e^{j\omega}) = e^{-j\omega \frac{N-1}{2}} e^{j\frac{\pi k}{N}} \frac{\sin(N\omega/2)}{\sin(\frac{\omega}{2} + \frac{\pi k}{N})} = e^{-j\omega \frac{N-1}{2}} e^{j\frac{\pi k}{N}} f_N(\omega, k). \quad (3.3)$$

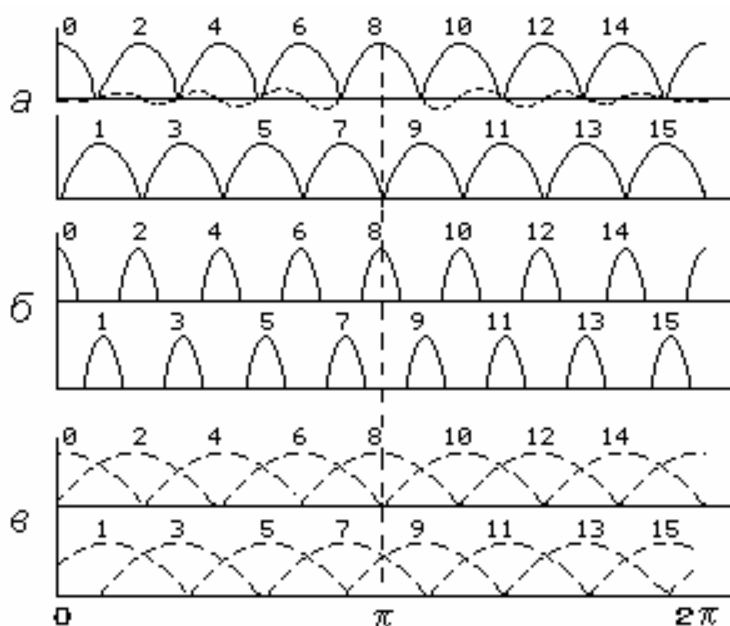


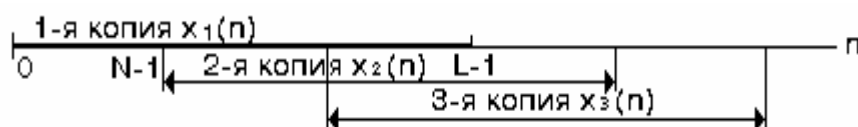
Рис. 3.2

На рис. 3.2, а изображены графики функций $f_N(\omega, k)$ для случая, соответствующего рис. 1.4, з, т. е. для 16 точек, равномерно распределенных на единичной окружности. Частотные характеристики фильтров с четными номерами k показаны на рис. 3.2, а сверху, а с нечетными значениями k — внизу. Чтобы избежать путаницы, для всех фильтров, за исключением 8-го, изображены лишь главные лепестки.

Видно, что скользящее БПФ эквивалентно довольно грубому набору фильтров с относительно большими боковыми лепестками и существенным перекрытием между соседними фильтрами.

Предположим теперь, что число отсчетов сигнала L больше числа спектральных отсчетов N . Пусть, например, $L = 32$ и $N = 16$. Самый простой способ выполнения таких спектральных измерений состоит в том, что вычисляется 32-точечное скользящее БПФ и просто отбрасываются измерения, соответствующие фильтрам, следующим через один. В результате получится набор фильтров с характеристиками, изображенными на рис. 3.2, б.

Пусть теперь $N = 16$, но имеется только $L = 8$ отсчетов сигнала. В этом случае спектр можно найти, добавив к L отсчетам сигнала $N-L$ нулевых отсчетов так, чтобы общее число преобразуемых отсчетов равнялось количеству спектральных отсчетов, и затем вычислив L -точечное БПФ. Результирующие частотные характеристики фильтров для случая $L = 8$, $N = 16$ приведены на рис. 3.2, в. Число каналов эквивалентной гребенки фильтров осталось прежним, но каждый фильтр стал шире. Таким образом, в рассматриваемом



примере главное изменение состоит в ухудшении разрешающей способно-

сти по частоте.

Метод выполнения анализа при $L = 2N$ становится очевидным, если обратиться к направленному графу БПФ (см., например, рис. 2.9). Видно, что выходные отсчеты с четными номерами располагаются в верхней половине графа. Это означает, что в алгоритме БПФ, предназначенном для получения только четных отсчетов спектра, достаточно лишь частично обработать все L отсчетов, чтобы получить верхнюю половину выходных отсчетов первого этапа БПФ. Как видно из рис. 2.9, только эта половина отсчетов дает все четные коэффициенты ДПФ. Этот подход можно развивать дальше, отметив, например, что восемь верхних выходных коэффициентов дают каждый четвертый коэффициент ДПФ, причем их можно определить, выполнив половину операций на первом этапе, половину – на втором, а также четырехточечное БПФ полученных на втором этапе восьми отсчетов. Пусть в общем случае число отсчетов сигнала L равно MN , где N – требуемое число спектральных отсчетов, а M – целое число, большее единицы. Искомое преобразование равно

$$X(k) = \sum_{n=0}^{N-1} \left[\sum_{m=0}^{M-1} x(n + mN) \right] e^{-j(2\pi/N)nk}, \quad k = 0, 1, \dots, N-1. \quad (3.4)$$

Из этой формулы следует, что фактически при вычислении спектра производится предварительное суммирование входных отсчетов, равноотстоящих друг от друга на N точек, после чего вычисляется ДПФ образованной новой последовательности. Смысл этой операции становится понятным, если вспомнить, что дискретизация в частотной области (ее и нужно было получить в данном случае) соответствует наложению во временной области, которое описывается формулой (3.4).

Обратное ДПФ от $X(k)$ даст не исходную последовательность $x(n)$, а набор ее сдвинутых на N отсчетов относительно друг друга копий $x_i(n)$, $n = 0, \dots, N-1$. Пусть $L/N = 3$. Из рисунка видно, что неискаженным будет лишь участок $n = 0, \dots, N-1$. Кроме того, общее число отсчетов увеличилось. Очевидно, что, чем ближе значения N к L (M ближе к единице), тем меньше будут искажения, связанные с наложением результатов преобразования.

Из вышеприведенных рассуждений следует, что использование последовательностей данных длиной L , большей числа спектральных отсчетов N , улучшает разрешение в N/L раз, но вносит большие искажения и, следовательно, неприменимо в случаях, когда полученная спектральная оценка затем обрабатывается при помощи ОДПФ.

3.2. Соотношение между "скачущим" БПФ и гребенкой фильтров

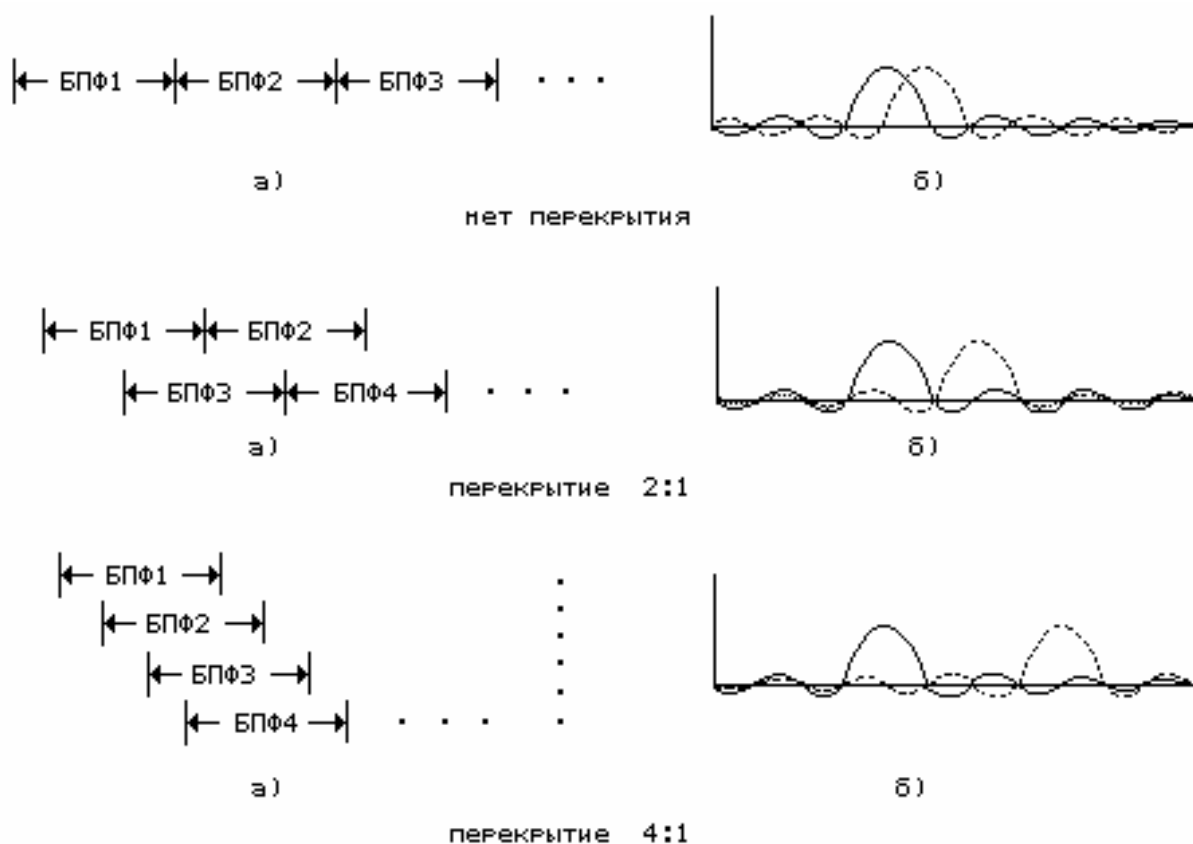


Рис. 3.3

Выше было показано, что скользящее БПФ в точности эквивалентно анализу с использованием гребенки фильтров, поэтому возникает вопрос, какой из этих двух методов эффективнее с точки зрения объема вычислений.

При скользящем БПФ для каждого входного отсчета приходится вычислять полное БПФ, т. е. $(N/2)\log_2 N$ базовых операций на отсчет. При анализе с помощью гребенки фильтров, состоящей из гребенчатого фильтра, к выходу которого параллельно подключена группа цифровых резонаторов (рис. 3.3), число комплексных умножений на отсчет равно N . Таким образом, в большинстве случаев, представляющих интерес, гребенка фильтров оказывается эффективнее скользящего БПФ, поскольку в последнем случае не используются возможности итеративного расчета ДПФ. Однако ниже будет показано, что при определенных условиях алгоритм БПФ может оказаться эффективнее эквивалентной ему гребенки фильтров.

Для стандартного метода, позволяющего использовать БПФ для приближенного вычисления текущего спектра, характерно то, что интервал вычисления БПФ не скользит, а скачет. Несколько вариантов *скачущего БПФ* приведено на рис. 3.4, а, где представлены три из возможных способов размещения интервалов вычисления БПФ: с шагом, равным размеру БПФ, поло-

вине и четверти размера БПФ.

Вычислительная эффективность скачущего БПФ растет пропорционально величине "скачка", однако важно представить, как такой подход влияет на характер спектральных измерений. Отметим прежде всего, что скачущие измерения являются всего лишь отсчетами скользящих измерений. Следовательно, эквивалентный результат можно было бы получить, отбра-

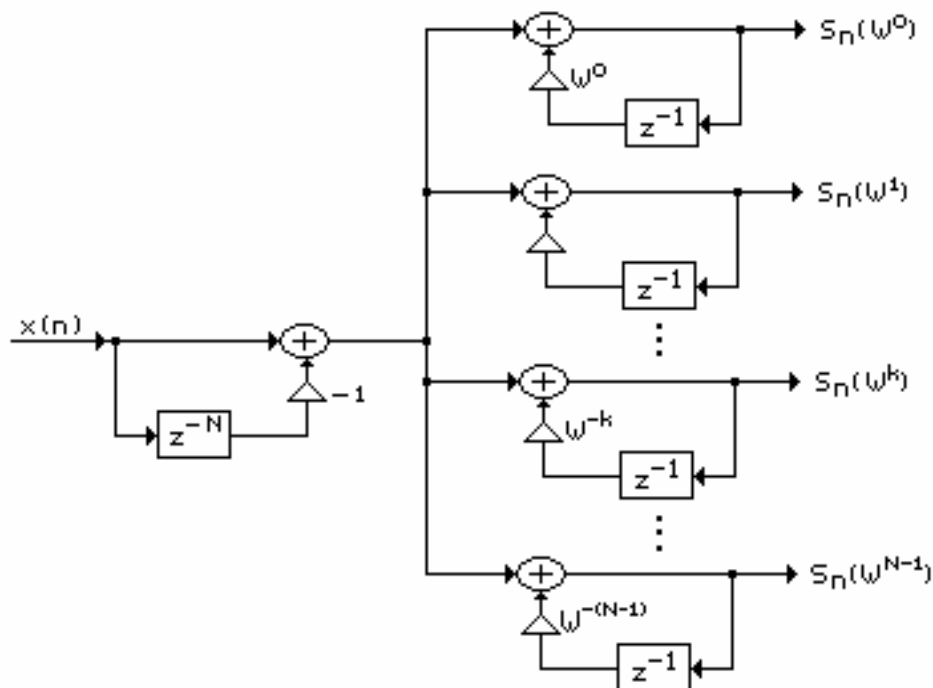


Рис. 3.4

сывая часть отсчетов на выходах гребенки фильтров.

Эффекты такого прореживания можно проанализировать, используя стандартную трактовку наложения частот, вспомнив, что наложение во временной области соответствует дискретизации в частотной. На рис. 3.4, б приближенно пред-

ставлены эффекты наложения частотных характеристик эквивалентных фильтров для трех случаев скачущего БПФ, приведенных на рис. 3.4, а. При отсутствии перекрытия интервалов вычисления БПФ наблюдается значительное наложение частотных характеристик, что может привести к серьезным искажениям спектральных измерений. При перекрытии 2:1 и особенно 4:1 эффект наложения значительно ослабляется, хотя и остается довольно заметным. Таким образом, при увеличении перекрытия интервалов анализа частотные характеристики эквивалентных фильтров анализатора оказываются более разнесенными по частоте.

На практике для борьбы с наложениями используется метод взвешивания, который позволяет уменьшить уровень боковых лепестков и тем самым ослабить эффекты наложения. В качестве приближенной оценки этого эффекта можно использовать уровень наложенного спектра (он изображен пунктирной линией) в пределах главного лепестка фильтра.

4. ИСПОЛЬЗОВАНИЕ ОКОН В СПЕКТРАЛЬНОМ АНАЛИЗЕ

Один из важных вопросов, который является общим для всех классических методов спектрального оценивания, связан с применением функции окна. Обработка с помощью окна (*windowing*) используется для управления эффектами, обусловленными наличием боковых лепестков в спектральных оценках. Заметим, что термин "взвешивание" (*weighting*) используется как синоним термина "обработка с помощью окна". Это контрастирует с практикой тех специалистов, которые употребляют термин "окно" только применительно к преобразованиям, связанным с применением весовой функции во временной области. Введение функции окна позволяет управлять формой частотных характеристик фильтров анализатора как в полосе пропускания, так и, что особенно важно, в полосе подавления. Функция $W(f)$ является частотным окном в том случае, когда она получена в результате дискретно-временного преобразования Фурье (ДВПФ) окна данных $w(n)$. Окна данных называются также обуживающими функциями (*tapering functions*), т. е. плавно спадающими к краям. Функция $\Omega(f)$ обозначает спектральное окно в том случае, когда она является ДВПФ корреляционного окна $\omega(m)$, применяемого к дискретно-временной автокорреляционной последовательности. Основное назначение окна данных – уменьшить величину смещения в периодограммных спектральных оценках. Основное назначение корреляционного окна – уменьшить дисперсию коррелограммной оценки СПМ.

4.1. Прямоугольное окно и БПФ

Заметим, что имеющуюся конечную запись данных или имеющуюся конечную корреляционную последовательность удобно рассматривать как некоторую часть соответствующей бесконечной последовательности, видимую через применяемое окно. Например, последовательность наблюдаемых данных $x_0(n)$ из N отсчетов математически можно записать как произведение прямоугольной функции единичной амплитуды

$$rect(n) = \begin{cases} 1, & 0 \leq n \leq N-1; \\ 0, & \text{при других } n \end{cases} \quad (4.1)$$

и бесконечной последовательности $x(n)$:

$$x_0(n) = x(n) \cdot rect(n). \quad (4.2)$$

При этом принимается очевидное допущение о том, что все ненаблюдаемые отсчеты равны нулю независимо от того, так это на самом деле или нет. Дискретно-временное преобразование Фурье взвешенной окном последователь-

ности, выраженное через преобразования последовательности $x(n)$ и прямоугольного окна $\text{rect}(n)$, равно свертке этих преобразований

$$X_0(f) = X(f) * D_N(f), \quad (4.3)$$

где

$$D_N(f) = T \exp(-j2\pi fT(N-1)) \frac{\sin(\pi fTN)}{\sin(\pi fT)}. \quad (4.4)$$

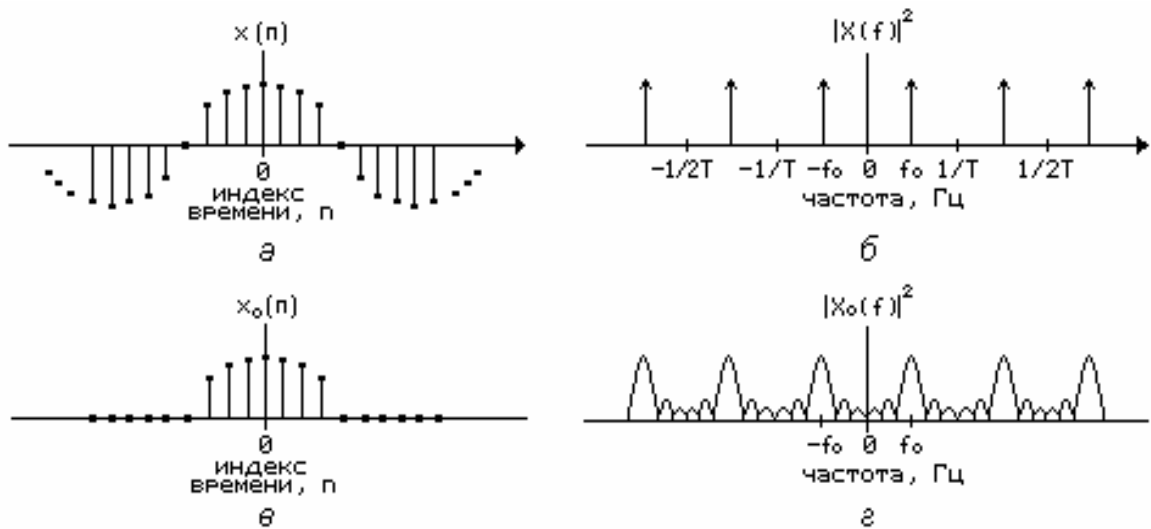


Рис. 4.1

Функция $D_N(f)$, называемая дискретной функцией sinc, или ядром Дирихле, представляет собой ДВПФ прямоугольной функции. Преобразование наблюдаемой конечной последовательности является искаженной версией преобразования бесконечной последовательности. Влияние прямоугольного окна на дискретно-временную косинусоиду с частотой f_0 иллюстрирует рис. 4.1. Здесь: *a* – исходная дискретно-временная косинусоидальная последовательность; *б* – модуль периодической ДВПФ косинусоидальной последовательности; *в* – взвешенная косинусоидальная последовательность; *г* – модуль ДВПФ взвешенной последовательности. Из рис. 4.1 видно, что острые спектральные пики ДВПФ бесконечной косинусоидальной последовательности расширились за счет воздействия копий преобразования окна. Таким образом, минимальная ширина спектральных пиков взвешенной окном последовательности ограничена шириной, определяемой главным лепестком преобразования этого окна, и не зависит от исходных данных. Боковые лепестки преобразования окна, иногда называемые просачиванием, будут изменять амплитуды соседних спектральных пиков. Поскольку ДВПФ – периодическая функция, то наложение боковых лепестков от соседних спектральных периодов может привести к дополнительному смещению. Увеличение частоты отсчетов позволяет ослабить эффект наложения боковых лепестков. Аналогичные искажения будут наблюдаться и в случае несинусоидальных сиг-

налов.

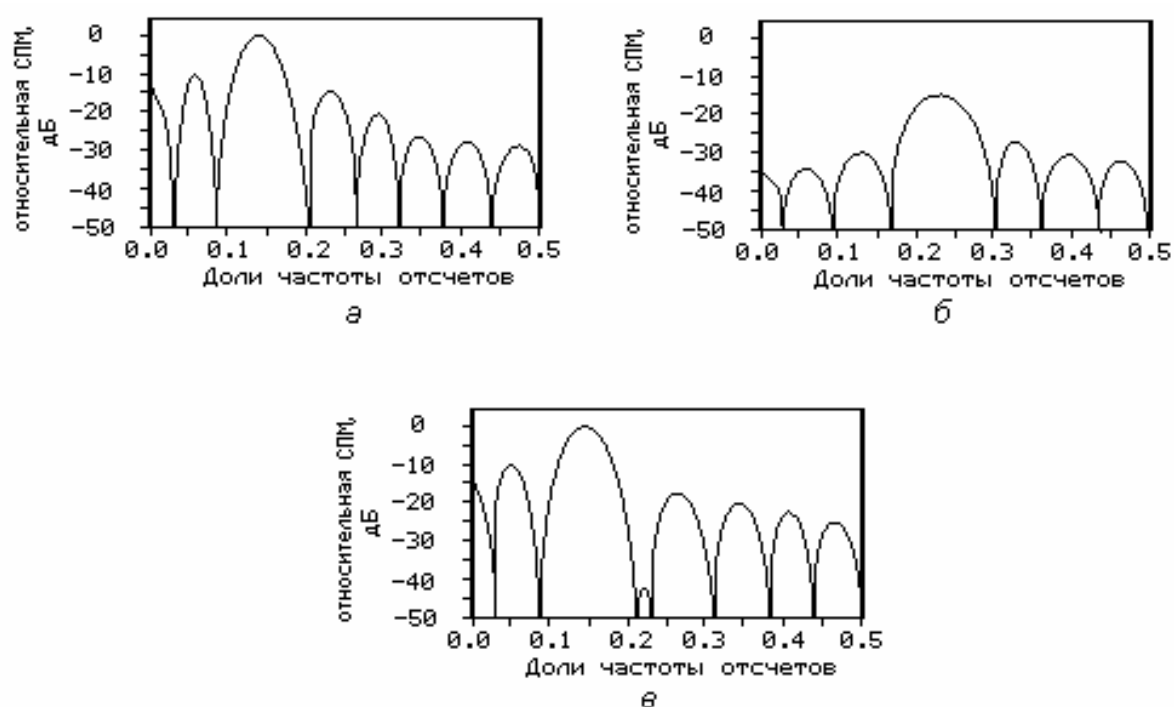


Рис. 4.2

Просачивание приводит не только к появлению амплитудных ошибок в спектрах дискретных сигналов, но может также маскировать присутствие слабых сигналов и, следовательно, препятствовать их обнаружению. Рассмотрим ДВПФ 16 отсчетов сигнала, состоящего из двух синусоид (рис. 4.2, в). Мощность более слабой синусоиды (рис. 4.2, б) в этом сигнале примерно на 17 дБ меньше мощности более сильной синусоиды (рис. 4.2, а). В этом экстремальном случае боковые лепестки более сильной синусоиды почти полностью подавляют главный лепесток более слабой синусоиды.

4.2. Наиболее распространенные функции окна и их характеристики

Можно предложить ряд других функций окна, применение которых позволяет снизить уровень боковых лепестков по сравнению с тем их уровнем, который они имеют в случае прямоугольного окна. Снижение уровня боковых лепестков будет уменьшать смещение. Однако это дается ценой расширения главного лепестка спектра окна, что, естественно, приводит к ухудшению разрешения. Следовательно, должен выбираться какой-то компромисс между шириной главного лепестка и уровнем подавления боковых лепестков.

Для классификации функций окна используется несколько показателей оценки их качества.

Ширина полосы частот главного лепестка позволяет судить о частотном разрешении. Для количественной оценки ширины полосы главного лепестка используются два показателя.

Традиционным показателем является **ширина полосы на уровне половинной мощности**, т. е. на уровне, который на 3 дБ ниже максимума главного лепестка.

Второй показатель – **эквивалентная ширина полосы B_e дискретно-временного преобразования Фурье $X(f)$ сигнала $x(n)$** – определяется как

$$B_e = \frac{\int_{-T/2}^{T/2} X(f) df}{X(0)}. \quad (4.5)$$

Два показателя используются и для оценки характеристик боковых лепестков.

Один из них – это **пиковый (или максимальный) уровень боковых лепестков**, который позволяет судить о том, насколько хорошо окно подавляет просачивание.

Второй – это **скорость спадания уровня боковых лепестков**, который характеризует скорость, с которой снижается уровень боковых лепестков, ближайших к главному лепестку. По сути дела, скорость спадания уровня боковых лепестков зависит от числа используемых отсчетов N и с увеличением N стремится к некоторой асимптотической величине, которую принято выражать в децибелах на октаву изменения ширины полосы частот.

В табл. 4.1 даны определения некоторых наиболее употребительных дискретно-временных функций окна из числа предложенных в разное время для использования при спектральном оценивании.

На рис. 4.3 – 4.8 показаны типичные 51-точечные окна и их частотные характеристики, полученные посредством вычисления модуля ДВПФ каждого окна.

В табл. 4.1 определены N -точечные симметричные окна с двумя различными начальными точками отсчета. Для окон данных начальной точкой отсчета является $n = 0$. Для корреляционных окон этой точкой является значение $-(N-1)/2$. В качестве корреляционных окон используются только окна нечетной длины, поскольку точка симметрии таких окон приходится на средний элемент окна. В случае корреляционного окна значение этого среднего (или центрального) элемента всегда равно единице.

