Министерство образования Республики Беларусь

Учреждение образования
«Полоцкий государственный университет»

С. В. МАЛЬЦЕВ

ОСНОВЫ ДИСКРЕТНОЙ МАТЕМАТИКИ

Учебно-методический комплекс
для студентов специальности 1-39 01 01 «Радиотехника»

Новополоцк
ПГУ
2009

УДК 519.85(075.8)
ББК 22.1я73
М21

Рекомендован к изданию методической комиссией радиотехнического факультета в качестве учебно-методического комплекса (протокол № 5 от 23.05.2008)

РЕЦЕНЗЕНТЫ:

главный инженер РПУП «Новополоцкий завод “Измеритель”» Е. В. ГЛУШКО;
заведующий кафедрой технической кибернетики УО «ПГУ» Р. П. БОГУШ

Мальцев, С. В.

М21 Основы дискретной математики : учеб.-метод. комплекс для студентов спец. 1-39 01 01 / С. В. Мальцев. – Новополоцк : ПГУ, 2009. – 96 с.

ISBN 978-985-418-981-9.

Приведены темы изучаемого курса, их объем в часах лекционных и практических занятий, изложены теоретические и практические основы дискретной математики, применяемые в области радиоэлектроники. Представлены задания для практических занятий, рекомендации по организации рейтингового контроля изучения дисциплины, вопросы к зачету.

Предназначен для студентов вузов, обучающихся по специальности «Радиотехника».

УДК 519.85(075.8)
ББК 22.1я73

ISBN 978-985-418-981-9

РАБОЧАЯ ПРОГРАММА

Дисциплина «Основы дискретной математики» введена в учебный план специальности 1-39 01 01 «Радиотехника» как общеобразовательная.

Целью дисциплины является получение знаний, необходимых для решения задач анализа и построения алгоритмов цифровой обработки сигналов.

В результате изучения дисциплины студенты должны

- *знать:*

– математические основы современных методов обработки дискретных сигналов;

– алгебраические принципы построения алгоритмов обработки дискретных сигналов применительно к радиотехническим задачам;

– методы построения эффективных алгоритмов обработки дискретных сигналов с использованием аппарата дискретной математики;

- *уметь:*

– формализовать различные радиотехнические задачи для использования аппарата дискретной математики при их решении;

– разрабатывать эффективные алгоритмы обработки информации;

- *иметь представление:*

– о вычислительной сложности алгоритмов обработки информации;

– о направлениях, перспективах и проблемах развития теории дискретной математики для решения современных информационных задач.

Перечень дисциплин, которые изучаются в продолжение данной: «Цифровая обработка сигналов», «Прикладное кодирование», «Радиотехнические системы».

Общее количество часов, отводимых на изучение дисциплины, составляет 90 часов, из них 54 аудиторных часа. Распределение по видам занятий составляет: 36 часов лекций, 18 часов практических занятий, 36 часов для самостоятельной работы.

Тематика практических занятий

Наименование практического занятия	Количество часов
Дискретные сигналы и способы их представления.	2
Введение в теорию групп, колец и полей.	2
Поля Галуа, понятие первообразной.	2
Теория вычетов, основные понятия.	2
Основы дискретной арифметики.	2
Формирование сигналов с идеальными корреляционными функциями.	2
Преобразования Фурье в дискретном базисе.	2
Преобразование спектров. Спектры некоторых сигналов.	2
Быстрое преобразование Фурье в базисе ВКФ.	2

Оценка знаний студентов

Для оценки работы и знаний студентов в рамках курса «Основы дискретной математики» используется накопительная система. Зачет выставляется по сумме баллов, которые студент набирает в течение всего учебного семестра, а также в результате выходного итогового контроля – контрольной работы.

Для получения аттестации (100 баллов максимум) необходимо:

– по практическим работам набрать не менее 75 процентов от максимально возможного количества баллов на момент аттестации по данному виду занятий;

– выполнить контрольную работу (два зачетных задания).

Распределение баллов по видам занятий

Вид занятий	Форма оценки учебной активности студента	Максимальное количество баллов по каждой форме оценки	Максимальное количество баллов по каждому виду занятий
Практические работы	Устные ответы на вопросы	5	20 × 9 = 180
	Решение задач	10	
	Выполнение тестовых заданий	5	

Дополнительные баллы предусматриваются:

– за выполнение научно-исследовательской работы по основам дискретной математики (до 200 баллов);

– за подготовку докладов и / или рефератов по темам, не включенным в рабочую программу курса «Основы дискретной математики» (до 100 / 50 баллов).