Таблица 4.1

Название окна	Дискретно-временная функция $w(n)$ или $\omega(n)$	Частотная характеристика $W(f)$ или $\Omega(f)$
Прямоугольное (равномерное)	$1(n)$	$D_N(f)$
Треугольное (окно Бартлетта)	$1 - 2 t(n) $	$(2/N)D_N^2(f/2)$
Косинус-квадрат (окно Ханна)	$\cos^2(\pi t(n)) = 0.50 + 0.50\cos(2\pi t(n))$	$0.50D_N(f) + 0.25[D_N(f - \frac{1}{NT}) + D_N(f + \frac{1}{NT})]$
Приподнятый косинус (окно Хэмминга)	$0.54 + 0.46\cos(2\pi t(n))$	$0.54D_N(f) + 0.23[D_N(f - \frac{1}{NT}) + D_N(f + \frac{1}{NT})]$
Взвешенные косинусы (окно Наттола, R=3)	$\sum_{r=0}^R a_r \cos(2\pi r t(n))$	$\sum_{r=0}^R 0.5a_r [D_N(f - \frac{r}{NT}) + D_N(f + \frac{r}{NT})]$
Усеченное гауссовское ($\alpha=2,5$)	$\exp[-\frac{1}{2}(2\alpha t(n))^2]$	$(\sqrt{2\pi}/2\alpha)\exp[-\frac{1}{2}(2\pi f T / \alpha)^2] \cdot D_N(f)$

Здесь T – интервал отсчетов. Окно данных $w(n)$ определено в диапазоне индексов времени $0 \leq n \leq N-1$, где N может быть четным или нечетным. Корреляционное окно $\omega(m)$ определено в диапазоне индексов временного сдвига $-(N-1)/2 \leq m \leq (N-1)/2$, где N всегда нечетно. Функция $t(n) = (n - [N-1]/2)/[N-1]$ для $w(n)$ и $t(m) = m/N$ для $\omega(m)$. Функция $D_N(f) = T \operatorname{sinc}(fNT)$ для $w(n)$ и $D_N(f) = T \exp(j2\pi f T) \operatorname{sinc}(fNT)$ для $\omega(m)$.

У окон четной длины точка симметрии находится в центре окна, точно посередине между двумя элементами. В литературе часто используются и несколько отличные определения функций окна. Например, некоторые авторы предпочитают опускать концевые точки окна с нулевыми значениями. Для операций с дискретно-временными рядами Фурье, не требующих дополнения нулевыми отсчетами (нулями), точка, соответствующая наибольшему временному индексу, обычно опускается, с тем чтобы обеспечить корректность периодического продолжения окна. Характеристики окон, описанных в табл. 4.1, приведены в табл. 4.2. Величины, приведенные в колонке "эквивалентная ширина полосы", нормированы относительно частотного разрешения ДВПФ, равного $1/NT$ Гц.

Из всех приведенных в табл. 4.2 окон самый узкий главный лепесток имеет частотная характеристика прямоугольного окна, но зато у него и самый высокий уровень боковых лепестков. Автокорреляционной функцией прямоугольного окна является треугольное окно, ширина которого равна удвоенной ширине прямоугольного окна. Треугольное окно впервые было описано Бартлеттом в связи с проводимым им анализом периодограммных

спектральных оценок, и поэтому часто называется его именем. Окно типа "косинус квадрат" названо в честь австрийского метеоролога Юлиуса фон Ханна. Это окно часто ошибочно называют окном Хэннинга. Оно примечательно тем, что его легко реализовать в частотной области всего лишь с помощью трех операций сложения и двух операций сдвига, что, по сути дела, сводится к умножению на коэффициенты $1/2$ и $1/4$ на каждой частоте.

Таблица 4.2

Характеристики окон, приведенных в таблице 4.1

Окно	Максимальный уровень боковых лепестков, дБ	Асимптотическая скорость спада боковых лепестков, дБ/октава	Эквивалентная ширина полосы	Ширина полосы по уровню половинной мощности
Прямоугольное	-13,3	-6	1,00	0,89
Треугольное	-26,5	-12	1,33	1,28
Окно Ханна	-31,5	-18	1,50	1,44
Окно Хэмминга	-43	-6	1,36	1,30
Окно Наттола ($r=3$)	-98	-6	1,80	1,70
Гауссовское	-42	-6	1,39	1,33
Равноволновое	-50	0	1,39	1,33

Окно типа приподнятой косинусоиды было введено Р. У. Хэммингом и часто называется его именем. Множители 0.54 и 0.46 были выбраны для того, чтобы практически полностью устранить максимальный боковой лепесток функции $D_N(f)$.

Этот подход можно обобщить на все семейство взвешивающих косинусных окон. Веса косинусов могут быть оптимизированы относительно ряда условий, включая требование минимального уровня боковых лепестков, максимальной скорости их спада и максимальной гладкости (наибольшее число производных без нарушения непрерывности). Например, окно Наттола, показанное на рис. 4.7, является четырехчленным косинусным окном, которое при $a_0 = 0.3635819$, $a_1 = 0.4891775$, $a_2 = 0.1365995$ и $a_3 = 0.0106411$ имеет минимальный уровень боковых лепестков, равный -98 дБ.

Гауссовское окно имеет наименьшую величину произведения длительности на ширину полосы из всех приведенных функций окна, хотя это не справедливо в случае усеченного (имеющего конечную ширину) гауссовского окна. В окне, показанном на рис. 4.8, использован параметр $a = 2.5$.

У чебышевского (или равноволнового) окна все боковые лепестки имеют один уровень. Это окно впервые было использовано в теории антенных решеток и обладает тем свойством, что из всех N -точечных дискретных окон с уровнем боковых лепестков, равным или не превосходящим некоторый заданный уровень, имеет самый узкий главный лепесток.

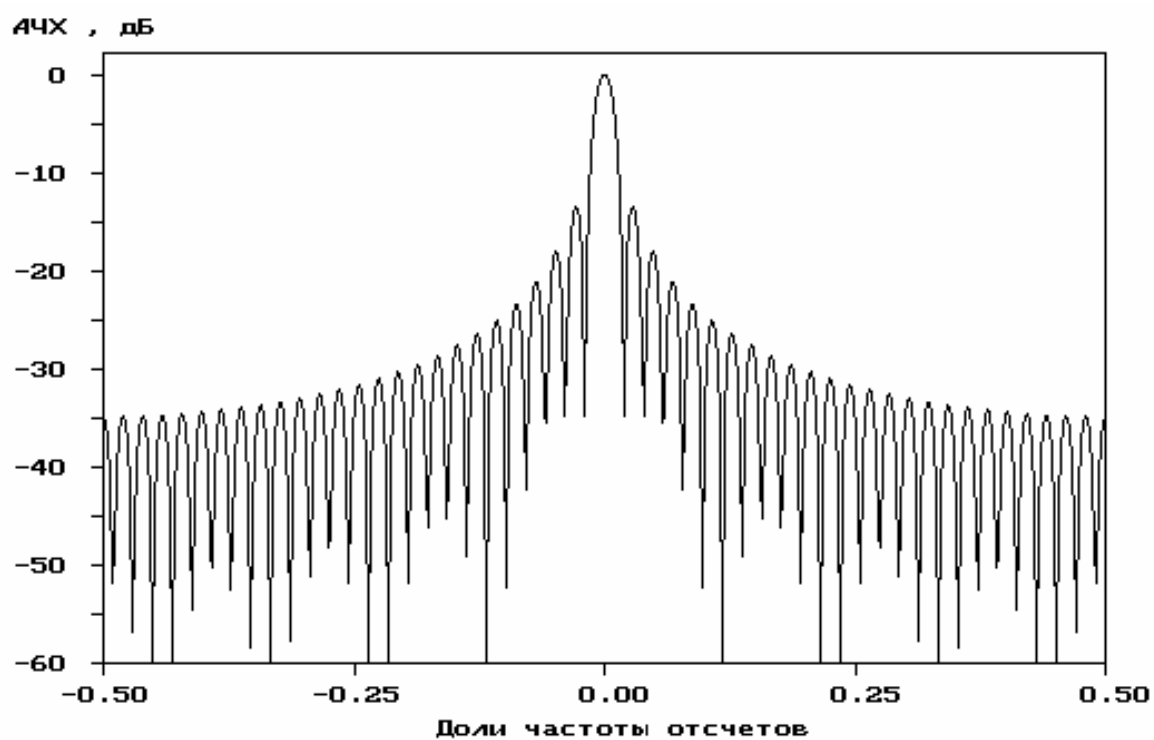
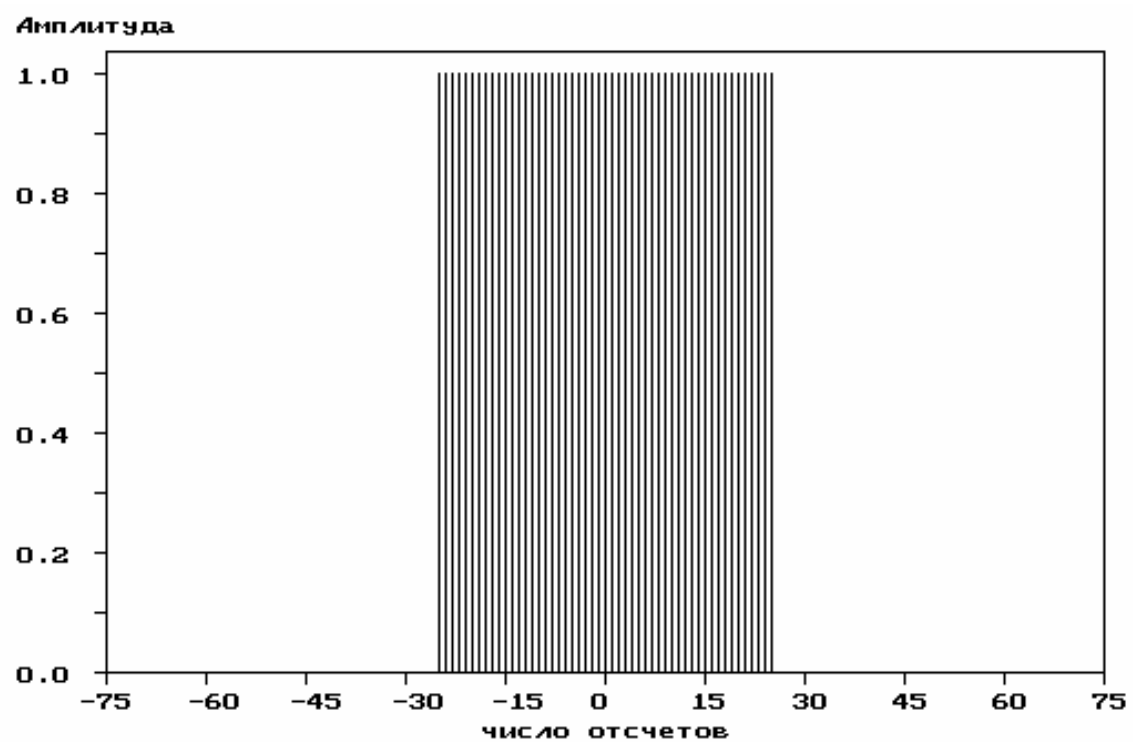


Рис. 4.3. Прямоугольное окно (а) и его частотная характеристика (б)

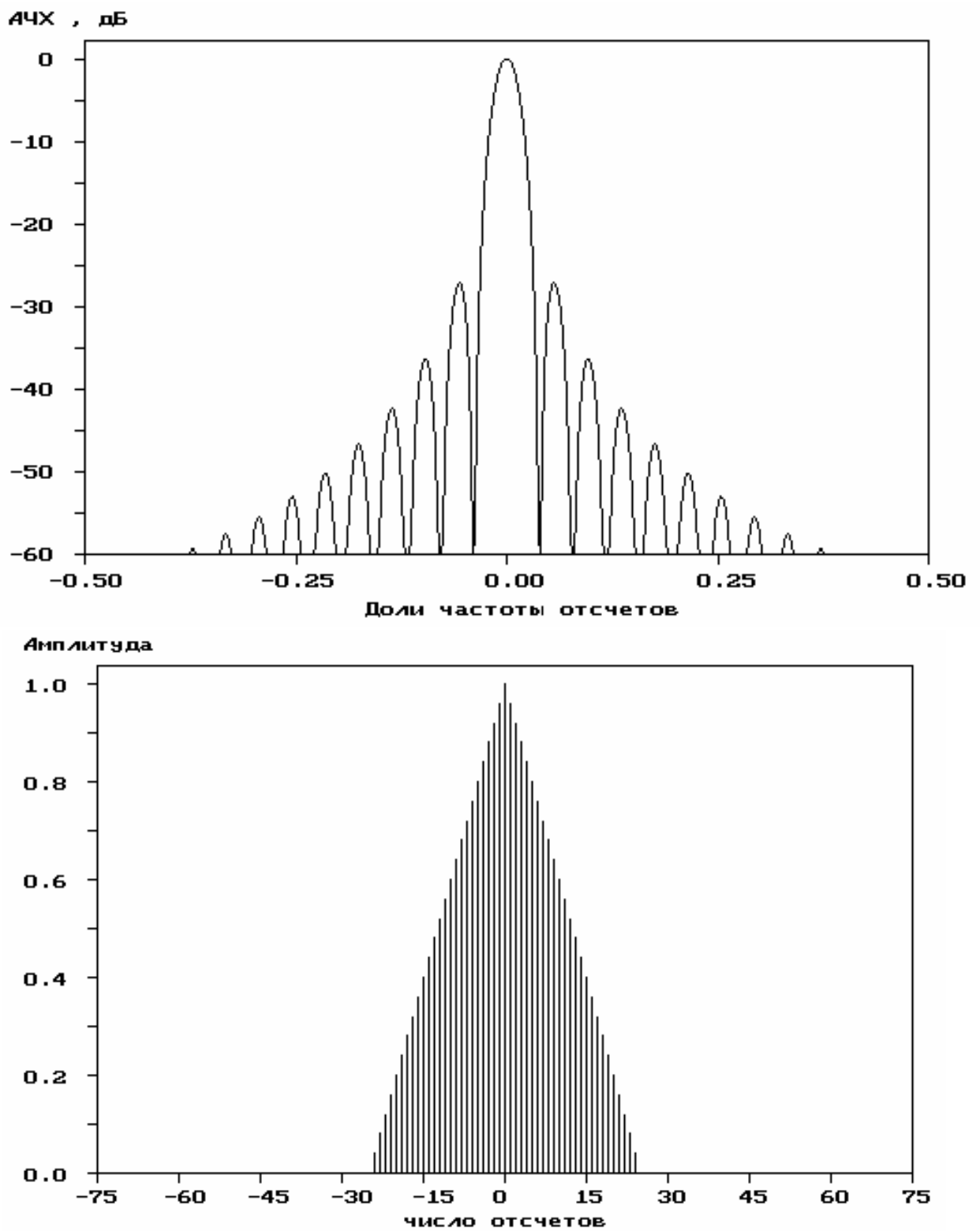


Рис. 4.4. Треугольное окно (а) и его частотная характеристика (б)

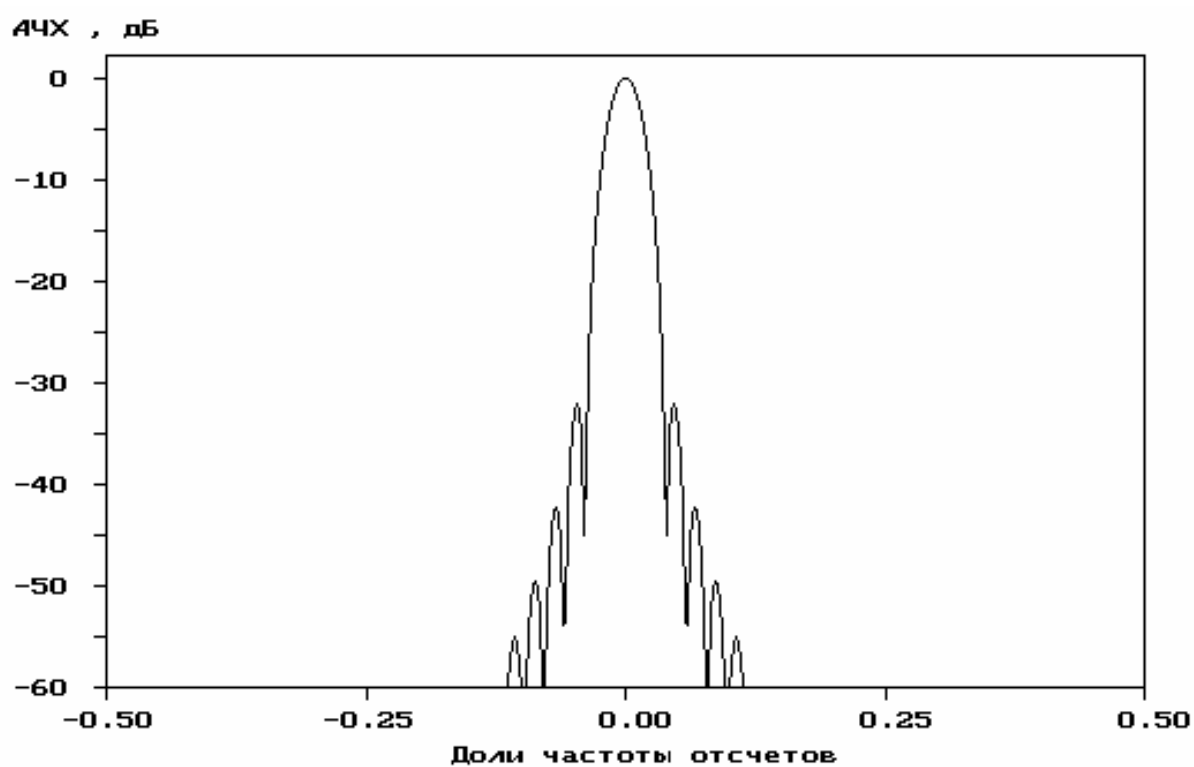
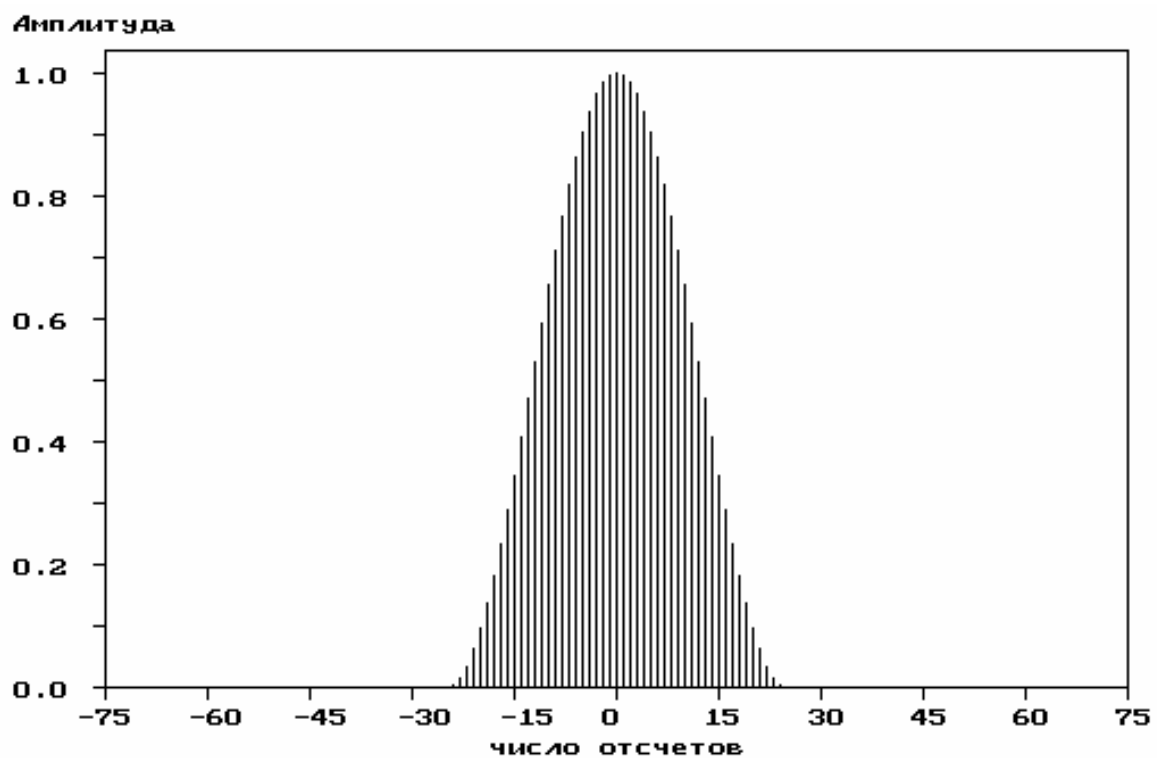


Рис. 4.5. Окно Ханна (а) и его частотная характеристика (б)

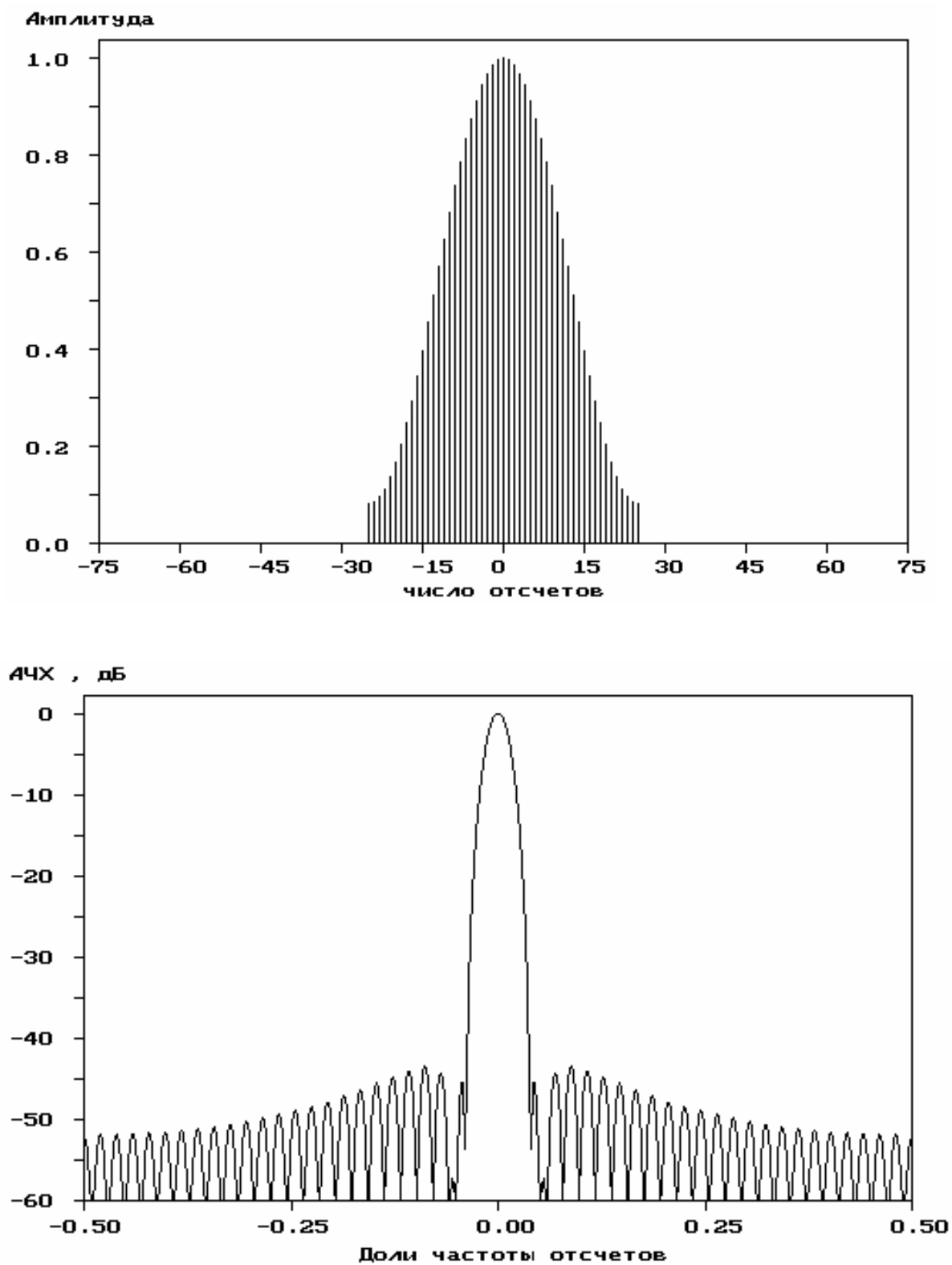


Рис. 4.6. Окно Хэмминга (а) и его частотная характеристика (б)

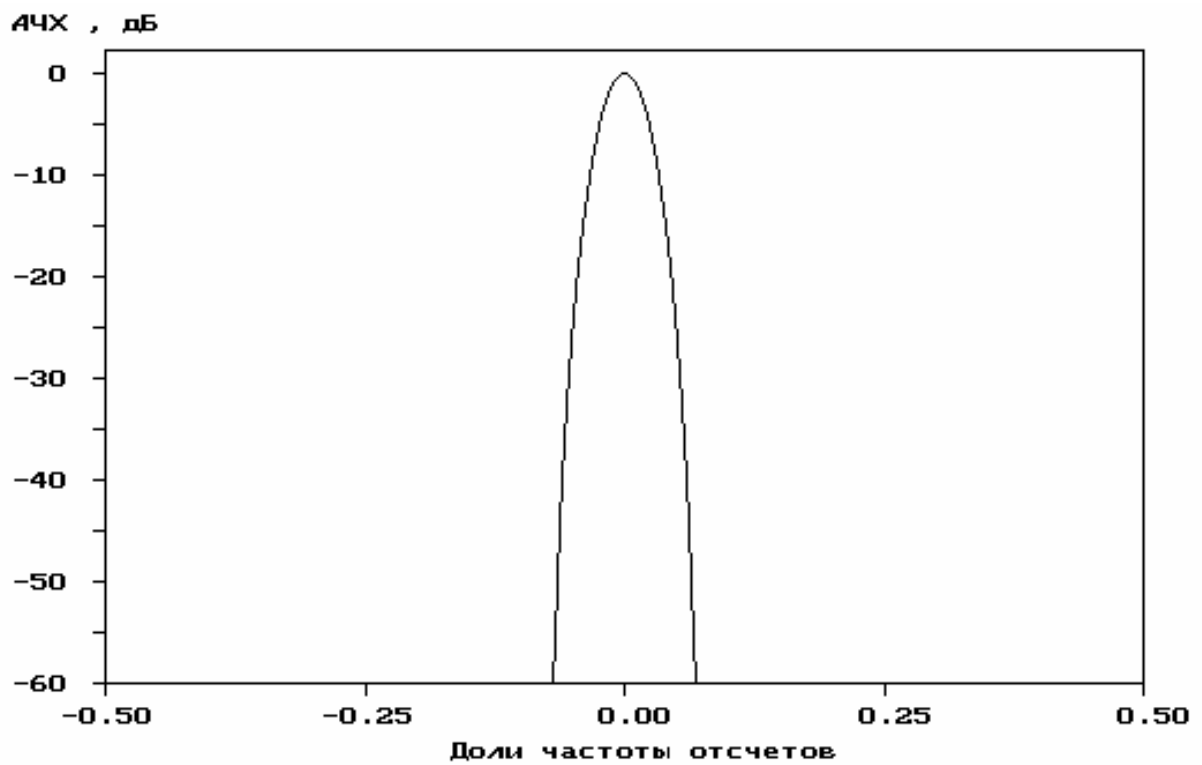
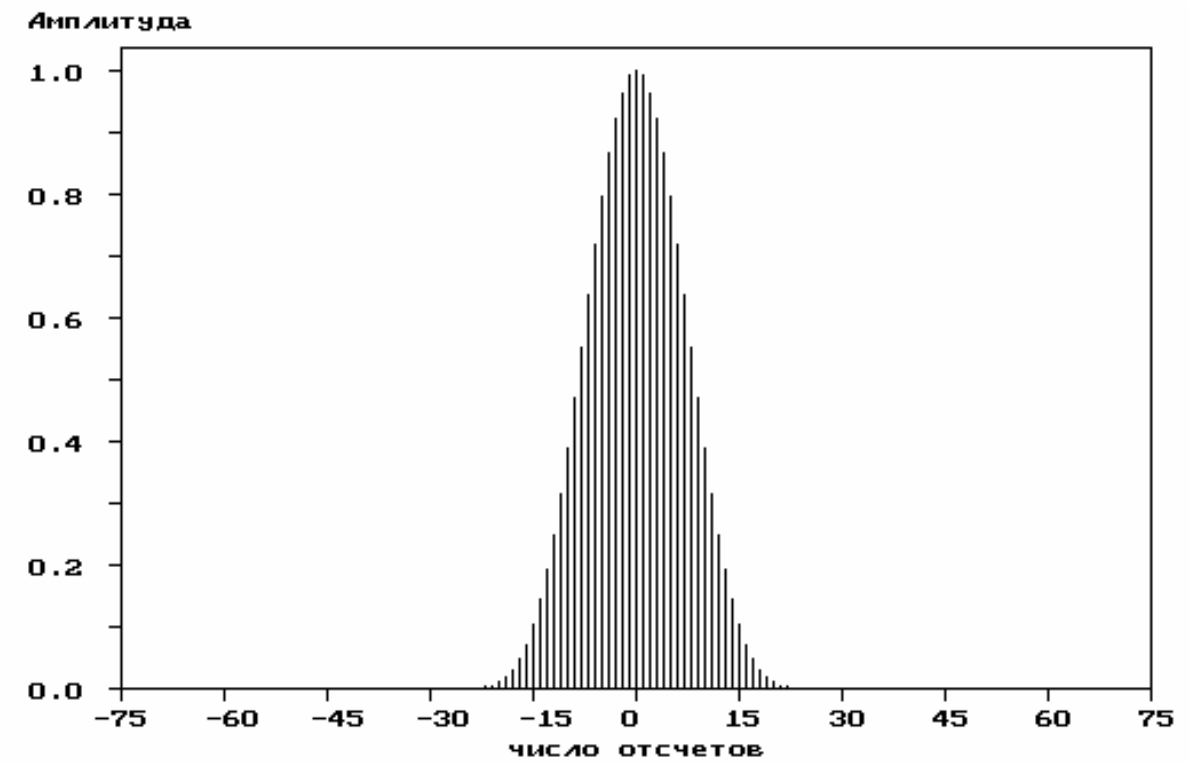


Рис. 4.7. Окно Наттола (а) и его частотная характеристика (б)

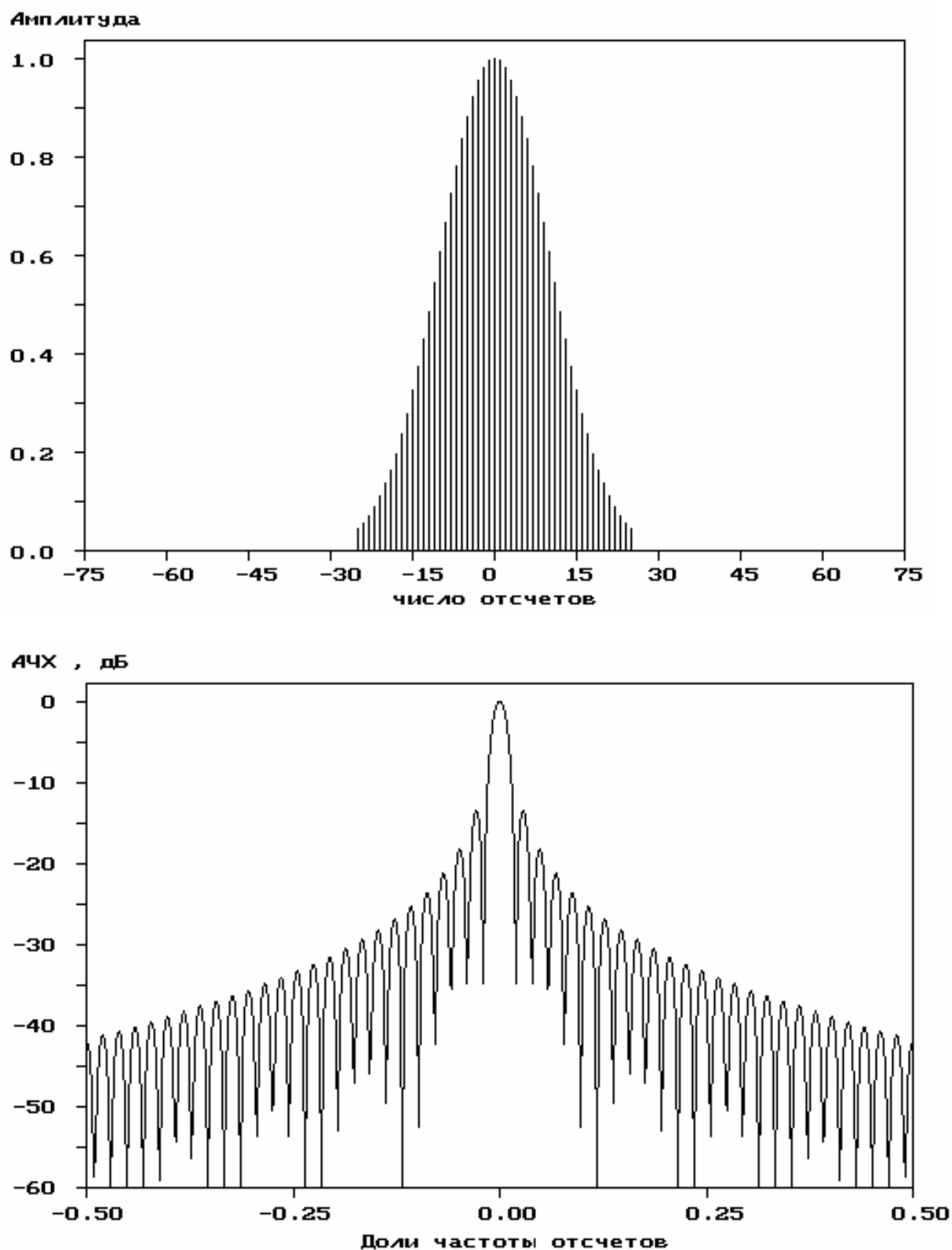


Рис. 4.8. Гауссовское окно (а) и его частотная характеристика (б)



Рис. 4.9

Стратегия выбора окон диктуется компромиссом между смещением из-за помех в области близких боковых лепестков и смещением из-за помех в области дальних боковых лепестков.

Например, если достаточно сильные компоненты сигнала расположены вблизи и на отдалении от слабой компоненты сигнала, то следует выбрать окно с одинаковым уровнем боковых лепестков около главного лепестка, с тем чтобы обеспечить малое смещение (рис. 4.9, а).

Если же имеется одна сильная компонента, удаленная от слабой компоненты сигнала, как показано на рис. 4.9, б, то следует выбирать окно с быстро спадающим уровнем боковых лепестков, причем их уровень в непосредственной близости к главному лепестку в данном случае не имеет значения.

В том случае, когда необходимо обеспечить высокое разрешение между очень близкими компонентами сигнала и удаленные компоненты отсутствуют, вполне приемлемым может оказаться окно даже с увеличивающимся уровнем боковых лепестков, но зато с очень узким главным лепестком, что иллюстрирует рис. 4.9, в.

Если динамический диапазон сигнала ограничен, то характеристики боковых лепестков не имеют особого значения, и поэтому можно выбрать окно, которое проще для численной реализации.

Если спектр сигнала относительно гладок, то можно вообще не применять окна, т. е. использовать прямоугольное окно.

Существуют окна, которые могут сами в некотором смысле регулироваться, или "адаптироваться", к параметрам данных.

5. КЛАССИЧЕСКИЕ МЕТОДЫ СПЕКТРАЛЬНОГО ОЦЕНИВАНИЯ

Существуют два формальных, но эквивалентных типа методов определения спектральной плотности мощности (СПМ). Прямые методы основаны на вычислении квадрата модуля преобразования Фурье для бесконечной последовательности данных с использованием соответствующего статистического усреднения и исторически получили название **периодограммных методов**. Косвенные методы основаны на использовании бесконечной последовательности значений данных для расчета автокорреляционной последовательности (АКП), преобразование Фурье которой дает искомую оценку СПМ. Они называются **коррелограммными методами** спектральной оценки. Эти методы основаны на связи АКП и СПМ дискретным преобразованием Винера – Хинчина.

Оба типа методов применяются сравнительно давно и поэтому получили название **классических методов** спектрального оценивания. Все эти методы обычно используются применительно к дискретно-временным последовательностям отсчетов данных.

Мы говорим об оценке СПМ, так как на практике в спектральном анализе используют конечные последовательности данных вместо бесконечных, что обуславливает невозможность получения точного спектра. Оценка характеризуется следующими параметрами.

Смещением $B(\hat{\alpha})$ оценки $\hat{\alpha}$ параметра α называется разность между истинным значением этого параметра и математическим ожиданием этой оценки:

$$B(\hat{\alpha}) = \alpha - \mathcal{E}(\hat{\alpha}).$$

Несмещенной оценкой называется такая, для которой $B(\hat{\alpha}) = 0$.

Дисперсией оценки называется средний квадрат отклонения оценки от ее среднего значения

$$\text{var}\{\hat{\alpha}\} = \mathcal{E}\{|\hat{\alpha}|^2\} - |\mathcal{E}\{\hat{\alpha}\}|^2.$$

Говорят, что оценка **состоятельна**, если с увеличением числа наблюдений смещение и дисперсия стремятся к нулю.

Классические методы спектрального оценивания относятся к числу наиболее устойчивых (робастных) методов спектрального оценивания. Они применимы почти ко всем классам стационарных сигналов и шумов, тогда как методы высокого разрешения (основанные на линейном моделировании) оказываются робастными только в случае ограниченного класса стационарных сигналов.

В классических методах, как правило, используются алгоритмы БПФ, поэтому они оказываются наиболее вычислительно эффективными. В полу-

чаемых с их помощью оценках СПМ (в отличие от методов высокого разрешения) высота спектральных пиков линейно зависит от мощности синусоид, присутствующих в данных.

Основной недостаток классических методов спектрального оценивания обусловлен искажающим воздействием просачивания по боковым лепесткам из-за использования конечных последовательностей данных. Взвешивание позволяет ослабить влияние боковых лепестков, но лишь за счет ухудшения спектрального разрешения.

Разрешение, обеспечиваемое классическими методами, не может превышать величины, обратной длине записи данных, и не зависит от характеристик анализируемых данных. Для улучшения статистической устойчивости спектральных оценок может быть использовано псевдоусреднение по ансамблю за счет сегментации данных, но это также ухудшает спектральное разрешение. Обработка конечных записей данных требует принятия компромиссов относительно разрешения, устойчивости (минимизации дисперсии оценки) и подавления просачивания.

И периодограммные, и коррелограммные методы дают оценки, которые обладают аналогичными статистическими характеристиками и в целом выглядят примерно одинаково, если не считать небольших визуальных различий в тонких деталях формы спектров. На практике предпочтение отдается наиболее эффективному в вычислительном отношении методам. До широкого распространения ЭВМ в основном применялся корреляционный метод. С появлением алгоритмов БПФ и спецпроцессоров ЦОС предпочтение было отдано периодограммным методам.

5.1. Периодограммные методы спектрального анализа

Термин "*периодограмма*" впервые был использован в работах А. Шустера в 1897 г. Идея предложенного им метода заключалась в усреднении спектральной оценки, полученной для конечной реализации данных, по нескольким оценкам для различных реализаций (периодограмм). Метод получил развитие, и в настоящее время имеется ряд его модификаций.

Одним из формальных определений СПМ, основанных на допущении об эргодичности, является следующая дискретно-временная форма:

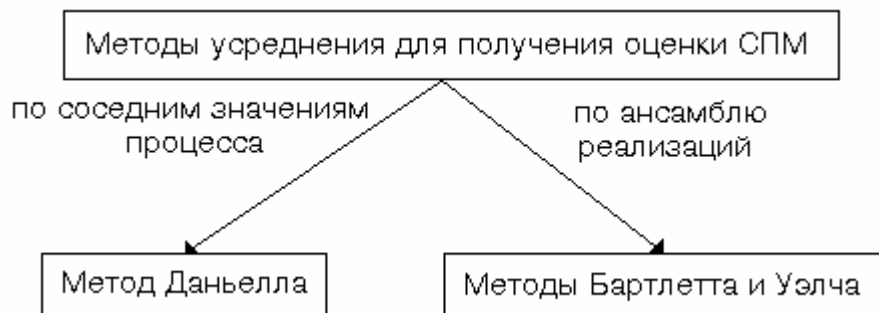
$$P(f) = \lim_{N \rightarrow \infty} \mathcal{E} \left\{ \frac{1}{(2N+1)T} \left| T \sum_{n=-N}^N x(n) \exp(-j2\pi f n T) \right|^2 \right\}. \quad (5.1)$$

Пренебрегая операцией вычисления математического ожидания $\mathcal{E}\{\}$ и полагая, что конечное множество данных $\{x(0), \dots, x(N-1)\}$ содержит N отсче-

тов, получим выборочный спектр

$$\tilde{P}(f) = \frac{1}{NT} \cdot \left| T \sum_{n=0}^{N-1} x(n) \exp(-j2\pi fnT) \right|^2 = \frac{T}{N} \left| \sum_{n=0}^{N-1} x(n) \exp(-j2\pi fnT) \right|^2, \quad (5.2)$$

который может быть вычислен по конечной последовательности данных. Это – исходная немодифицированная форма периодограммной оценки СПМ. Выборочный спектр будет давать статистически несостоятельные (т. е. неустойчивые) оценки СПМ, поскольку была опущена операция вычисления математического ожидания, предусмотренная выражением (5.1). Поэтому для сглаживания периодограммной оценки необходимо применять что-то вроде псевдоусреднения по ансамблю. Предложены три основных типа методов сглаживания. В методе Даньелла (1946 г.) периодограмма сглаживается путем усреднения по соседним ячейкам частотного разрешения. В методе Бартлетта (1948 г.) производится усреднение по множеству периодограмм, полученных по неперекрывающимся сегментам исходной последовательности данных. В методе Уэлча подход Бартлетта применяется к перекрывающимся сегментам и вводится окно данных для уменьшения смещения оценок из-за эффекта просачивания.



Классификация перечисленных методов оценки СПМ отражена на рис. 5.1.

Рис. 5.1

В литературе описано также много других оптимальных методов сглаживания периодограмм, однако все они оптимальны только применительно к некоторым ограниченным классам сигналов. Три перечисленных метода, которые мы рассмотрим, возможно, не всегда оптимальны, но практика доказала их статистическую устойчивость для многих классов сигналов.

5.1.1. Периодограмма Даньелла

Даньелл предположил, что для сглаживания быстрых флуктуаций выборочного спектра можно использовать усреднение по соседним спектраль-

ным частотам. Если для вычисления выборочного спектра $P(f)$ на сетке частот $f_k = k / NT$, $0 \leq k \leq N - 1$ используется алгоритм БПФ, то модифицированная оценка периодограммы на частоте f_i может быть получена посредством усреднения в P точках значений $P(f)$ с каждой стороны от этой частоты:

$$\bar{P}_D(f_i) = \frac{1}{2P+1} \sum_{n=i-P}^{i+P} \tilde{P}(f_n). \quad (5.3)$$

Обобщением этого подхода является обработка выборочного спектра с помощью фильтра нижних частот с частотной характеристикой $H(f)$. Периодограмму Даньелла можно в этом случае записать в виде свертки выборочного спектра и частотной характеристики фильтра нижних частот

$$\bar{P}_D(f) = \tilde{P}(f) \otimes H(f). \quad (5.4)$$

Подстрочечный индекс D в этих выражениях использован для того, чтобы отличить периодограммную оценку Даньелла от других оценок СПМ.

5.1.2. Периодограмма Бартлетта

Сглаживание выборочного спектра по методу Бартлетта основано на создании псевдоансамбля периодограмм за счет деления последовательности из N отсчетов данных на P неперекрывающихся сегментов по D отсчетов в каждом, так что $PD \leq N$. Тогда p -й сегмент будет состоять из отсчетов

$$x^{(p)}(n) = x(pD + n), \quad 0 \leq n \leq D - 1,$$

где надстрочечный индекс (p) обозначает номер сегмента. По каждому сегменту $0 \leq p \leq P - 1$ независимо, в диапазоне $-1/2T \leq f \leq 1/2T$ вычисляется выборочный спектр

$$\tilde{P}^{(p)}(f) = \frac{1}{DT} \left| T \sum_{n=0}^{D-1} x^{(p)}(n) \exp(-j2\pi f n T) \right|^2, \quad (5.5)$$

для чего может быть использован алгоритм БПФ на некоторой сетке частот. Затем на каждой частоте, представляющей интерес, P отдельных немодифицированных периодограмм усредняются, чтобы получить усредненную пе-

риодограмму Бартлетта $P_B(f)$

$$\bar{P}_B(f) = \frac{1}{P} \sum_{p=0}^{P-1} \tilde{P}^{(p)}(f). \quad (5.6)$$

Смещение среднего значения этой периодограммы определяется выражением

$$\mathcal{E}\{\bar{P}_B(f)\} = \frac{1}{P} \sum_{p=0}^{P-1} \mathcal{E}\{\tilde{P}^{(p)}(f)\} = \mathcal{E}\{\tilde{P}(f)\}. \quad (5.7)$$

Таким образом, смещение оценки $P_B(f)$ обусловлено воздействием окна, которое смещает и выборочный спектр периодограммы каждого отдельного сегмента. Если периодограммы P сегментов статистически независимы (что приближенно выполняется в том случае, когда значения автокорреляции

$$r(m) = \mathcal{E}\{x(n+m)x^*(n)\}$$

малы при $m > D$, то $P_B(f)$ можно рассматривать как выборочное среднее значение некоторой совокупности из P независимых наблюдений выборочного спектра $P(f)$. Можно показать, что величина дисперсии усредненной периодограммы Бартлетта обратно пропорциональна числу используемых сегментов, т. е.

$$\text{var}\{\bar{P}_B(f)\} \sim \frac{P^2(f)}{P}. \quad (5.8)$$

Следовательно, приближенно можно считать, что устойчивость этой спектральной оценки улучшается как величина, обратная числу сегментов P . Уменьшение дисперсии с увеличением P будет меньше в том случае, когда периодограммы сегментов оказываются статистически зависимыми. Что же касается разрешения, то оно в результате разбиения последовательности данных на сегменты по D отсчетов в каждом, где $D < N$, будет, естественно, уменьшаться, поскольку результирующее эффективное спектральное окно в этом случае имеет более широкий главный лепесток. При фиксированном значении $N = PD$ соблюдается обычное компромиссное соотношение между высоким спектральным разрешением (при максимально возможном значении D) и минимальной дисперсией оценки (при максимально возможном значении P).

5.1.3. Периодограмма Уэлча

Уэлч модифицировал основную схему метода сегментирования и усреднения Бартлетта за счет применения окна данных и использования частично перекрывающихся сегментов. Перед вычислением периодограммы каждого сегмента этот сегмент обрабатывается с помощью окна данных. Цель

применения окна – за счет незначительного ухудшения разрешения ослабить эффекты, возникающие из-за боковых лепестков и уменьшить смещение оценок. Цель перекрытия сегментов – увеличить число усредняемых сегментов при заданной длине записи данных и тем самым уменьшить дисперсию оценки СПМ. На основе БПФ Уэлч разработал также эффективную вычислительную процедуру [1, с. 211] для реализации своего метода усреднения периодограмм взвешенных и перекрывающихся сегментов данных. Именно это сделало метод Уэлча самым популярным периодограммным методом спектрального оценивания в наши дни.

Если запись комплексных данных $\{x(0), \dots, x(N-1)\}$ из N отсчетов разбита на P сегментов по D отсчетов в каждом со сдвигом S отсчетов между соседними сегментами ($S \leq D$, то максимальное число сегментов P будет определяться целой частью числа $(N-D)/S+1$. В частном случае, когда N – степень двух, число сегментов $P = (N/S)-1$. Взвешенный окном $w(n)$ p -й сегмент

$$x^{(p)}(n) = w(n)x(n + pS)$$

будет состоять из D отсчетов ($0 \leq n \leq D-1$), а номер сегмента p лежит в интервале $0 \leq p \leq P-1$.

Выборочный спектр взвешенного p -го сегмента в диапазоне частот $-1/2T \leq f \leq 1/2T$ определяется выражением:

$$\tilde{P}^{(p)}(f) = \frac{1}{UDT} X^{(p)}(f) [X^{(p)}(f)]^* = \frac{1}{UDT} |X^{(p)}(f)|^2, \quad (5.9)$$

где

$$X^{(p)}(f) = T \sum_{n=0}^{D-1} x^{(p)}(n) \exp(-j2\pi f n T)$$

– дискретно-временное преобразование p -го сегмента, а $U = \sum_{n=0}^{D-1} w^2(n)$ –

энергия дискретно-временного окна. Среднее значение периодограммы взвешенных сегментов дает оценку периодограммы Уэлча $P_W(f)$:

$$\bar{P}_w(f) = \frac{1}{P} \sum_{p=0}^{P-1} \tilde{P}^{(p)}(f). \quad (5.10)$$

Множитель $1/U$ в (5.9) устраняет эффект влияния энергии окна на смещение в оценке СПМ $P_W(f)$. Можно показать (доказать это утверждение студентам предоставляется самостоятельно), что смещение периодограммы Уэлча представимо в следующем виде:

$$\varepsilon\{\bar{P}_w(f)\} = \frac{1}{P} \sum_{p=0}^{P-1} \varepsilon\{\tilde{P}^{(p)}(f)\} = \frac{1}{U} P(f) \otimes |W(f)|^2, \quad (5.11)$$

где

$$W(f) = T \sum_{n=0}^{D-1} w(n) \exp(-j2\pi f n T)$$

– дискретно-временное преобразование Фурье окна данных, а $P(f)$ – истинная СПМ.

ОДВПФ функции (5.11) дает ожидаемое значение эффективной автокорреляционной последовательности периодограмм Уэлча

$$\mathcal{E}\{\bar{r}^w(m)\} = r(m)\Phi(m)/U,$$

где

$$\Phi(m) = w(m)w^*(-m) = T \sum_{n=0}^{D-m-1} w(n+m)w(n)$$

– автокорреляционная последовательность окна данных.

Полная мощность в исходной последовательности отсчетов данных определяется членом автокорреляционной последовательности, соответствующим нулевому временному сдвигу. Поскольку $U = \Phi(0)$, то

$$\mathcal{E}\{\bar{r}^w(0)\} = r(0)\Phi(0)/U = r(0),$$

а это означает, что $\bar{r}^w(0)$ – несмещенная оценка мощности.

До сих пор мы оценивали автоспектральную плотность, причем в обозначениях СПМ были опущены подстрочные индексы xx . Процедура Уэлча для оценивания взаимной спектральной плотности во многом аналогична процедуре оценки спектральной плотности. В этом случае N -точечные последовательности $x(n)$ и $y(n)$ сначала сегментируются и взвешиваются, в результате чего получаются последовательности

$$x^{(p)}(n) = w(n)x(n + ps), \quad y^{(p)}(n) = w(n)y(n + ps),$$

где $0 \leq n \leq D-1$, $0 \leq p \leq P-1$, по которым затем вычисляется выборочный взаимный спектр

$$\tilde{P}_{xy}^{(p)}(f) = \frac{1}{UDT} X^{(p)}(f) [Y^{(p)}(f)]^*,$$

где

$$X^{(p)}(f) = T \sum_{n=0}^{D-1} x^{(p)}(n) \exp(-j2\pi f n T), \quad Y^{(p)}(f) = T \sum_{n=0}^{D-1} y^{(p)}(n) \exp(-j2\pi f n T)$$

– ДВПФ, а U – энергия дискретно-временного окна. Усреднение по периодо-

граммам всех сегментов дает окончательную оценку периодограммы

$$\bar{P}_{xy}(f) = \frac{1}{P} \sum_{p=0}^{P-1} \hat{P}_{xy}^{(p)}(f).$$

В случае гауссовских процессов характеристики среднего значения и дисперсии этой взаимной спектральной оценки во многом сходны с характеристиками аналогичных величин для спектральной оценки.

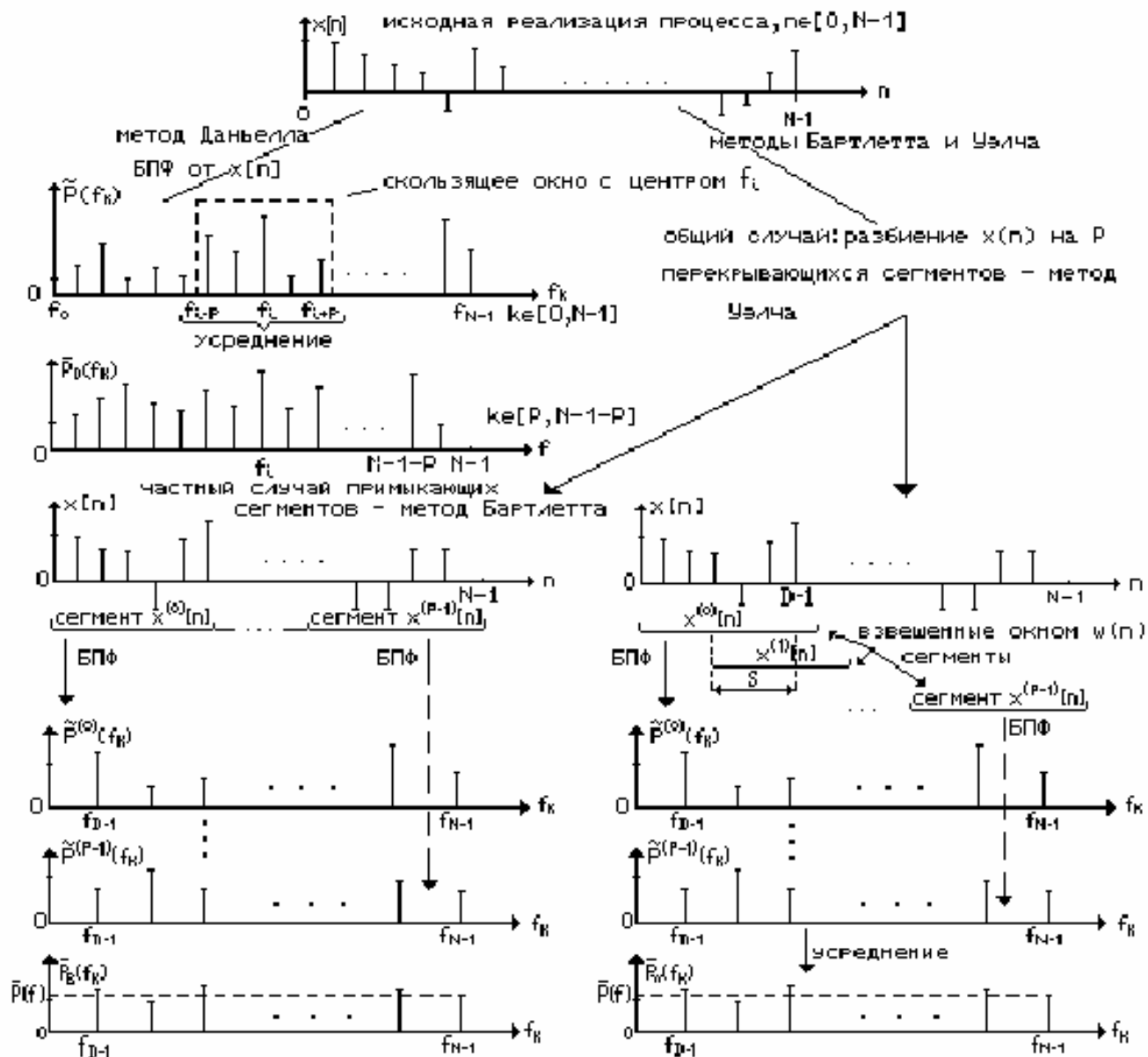
Уэлч, в частности, предложил использовать окно Ханна и 50%-ное перекрытие сегментов, которое обеспечивало очень эффективные реализации его метода на основе алгоритма БПФ при вычислении оценки (5.10) на некоторой сетке частот. Кроме того, при 50%-ном перекрытии сегментов все данные используются дважды, за исключением $D/2$ отсчетов на каждом конце исходной N -точечной последовательности данных, а это выравнивает обработку большинства отсчетов данных, поскольку те отсчеты, которые имеют малые веса на одном сегменте, получают большие веса на следующем сегменте.