Для получения зачета студент должен набрать не менее 260 баллов, а для этого необходимо:

– получить две аттестации – 200 баллов минимум;

– получить не менее 60 баллов за итоговую контрольную работу (из возможных 100).

1. ДИСКРЕТНЫЕ СИГНАЛЫ

Дискретная математика представляет собой область математики, занимающуюся изучением дискретных структур, которые возникают как в пределах самой математики, так и в ее приложениях. Сама по себе дискретная математика в основном изучает такие структуры, как конечные группы, конечные графы, некоторые математические модели преобразователей информации, конечные автоматы, машины Тьюринга и т. д.. Основные разделы дискретной математики:

- математическая логика;
- математическая кибернетика;
- комбинаторика;
- теория графов;
- машинная арифметика;
- теория алгоритмов;
- теория игр;
- теория кодирования;
- теория конечных автоматов;
- теория множеств;
- теория формальных грамматик;
- вычислительная геометрия;
- теория булевых функций;
- логическое программирование;
- функциональное программирование;
- λ-исчисление;
- булева алгебра;
- комбинаторная логика;
- математическая лингвистика;
- теория искусственного интеллекта;
- прямоугольная система линейных алгебраических уравнений.

Современному специалисту в области радиотехники необходима серьезная подготовка в области цифровой обработки сигналов (ЦОС). ЦОС основывается на таких приложениях дискретной математики, как теория дискретных сигналов, теория полей Галуа, теории дискретных преобразований.

1.1. Актуальность цифровой обработки дискретных сигналов

При цифровой обработке используется представление сигналов в виде последовательностей чисел или символов. Цель такой обработки может заключаться как в оценке характерных параметров сигнала, так и в преобразовании сигнала в форму, которая в некотором смысле более удобна для последующей работы с ним. Формулы классического численного анализа, например, формулы для интерполяции[1], интегрирования и дифференцирования, являются алгоритмами цифровой обработки. Наличие быстродействующих цифровых ЭВМ благоприятствовало развитию все более сложных

[1]Интерполяция – способ нахождения промежуточных значений величины по имеющемуся дискретному набору известных значений.

и рациональных алгоритмов обработки сигналов; последние же успехи в технологии интегральных схем дают высокую экономичность построения сложных систем цифровой обработки сигналов внутри одного кристалла – чипа, что позволяет реализовывать сложные системы ЦОС в миниатюрном виде.

Цифровая обработка сигналов применяется в различных областях человеческой деятельности, например, в биомедицине, акустике, звуковой и радиолокации, сейсмологии, системах связи, системах передачи данных и многих других. Так, при анализе электроэнцефалограмм, электрокардиограмм, а также распознавании речи требуется выделять некоторые характерные параметры сигнала – частоту, фазу, период, искажения формы, девиацию параметра и т. д. В некоторых приложениях ЦОС возникает необходимость отделения помехи типа шума от сигнала или приведения сигнала к виду, который наиболее удобен для пользователя. В качестве другого примера обработки сигналов можно привести случай, когда передаваемый через канал связи сигнал подвергается различным искажениям, а приемник пытается компенсировать их.

До конца прошлого века обработка сигналов в основном выполнялась при помощи аналоговых устройств. Первые исключения из этого правила встречаются с 50-х годов XX века в первую очередь в тех применениях, где требовалась реализация сложного алгоритма обработки сигналов. Одним из первых примеров практического применения ЦОС стал анализ некоторых геофизических данных, которые записывались на носитель информации для последующей обработки на больших цифровых ЭВМ того времени. Из-за большого объема вычислений обработка сигналов не всегда могла быть выполнена в реальном времени (realtime processing) – для обработки данных часто требовались минуты, часы или даже сутки машинного времени. Даже в те времена, когда вычислительная техника не была такой быстрой, как на сегодняшний день, универсальность цифровой ЭВМ обеспечивала высокую эффективность метода.