Анализ поведения дисперсии периодограммы Уэлча для гауссовских процессов показал, что минимальная дисперсия для окна Ханна достигается при 65%-ном перекрытии, при этом величина дисперсии увеличивается приблизительно на 8% при использовании 50%-ного перекрытия сегментов. Так же как и дисперсия периодограммы Бартлетта, дисперсия периодограммы Уэлча примерно обратно пропорциональна числу сегментов, т. е.

$$\text{var}\{\bar{P}_w(f)\} \sim \frac{P^2(f)}{P} \quad (5.12)$$

в предположении независимости сегментов (хотя перекрытие сегментов приводит, конечно, к некоторой их зависимости). Благодаря перекрытию сегментов при одинаковой длине исходной последовательности данных, можно сформировать большее число сегментов, чем в методе Бартлетта, а это уменьшает величину дисперсии периодограммы Уэлча по сравнению с дисперсией периодограммы Бартлетта.

В ряде опубликованных исследований утверждается, что применение окон данных лишь придает одним отсчетам большую значимость перед другими отсчетами, а также ухудшает разрешение (уширяет главный лепесток частотного окна) без какого-либо компенсирующего уменьшения дисперсии. Эти утверждения, несомненно, верны в отношении выборочного спектра, получаемого сразу по всей последовательности отсчетов сигнала с относительно плоской спектральной характеристикой. Однако по отношению к периодограмме Уэлча они не верны, поскольку перекрытие сегментов как раз и используется для выравнивания обработки (т. е. выравнивания значимости отсчетов) данных, а увеличение числа сегментов – для уменьшения дисперсии оценки СПМ. Не верны они и в отношении спектров сигналов с большим



различием амплитуд составляющих.

Рис.5.2

Рассмотренные периодограммные методы анализа отображены на рис. 5.2. Следует заметить, что при выборе того или иного метода необходимо учитывать его ограничения и особенности. Так, очевидно, метод Даниелла неприменим при значительных флуктуациях $P(f)$. Однако, по-видимому, ему следует отдать предпочтение, если априорно известно, что $P(f) = \text{const}$ в диапазоне анализа. В этом случае выражение (5.3) представимо в виде

$$\bar{P}_D(f) = \frac{2}{N-2} \sum_{k=0}^{N/2-1} \tilde{P}(f_k), \quad k \in [0, N/2).$$

Сегментирование исходной последовательности $x(n)$, используемое в методах Бартлетта и Уэлча, приводит к ухудшению разрешающей способности в P раз, поэтому при выборе числа элементов в сегменте также необходимо иметь априорную информацию о спектре флуктуаций $P(f)$ в диапазоне анализа.

5.2. Корреляционный анализ и спектральная оценка

Одной из важнейших временных характеристик сигнала является автокорреляционная функция (АКФ), позволяющая судить о степени связи сигнала с его сдвинутой по времени копией. Для действительной последовательности $x(n)$ АКФ определяется в единицах мощности:

$$r_{xx}(m) = P_{xx}(m) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1-m} x(n)x(n+m), n \in [0, N-1]. \quad (5.13)$$

При $m = 0$ значение АКФ достигает максимума и равно средней мощности сигнала $x(n)$:

$$r_{xx}(0) = P_{xx}(0) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} |x(n)|^2, \quad n \in [0, N-1].$$

Для оценки степени подобия двух сигналов $x(n)$ и $y(n)$ используется функция взаимной корреляции (ФВК):

$$r_{xy}(m) = P_{xy}(m) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1-m} x(n)y(n+m), \quad n \in [0, N-1]. \quad (5.14)$$

При $m = 0$ значение ФВК определяет взаимную мощность сигналов $x(n)$ и $y(n)$:

$$r_{xy}(0) = P_{xy}(0) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n)y(n), \quad n \in [0, N-1].$$

Функцию взаимной корреляции измеряют в случаях, когда последовательности $x(n)$ и $y(n)$ связаны операциями фильтрации и задержки. Например, в радиолокационных системах $y(n)$ – это отфильтрованный, зашумленный и задержанный зондирующий сигнал $x(n)$. В таких случаях положение максимума ФВК дает величину задержки входного сигнала, что является основой для решения многих задач радиолокации.

Пусть, например, $\{x(n)\} = (0, 1, 3, 6, 2, 1, 0, 0)$ и $y(n)$ сдвинута на два отсчета: $\{y(n)\} = (3, 6, 2, 1, 0, 0)$. Найдем ФВК: $\{z(m)\} = (18/6, 36/6, 50/6, 32/6, 12/6, 3/6)$. Анализируя ФВК, находим ее максимум: $z(2) = 50/6$. Номер ($m=2$) максимума

соответствует сдвигу $y(n)$ относительно $x(n)$.

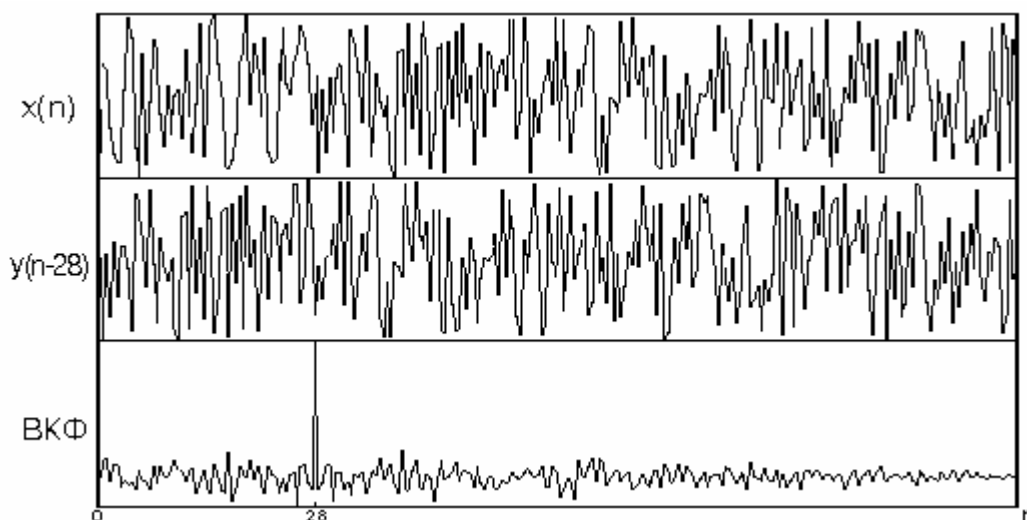


Рис. 5.3

На рис. 5.3 представлен пример двух сдвинутых на 28 отсчетов последовательностей и их функция взаимной корреляции.

Спектральная плотность АКФ равна спектральной плотности мощности [6, с. 77]. Таким образом, спектральная плотность мощности (СПМ) определяется как ДВПФ автокорреляционной последовательности:

$$P_{xx}(f) = \sum_{m=-\infty}^{\infty} r_{xx}(m) \exp(-j2\pi fm);$$

$$r_{xx}(m) = \frac{1}{N} \sum_{n=0}^{N-1-m} x(n)x(n+m),$$
(5.15)

где $r_{xx}(m)$ – функция автокорреляции, а N – число выборок $x(n)$, $0 \leq n \leq N-1$.

Взаимная СПМ двух процессов $x(n)$ и $y(n)$ определяется как ДВПФ взаимной корреляционной последовательности:

$$P_{xy}(f) = \sum_{m=-\infty}^{\infty} r_{xy}(m) \exp(-j2\pi fm);$$

$$r_{xy}(m) = \frac{1}{N} \sum_{n=0}^{N-1-m} x(n)y(n+m),$$
(5.16)

где $r_{xy}(m)$ – функция взаимной корреляции, а N – число выборок $x(n)$ и $y(n)$.

Функция СПМ описывает распределение мощности случайного процесса по частоте. В силу того, что функции АКФ и ВКФ имеют конечную длину, бесконечные пределы суммирования заменяют на конечные. Если принять во внимание, что $x(n)$ и $y(n)$ имеют большое число выборок

($N \rightarrow \infty$), то основная формула для вычислений АКФ и ФВК:

$$r_{xx}(m) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n)x(n+m), \quad -\infty < m < \infty, \quad (5.17)$$

$$r_{xy}(m) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n)y(n+m), \quad -\infty < m < \infty. \quad (5.18)$$

Предельные выражения показывают, что средние значения оценок сходятся при больших N к некоторому определенному постоянному числу. Формулы (5.17) и (5.18) хорошо работают при $N > 1024$. Однако дисперсия оценки $P_{xx}(f)$ и $P_{xy}(f)$ будет сходиться к нулю лишь при выполнении дальнейшего усреднения этих оценок.

Рассмотрим более подробно

- 1) метод получения оценки функции корреляции;
- 2) метод получения статистически устойчивых оценок СПМ на основе функции корреляции, получивший название корреляционного или коррелограмного метода оценки СПМ.

5.2.1. Метод получения оценки ФВК

Рассмотрим последовательности $x(n)$ и $y(n)$. Их ДПФ:

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp(-j2\pi nk / N), \quad Y(k) = \sum_{n=0}^{N-1} y(n) \exp(-j2\pi nk / N).$$

Умножение $X(k)Y(k)$ соответствует свертке $x(n)$ и $y(n)$:

$$X(k)Y(k) \leftrightarrow \sum_{n=0}^{N-1} x(n)y(n-m) = \sum_{n=0}^{N-1} x(n-m)y(n).$$

Таким образом, ФВК можно найти как ОДПФ от произведения ДПФ исходных последовательностей:

$$r_{xy}(m) = \sum_{n=0}^{N-1} x(n)y(n+m) = \text{ОДПФ} \{X(k)Y^*(k)\},$$

$$r_{xx}(m) = \sum_{n=0}^{N-1} x(n)x(n+m) = \text{ОДПФ} \{X(k)X^*(k)\}.$$

Причем $x(n)$ и $y(n)$ во всех выкладках считаются действительными.

Алгоритм получения оценки АКФ и ФВК опирается на метод быстрой свертки и для последовательностей $x(n)$ и $y(n)$ ($0 \leq n \leq N-1$) выглядит сле-

дующим образом.

- 1) Вычислить $x'(n) = x(n)w(n)$, $y'(n) = y(n)w(n)$.
- 2) Вычислить $X(k) = \text{БПФ}\{x'(n)\}$, $Y(k) = \text{БПФ}\{y'(n)\}$.
- 3) Вычислить оценку $S_{xy}(k) = X(k)Y^*(k)$.
- 4) Получить оценку ФВК, выполнив $r_{xy}(m) = \text{ОБПФ}\{S_{xy}(k)\}$.



Рис. 5.4

Схема алгоритма изображена на рис. 5.4. В этом методе алгоритм БПФ используется непосредственно для вычисления спектральной плотности мощности $S_{xy}(k)$ на L дискретных частотах. Для уменьшения эффектов просачивания каждая из последовательностей умножается на сглаживающее окно $w(n)$. После выполнения ОБПФ имеем искомую ФВК.

5.2.2. Коррелограммный метод оценки СПМ

Суть алгоритма коррелограммного метода оценки СПМ заключается в усреднении функции корреляции. Вся реализация $x(n)$ длиной N разбивается на M кусков с перекрытием. Для каждого куска вычисляется своя АКФ $r_i(m)$, далее производят суммирование по всем $i \in [1, M]$:

$$\bar{r}_{xx}(m) = \sum_{i=1}^M r_i(m), \quad 0 < m < M - 1.$$

Оценку СПМ определяют как ДПФ от $r_{xx}(m)$:

$$\bar{P}_{xx}(f) = \sum_{m=0}^{L-1} \bar{r}_{xx}(m) \exp(-j2\pi fm). \quad (5.19)$$

Для того чтобы дисперсия оценки сходилась к нулю, необходимо выполнение соотношения $L \ll N$. В результате оценка СПМ вычисляется через преобразование последовательности, усеченной во временной области. Усе-

чение во временной области может вызвать большие пульсации в частотной области, для их подавления используются окна. На выбор окна накладываются некоторые ограничения. Если ДПФ окна не является положительной функцией, то возможно, что СПМ будет отрицательной, что лишено физического смысла. Существуют окна, ДПФ которых положительно на всех частотах, например, треугольные окна, симметричные относительно начальной точки отсчета. Если корреляционная последовательность определена на бесконечном интервале временных сдвигов, то окно следует определять симметричным относительно нулевого сдвига; если же – на положительной полуоси, то и окно определяется на этой полуоси.

Схема алгоритма коррелограммного метода приведена на рис. 5.5.

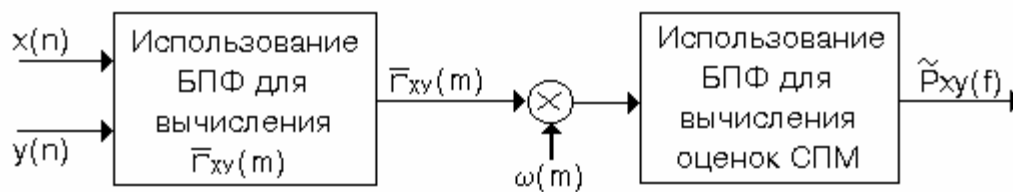


Рис. 5.5

Для корреляционного метода оценки СПМ были разработаны быстрые алгоритмы с применением БПФ. Рассмотрим *первый алгоритм*. Пусть имеется N -точечная последовательность $x(n)$ ($N = 20$) и необходимо найти первые L отсчетов АКФ ($L = 5$). $\{x(n)\} = \{x_0, x_1, x_2, x_3, x_4, \dots, x_{17}, x_{18}, x_{19}\}$.

Алгоритм выглядит следующим образом:

- 1) Вычислить

$$U_0(k) = \text{ДПФ}\{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\},$$

$$V_0(k) = \text{ДПФ}\{x_0, x_1, x_2, x_3, x_4, 0, 0, 0, 0\},$$

$$U_1(k) = \text{ДПФ}\{x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}\},$$

$$V_1(k) = \text{ДПФ}\{x_5, x_6, x_7, x_8, x_9, 0, 0, 0, 0\},$$

$$U_2(k) = \text{ДПФ}\{x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}, x_{18}\},$$

$$V_2(k) = \text{ДПФ}\{x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, 0, 0, 0, 0\},$$

$$U_3(k) = \text{ДПФ}\{x_{15}, x_{16}, x_{17}, x_{18}, x_{19}, 0, 0, 0, 0\},$$

$$V_3(k) = \text{ДПФ}\{x_{15}, x_{16}, x_{17}, x_{18}, x_{19}, 0, 0, 0, 0\}.$$
- 2) Вычислить $X_i(k) = U_i(k)V_i^*(k), \quad i \in [0, 3].$
- 3) Вычислить $X(k) = X_0(k) + X_1(k) + X_2(k) + X_3(k).$
- 4) Вычислить $r(m) = \text{ОДПФ}\{X(k)\}.$

Первые 5 отсчетов будут искомым результатом. Из описанной методики видно, что ДПФ должно приблизительно в два раза превышать требуемое число отсчетов корреляционной функции. А требуемое число преобразований ДПФ равно $2K+1$, где $K = N/L$ (N – общее число отсчетов сигнала, а L – требуемое количество отсчетов функции корреляции). Затем, дополнив $r(m)$ нулями и выполнив ДПФ, получим оценку СПМ с любым частотным разре-

шением.

Второй алгоритм эффективнее первого за счет того, что значение $U_i(k)$ можно выразить через $V_i(k)$ и $V_{i+1}(k)$, но необходимо, чтобы интервал разбиения содержал четное число отсчетов, желательно кратное 2, и перекрытие интервалов было в точности равно 2:1. Для демонстрации алгоритма возьмем ту же последовательность: $N = 20$.

1) Вычислить все $V_i(k)$, где

$$V_0(k) = \text{ДПФ}\{x_0, x_1, x_2, x_3, 0, 0, 0, 0\},$$

$$V_1(k) = \text{ДПФ}\{x_4, x_5, x_6, x_7, 0, 0, 0, 0\},$$

$$V_2(k) = \text{ДПФ}\{x_8, x_9, x_{10}, x_{11}, 0, 0, 0, 0\},$$

$$V_3(k) = \text{ДПФ}\{x_{12}, x_{13}, x_{14}, x_{15}, 0, 0, 0, 0\},$$

$$V_4(k) = \text{ДПФ}\{x_{16}, x_{17}, x_{18}, x_{19}, 0, 0, 0, 0\}.$$

2) Вычислить $X_i(k) = V_i(k)(V_i^*(k) + (-1)^k V_{i+1}(k))$, причем

$$X_4(k) = V_4(k)V_4^*(k).$$

3) Вычислить $X[k] = \sum_{i=1}^4 X_i[k]$, $k \in [0, 7]$.

4) Вычислить $r(m) = \text{ОДПФ}\{X(k)\}$.

5) Получить оценку СПМ, выполнив ДПФ от $r(m)\omega(m)$.

Количество преобразований, которое необходимо выполнить, равно $K+1$, где K – количество интервалов.

В заключение необходимо отметить, что в быстром алгоритме корреляционного метода применяют БПФ, причем необходимую длину последовательности получают дополнением ее нулями.

С увеличением числа значений автокорреляции корреляционный метод дает асимптотически несмещенные оценки СПМ. Блэкман и Тьюки рекомендовали использовать такое число значений оценки автокорреляций, чтобы стандартное отклонение оценки СПМ не превышало $1/3$ ее среднего значения, т. е. число оцениваемых значений АКП должно быть равно примерно 10 процентам числа имеющихся отсчетов данных.

Для того чтобы свести смещение оценок взаимной СПМ к минимуму, следует совместить ВКП и корреляционное окно так, чтобы пики взаимной корреляции соответствовали точке нулевого временного сдвига. Тогда статистические характеристики оценки взаимной СПМ будут аналогичны статистическим характеристикам автоспектральной оценки.

5.3. Критерии качества классической спектральной оценки

Статистический показатель качества оценки определяется как отношение дисперсии оценки СПМ к квадрату математического ожидания

этой оценки

$$Q = \frac{\text{var}\{\hat{P}(f)\}}{[\mathcal{E}\{\hat{P}(f)\}]^2}. \quad (5.20)$$

Показатель качества Q – это, по сути дела, инвертированное отношение сигнал/шум, которое непосредственно связано со статистической устойчивостью спектральной оценки. Значения Q , много меньшие единицы, соответствуют гладким спектральным оценкам с малыми флуктуациями (малой дисперсией). Значения Q , много большие единицы, соответствуют весьма зашумленным спектральным оценкам с большой дисперсией.

Существует очень важное соотношение между следующими величинами.

B_s – *эффективная статистическая ширина полосы*, характеризует ширину полосы эквивалентного прямоугольного окна с тем же значением отношения дисперсии к квадрату среднего значения на его выходе, как и у используемого окна $\Omega(f)$ в том случае, когда на вход подан белый шумовой процесс, и равна

$$B_s = \frac{\left[\int_{-1/2T}^{1/2T} \Omega(f) df \right]^2}{\int_{-1/2T}^{1/2T} \Omega^2(f) df}. \quad (5.21)$$

Полоса B_s связана с эквивалентной детерминированной шириной полосы B_e соотношением $B_s = aB_e$, где коэффициент a равен примерно 0.8 для прямоугольного окна, 1.3 для окна Ханна и 1.4 для окна Хэмминга.

$T_e = NT$ секунд – *длительность* временного сигнала.

Если значение произведения $T_e B_s$ выбрано значительно большим единицы, то для всех классических спектральных оценок будет выполняться соотношение

$$QB_s T_e \approx 1. \quad (5.22)$$

Произведение "устойчивость×длительность×ширина полосы" устанавливает взаимосвязь между тремя фундаментальными параметрами, от которых зависят характеристики спектральных оценок. Для данной записи данных продолжительностью T_e невозможно получить оценки, которые одновременно обладают высоким разрешением (малыми значениями B_s) и высокой устойчивостью (малыми значениями Q). Заметим, что увеличение числа отсчетов за счет увеличения частоты дискретизации при неизменном интервале T_e не будет влиять на максимально достижимое разрешение, так как на него влияет

только длина записи данных, но не число отсчетов.

Если неизвестно, как выбирать компромиссное решение относительно устойчивости и разрешения оценок, то рекомендуется начать с выбора максимальной устойчивости, а это значит – с минимального разрешения, а затем постепенно увеличивать длину сегментов (в случае периодограммы) или число корреляционных членов (в случае коррелограммы). Следует продолжать эту процедуру до тех пор, пока достигнутое разрешение не позволит передать наиболее тонкие и важные детали спектра. Если информация о важности деталей отсутствует, то следует остановиться тогда, когда дальнейшее увеличение длины сегментов данных или величины временных сдвигов к существенным изменениям в спектральных оценках уже не приводит.

5.4. Комбинированные периодограммно-коррелограммные оценки

Основной комбинированный метод состоит из четырех этапов.

1) Вычисляется обычная периодограмма Уэлча $P_W(f)$ с произвольным выбором перекрытия сегментов S и окна данных $w(n)$.

2) Вычисляется ОДВПФ для $P_W(f)$ с целью получения симметричной оценки АКФ $r_{xx}^w(m)$ для $2D+1$ временных сдвигов. Оценка АКФ $r_{xx}^w(m)$ охватывает значения $|m| \leq D$ вследствие воздействия АКФ окна данных $\Phi(m) = w(n) \otimes w(n)$.

3) Эта оценка обрабатывается с помощью действительного симметричного корреляционного окна $\omega(m)$ нечетной длины $2L+1$ и получается взвешенная оценка АКФ:

$$\hat{r}_{xx}^c(m) = \omega(m) \hat{r}_{xx}^w(m).$$

Ширина этого корреляционного окна выбирается меньше ширины автокорреляционной функции $\Phi(m)$ окна данных ($L < D$).

4) Вычисляется ДВПФ для $r_{xx}^c(m)$, что и дает искомую оценку СПМ:

$$\hat{P}_c(f) = \hat{P}_w(f) \otimes \Omega(f),$$

где $\Omega(f)$ – преобразование Фурье корреляционного окна $\omega(m)$.

Комбинированное временное и корреляционное взвешивание позволяют управлять уровнем боковых лепестков. Устойчивость оценки обеспечивается за счет усреднения по сегментам и частотного сглаживания (корреляционного взвешивания).

Среднее значение оценки СПМ

$$\mathcal{E}\{\hat{P}_c(f)\} = \hat{P}_{xx}(f) \otimes \Omega_e(f),$$

где $\Omega_e(f)$ – преобразование Фурье эффективного корреляционного окна $\omega_e(m)$:

$$\omega_e(m) = \omega(m)\Phi(m)/U = \omega(m)\Phi(m)/\Phi(0), \quad \Omega_e(f) = \Omega(f)\left|W(f)^2\right|.$$

Если эффективное корреляционное окно выбрано заранее, тогда при $L < D$ окно данных можно выбирать произвольно, так как его влияние может быть скомпенсировано выбором корреляционного окна следующего вида:

$$\omega(m) = \frac{\omega_e(m)}{\Phi(m)/U}, \quad |m| \leq L.$$

Чтобы оценка $r_{xx}^c(0)$ была несмещенной оценкой полной мощности, нормируют корреляционные окна так, чтобы $\omega_e(0) = 1$, $\omega(0) = 1$.

Для получения с помощью этого метода оценки Блэкмана – Тьюки (крайний коррелограммный частный случай) используем один сегмент ($P = 1$), корреляционное окно шириной много меньше числа отсчетов данных ($L \ll D = N$) и прямоугольное окно данных ($w(n) = 1$, $n = 0, \dots, N-1$). Наттол показал, что в этом случае $QT_eB_S \gg 1$.

Для получения периодограммы Уэлча (крайний периодограммный частный случай) используем $P = 2N/D - 1$ сегментов, 62 %-ное перекрытие ($S = 0,38D$), окно Ханна данных и не применяем корреляционного взвешивания ($\omega(m) = 1$, $|m| \leq L = D$). И в этом случае $QT_eB_S \gg 1$. При использовании 50 %-ного перекрытия $QT_eB_S \gg 1,08$.

Очень интересен частный случай, называемый **методом Наттола – Картера**, при котором используются прямоугольное окно данных ($w(n) = 1$) без перекрытия ($S = D$), число сегментов, вдвое меньшее, чем для периодограммы Уэлча, и корреляционное окно, подбираемое для получения эффективного корреляционного окна Ханна. Здесь $QT_eB_S \gg 1$, метод имеет те же статистические характеристики, что и метод Уэлча. Однако вычислительные затраты для него вдвое ниже, чем для метода Уэлча, так как БПФ в нем вычисляется для вдвое меньшего числа сегментов, к тому же в нем отсутствуют операции умножения, связанные с обработкой сегментов с помощью окна данных.

6. АЛГОРИТМЫ СПЕКТРАЛЬНОГО АНАЛИЗА В ОГРАНИЧЕННОМ СЕКТОРЕ Z-ПЛОСКОСТИ

6.1. Алгоритм анализа с использованием БПФ

Пусть заданы N отсчетов сигнала, и необходимо найти z -преобразование в точках, расположенных на дуге окружности в z -плоскости. Пусть $N = 64$, а количество искоемых спектральных отсчетов, положение которых показано на рис. 6.1, а, равно 16. В данном случае выражение для z -преобразования имеет следующий вид:

$$X(z_k) = \sum_{n=0}^{N-1} x(n)z_k^{-n}, \quad k = 0, 1, \dots, 15, \quad \text{где} \quad z_k = re^{j[\theta + 2\pi k / N]}, \quad (6.1)$$

так что

$$X(z_k) = \sum_{n=0}^{N-1} [x(n)r^{-n}e^{-jn\theta}]e^{-j2\pi k / N}. \quad (6.2)$$

Из формулы (6.2) следует, что после предварительного умножения $x(n)$ на $r^{-n}e^{-jn\theta}$ задача расчета z -преобразования сводится к задаче спектральных измерений на дуге единичной окружности, которая изображена на рис. 6.1, б.

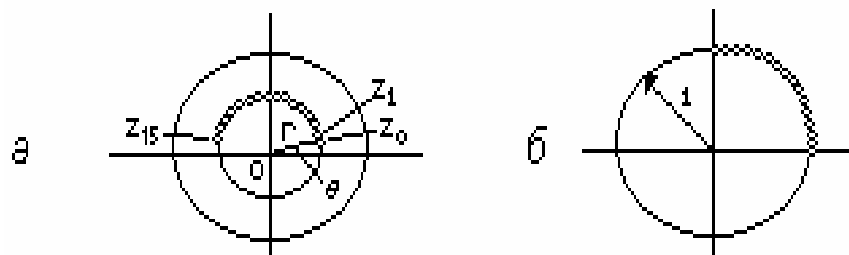


Рис. 6.1: а – круговой сектор в z -плоскости; б – преобразование кругового сектора в дугу окружности с помощью предварительного умножения массива сигнала

Для эффективности расчета спектра методом БПФ точки z_k следует перераспределить таким образом, чтобы они равномерно располагались на единичной окружности. Этого можно достичь, прореживая отсчеты сигнала и комбинируя затем получающиеся спектры. Перечислим последовательность выполняемых при этом операций.

- 1) Вычислить 16-точечное БПФ отсчетов сигнала с номерами 0, 4, 8, ..., 60, что даст $X_0(k)$.
- 2) Вычислить 16-точечное БПФ отсчетов с номерами 1, 5, 9, ..., 61, что даст $X_1(k)$.
- 3) Повторить то же для отсчетов 2, 6, 10, ..., 62 и 3, 7, 11, ..., 63, что даст

$X_2(k)$ и $X_3(k)$ соответственно.

4) Сложить коэффициенты БПФ по формуле

$$X(k) = X_0(k) + W_N^k X_1(k) + W_N^{2k} X_2(k) + W_N^{3k} X_3(k), \quad N = 14, \quad W_N = e^{-j2\pi/N}, \quad (6.3)$$

где $k \in [0, 15]$.

Этот же результат можно получить иначе, вычислив 64-точечное БПФ и взяв первые 16 коэффициентов БПФ. В этом случае потребовалось бы выполнить $(N/2)\log_2 N = 192$ базовых операций, тогда как при использовании четырех 16-точечных преобразований их нужно было бы 128. Таким образом, в зависимости от параметров системы можно использовать по крайней мере два различных метода вычисления спектра вдоль дуги в z -плоскости, изображенной на рис. 6.1, а.

6.2. Алгоритм Блестейна

Алгоритм БПФ позволяет существенно уменьшить время вычисления N -точечного ДПФ, но при условии, что N раскладывается на простые сомножители. Существуют другие эффективные алгоритмы расчета ДПФ последовательностей (также $N\log_2 N$ операций). Один из них, называемый алгоритмом Блестейна, дает способ вычисления ДПФ N -точечной последовательности с помощью весовой обработки выходных отсчетов на выходе ЛЧМ-фильтра (его импульсная характеристика имеет вид колебания с линейной частотной модуляцией). Он применяется при любых N и основан на цифровой фильтрации, эквивалентной вычислению ДПФ.

Рассмотрим цифровой фильтр с импульсной характеристикой $h(n)$ вида

$$h(n) = \begin{cases} 0, & n < 0, \\ e^{j\pi n^2/N}, & 0 \leq n \leq 2N-1, \\ 0, & n > 2N. \end{cases} \quad (6.4)$$

Фильтр с такой импульсной характеристикой обычно называют **ЛЧМ-фильтром** из-за его сходства с аналоговым фильтром, согласованным с ЛЧМ-сигналом. При подаче на его вход N -точечной последовательности $x(n)$ (отсчеты $x(n)$ отличны от нуля только на интервале $[0, N-1]$) выходная последовательность $y(n)$ ($n \in [N, 2N-1]$) будет равна

$$y(n) = \sum_{r=0}^{N-1} x(r)h(n-r). \quad (6.5)$$

Введя в эту формулу новую переменную $k = n - N$, получим:

$$y(k) = \sum_{r=0}^{N-1} x(r)h(k + N - r) = \sum_{r=0}^{N-1} x(r)e^{j(\pi/N)(k+N-r)^2}, \quad (6.6)$$

откуда

$$y(k) = e^{j\pi k^2/N} e^{j\pi N} \sum_{r=0}^{N-1} \underbrace{[x(r)e^{j\pi r^2/N}]_{w(r)}}_{w(r)} e^{-j(2\pi/N)rk}. \quad (6.7)$$

Из формулы (6.7) следует, что выходная последовательность $y(k)$ равна взвешенным с весами $e^{j\pi k^2/N} e^{j\pi N}$ отсчетам ДПФ N -точечной последовательности вида $x(r)e^{j\pi r^2/N}$. Следовательно, чтобы получить N -точечное ДПФ последовательности $f(n)$, необходимо в соответствии с (6.7) подать на вход фильтра последовательность $x(n) = f(n)e^{j2\pi n/N}$, а отсчеты выходной последовательности $y(n)$ с номерами $N \leq n \leq 2N - 1$ умножить на весовые коэффициенты $e^{-j\pi(n-N)^2/N} e^{j\pi N}$.

Все эти операции, необходимые для вычисления N -точечного ДПФ последовательности $f(n)$, представлены на рис. 6.2.

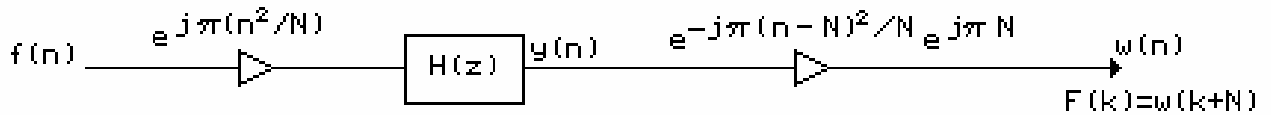


Рис. 6.2

Для некоторых значений N (а именно, если N равно квадрату целого числа) количество операций, используемых в цифровом фильтре, может быть пропорционально $N^{1.5}$.

Основной смысл алгоритма Блустейна состоит в том, что с его помощью показана возможность получения ДПФ последовательности посредством ее фильтрации ЛЧМ-фильтром.

6.3. Алгоритм z -преобразования с использованием ЛЧМ-фильтрации

Алгоритм z -преобразования с ЛЧМ-фильтрацией является вторым эффективным алгоритмом вычисления z -преобразования конечной последовательности вдоль определенного контура на z -плоскости. Этот алгоритм также используется для вычисления ДПФ последовательности даже при простых N .

Пусть заданная N -точечная последовательность $\{x(n)\}$, $0 \leq n \leq N-1$, имеет z -преобразование

$$X(z) = \sum_{n=0}^{N-1} z^{-n}. \quad (6.8)$$

ДПФ $X(k)$ заданной последовательности по определению равно

$$X(k) = X(z) \Big|_{z=e^{j(2\pi/N)k}}, \quad k=0,1,\dots,N-1. \quad (6.9)$$

Используя алгоритм z -преобразования с ЛЧМ-фильтрацией, можно найти z -преобразование в соответствии с формулой (6.3) вдоль контура более общего вида:

$$z_k = AW^{-k}, \quad k=0,1,\dots,M-1. \quad (6.10)$$

Здесь M – произвольное целое число (не обязательно равное N), а A и W – произвольные комплексные числа, равные

$$A = A_0 e^{j2\pi\theta_0}, \quad (6.11)$$

$$W = W_0 e^{j2\pi\varphi_0}. \quad (6.12)$$

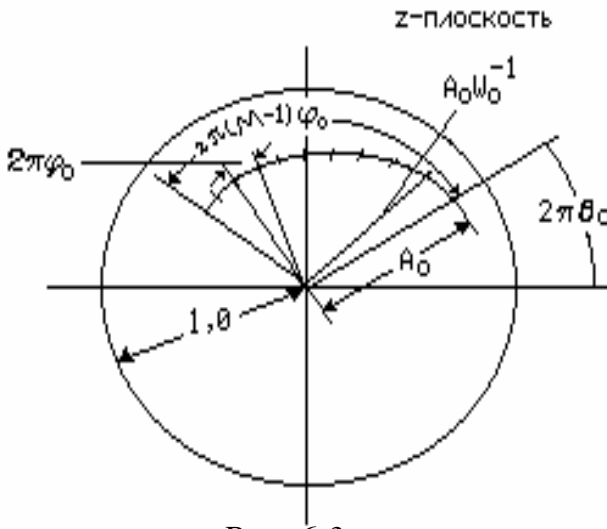


Рис. 6.3

На рис. 6.3 изображен контур в z -плоскости, удовлетворяющий формуле (6.10), т.е. контур расчета z -преобразования с ЛЧМ-фильтрацией в z -плоскости, и графически представлены величины A_0 , W_0 , φ_0 и θ_0 .

Как следует из рис. 6.3 и формул (6.10) – (6.12), при $A=1$, $M=N$ и $W=e^{j2\pi/N}$ имеем $z_k=e^{j2\pi n/N}$, т.е. точки вдоль контура будут соответствовать положениям спектральных отсчетов при вычислении N -точечного ДПФ последовательности.

Постоянная W_0 определяет скорость ухода контура внутрь окружности радиуса A_0 или наружу от нее. При $W_0>1$ контур будет скручиваться, а при $W_0<1$ – раскручиваться. Несколько проще интерпретировать скручивание контура в z -плоскости, рассматривая эквивалентный контур в s -плоскости. Введя подстановку $z=e^{sT}$, получим $z_0=A=e^{s_0T}$ или

$$s_0 = \ln A/T = \sigma_0 + j\omega_0 = \ln A_0/T + j2\pi\theta_0/T.$$

Далее, поскольку $z_k = e^{s_k T} = AW^{-k}$, то

$$s_k = \ln(AW^{-k})/T = \ln A/T + \ln(W^{-k})/T = s_0 - k \ln(w)/T.$$

На рис. 6.4, *в* в s -плоскости изображен контур, эквивалентный контуру, показанному на рис. 6.1, *а*. Таким образом, спиралеобразные контуры в z -плоскости соответствуют прямым линиям в s -плоскости, причем скорость закручивания определяет угол наклона прямой в s -плоскости.

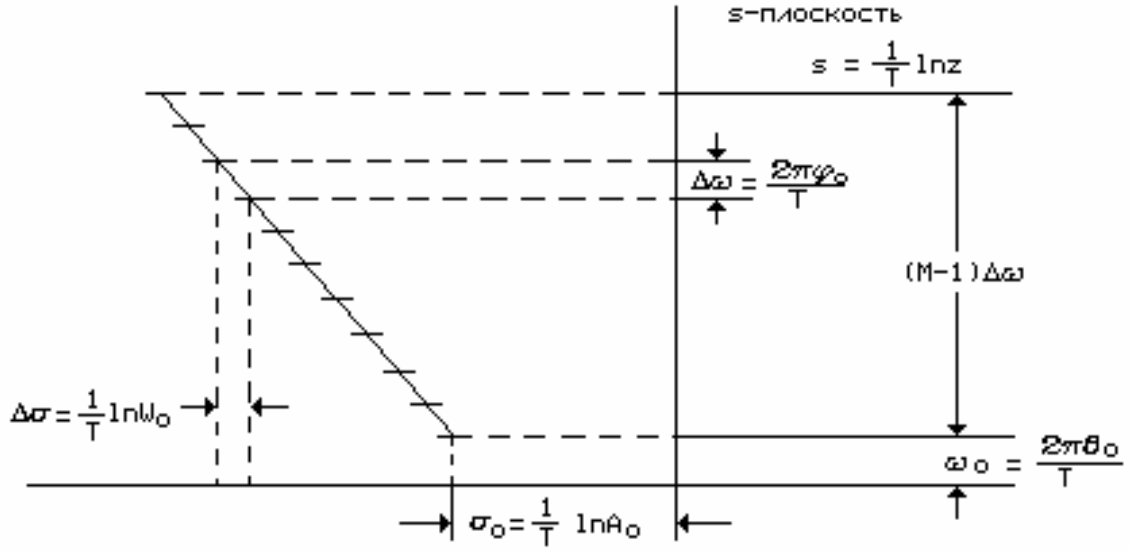


Рис. 6.4. Контур расчета z -преобразования с ЛЧМ-фильтрацией в s -плоскости

Покажем теперь, как эффективно рассчитывать z -преобразование последовательности согласно формуле (6.8) вдоль контура (6.10). Обозначив через X_k искомые значения z -преобразования при $z = z_k$, т. е.

$$X_k = \sum_{n=0}^{N-1} x(n)z_k, \quad (6.13)$$

и учитывая (6.10), находим

$$X_k = \sum_{n=0}^{N-1} x(n)A^{-n}W^{nk}, \quad k = 0, 1, \dots, M-1. \quad (6.14)$$

Подстановка в формулу (6.14) выражения

$$nk = \frac{n^2 + k^2 - (k-n)^2}{2} \quad (6.15)$$

дает

$$\begin{aligned}
 X_k &= \sum_{n=0}^{N-1} x(n) A^{-n} W^{[n^2+k^2-(k-n)^2]/2} = \\
 &= \sum_{n=0}^{N-1} [x(n) A^{-n} W^{n^2/2}] (W^{k^2/2}) W^{-(k-n)^2/2}
 \end{aligned}
 \quad (6.16)$$

или

$$X = W^{k^2/2} \sum_{n=0}^{N-1} y(n) v(k-n), \quad (6.17)$$

где

$$y(n) = x(n) A^{-n} W^{n^2/2}, \quad (6.18)$$

$$v(n) = W^{n^2/2}. \quad (6.19)$$

Из (6.17) видно, что значения X_k можно найти, рассчитав взвешенную свертку последовательностей $y(n)$ и $v(n)$; последнее можно эффективно провести, используя алгоритм быстрой свертки на основе БПФ. На рис. 6.5 приведены все основные операции алгоритма z -преобразования с ЛЧМ-фильтрацией.



Рис. 6.5

Прежде чем перейти к детальному описанию алгоритма z -преобразования с ЛЧМ-фильтрацией, перечислим преимущества этого алгоритма по сравнению со стандартным алгоритмом БПФ.

1) Число отсчетов входной последовательности N не обязательно должно быть равно числу точек M , в которых рассматривается z -преобразование.

2) Ни N , ни M могут не быть составными числами; в частности, они могут быть простыми.

3) Угловое смещение точек z_k может быть произвольным, так что и частотное разрешение может быть любым.

4) Контур не обязательно должен быть окружностью в z -плоскости. В частности, использование спиралевидного контура в анализаторе речи дает некоторые преимущества.

5) Начальная точка контура в z -плоскости произвольна. Это свойство особенно полезно при анализе в узкой полосе, когда высокое частотное разрешение (п. 3) сочетается с произвольной начальной частотой.

6) При $A = 1$, $M = N$ и $W = e^{-j2\pi/N}$ z -преобразования с ЛЧМ-фильтрацией выполняются в несколько следующих этапов, иллюстрируемых с помощью рис. 6.6.

Этап 1. Выбор L – наименьшего целого числа, большего или равного $(N+M-1)$, которое можно было бы использовать в обычном алгоритме БПФ. Величина L определяет размер преобразований, которые выполняются при расчете быстрой свертки в системе, изображенной на рис. 6.6.

Этап 2. Формирование L -точечной последовательности $y(n)$ вида

$$y(n) = A^{-n} W^{n^2/2} x(n),$$

при $n = 1, 0, \dots, N-1$, и равной нулю для остальных точек диапазона, т. е. $n \in [N, L-1]$. Типичные последовательности $x(n)$ и $y(n)$ изображены на рис. 6.6, а и б.

Этап 3. Расчет L -точечного ДПФ последовательности $y(n)$ с помощью обычного алгоритма БПФ, результаты расчета обозначаются через Y_r . На рис. 6.6, в изображена типичная последовательность Y_r .

Этап 4. Формирование L -точечной последовательности $v(n)$ по формуле

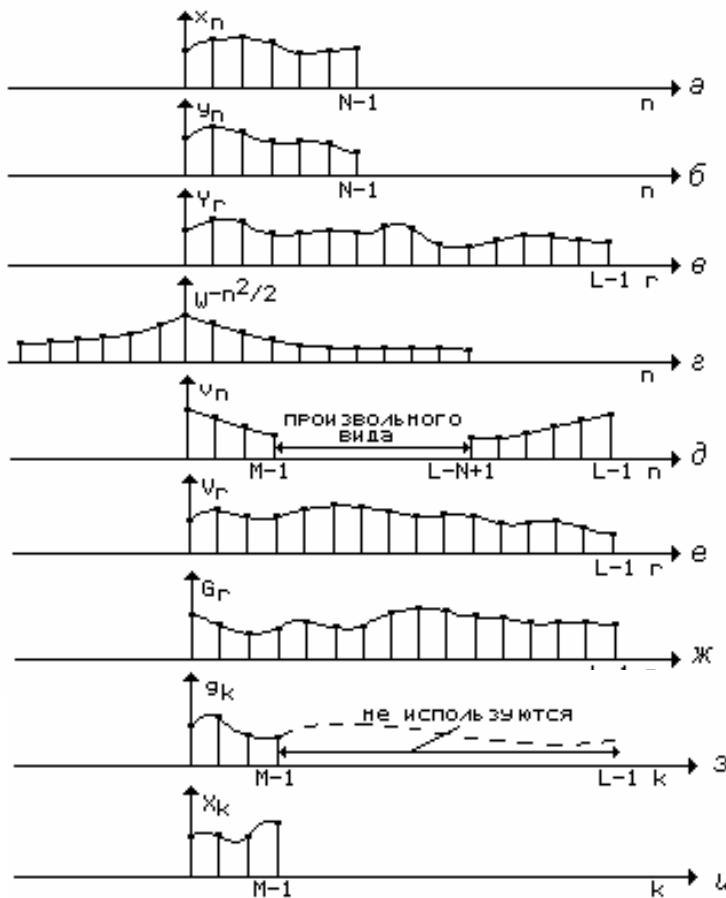


Рис. 6.6

$$v(n) = \begin{cases} W^{-n^2/2}, & 0 \leq n \leq M-1, \\ W^{-(L-n)^2/2}, & L-N+1 \leq n < L, \\ \text{произвольно} & \text{при других } n. \end{cases}$$

Последовательности $W^{-n^2/2}$ и $v(n)$ изображены на рис. 6.6, г и д соответственно.

Этап 5. Расчет L -точечного ДПФ последовательности $v(n)$, результаты

которого обозначаются через V_r (рис. 6.6, *е*).

Этап 6. Почленное умножение последовательностей V_r и Y_r , что дает $G_r = V_r Y_r$ (рис. 6.6, *ж*).

Этап 7. Расчет обратного ДПФ последовательности G_r , результаты которого обозначаются через g_k (рис. 6.6, *з*).

Этап 8. Умножение g_k на $W^{k^2/2}$, что дает X_k , т. е.

$$X_k = g_k W^{k^2/2}, \quad k = 0, 1, \dots, M-1.$$

Значения g_k при $k \geq M$ не имеют смысла и отбрасываются. Последовательность X_k изображена на рис. 6.6, *и*.

Остается показать, что алгоритм z -преобразования с ЛЧМ-фильтрацией эффективнее прямого вычисления z -преобразования заданной последовательности. Для этого достаточно подсчитать число комплексных умножений, используемых при выполнении z -преобразования с ЛЧМ-фильтрацией. Подсчет показывает, что основными операциями при вычислении z -преобразования с ЛЧМ-фильтрацией являются три (иногда два) L -точечных БПФ, поэтому время вычисления будет пропорционально $L \log_2 L$ (напомним, что $L \leq N + M - 1$), но только коэффициент пропорциональности будет в два-три раза больше, чем при БПФ. Таким образом, рассматриваемый алгоритм является эффективным инструментом расчета z -преобразования последовательности вдоль определенного контура в z -плоскости.

7. СПЕКТРАЛЬНЫЙ АНАЛИЗ, ОСНОВАННЫЙ НА ЛИНЕЙНОМ МОДЕЛИРОВАНИИ ИССЛЕДУЕМОГО СИГНАЛА

В разделах 3 – 5 СПМ была определена как дискретно-временное преобразование Фурье (ДВПФ) бесконечной автокорреляционной последовательности (АКП). Это соотношение между СПМ и АКП можно рассматривать как *непараметрическое* описание статистик второго порядка случайного процесса. Частным случаем такого описания можно считать оценку СПМ в одной точке z -плоскости (раздел 6).

К *параметрическому* описанию статистик второго порядка можно прийти, рассматривая модель временного ряда, соответствующего анализируемому случайному процессу. В этом случае СПМ модели временного ряда будет прежде всего некоторой функцией параметров этой модели, а не АКП. Методы, рассматриваемые в данном разделе, применимы только к частному классу моделей, возбуждаемых белым шумовым процессом и обладающих рациональной системной функцией. Этот класс включает модель авторегрессионного процесса (АР), модель процесса скользящего среднего (СС) и модель процесса авторегрессии-скользящего среднего (АРСС). Выходные процессы моделей этого класса имеют спектральные плотности мощности, которые полностью описываются с помощью параметров модели и дисперсии белого шумового процесса. Значения этих параметров и дисперсии белого шума определяются из выборки исследуемого процесса при помощи методов, основанных на линейном моделировании.

Причины применения параметрических моделей: возможность получения на их основе более точных оценок СПМ и более высокого спектрального разрешения, чем это возможно с помощью классических методов спектрального оценивания. Степень улучшения разрешения и повышения достоверности спектральных оценок определяется соответствием выбранной модели анализируемому процессу и возможностью аппроксимации исходных данных или АКП (известной или оцененной по этим данным) с помощью нескольких параметров модели.

Методика спектрального оценивания, основанная на моделировании исследуемого процесса, состоит из трех этапов: выбора параметрической модели фильтра; определения параметров модели; оценки СПМ модели процесса.

7.1. Модели случайных процессов

Многие последовательности, представляющие собой сумму детерминированных и случайных последовательностей, могут быть достаточно хорошо аппроксимированы выходным сигналом линейного дискретного фильтра с рациональной системной функцией, возбуждаемого белым шумом с нулевым средним и дисперсией σ^2 .

При решении задачи моделирования требуется обеспечить устойчивость и (или) каузальность применяемого фильтра, чтобы получить однозначно определенную модель.

Существует три типа моделей, описывающих такой фильтр.

1) Самая общая модель – *модель авторегрессии-скользящего среднего (АРСС)* описывается линейным разностным уравнением с комплексными коэффициентами

$$x(n) = - \sum_{k=1}^p a(k)x(n-k) + \sum_{k=0}^q b(k)u(n-k) = \sum_{k=0}^{\infty} h(k)u(n-k), \quad (7.1)$$

где $x(n)$ – выходная последовательность моделирующего фильтра, $u(n)$ – возбуждающая шумовая последовательность. Без потери общности можно положить $b(0) = 1$ и суммировать с $k = 1$, так как $u(n)$ всегда можно соответствующим образом промасштабировать, чтобы учесть любой коэффициент усиления фильтра.

Функциональная схема АРСС-модели приведена на рис. 7.1, а. Здесь параметры $a(k)$ характеризуют авторегрессионную часть модели, а $b(k)$ – часть, соответствующую скользящему среднему.

Системная функция фильтра определяется выражением

$$H(z) = \frac{B(z)}{A(z)} = \left(1 + \sum_{k=1}^q b(k)z^{-k} \right) / \left(1 + \sum_{k=1}^p a(k)z^{-k} \right) = 1 + \sum_{k=1}^{\infty} h(k)z^{-k}. \quad (7.2)$$

Предполагается, что нули полиномов $A(z)$ и $B(z)$ расположены внутри единичной окружности в z -плоскости, что гарантирует принадлежность функции $H(z)$ устойчивому фильтру. Подставив сюда $z = \exp(j2\pi fT)$, после масштабирования интервалом отсчетов T получим СПМ для АРСС-модели:

$$P_{\text{АРСС}}(f) = T\sigma^2 \left| \frac{B(f)}{A(f)} \right|^2, \quad (7.3)$$

где полиномы $A(f)$ и $B(f)$ определяются выражениями

$$A(f) = 1 + \sum_{k=1}^p a(k) \exp(-j2\pi fkT), \quad B(f) = 1 + \sum_{k=1}^q b(k) \exp(-j2\pi fkT),$$

а векторы параметров \mathbf{a} и \mathbf{b} имеют вид

$$\mathbf{a} = (1 \quad a(1) \quad . \quad . \quad . \quad a(p))^T, \quad \mathbf{b} = (1 \quad b(1) \quad . \quad . \quad . \quad b(q))^T.$$

СПМ АРСС-процесса вычисляется в диапазоне частот $-1/2T \leq f \leq 1/2T$.

В литературе часто используется обозначение $\text{АРСС}(p, q)$ – АРСС-модель с АР-параметрами порядка p и СС-параметрами порядка q . Задание АР-, СС-параметров и дисперсии белого шума σ^2 полностью характеризует СПМ АРСС-процесса $x(n)$.