В дальнейшем использование цифровых ЭВМ в обработке сигналов шло различными путями. Благодаря своей гибкости цифровые ЭВМ были полезны для моделирования систем обработки сигналов до их конкретной технической реализации. При таком подходе новые алгоритмы обработки сигналов или целые системы могли быть изучены еще в экспериментальных условиях без расходования экономических и технических ресурсов.

Применение цифровых ЭВМ давало большой выигрыш из-за их гибкости и универсальности. Однако обработка не всегда могла быть выполнена в реальном времени. Следовательно, цифровая ЭВМ использовалась в

основном для аппроксимаций или моделирования аналоговых систем обработки, как правило – решения задачи фильтрации. В соответствии с этим, вначале задача цифровой фильтрации в основном сводилась к программированию фильтра на цифровой ЭВМ так, чтобы при аналого-цифровом преобразовании сигнала с последующей цифровой фильтрацией и цифро-аналоговым преобразованием система аппроксимировала хороший аналоговый фильтр. Представление о том, что цифровые системы могут в действительности быть практичны для непосредственной обработки в реальном времени сигналов в радиосвязи, радиолокации или во многих других сферах приложений, казалось маловероятным. Быстродействие, стоимость и размеры были, конечно, тремя важными факторами, говорившими в пользу применения аналоговых устройств.

По мере того как обработка сигналов осуществлялась на цифровых ЭВМ, естественной тенденцией было исследование все более сложных алгоритмов обработки сигналов. Некоторые из этих алгоритмов были разработаны с учетом больших возможностей цифровой ЭВМ, однако из-за сложности (больших объемов промежуточных данных, большой сложности вычислений) не реализовывались в аналоговой аппаратуре, т. е. многие из этих алгоритмов оказывались интересными, но до некоторой степени непрактичными. Примером класса алгоритмов этого типа был ряд алгоритмов, названных анализом кепстра[2] и гомоморфной фильтрацией[3]. На цифровых ЭВМ было ясно продемонстрировано, что эти алгоритмы могли быть успешно применены в системах полосового сжатия речи, развертки и устранения эхо-сигналов. Использование этих алгоритмов требует точной оценки обратного преобразования Фурье, логарифма преобразования Фурье входного сигнала. При этом требования к точности и разрешающей способности были таковы, что обычные аналоговые анализаторы спектра оказывались непрактичными, а построение специальных анализаторов было затратно. Развитие таких алгоритмов обработки сигналов сделало привлекательной идею построения полностью цифровых систем обработки сигналов. Активная работа началась с исследования цифровых вокодеров[4],

[2]Кепстр представляет собой спектр логарифма спектра исходного сигнала. Логарифмирование не имеет отношения к существу метода, поэтому для простоты считают, что кепстр – это спектр спектра. Часто применяется в кепстральном анализе.

[3]В общем случае под гомоморфной фильтрацией понимается сведение задачи нелинейной фильтрации к линейной с помощью каких-либо преобразований.

[4] Вокодер (англ. voice coder – кодировщик голоса) – устройство синтеза речи на основе произвольного сигнала с богатым спектром.

цифровых анализаторов спектра и других полностью цифровых систем в предположении, что со временем такие системы станут практичными.

Развитие новой точки зрения на цифровую обработку сигналов в дальнейшем было ускорено открытием в 1965 г. эффективных алгоритмов для вычислений преобразований Фурье. Этот класс алгоритмов стал известен как быстрое преобразование Фурье (БПФ). Возможности БПФ были значительными с нескольких точек зрения. Многие алгоритмы обработки сигналов на цифровых ЭВМ требовали временных интервалов на несколько порядков бóльших по сравнению с реальным временем изменения сигнала. Часто это было связано с тем, что спектральный анализ был важной составной частью обработки сигналов, а эффективные средства для его выполнения не были известны. Алгоритм быстрого преобразования Фурье значительно уменьшил время вычисления спектрального преобразования. Это позволило реализовывать и создавать более сложные алгоритмы обработки сигналов в реальном времени. Кроме того, с учетом возможностей действительной реализации алгоритма быстрого преобразования Фурье в специализированном цифровом устройстве многие алгоритмы обработки сигналов, ранее бывшие непрактичными, стали находить воплощение в специализированных устройствах.