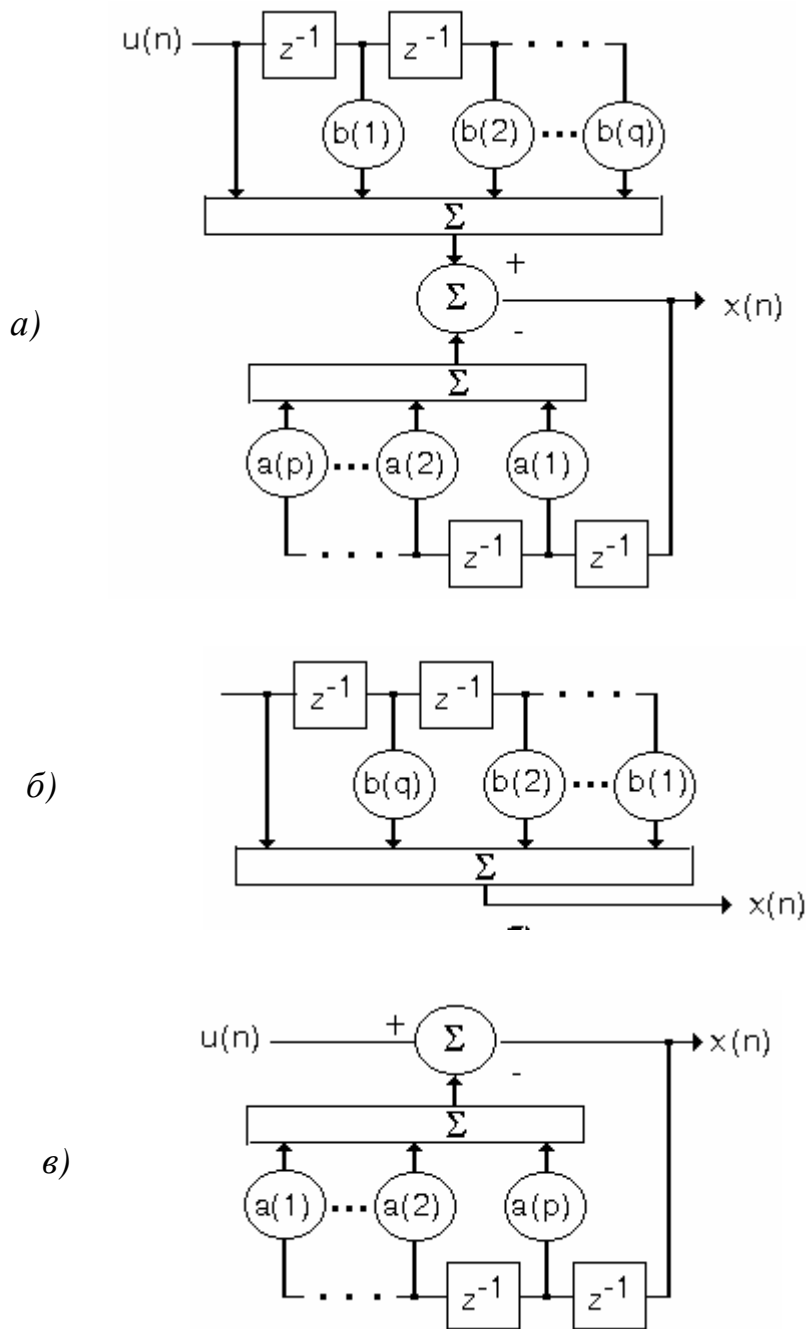


Рис. 7.1

2) **Модель процесса скользящего среднего порядка q** , или $CC(q)$ -модель, получается из $ARCC$ -модели при нулевых (за исключением $a(0) = 1$) AR -параметрах:

$$x(n) = \sum_{k=1}^q b(k)u(n-k) + u(n). \quad (7.4)$$

СПМ CC -процесса

$$P_{CC}(f) = T\sigma^2 |B(f)|^2. \quad (7.5)$$

Функциональная схема СС-процесса показана на рис. 7.1, б.

3) **Модель авторегрессионного процесса порядка p** , или $AR(p)$ -модель получается из АРСС-модели при нулевых (за исключением $b(0) = 1$) СС-параметрах:

$$x(n) = - \sum_{k=1}^p a(k)x(n-k) + u(n). \quad (7.6)$$

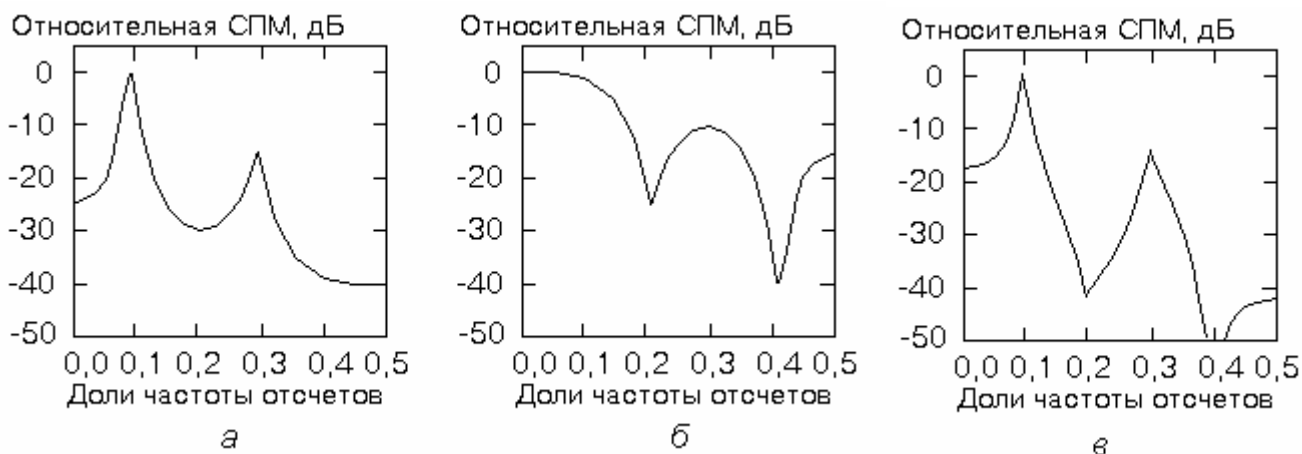
СПМ АР-процесса

$$P_{AR}(f) = \frac{T\sigma^2}{|A(f)|^2}. \quad (7.7)$$

Функциональная схема АР-модели приведена на рис. 7.1, в.

Рис. 7.2: а - АР-спектр, б - СС-спектр, в - АРСС-спектр

На рис. 7.2 показаны спектры типичных АРСС-, СС- и АР-процессов. Выбор одной из трех моделей требует некоторых предварительных сведений о возможной форме спектральной оценки. Если необходимы спектры с острыми пиками, но без глубоких нулей, то подходит АР-модель. Если, наоборот, необходимы спектры с глубокими нулями, но без острых пиков, лучше взять СС-модель. АРСС-модель пригодна в обоих случаях. В спорных ситуациях следует использо-



вать ту модель, которая имеет наименьшее число параметров. Этот **принцип экономии** был предложен Боксом и Дженкинсом и основан на том факте, что получить оценки с хорошими статистическими свойствами можно, как правило, в том случае, если число оцениваемых параметров минимально. В дальнейшем речь пойдет об АР-моделях, что обусловлено двумя причинами:

1) АР-спектры имеют острые пики, а это связано со спектральным разрешением;

2) оценки АР-параметров можно получить как решения линейных уравнений (в отличие от СС-параметров), следствием чего явилось множество хорошо разработанных алгоритмов их нахождения.

Существует **теорема декомпозиции**: Если задана АР-, СС- или АРСС-модель с конечным числом параметров, то ее можно выразить через две другие

модели.

АРСС-и СС-процессы можно записать с помощью одной АР-модели в общем случае бесконечного порядка.

Так, пусть

$$C(z) = 1 + \sum_{k=1}^{\infty} c(k)z^{-k}$$

– полином знаменателя АР(∞)-модели, эквивалентной заданной АРСС(p, q)-модели. Параметры АР-модели находятся с помощью равенства системных функций АРСС(p, q)- и АР(∞)-моделей:

$$\frac{B(z)}{A(z)} = 1 / C(z), \quad C(z)B(z) = A(z).$$

Отсюда параметры АР(∞)-модели с начальными условиями $c(-1) = \dots = c(-q) = 0$:

$$c(n) = \begin{cases} 1, & n = 0; \\ - \sum_{k=1}^q b(k)c(n-k) + a(n), & 1 \leq n \leq p; \\ - \sum_{k=1}^q b(k)c(n-k) + a(n), & n > p \end{cases} \quad (7.8)$$

Замечание. Рассмотрим z -преобразование АР-процесса

$$P_{AP}(z) = \sigma^2 H(z)H^*(1/z^*) = \frac{\sigma^2}{A(z)A^*(1/z^*)}.$$

Полюсы $P_{AP}(z)$ будут комплексно-сопряженными взаимно обратными парами. С точки зрения оценки СПМ можно не требовать, чтобы полином $A(z)$ являлся минимально-фазовым полиномом (все нули и полюсы которого лежат в единичной окружности в z -плоскости). Но так как для реализации фильтра нам приходится использовать оценки его коэффициентов, то встает вопрос об устойчивости, и полином $A(z)$ должен являться минимально-фазовым, что может быть сделано посредством инверсного переноса всех его полюсов, расположенных вне единичной окружности, внутрь ее.

7.2. Методы оценки параметров АР-моделей. Блочные алгоритмы

Методы оценки АР-параметров можно разбить на две категории: алгоритмы для обработки **блоков** данных и алгоритмы для обработки **последовательных** данных.

Алгоритмы блочной обработки предназначены для обработки целых блоков накопленных отсчетов данных некоторого фиксированного объема. Блочные методы можно кратко описать как алгоритмы с фиксированным временем, рекурсивные относительно порядка в том смысле, что позволяют рекурсивным образом получать оценки параметров АР-модели более высокого порядка по оценкам параметров модели более низкого порядка. Такие алгоритмы целесообразно применять в тех случаях, когда порядок требуемой модели неизвестен и необходимо испытывать модели различных порядков и сравнивать результаты.

АР(p)-процесс имеет три эквивалентных представления:

- 1) в виде бесконечно протяженной АКП;
- 2) в виде конечной последовательности АР-параметров;
- 3) в виде конечной последовательности коэффициентов отражения.

Коэффициентами отражения называют коэффициенты $k_m = a_m f(m)$, $1 \leq m \leq p$ фильтра линейного предсказания (ЛП) вперед.

Оценка и ошибка линейного предсказания вперед:

$$\hat{x}^f(n) = - \sum_{k=1}^m a^f(k) x(n-k), \quad e^f(n) = x(n) - \hat{x}^f(n), \quad (7.9)$$

где ЛП вперед понимается в том смысле, что результат предсказания для текущего отсчета данных является взвешенной суммой из p предшествующих отсчетов.

Оценка и ошибка ЛП назад:

$$\hat{x}^b(n) = - \sum_{k=1}^m a^b(k) x(n+k), \quad e^b(n) = x(n-m) - \hat{x}^b(n-m), \quad (7.10)$$

где ЛП назад понимается в том смысле, что результат предсказания для текущего отсчета данных является взвешенной суммой из p последующих отсчетов.

Коэффициенты k_m являются параметрами на каждой ступени решетчатого фильтра ЛП ошибки и интерпретируются как физические параметры акустической трубы, используемой в модели речи, или слоистых моделей Земли при обработке сейсмических данных. В структуре решетчатых фильтров одновременно распространяются ошибки ЛП вперед и назад, взаимно ортогональные на каждой ступени фильтра. Решетчатый фильтр мало чувствителен к шуму округления и флуктуациям значений коэффициентов.

Для того чтобы все три отображения АР-процесса были возможны, необходимо выполнить следующие три эквивалентных условия:

- 1) $r_{xx}(0), \dots, r_{xx}(p)$ – положительно-определенная АКП;

2) полином $A(z) = 1 + \sum_{k=1}^p a_p(k)z^{-k}$ – минимально-фазовый;

3) модуль коэффициента отражения $|k_m| < 1$ и дисперсия на m -й ступени $\sigma_m^2 > 0, 1 \leq m \leq p$.

Соответствующие этим представлениям методы оценки параметров представлены на рис. 7.3. В отдельную группу выделены методы, основанные на линейном предсказании по методу наименьших квадратов.

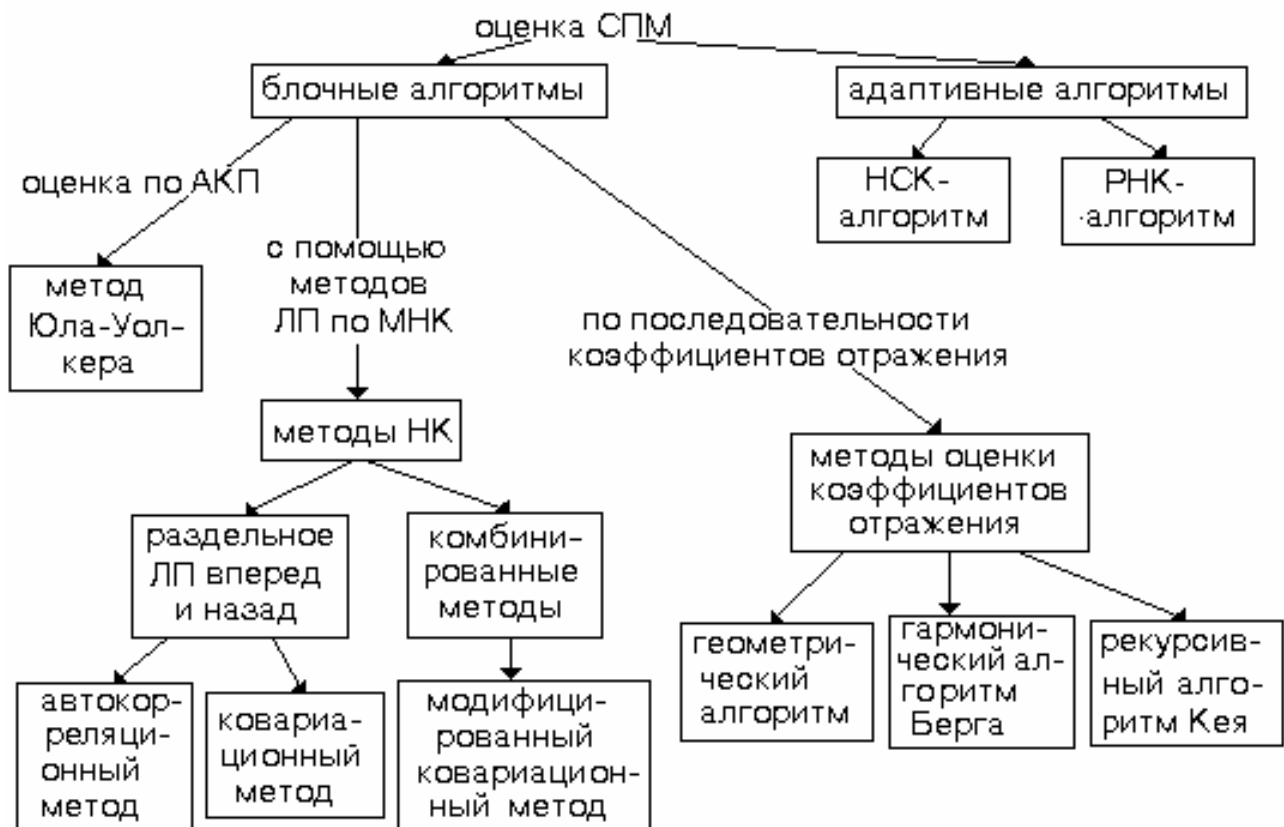


Рис. 7.3

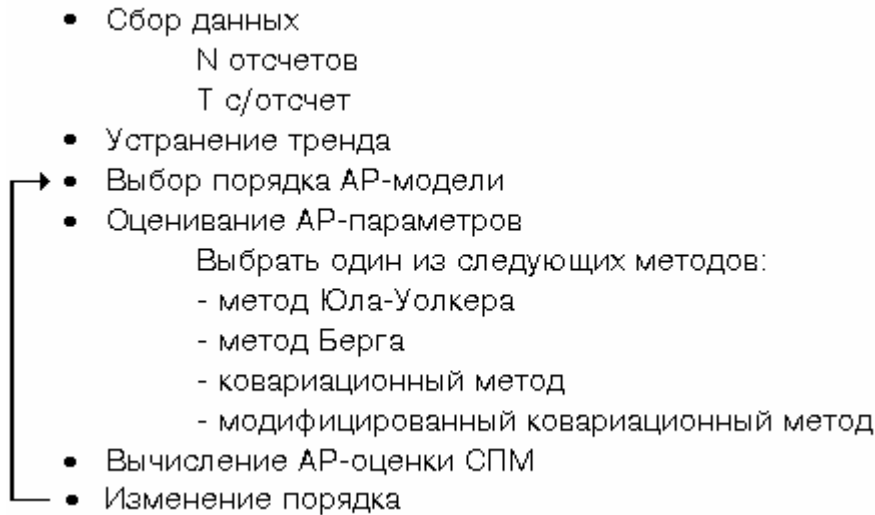


Рис. 7.4

Если анализируется процесс $x(t) = m(t) + n(t)$, где $n(t)$ – стационарный случайный процесс со спектром, ширина которого значительно превышает ширину спектра детерминированной функции $m(t)$, то функция $m(t)$ характеризует отклонение процесса $x(t)$ от стационарности и называется **трендом**.

Блок-схема блочных алгоритмов АР-оценивания СПМ приведена на рис. 7.4.

7.3. Оценка параметров АР-моделей по автокорреляционной последовательности. Метод Юла-Уолкера

АР-модель описывается уравнением (7.6):

$$x(n) = - \sum_{k=1}^p a(k)x(n-k) + u(n) = \sum_{k=0}^{\infty} h(k)u(n-k).$$

Умножая обе его части на $x^*(n-m)$ и усредняя (операция $\langle \rangle$), получим

$$\langle x(n)x^*(n-m) \rangle = - \sum_{k=1}^p a(k) \langle x(n-k)x^*(n-m) \rangle + \langle u(n)x^*(n-m) \rangle$$

или

$$r_{xx}(m) = - \sum_{k=1}^p a(k)r_{xx}(m-k) + r_{ux}(m). \quad (7.11)$$

Взаимную корреляцию $r_{ux}(m)$ найдем из второго равенства (7.6):

$$r_{ux}(m) = \langle u(n+m)x^*(n) \rangle = \sum_{k=0}^{\infty} h^*(k) \langle u(n+m)x^*(n-k) \rangle = \sum_{k=0}^{\infty} h^*(k) r_{uu}(m+k).$$

Так как $u(n)$ – белый шум, то

$$r_{uu}(m+k) = \begin{cases} \sigma^2, & \text{при } k_0 : n+m = n-k_0, \quad k_0 = -m; \\ 0, & k \neq k_0. \end{cases} \quad (7.12)$$

При $m > 0$ $k_0 < 0$, $h(k_0) = 0$ и $r_{ux}(m) = 0$; при $m \leq 0$ $k_0 \geq 0$ $h(k_0) = h^*(-m)$ и $r_{ux}(m) = \sigma^2 h^*(-m)$. При $m < 0$ АКП по определению $r_{xx}(m) = r_{xx}^*(-m)$. Подставляя значения $r_{ux}(m)$ в (7.11), получим

$$r_{xx}(m) = \begin{cases} -\sum_{k=1}^p a(k) r_{xx}(m-k), & m > 0, \\ -\sum_{k=1}^p a(k) r_{xx}(-k) + \sigma^2, & m = 0, \\ r_{xx}^*(-m), & m < 0. \end{cases} \quad (7.13)$$

Это выражение можно записать для $p+1$ значений индекса временного сдвига $0 \leq m \leq p$ в матричной форме

$$\begin{pmatrix} r_{xx}(0) & r_{xx}(-1) & \cdot & \cdot & \cdot & r_{xx}(-p) \\ r_{xx}(1) & r_{xx}(0) & \cdot & \cdot & \cdot & r_{xx}(-p+1) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{xx}(p) & r_{xx}(p-1) & \cdot & \cdot & \cdot & r_{xx}(0) \end{pmatrix} \begin{pmatrix} 1 \\ a(1) \\ \cdot \\ \cdot \\ \cdot \\ a(p) \end{pmatrix} = \begin{pmatrix} \sigma^2 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}. \quad (7.14)$$

Таким образом, если задана АКП для $0 \leq m \leq p$, то АР-параметры можно найти в результате решения уравнений (7.14), которые называются **уравнениями Юла - Уолкера для АР-процесса**.

Процедура оценивания СПМ по АР-параметрам, получающимся в результате решения уравнений (7.14), называется **алгоритмом Юла - Уолкера**. Если задана последовательность данных, то в уравнения (7.14) вместо значений неизвестной автокорреляционной функции подставляются их оценки, смещенные или не смещенные. Гарантию устойчивости АР-фильтра дает использование смещенных автокорреляционных оценок.

Для решения системы (7.14) можно воспользоваться алгоритмом Левинсона [1, с. 111]. Количество необходимых для этого вычислительных операций пропорционально величине p^2 .

СПМ АР-процесса получается следующим образом:

$$P_{\text{АР}}(f) = \frac{T\sigma^2}{|A(f)|^2} = T \sum_{k=-\infty}^{\infty} r_{xx}(k) \exp(-j2\pi f k T). \quad (7.15)$$

Заметим, что значения автокорреляции, соответствующие значениям индекса временного сдвига от 0 до p , однозначно описывают АР(p)-процесс, так как значения автокорреляции при $|k| > p$ получаются рекурсивно (см. формулу (7.13)).

7.4. Оценка параметров АР-моделей по последовательности оценок коэффициентов отражения

Рекурсивное решение уравнений Юла - Уолкера методом Левинсона связывает АР-параметры порядка p с параметрами порядка $p-1$ выражениями [1, с. 239]:

$$a_p(n) = a_{p-1}(n) + k_p a_{p-1}^*(p-n), \quad 1 \leq n \leq p-1, \quad (7.16)$$

а коэффициент отражения определяется по известным значениям АКФ, соответствующим корреляционным сдвигам от 0 до $p-1$:

$$k_p = a_p(p) = \frac{-\sum_{n=0}^{p-1} a_{p-1}(n) r_{xx}(p-n)}{\sigma_{p-1}^2}. \quad (7.17)$$

Рекурсивное выражение для дисперсии возбуждающего белого шума:

$$\sigma_p^2 = \sigma_{p-1}^2 (1 - |k_p|^2), \quad \sigma_0^2 = r_{xx}(0). \quad (7.18)$$

Одна из процедур получения АР-оценки СПМ в том случае, когда имеется некоторый блок отсчетов данных, может быть основана на оценивании коэффициента отражения по этим отсчетам на каждом шаге рекурсии Левинсона. Эта идея реализована в трех следующих методах.

7.4.1. Геометрический алгоритм

Ошибки ЛП вперед и назад определяются соответственно выражениями

$$\begin{aligned} e_p^f(n) &= x(n) + \sum_{m=1}^p a_p^f(m)x(n-m), \\ e_p^b(n) &= x(n-p) + \sum_{m=1}^p a_p^{f*}(m)x(n+m-p). \end{aligned} \quad (7.19)$$

Подставляя сюда значение $a_p(n)$ из (7.16), получаем рекурсивные соотношения

$$\begin{aligned} e_p^f(n) &= e_{p-1}^f(n) + k_p e_{p-1}^b(n-1), \\ e_p^b(n) &= e_{p-1}^b(n-1) + k_p^* e_{p-1}^f(n), \end{aligned} \quad (7.20)$$

которые связывают ошибки предсказания порядка p с ошибками предсказания порядка $p-1$.

Коэффициент отражения можно рассматривать как взятый со знаком минус нормированный коэффициент корреляции между ошибками ЛП вперед и назад с единичным временным сдвигом [1, с. 241]:

$$\begin{aligned} k_p &= - \frac{\langle e_{p-1}^f(n) + k_p e_{p-1}^{b*}(n-1) \rangle}{\langle \sqrt{|e_{p-1}^f(n)|^2} \rangle \langle \sqrt{|e_{p-1}^b(n-1)|^2} \rangle} = \\ &= - \frac{\langle e_{p-1}^f(n) + k_p e_{p-1}^{b*}(n-1) \rangle \cdot \langle e_{p-1}^f(n) + k_p e_{p-1}^{b*}(n-1) \rangle}{\sqrt{\left(\langle \sqrt{|e_{p-1}^f(n)|^2} \rangle \right)^2} \sqrt{\left(\langle \sqrt{|e_{p-1}^b(n-1)|^2} \rangle \right)^2}} = -\sqrt{a \cdot b}. \end{aligned} \quad (7.21)$$

Это выражение определяет *геометрическое среднее* между нормированной корреляцией ошибки предсказания вперед (a) и нормированной корреляцией ошибки предсказания назад (b). Подставляя в (7.21) оценки взаимной корреляции и автокорреляции ошибок предсказания вперед и назад

$$\begin{aligned} \frac{1}{N} \sum_{n=p+1}^N e_{p-1}^f(n) e_{p-1}^{b*}(n-1), \\ \frac{1}{N} \sum_{n=p+1}^N \left| e_{p-1}^f(n) \right|^2, \\ \frac{1}{N} \sum_{n=p+1}^N \left| e_{p-1}^b(n-1) \right|^2, \end{aligned}$$

получим оценку коэффициентов отражения:

$$\hat{k}_p = - \frac{\sum_{n=p+1}^N e_{p-1}^f(n) e_{p-1}^{b*}(n-1)}{\sqrt{\sum_{n=p+1}^N |e_{p-1}^f(n)|^2} \sqrt{\sum_{n=p+1}^N |e_{p-1}^b(n-1)|^2}}. \quad (7.22)$$

Заметим, что так как масштабирующий множитель $1/N$ сократился, то не имеет значения, используются смещенные или несмещенные оценки корреляции. Индексы ошибок предсказания лежат в диапазоне от $n = p + 1$ до $n = N$, так как используются только имеющиеся отсчеты данных. Можно показать, что $|k_p| \leq 1$, что гарантирует устойчивость АР-фильтра.

Алгоритм вычислений таков. Оценки коэффициентов отражения находятся по формуле (7.22) с использованием рекурсий для ошибок предсказания (7.20), начиная с рекурсии нулевого порядка

$$e_0^f(n) = e_0^b(n) = x(n), \quad 1 \leq n \leq N. \quad (7.23)$$

В качестве начального значения дисперсии ошибки ЛП используется оценка

$$\sigma_0^2 = \frac{1}{N} \sum_{n=1}^N |x(n)|^2$$

– мощность, содержащаяся в N отсчетах данных. Параметры модели находятся по оценкам коэффициентов отражения с помощью формул (7.16).

7.4.2. Гармонический алгоритм Берга

Один из самых первых и наиболее известных алгоритмов АР спектрального оценивания был предложен Дж. Бергом. Алгоритм Берга отличается от геометрического алгоритма лишь последним этапом: в нем используется другой тип оценки коэффициента отражения – оценка, определяемая по методу наименьших квадратов. При каждом значении p в нем минимизируется **арифметическое среднее** мощности ошибок ЛП вперед и назад (7.20) (выборочных дисперсий ошибок предсказания):

$$\sigma_p^{2fb} = \frac{1}{2} \left[\frac{1}{N} \sum_{n=p+1}^N |e_p^f(n)|^2 + \frac{1}{N} \sum_{n=p+1}^N |e_p^b(n)|^2 \right]. \quad (7.24)$$

Суммирование ведется только по имеющимся отсчетам данных. Величина σ_p^{fb}

является функцией только комплексного коэффициента отражения k_p , так как ошибки предсказания, начиная с порядка $p-1$, известны.

Найдем минимум σ_p^{2fb} , приравнявая комплексную производную от нее нулю:

$$\frac{\partial \sigma_p^{2fb}}{\partial \operatorname{Re}\{k_p\}} + j \frac{\partial \sigma_p^{2fb}}{\partial \operatorname{Im}\{k_p\}} = 0. \quad (7.25)$$

Решая это уравнение относительно k_p , получим оценку по МНК

$$\hat{k}_p = \frac{-2 \sum_{n=p+1}^N e_{p-1}^f(n) e_{p-1}^{b*}(n-1)}{\sum_{n=p+1}^N |e_{p-1}^f(n)|^2 + \sum_{n=p+1}^N |e_{p-1}^b(n-1)|^2}. \quad (7.26)$$

Эта оценка представляет собой гармоническое среднее $(2ab/(a+b))$ коэффициентов нормированной корреляции ошибок предсказания вперед (a) и назад (b). Модуль $|\hat{k}_p| \leq 1$, и АР-фильтр будет устойчив.

Н. Андерсен предложил следующее рекурсивное выражение, которое упрощает вычисление знаменателя (*denominator*, DEN) формулы (7.27):

$$DEN_p = \left(1 - |\hat{k}_{p-1}|^2\right) DEN_{p-1} - |e_{p-1}^f(p)|^2 - |e_{p-1}^b(N)|^2. \quad (7.27)$$

Реализация алгоритма Берга требует выполнения $2Np - p^2$ комплексных сложений и умножений и p действительных делений плюс память для хранения $3N+p$ комплексных чисел.

Для уменьшения смещения оценки частоты был предложен ряд модификаций метода Берга, взвешивающих средний квадрат ошибки предсказания

$$\sigma_p^{2fb} = \frac{1}{2} \left[\frac{1}{N} \sum_{n=p+1}^N w_p(n) \left(|e_p^f(n)|^2 + |e_p^b(n)|^2 \right) \right], \quad (7.28)$$

что дает оценку коэффициента отражения

$$\hat{k}_p = \frac{-2 \sum_{n=p+1}^N w_{p-1}(n) e_{p-1}^f(n) e_{p-1}^{b*}(n-1)}{\sum_{n=p+1}^N w_{p-1}(n) \left(|e_{p-1}^f(n)|^2 + |e_{p-1}^b(n-1)|^2 \right)}. \quad (7.29)$$

В алгоритме Берга использовалась равномерная весовая функция $W_{p-1}(n) = 1/N$. Д. Суинглер предложил использовать окно Хэмминга, М. Кавех и Дж. Липперт – специальное квадратичное окно, Б. Хелме и К. Никиас – адаптивную к данным весовую функцию вида

$$w_{p-1}(n) = \sum_{k=n-p+1}^{n-1} |x(k)|^2, \quad p \geq 2. \quad (7.30)$$

7.4.3. Рекурсивное оценивание по методу максимального правдоподобия

Алгоритм для действительных данных разработан С. Кеем. В нем для получения коэффициентов отражения и обновления порядка АР-коэффициентов, получаемых с помощью рекурсии Левинсона, используется метод максимального правдоподобия.

Если плотность вероятности гауссовского АР-процесса с нулевым средним максимизируется относительно коэффициента отражения k_p и дисперсии возбуждающего шума σ_p^2 в предположении, что АР-коэффициенты порядка $p-1$ уже вычислены, то оценки k_p и σ_p^2 , максимизирующие эту плотность вероятности для заданной последовательности данных $\{x(0), x(1), \dots, x(N-1)\}$, получают-ся как решения кубического уравнения

$$\hat{k}_p^3 + \left(\frac{N-2p}{N-p} \right) \left(\frac{\alpha_p}{\beta_p} \right) \hat{k}_p^2 - \left(\frac{p\delta_{p-1} + N\alpha_p}{[N-p]\beta_p} \right) \hat{k}_p - \left(\frac{N}{N-p} \right) \left(\frac{\alpha_p}{\beta_p} \right) = 0, \quad (7.31)$$

для которого определяется его действительный корень на интервале $[-1, 1]$, и линейных уравнений

$$\delta_p = \delta_{p-1} + 2\alpha_p \hat{k}_p + \beta_p \hat{k}_p^2, \quad \hat{\sigma}_p^2 = \delta_p / N, \quad \delta_0 = \sum_{n=0}^{N-1} x^2(n). \quad (7.32)$$

Действительные коэффициенты α_p и β_p определяются с помощью квадратичных форм

$$\alpha_p = \begin{pmatrix} \mathbf{a}_{p-1}^T & 0 \end{pmatrix} \mathbf{S}_p \begin{pmatrix} \mathbf{J} \mathbf{a}_{p-1} \\ 0 \end{pmatrix}, \quad \beta_p = \begin{pmatrix} 0 & \mathbf{a}_{p-1}^T \mathbf{J} \end{pmatrix} \mathbf{S}_p \begin{pmatrix} \mathbf{J} \mathbf{a}_{p-1} \\ 0 \end{pmatrix},$$

где $\mathbf{a}^T = (1 \ a_{p-1}(1) \ \dots \ a_{p-1}(p-1))^T$ – вектор авторегрессионных коэффициентов;

\mathbf{J} – $(p \times p)$ -матрица отражения;

S_p – $(p+1) \times (p+1)$ -матрица с элементами, определенными при $0 \leq i, j \leq p$:

$$S_p(i, j) = \sum_{n=p}^{N-1+i+j-p} x(n-i)x(n-j)$$

Алгоритм запускается вычислением δ_0 . Найденные из (7.32) оценки k_p подставляются в рекурсию Левинсона (7.16) для определения АР-параметров.

7.5. Оценка параметров АР-моделей, основанная на линейном предсказании по МНК

Налагая на АР-коэффициенты ограничения (их взаимосвязь по рекурсии Левинсона), Берг оптимизировал по МНК единственный параметр – коэффициент отражения. Отбросив данное ограничение, можно оптимизировать по МНК одновременно все коэффициенты ЛП.

Существует два типа алгоритмов спектрального оценивания по МНК.

1) Алгоритмы с раздельными оценками коэффициентов ЛП вперед и назад.

Предположим, что для оценивания АР-параметров порядка p имеется N -точечная последовательность данных $\{x(1), \dots, x(N)\}$. Ошибка ЛП вперед дается выражением

$$e_p^f(n) = x(n) + \sum_{k=1}^p a_p^f(k)x(n-k),$$

где $e_p^f(n)$ не является белым шумом, так как последовательность данных конечна. Поэтому, чтобы воспользоваться АР-моделью, будем полагать, что ошибка предсказания является отбеленным процессом, что позволит приравнять АР-параметры и коэффициенты ЛП.

Ошибку ЛП вперед можно определить в диапазоне временных индексов от $n=1$ до $n=N+p$, если предположить, что данные до первого и после последнего отсчетов равны нулю ($x(n)=0$ при $n<1$ и $n>N$).

Необходимо минимизировать модуль среднего квадрата ошибки ЛП

$$\sigma_p^{2f} = \sum_n \left| e_p^f(n) \right|^2 \quad (7.33)$$

или σ_p^{2f} / N – выборочную дисперсию.

Рассмотрим три случая спектрального оценивания:

а) выбирая полный диапазон суммирования от $e_p^f(1)$ до $e_p^f(N+p)$, получаем так называемый **взвешенный случай**, так как он включает пред- и постобработку данных с помощью окна (приравнивание отсутствующих данных к ну-

лю). Этот случай получил название **автокорреляционного метода ЛП**. Метод эквивалентен методу Юла – Уолкера для оценивания АР-параметров по смещенным корреляционным оценкам. Обработка данных с помощью окна ухудшает разрешение, поэтому в случае коротких записей данных метод редко применяется на практике;

б) выбирая диапазон суммирования от $e_p^f(1)$ до $e_p^f(N)$, получаем **предвзвешенный случай**, так как полагается, что значения данных, предшествующих отсчету $x(1)$, равны нулю. Метод предвзвешивания целесообразно использовать в процедурах ЛП по МНК, основанных на последовательной обработке данных;

в) диапазон суммирования от $e_p^f(p+1)$ до $e_p^f(N)$ соответствует **невзвешенному случаю**, так как используются только имеющиеся отсчеты данных. Метод называется **ковариационным**. Для этого метода существенно ограничение: выбранный порядок модели не должен превышать половины записи данных: $p \leq N/2$. В случае гауссовских процессов ковариационный метод дает для АР-параметров оценку, приближающуюся к оценке максимального правдоподобия. Вычислительные затраты пропорциональны величине p^2 .

Все три метода могут быть рассмотрены и применительно к оценке ЛП назад.

Следует отметить, что коэффициенты ЛП вперед и назад, определяемые с помощью ковариационного метода, не гарантируют получение устойчивого фильтра. В задаче спектрального оценивания это не приводит к каким-либо затруднениям. Более того, спектральные оценки, полученные по оценкам АР-коэффициентов, определяемых с помощью ковариационного метода, обычно имеют меньшие искажения, чем спектральные оценки, получаемые с помощью методов, гарантирующих устойчивость фильтра ЛП, например с помощью автокорреляционного метода.

В одном из вариантов ковариационного метода (К. Никиас и П. Скотт) коэффициенты ЛП вперед и назад выбираются на основе взвешенных квадратов ошибок

$$\sigma_p^{2f} = \sum_{n=p+1}^N w_p(n) |e_p^f(n)|^2, \quad \sigma_p^{2b} = \sum_{n=p+1}^N w_p(n) |e_p^b(n)|^2, \quad (7.34)$$

где веса характеризуют энергию отсчетов данных, используемых для формирования ошибок предсказания в пределах окна

$$w_p(n) = \sum_{k=n-p+1}^n |x(k)|^2. \quad (7.35)$$

2) Комбинированные алгоритмы ЛП вперед и назад.

Поскольку оба направления предсказания обеспечивают получение одинаковой статистической информации, целесообразно объединить статистики ошибок ЛП вперед и назад, чтобы получить большее число точек, в которых определяются ошибки.

Минимизируется среднее значение квадратов ошибок предсказания вперед

и назад

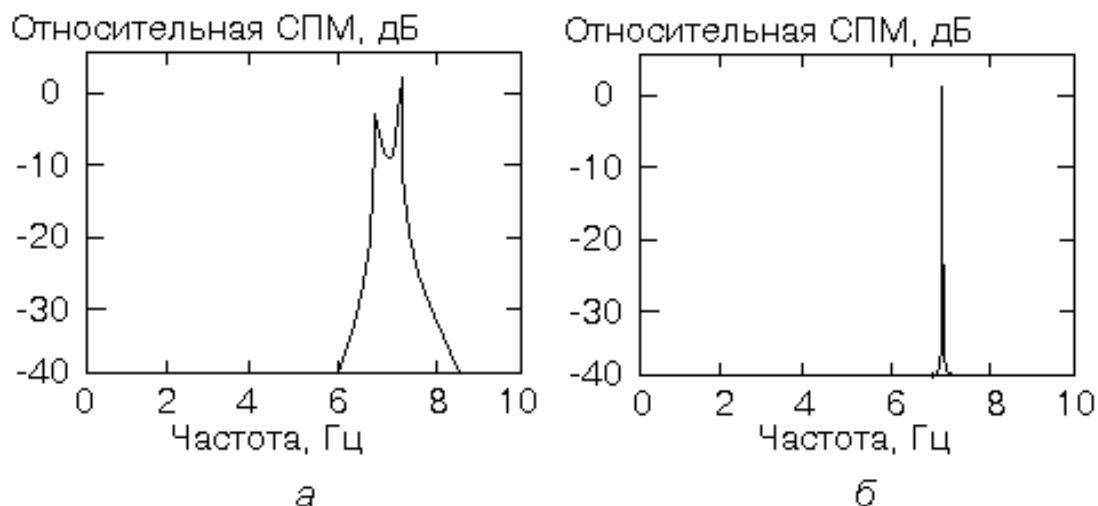


Рис. 7.6. а - оценка методом Берга; б - оценка модифицированным ковариационным методом

$$\sigma_p^{2fb} = \frac{1}{2} \left[\sum_{n=p+1}^N \left\{ \left| e_p^f(n) \right|^2 + \left| e_p^b(n) \right|^2 \right\} \right] \quad (7.36)$$

по имеющимся данным. Так как диапазон суммирования идентичен диапазону суммирования в ковариационном методе, то метод получил название **модифицированного ковариационного метода**. Метод был независимо разработан Бергом, Ульрихом, Клейтоном и Наттолом. Значение выбранного порядка модели не должно превышать двух третей длины записи данных: $p \leq 2N/3$.

В отличие от метода Берга, модифицированный ковариационный метод не гарантирует получение устойчивого фильтра ЛП. Но модули коэффициентов отражения всегда не превосходят 1, поэтому мы будем получать устойчивые решетчатые фильтры.

7.6. Сравнительные характеристики СПМ, получаемых с помощью блочных методов АР-оценивания

Если выбран большой порядок АР-модели относительно имеющегося числа отсчетов данных, то дополнительные полюсы, порожденные лишними АР-параметрами, приводят к появлению **ложных спектральных пиков**. Уменьшение же порядка модели снижает разрешение спектральной оценки. При заданном порядке модели более высокое разрешение обеспечивают метод Берга, ковариационный и модифицированный ковариационный методы. Это связано с отсутствием в них использования окон (существенного момента всех классических методов, см. главу 5) и всех сопутствующих им эффектов. Если в методах с использованием окна за пределами этого окна применяется нулевая, то в методе Берга и др. –

ненулевая экстраполяция АКП, определяемая выражением

$$r_{xx}(n) = - \sum_{k=1}^p a(k)r_{xx}(n-k), \quad n > p.$$

Это иллюстрируется рисунком 7.5.

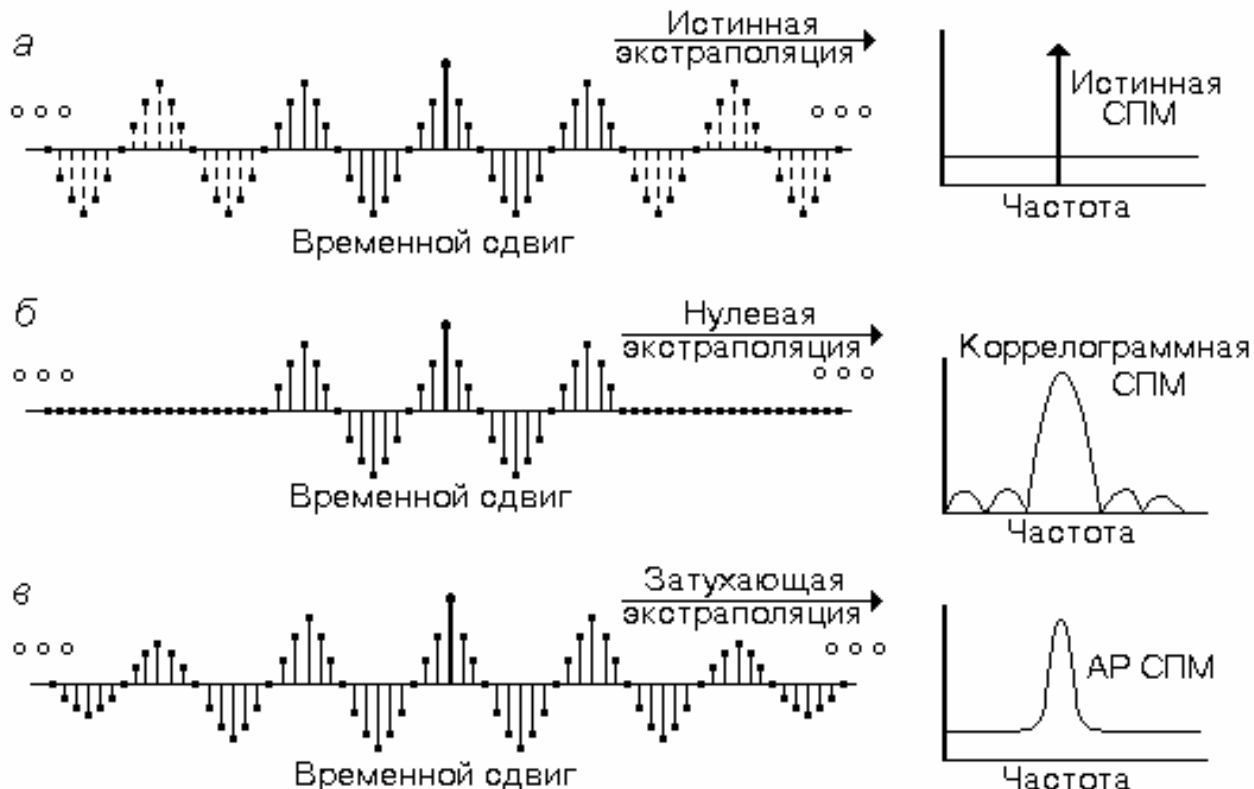


Рис. 7.5. Экстраполяция автокорреляционной последовательности:
а – исходная бесконечная АКП и истинный спектр процесса, состоящего из одной синусоиды в белом шуме; б – экстраполяция в коррелограммном методе; в – экстраполяция при АР-оценивании СПМ

При использовании метода геометрического среднего, метода Берга и автокорреляционного метода может возникать **расщепление спектральных линий** (рис. 7.6). Фужер, впервые описавший это явление, отмечал, что при использовании метода Берга расщепление спектральных линий наиболее вероятно в тех случаях, когда

- 1) велико отношение сигнал/шум;
- 2) начальные фазы синусоидальных компонент нечетно кратны углу 45° ;
- 3) протяженность последовательности данных во времени такова, что синусоидальные компоненты имеют нечетное число четвертей периодов;
- 4) процентное соотношение между числом оцениваемых АР-параметров и числом используемых для этой цели отсчетов данных относительно велико.

Расщепление спектральных линий отмечалось при использовании как действительных, так и комплексных данных. Спектры с расщеплением линий, как правило, содержат много ложных спектральных пиков. С ростом длины записи данных вероятность расщепления спектральных линий быстро уменьшается. Уст-

ранить расщепление линий в случае метода Юла – Уолкера можно, используя смещенную оценку АКП; в случае метода Берга – используя взвешенные квадраты ошибок. Р. Херринг полагает, что данное явление обусловлено смещением между положительными и отрицательными компонентами действительных синусоид. Не дает расщепления модифицированный ковариационный метод.

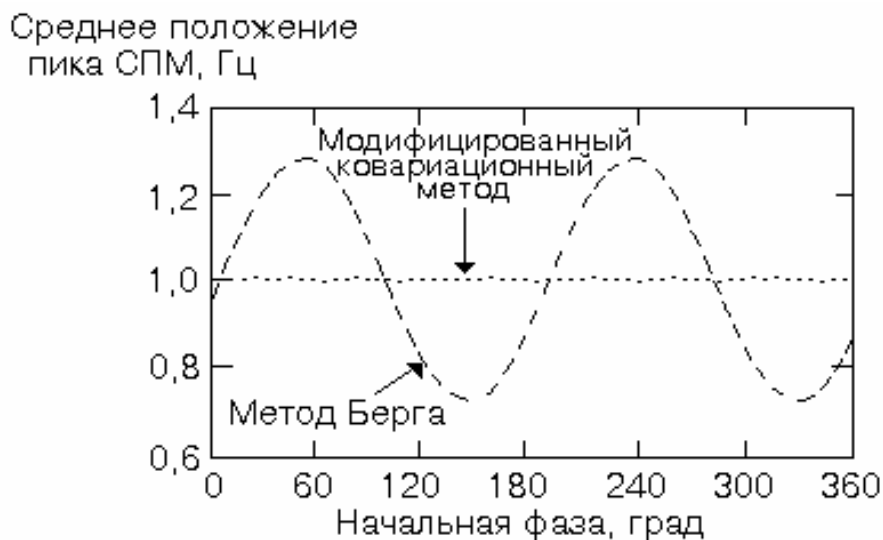


Рис. 7.7. Зависимости среднего по ансамблю положения спектральных пиков от начальной фазы косинусоидального сигнала в белом шуме

В случае процесса, состоящего из смеси одной или двух синусоид и аддитивного шума, спектральные пики авторегрессионной спектральной оценки по методу Берга оказываются сдвинутыми, причем величина их сдвига зависит от начальной фазы этих синусоид примерно по синусоидальному закону (рис. 7.7). Суинглер показал, что **смещение спектральных пиков** может достигать 16% величины ячейки частотного разрешения (она равна $1/NT$ Гц); ослабляет фазовую зависимость частотных оценок по методу Берга использование взвешенных квадратов ошибок и комплексного сигнала.

Проведенные Наттолом исследования показали, что в случае несинусоидальных процессов лишь метод Берга и модифицированный ковариационный метод дают оценки СПМ и частоты с минимальной дисперсией. Спектральные оценки, получаемые с помощью методов, основанных на ЛП, в целом характеризуются меньшей вариабельностью склонов, но большей амплитудой выбросов вблизи частот, соответствующих истинным пикам спектра. Они дают оценки с несколько меньшей дисперсией частоты, но с большей дисперсией СПМ, чем методы, основанные на оценивании коэффициента отражения. Большая дисперсия СПМ объясняется тем, что в случае методов, основанных на оценивании коэффициента отражения, фильтр должен быть устойчивым, а в случае методов ЛП это необязательно. Наихудшую оценку спектра АР-процесса дает метод Юла – Уолкера.

Если сигналы имеют большие уровни постоянных составляющих или же характеризуются заметным линейным трендом, то их АР СПМ будут искажены, особенно в НЧ области спектра. Поэтому такие составляющие должны оцени-

ваться и удаляться до выполнения процедуры АР-спектрального оценивания.

Х. Сакаи и Р. Килер показали, что дисперсия частоты АР спектральной оценки оказывается обратно пропорциональной длине записи и либо квадрату отношения сигнал/шум (для аддитивной смеси синусоид и шума), либо отношению сигнал/шум (для несинусоидальных процессов).

Г. Акаике, Р. Кроумер и К. Берк показали, что среднее значение АР-оценки СПМ оказывается в пределе равным истинному среднему значению этой оценки, а ее дисперсия стремится к значению, пропорциональному величине $(4p/N)P_{AP}^2(f)$.

7.7. Авторегрессионное спектральное оценивание: алгоритмы обработки последовательных данных

Алгоритмы последовательной обработки данных по мере их поступления можно рассматривать как алгоритмы с фиксированным порядком, рекурсивные относительно времени в том смысле, что они применяются для последовательной обработки данных с целью обновления оценок параметров АР-модели фиксированного порядка. Такие алгоритмы целесообразно применять в тех случаях, когда необходимо осуществлять слежение за спектром, медленно изменяющимся во времени, т. е. адаптироваться к нему, поэтому их иногда называют *адаптивными алгоритмами*.

Последовательные алгоритмы делятся на две категории.

1) Простейшие алгоритмы на основе градиентной аппроксимации, в том числе и *метод наименьших средних квадратов (МНСК)*, в котором градиентная адаптивная процедура наискорейшего спуска позволяет рекурсивно оценивать p -компонентный вектор АР-параметров $\mathbf{a}_p^T = (a_p(1), \dots, a_p(p))$ в момент времени, соответствующий индексу $N + 1$, по предыдущей его оценке, соответствующей временному индексу N . Входными параметрами являются порядок модели p и постоянная времени адаптации m . Скорость сходимости в среднем адаптивной процедуры обратно пропорциональна величине m . Компромисс между скоростью сходимости алгоритма и точностью оценок АР-параметров достигается выбором m .

Адаптивный НСК-алгоритм был предложен Л. Гриффитсом. Он может использоваться для слежения за изменениями во времени статистик входного сигнала, если они изменяются достаточно медленно по сравнению со скоростью сходимости алгоритма. Алгоритм более функционально устойчив и менее чувствителен к квантованию, чем РНК-алгоритм. Вычислительные затраты пропорциональны p .

2) Рекурсивные алгоритмы наименьших квадратов (РНК).

Обеспечивают более высокие характеристики, чем алгоритмы НСК, но за счет дополнительных вычислительных затрат.

РНК-алгоритм позволяет по мере поступления каждого нового отсчета данных получать точное рекурсивное решение наименьших квадратов для коэффициентов ЛП. Он отличается от НСК-алгоритма тем, что адаптивный коэффициент

усиления является уже не постоянной скалярной величиной, а изменяющейся во времени матрицей.

Для того чтобы начать рекурсию, необходимо задать начальные значения вектора АР-параметров \mathbf{a}_0 и матрицы коэффициентов усиления \mathbf{P}_0 . Квадраты ошибок взвешиваются экспоненциально с коэффициентом затухания w .

Оказалось, что при продолжительной работе РНК-алгоритма в адаптивном режиме его характеристики расходятся. Это обусловлено накоплением ошибок квантования. Вычислительные затраты пропорциональны p^2 . Для нормальной работы РНК-алгоритма для представления чисел необходимо использовать не менее 10 двоичных разрядов, тогда как для НСК-алгоритма достаточно 7 двоичных разрядов.

Последовательные алгоритмы позволяют сразу обновлять АР параметры, а не коэффициенты отражения, что исключает необходимость применения алгоритма Левинсона и уменьшает вычислительные затраты.

Блок-схема алгоритмов последовательного во времени АР-оценивания СПМ приведена на рис. 7.8.

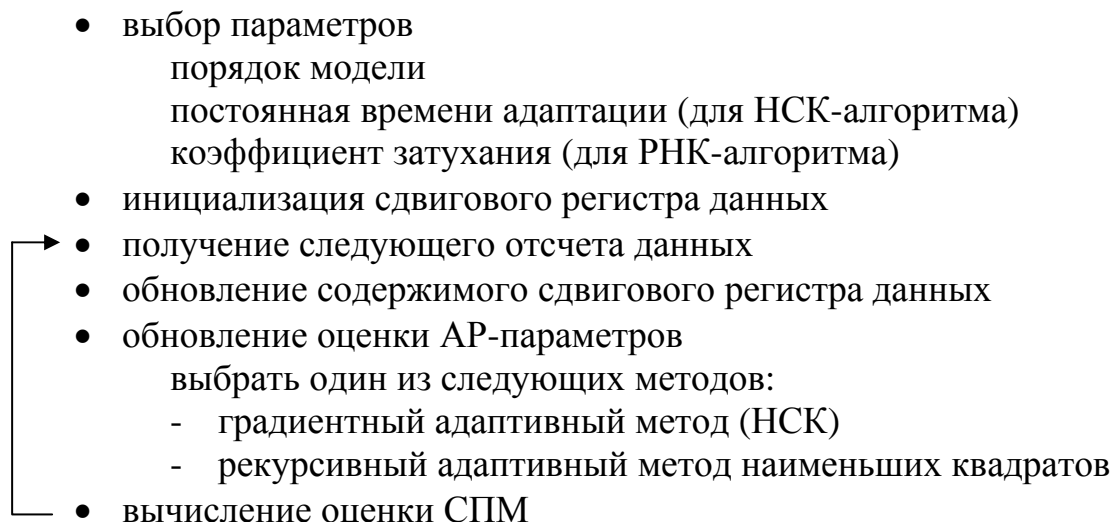


Рис. 7.8

7.8. Некоторые практические вопросы спектрального оценивания, основанного на линейном моделировании

7.8.1. Выбор порядка модели

Поскольку наилучшее значение порядка фильтра заранее, как правило, неизвестно, на практике обычно приходится испытывать несколько порядков модели. Если порядок модели выбран слишком малым, получаются сильно сглаженные спектральные оценки, если излишне большим – увеличивается разрешение, но в оценке появляются ложные спектральные пики. Дисперсия ошибки предсказания уменьшается с увеличением порядка модели монотонно и поэтому не может служить достаточным критерием окончания процедуры изменения порядка модели, пока неизвестно некоторое значение этого порядка, при дальнейшем изменении которого скорость изменения дисперсии ошибки резко снижается.

Для выбора порядка АР-модели предложено много различных критериев. Два критерия были предложены Акаике.

1) **Окончательная ошибка предсказания (ООП)**. Согласно этому критерию, выбор порядка модели осуществляется таким образом, чтобы минимизировать среднюю дисперсию ошибки на каждом шаге предсказания. Акаике рассматривает ошибку предсказания как сумму мощностей в непредсказуемой части анализируемого процесса и как некоторую величину, характеризующую неточность оценивания АР-параметров. ООП для АР-процесса определяется выражением

$$\text{ООП}(p) = \hat{\sigma}_p^2 \left(\frac{N + (p + 1)}{N - (p + 1)} \right), \quad (7.37)$$

где N – число отсчетов данных, p – порядок АР-процесса и $\hat{\sigma}_p^2$ – оценочное значение дисперсии белого шума, которая будет использоваться в качестве ошибки ЛП. Предполагается, что из данных вычтено выборочное среднее значение. Член в круглых скобках растет с увеличением порядка, характеризуя тем самым увеличение неопределенности оценки $\hat{\sigma}_p^2$ для дисперсии ошибки предсказания. Выбирается такое значение порядка, при котором величина ООП минимальна. Для идеальных АР-процессов критерий ООП обеспечивает идеальные результаты.

2) **Информационный критерий Акаике (ИКА)** основан на методике максимального правдоподобия. Согласно этому критерию, порядок модели определяется посредством минимизации некоторой теоретико-информационной функции. Если предположить, что исследуемый АР-процесс имеет гауссовы статистики, то ИКА будет определяться следующим выражением:

$$\text{ИКА}(p) = N \ln(\hat{\sigma}_p^2) + 2p. \quad (7.38)$$

Член $2p$ характеризует плату за использование дополнительных АР-коэффициентов, но это не приводит к значительному уменьшению дисперсии ошибки предсказания. Выбирается порядок модели, который минимизирует зна-

чение ИКА. При $\left(\frac{1}{N} \sum_{k=1}^p \left(\frac{\sigma_k^2}{\sigma_p^2} \right) \right)$ первый и второй критерии Акаике асимптотически эквивалентны.

М. Кавех и С. Бруццоне предложили вариант ИКА, который учитывает начальные условия при формулировке оценки максимального правдоподобия для АР-параметров. Отметим, что Р. Кашьяп считает ИКА статистически несостоятельным критерием в том смысле, что вероятность ошибки при выборе правильного порядка модели не стремится к нулю при $N \rightarrow \infty$, а это приводит к завышению порядка модели в том случае, когда длина записи данных возрастает. Для устранения этого недостатка Дж. Риссанен разработал другой вариант ИКА, который имеет следующую форму

$$\text{ДМО}(p) = N \ln(\hat{\sigma}_p^2) + p \ln(N), \quad (7.39)$$

где ДМО – длина минимального описания, о которой можно сказать, что она статистически состоятельна, поскольку величина $p \ln N$ растет с увеличением N быстрее, чем в случае с p .

3) **Критерий авторегрессионной передаточной функции (АПФК)** предложен Э. Парзенем. Порядок модели выбирается в этом случае равным порядку, при котором оценка разности среднего квадрата ошибок между истинным фильтром предсказания ошибки (его длина может быть бесконечной) и оцениваемым фильтром минимальна. Парзен показал, что эту разность можно вычислить, даже если истинный фильтр предсказания ошибки точно не известен

$$\text{АПФК}(p) = \left(\frac{1}{N} \sum_{j=1}^p (\bar{\sigma}_j^2)^{-1} \right) - (\bar{\sigma}_p^2)^{-1}, \quad (7.40)$$

где $\bar{\sigma}_j^2 = \sigma_j^2 N / (N - j)$. Значение p выбирается так, чтобы минимизировать АПФК(p).

Результаты оценивания спектра при использовании критериев ООП, ИКА и АПФК мало отличаются друг от друга, особенно в случае реальных данных, а не моделируемых АР-процессов. Т. Ульрих и Р. Клейтон показали, что в случае коротких записей данных ни один из этих критериев не обеспечивает удовлетворительных результатов. Для гармонических процессов в присутствии шума использование ООП и ИКА приводит к заниженной оценке порядка модели, особенно в тех случаях, когда отношение сигнал/шум велико. Ульрих и М. Уи экспериментально установили, что если при анализе коротких отрезков гармонических процессов с помощью ковариационного и модифицированного ковариационного методов порядок модели выбирать на интервале значений от $N/3$ до $N/2$, то во многих случаях будут получаться удовлетворительные результаты. С. Ланг и Дж. Макклеллан показали, что при использовании модифицированного ковариационного метода дисперсия положений спектральных пиков для случаев одной синусоиды или нескольких с хорошо различающимися частотами приближается к нижней границе, если порядок модели составляет примерно $1/3$ длины записи

данных. Тем не менее, окончательный выбор порядка модели для данных, получаемых из реальных записей неизвестных процессов, пока еще носит субъективный, а не точный научный характер. Поэтому все описанные критерии целесообразно использовать лишь для выбора начального значения порядка модели, поскольку они обеспечивают хорошие результаты в случае искусственных АР-сигналов, синтезированных с помощью ЭВМ, а в случае действительных данных результаты их применения зависят от того, насколько точно эти данные могут моделироваться с помощью того или иного АР-процесса.

7.8.2. Авторегрессионные процессы с шумом наблюдения

На практике отсчеты данных очень часто искажены шумом наблюдения, что приводит к ухудшению характеристик и разрешения спектральных оценок. Это ухудшение обусловлено тем, что используемая при спектральном анализе модель с одними лишь полюсами (чисто полюсная модель) при наличии шума наблюдения оказывается уже непригодной. Пусть $y(k) = x(k) + n(k)$ – искаженный шумом АР-процесс, где $x(k)$ – АР(p)-процесс, а $n(k)$ – шум наблюдения. Если $n(k)$ – белый шум с дисперсией σ_n^2 и если он не коррелирован с $x(k)$, то

$$\begin{aligned} P_{yy}(z) &= \frac{\sigma_w^2}{A(z)A^*(1/z^*)} + \sigma_n^2 = \frac{\sigma_w^2 + \sigma_n^2 A(z)A^*(1/z^*)}{A(z)A^*(1/z^*)} = \\ &= \frac{\sigma^2 [B(z)B^*(1/z^*)]}{A(z)A^*(1/z^*)}. \end{aligned} \quad (7.41)$$

Отсюда следует, что z -преобразование АКФ выходного процесса характеризуется как полюсами, так и нулями, а это означает, что $y(k)$ является АРСС(p, q)-процессом. Поэтому процессы этого типа следует анализировать с помощью методов на основе АРСС-моделей. Применительно к случаю, когда шум наблюдения является белым шумом, предложен ряд схем компенсации мощности этого шума, которые позволяют уменьшить связанное с ним смещение спектральных оценок. Поскольку мощность шума наблюдения влияет лишь на член АКП, соответствующий нулевому временному сдвигу,

$$r_{yy}(k) = \begin{cases} r_{xx}(0) + \sigma_n^2, & k = 0; \\ r_{xx}(k), & k \neq 0, \end{cases}$$

то, удаляя из оценки автокорреляционного члена $r_{yy}(0)$ оценку дисперсии σ_n^2 , можно устранить эффекты, обусловленные шумом наблюдения. Можно использовать метод гармонического разложения Писаренко. Аналогичные схемы компенсации можно разработать и для коэффициентов отражения. Следует, однако, отметить, что схемы компенсации шума, уменьшая смещение спектральных оценок,

одновременно увеличивают дисперсию этих оценок. Основной их недостаток состоит в том, что заранее неизвестно, какую долю мощности шума следует удалить; если оценка σ_n^2 слишком велика, то оцененный АР-спектр будет иметь более острые спектральные пики, чем истинный спектр.

7.8.3. Оценивание мощности синусоидальных компонент

АР-метод оценивания СПМ часто используется для того, чтобы выявить в данных наличие синусоидальных компонент. Мощности, соответствующие компонентам в оценке СПМ, можно точно вычислить, интегрируя площадь под кривой этой оценки. Однако это связано с большими вычислительными затратами. Для классических спектральных оценок надежным показателем относительной мощности служит высота пиков, так как в случае, когда анализируемый процесс состоит из аддитивной смеси синусоид и белого шума, высота спектральных пиков прямо пропорциональна мощности этих синусоид. Но для АР-оценок СПМ такой метод не применим. Р. Лакосс показал, что авторегрессионная m -го порядка СПМ для процесса, состоящего из одной синусоиды и белого шума, достигает своего максимума при $f = f_0$

$$P_{AP}(f_0) \approx T(mSNR)^2, \quad (7.42)$$

где $SNR = P / \sigma^2$ и предполагается $mSNR \gg 1$ (т. е. что отношение сигнал/шум велико). Таким образом, высота пика оказывается пропорциональной квадрату мощности, а следовательно, площадь под пиком прямо пропорциональна этой мощности.

Метод оценки мощности нескольких действительных синусоид по пикам в АР-спектре был предложен С. Джонсоном и Н. Андерсеном. Однако он дает хорошие результаты только в случае достаточно разнесенных спектральных пиков синусоид с высоким отношением сигнал/шум. Если АР-СПМ записать в виде z -преобразования

$$P_{AP}(z) = T\sigma^2 / A(z)A^*(1/z^*),$$

то мощность пика, наблюдаемого в АР-спектре на частоте f_k , будет оцениваться следующей величиной:

$$\text{Мощность}(f_k) = 2 \operatorname{Re} \left\{ \text{Вычет} \left(\frac{P_{AT}(z)}{z} \right) \right\} \quad \text{при} \quad z_k = \exp(j2\pi f_k T). \quad (7.43)$$

7.8.4. Частотное разрешение

Следует заметить, что степень улучшения разрешения оценки АР СПМ в случае процессов, состоящих из смеси синусоид и белого шума, зависит от вели-

чины отношения сигнал/шум. Степень разрешения синусоид равной мощности в случае известной автокорреляционной последовательности можно определить с помощью приближенной формулы С. Л. Марпла-мл.:

$$F = \frac{1.03}{T \cdot p[SNR(p+1)]^{0.31}}, \quad (7.44)$$

где F – разрешение в Гц, T – интервал отсчетов в секундах, p – максимальное значение индекса временного сдвига для АКП, SNR – линейное отношение сигнал/шум для отдельной синусоиды.

7.9. Метод Прони

Метод Прони – метод моделирования выборочных данных в виде линейной комбинации экспоненциальных функций. С помощью метода Прони осуществляется аппроксимация данных с использованием некоторой детерминированной экспоненциальной модели, в противоположность АР-методам, с помощью которых стремятся приспособить вероятностные модели представления статистик 2-го порядка для имеющихся данных.

В оригинальной статье Г. Прони описан метод точной подгонки, основанный на использовании такого большого числа полностью затухающих экспонент, сколько их необходимо для аппроксимации N имеющихся точек данных. Современный вариант метода Прони обобщен и на модели, состоящие из затухающих синусоид. В нем также используется анализ по методу наименьших квадратов для приближенной подгонки экспоненциальной модели в тех случаях, когда число точек данных превышает их число, необходимое для подгонки с помощью предполагаемого числа экспоненциальных членов. Одна из модификаций современного метода Прони позволяет использовать чисто синусоидальную модель с незатухающими компонентами.

Предположим, что имеется N комплексных отсчетов данных $\{x(1), \dots, x(N)\}$. Тогда метод Прони позволяет оценить $x(n)$ с помощью некоторой p -членной модели комплексных экспонент

$$\hat{x}(n) = \sum_{k=1}^p A_k \exp[(\alpha_k + j2\pi f_k)(n-1)T + j\theta_k], \quad (7.45)$$

где $1 \leq n \leq N$, T – интервал отсчетов в секундах, A_k и α_k – амплитуда и коэффициент затухания (в с^{-1}) k -й комплексной экспоненты, f_k и θ_k – частота (в Гц) и начальная фаза (в рад) k -й синусоиды. Значения всех этих параметров полностью произвольны. В случае отсчетов действительных данных комплексные экспоненты должны появляться комплексно-сопряженными парами равной амплитуды, что сводит экспоненциальное представление к

$$\hat{x}(n) = \sum_{k=1}^{p/2} 2A_k \exp[\alpha_k (n-1)T] \cos[2\pi f_k (n-1)T + \theta_k], \quad (7.46)$$

где $1 \leq n \leq N$. Если число p комплексных экспонент четно, то будем иметь $p/2$ затухающих косинусоид. Если p нечетно, то будем иметь $(p-1)/2$ затухающих косинусоид и одну полностью затухающую экспоненту.

7.9.1 Исходный метод Прони

Если число используемых отсчетов данных равно числу экспоненциальных параметров, то возможна точная подгонка экспонент под имеющиеся данные. Рассмотрим функцию дискретного времени, представляющую собой сумму p экспонент:

$$x(n) = \sum_{k=1}^p h_k z_k^{n-1}, \quad (7.47)$$

где комплексные константы h_k и z_k определяются выражениями

$$h_k = A_k \exp(j\theta_k), \quad z_k = \exp[(\alpha_k + j2\pi f_k)T] \quad (7.48)$$

или в матричной форме

$$\begin{pmatrix} z_1^0 & z_2^0 & \cdot & \cdot & \cdot & z_p^0 \\ z_1^1 & z_2^1 & \cdot & \cdot & \cdot & z_p^1 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ z_1^{p-1} & z_2^{p-1} & \cdot & \cdot & \cdot & z_p^{p-1} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \cdot \\ \cdot \\ \cdot \\ h_p \end{pmatrix} = \begin{pmatrix} x(1) \\ x(2) \\ \cdot \\ \cdot \\ \cdot \\ x(p) \end{pmatrix}. \quad (7.49)$$

Эти уравнения можно разделить по параметрам h_k и z_k , так как уравнение (7.49) является решением некоторого однородного линейного разностного уравнения с постоянными коэффициентами. Определим сначала полином $\Phi(z)$ с комплексными коэффициентами $a(m)$, $a(0)=1$, корнями которого являются экспоненты z_k

$$\Phi(z) = \prod_{k=1}^p (z - z_k) = \sum_{m=0}^p a(m) z^{p-m}. \quad (7.50)$$

Осуществляя сдвиг индекса от n к $n - m$ и умножая обе его части на параметр $a(m)$, получаем

$$a(m)x(n - m) = a(m) \sum_{k=1}^p h_k z_k^{n-m-1}.$$

Записывая аналогичные произведения $a(0)x(n)$, ..., $a(m - 1)x(n - m + 1)$ и осуществляя суммирование, получим

$$\sum_{m=0}^p a(m)x(n - m) = \sum_{i=0}^p h_i \sum_{m=0}^p a(m) z_i^{n-m-1}, \quad p + 1 \leq n \leq 2p.$$

Осуществляя подстановку $z_i^{n-m-1} = z_i^{n-p} z_i^{p-m-1}$, получим

$$\sum_{m=0}^p a(m)x(n - m) = \sum_{i=0}^p h_i z_i^{n-p} \sum_{m=0}^p a(m) z_i^{p-m-1} = 0. \quad (7.51)$$

Сумму в правой части (7.51) можно рассматривать как полином, определяемый уравнением (7.50), который записан через свои корни, что и обеспечивает в (7.51) равенство нулю. Уравнение (7.51) – это линейное разностное уравнение, однородное решение которого выражается формулой (7.47). Полином (7.50), ассоциированный с этим линейным разностным уравнением, называется характеристическим.

p уравнений, представляющих истинные значения коэффициентов $a(m)$, удовлетворяющих уравнению (7.51), можно записать в виде следующего $(1 \leq m \leq p)$ -матричного уравнения:

$$\begin{pmatrix} x(p) & x(p-1) & . & . & . & x(1) \\ x(p+1) & x(p) & . & . & . & x(2) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ x(2p-1) & x(2p-2) & . & . & . & x(p) \end{pmatrix} \begin{pmatrix} a(1) \\ a(2) \\ . \\ . \\ . \\ a(p) \end{pmatrix} = - \begin{pmatrix} x(p+1) \\ x(p+2) \\ . \\ . \\ . \\ x(2p) \end{pmatrix}. \quad (7.52)$$

Процедуру Прони для подгонки p экспонент к $2p$ отсчетам можно теперь представить в виде следующих трех этапов. На первом этапе получается решение уравнения (7.52) для коэффициентов полинома. На втором этапе вычисляются корни полинома, определяемого уравнением (7.50). Используя корень z_i , можно определить коэффициент затухания α_i и частоту синусоиды f_i с помощью соотношений

$$\alpha_i = \ln|z_i|/T \quad (c^{-1}), \quad f_i = \arctg[\text{Im}\{z_i\}/\text{Re}\{z_i\}]/2\pi T \quad (Гц). \quad (7.53)$$

На третьем этапе корни полинома, вычисленные на втором этапе, используются для формирования элементов матрицы уравнения (7.49), которое затем решается относительно p комплексных параметров h_1, \dots, h_p . Амплитуда A_i и начальная фаза θ_i находятся как:

$$A_i = |h_i|, \quad \theta_i = \arctg[\operatorname{Im}\{h_i\} / \operatorname{Re}\{h_i\}] \quad (pad). \quad (7.54)$$

Краткая запись метода экспоненциального оценивания Прони и вычисления СПЭ по методу Прони приведена на рис. 7.9.

- Сбор данных
 - N отсчетов
 - T с/отсчет
- Выбор входных параметров
 - p - число комплексных экспонент
 - метод - обобщенный или модифицированный
 - (для модифицированного p выбирается четным)
- Оценивание параметров комплексных экспонент
- Вычисление параметров действительных экспонент
(амплитуда, фаза, коэффициент затухания, частота)
- Вычисление спектральной плотности энергии Прони

Рис. 7.9

7.9.2. Метод наименьших квадратов Прони

На практике число отсчетов данных N обычно превышает то минимальное их количество, которое необходимо для подгонки модели из p экспонент, т. е. $N > 2p$. В этом переопределенном случае последовательность отсчетов данных может быть аппроксимирована лишь приблизительно с ошибкой аппроксимации $\hat{\varepsilon}(n) = x(n) - \hat{x}(n)$:

$$\hat{x}(n) = \sum_{k=1}^p h_k z_k^{n-1}, \quad 1 \leq n \leq N. \quad (7.55)$$

Одновременное отыскание порядка p и параметров $\{h_k, z_k\}$, $1 \leq k \leq p$, минимизирующих сумму квадратов ошибки

$$\sigma^2 = \sum_{n=1}^N |\varepsilon(n)|^2 \quad (7.56)$$

представляет собой трудную нелинейную задачу. Используя на первом и втором этапах метода Прони линейные процедуры НК, получим **обобщенный метод Прони**. В переопределенном случае линейное разностное уравнение (7.51) имеет вид:

$$\sum_{m=1}^p a(m)x(n-m) = e(n), \quad (7.57)$$

где $p+1 \leq n \leq N$. Член $e(n)$ характеризует ошибку аппроксимации на основе ЛП в отличие от ошибки экспоненциальной аппроксимации $\varepsilon(n)$. Уравнение (7.57) идентично уравнению для ошибки ЛП вперед, и теперь параметры $a(m)$ из уравнения (7.51) можно выбирать как параметры, которые минимизируют не сумму (7.56), а сумму квадратов ошибок ЛП

$$\sum_{n=p+1}^N |e(n)|^2.$$

Получим ковариационный метод ЛП. Существенно ограничение порядка: $p \leq N/2$.

Метод Прони не позволяет получить оценку шума, отдельную от сигнала. Модель Прони, учитывающая присутствие аддитивного шума, будет иметь форму

$$x(n) = \sum_{k=1}^p h_k z_k^{n-1} + \varepsilon(n). \quad (7.58)$$

Если в исходной процедуре Прони вместо $x(n)$ использовать $x(n) - \varepsilon(n)$, то линейное разностное уравнение, которое описывает процесс, состоящий из суммы экспонент и аддитивного белого шума, будет иметь вид

$$x(n) = - \sum_{k=1}^p a(m)x(n-m) + \sum_{k=0}^p a(m)\varepsilon(n-m). \quad (7.59)$$

На первом этапе метода Прони используется уравнение линейного предсказания

$$x(n) = - \sum_{k=1}^p a(m)x(n-m) + e(n)$$

и попытка "отбеливания" $e(n)$. Сравнивая два последних уравнения, видим, что отбеленный процесс $e(n)$ совершенно не соответствует небелому СС-процессу, представленному выражением

$$\sum_{k=0}^p a(m)\varepsilon(n-m).$$

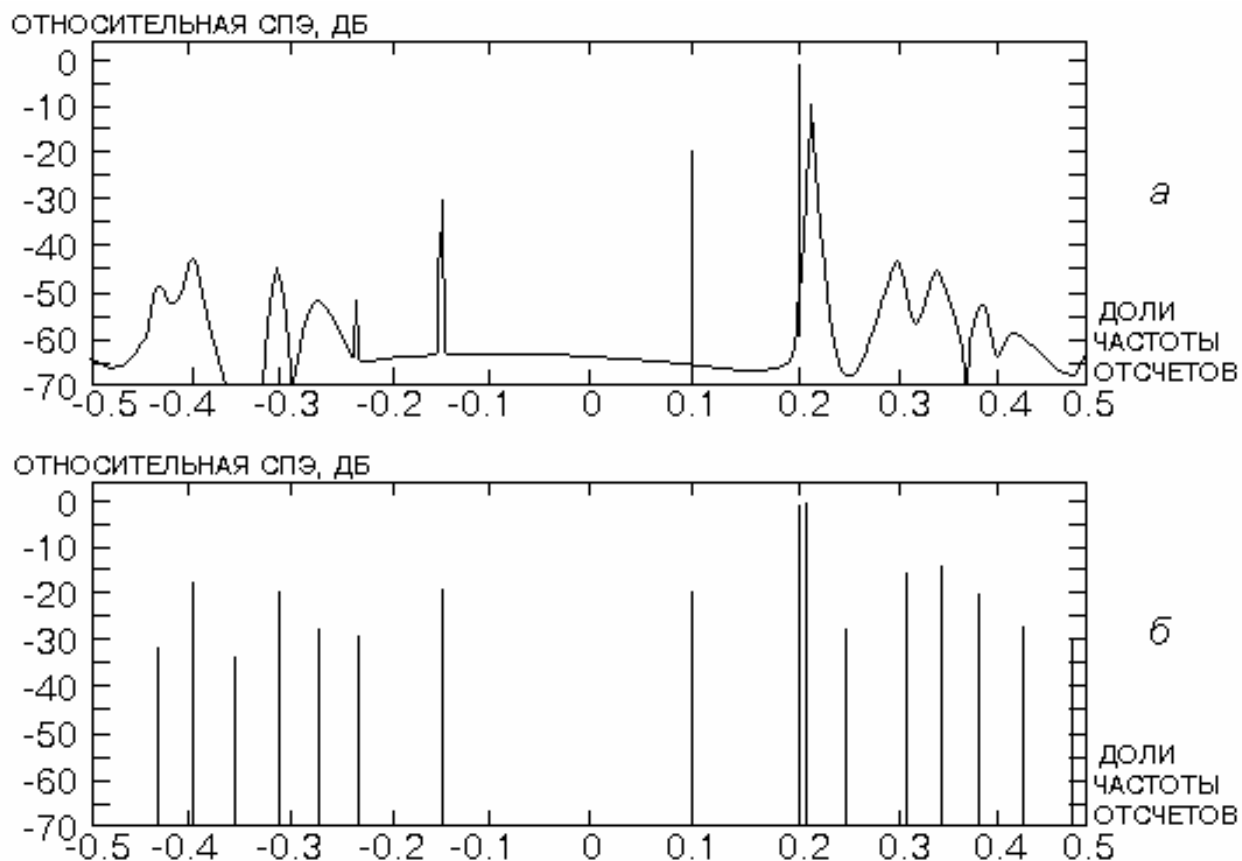


Рис. 7.10: а - обобщенный метод Прони, $p=15$;
 б - модифицированный метод Прони, $p=16$

Метод Прони не позволяет учесть наличие небелого шума в анализируемом процессе, поэтому при наличии сильного аддитивного шума получаются очень неточные, завышенные оценки коэффициентов затухания. Использование значений p , превышающих число действительно имеющихся полюсов, упрощает моделирование и позволяет учесть наличие шума.

Обычный МНК Прони может быть модифицирован для аппроксимации последовательности комплексных данных с помощью модели, состоящей из незатухающих комплексных синусоид – **модифицированный метод Прони**.

На рис. 7.10 показаны оценки по методу Прони СПЭ, полученные для 64-точечной тест-последовательности данных. Модифицированный метод Прони дает линейчатый спектр, так как в нем используется допущение о синусоидальной модели. Получаются очень точные оценки четырех действительных синусоид сигнала, но неточное представление окрашенного шума.

Литература

1. Марпл-мл. С. Л. Цифровой спектральный анализ и его приложения: Пер. с англ. – М.: Мир, 1990.
2. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов: Пер. с англ. – М.: Мир, 1978.
3. Гольденберг Л. М., Матюшкин Б. Д., Поляк М. Н. Цифровая обработка сигналов: Учеб. пособие для вузов. – М.: Радио и связь, 1990.
4. Трахтман А. М. Введение в обобщенную спектральную теорию сигналов. М.: Сов. радио, 1972.
5. Брюханов Ю. А., Кренев А. Н. Теория дискретных и цифровых сигналов и цепей: Учеб. пособие. Яросл. гос. ун-т. Ярославль, 1991.
6. Гоноровский И. С. Радиотехнические цепи и сигналы: Учебник для вузов. – М.: Радио и связь, 1986.
7. Трахтман А. М., Трахтман В. А. Основы теории дискретных сигналов на конечных интервалах. – М.: Сов. радио, 1975.
8. Кренев А. Н. Описание сигналов с помощью разрывных функций: Метод. указания. Яросл. гос. ун-т. Ярославль, 1989.

Учебное издание

К р е н ё в Александр Николаевич
А р т ё м о в а Татьяна Константиновна

Цифровой спектральный анализ

Редактор, корректор А.А. Антонова
Компьютерная верстка Т.К. Артёмовой

Подписано в печать 30.12.02. Формат 60х84/16.

Бумага тип. Усл. печ. л. 6,74. Уч.-изд.л. 7,51.

Тираж 200 экз. Заказ .

Оригинал-макет подготовлен редакционно-издательским отделом ЯрГУ
Ярославский государственный университет
150000 Ярославль, ул. Советская, 14

Отпечатано на ризографе.

ООО “Рио-Гранд”.
150000 Ярославль, ул. Свердлова, 18. Офис 34.