Другая важная особенность алгоритма быстрого преобразования Фурье связана с тем, что ему внутренне присуща концепция дискретного времени. Эта особенность касается непосредственно вычисления преобразования Фурье дискретного сигнала и заключается в ряде свойств математических операций, строго относящихся к дискретному времени. В связи с этим данный алгоритм не является просто аппроксимацией преобразования Фурье непрерывного сигнала. Это вызвало видоизменение многих понятий и алгоритмов обработки сигналов на основе математических методов для дискретного времени, которые затем привели к формулировке ряда четких соотношений для дискретного времени. Все это явилось отходом от представления, что обработка сигналов на цифровой ЭВМ является лишь аппроксимацией методов аналоговой обработки. При таком изменении точки зрения возник значительный интерес к новой или возрожденной области цифровой обработки.

Область применения цифровой обработки сигналов стремительно расширяется. Этому способствовало развитие больших интегральных схем и связанное с ним уменьшение стоимости и размеров цифровых устройств при одновременном увеличении их быстродействия. Цифровые фильтры специального назначения сейчас могут работать в мегагерцевом диапазоне тактовой частоты, экономически оправдываются процессоры специального

назначения для выполнения быстрого преобразования Фурье при высокой частоте входных данных, несложные цифровые фильтры выполняются на отдельных чипах; в настоящее время почти все вопросы, связанные с системами полосового сжатия речи, рассматриваются в плане построения полностью цифровых систем как наиболее практичных, цифровые процессоры также являются неотъемлемой частью многих современных радиолокационных и звуколокационных систем. В дополнение к развитию цифровых специализированных устройств обработки сигналов имеются цифровые программируемые ЭВМ специального назначения, архитектура которых приспособлена к задачам обработки сигналов. Такие ЭВМ находят применение при обработке сигналов в реальном времени так же, как и при моделировании в реальном времени на специализированных цифровых устройствах.

Область применения цифровой обработки сигналов постоянно расширяется. Методы цифровой обработки и в дальнейшем будут способствовать существенным изменениям в области науки и техники. Характерным примером является область телефонной связи, где цифровые методы обещают существенную экономию и гибкость при реализации систем переключения и передачи. Учитывая направление развития цифровой обработки сигналов, очевидно, что ее методы будут применяться скорее по своему прямому назначению, чем для аппроксимации аналоговых систем обработки.

1.2. Понятие дискретного сигнала

Сигнал представляет собой тот или иной физический процесс, содержащий в себе некоторую информацию. В радиотехнике чаще всего используются электрические сигналы. При этом основными носителями информации являются изменяющиеся во времени ток или напряжение. Электрические сигналы в общем случае легче обрабатывать, чем иные сигналы, они хорошо передаются на большие расстояния. Большинство сигналов имеют аналоговую природу, то есть изменяются непрерывно во времени и могут принимать любые значения на некотором интервале. Аналоговые сигналы описываются некоторой математической функцией времени. Однако развитие техники привело к тому, что значительная часть алгоритмов обработки сигналов реализуется в цифровом виде.

Итак, сигнал может быть определен как функция, переносящая информацию о состоянии или поведении физической системы. Сигнал зачастую может принимать форму колебаний, зависящих от времени или от пространственных координат. Математически сигналы представляются в виде функций одной или более независимых переменных. Так, например,

речевой сигнал математически представляется как функция времени, а изображение – как зависимость яркости от двух пространственных переменных – $Y(x, y)$. Как правило, при математическом представлении сигнала $s(t)$ независимой переменной считают время, хотя на самом деле эта переменная может иметь и другой смысл.

Независимая переменная в математическом представлении сигнала может быть как непрерывной, так и дискретной. Сигналы в непрерывном времени определяются на континууме моментов времени и, следовательно, представляются как функции от непрерывной переменной. Другими словами, непрерывные сигналы могут задаваться на несчетном множестве точек, например, [0; T) или [0; ∞). Такие сигналы также называют *аналоговыми*.

Дискретные сигналы (сигналы в дискретном времени) определяются в дискретные моменты времени и представляются последовательностями чисел (иногда употребляют термин решетчатой функцией). То есть такие сигналы задаются на счетном множестве точек, например, [0; $N - 1$]. На рис. 1.1 графически отображены сигнал непрерывного времени (черная сплошная линия) и дискретный сигнал, образованный его эквидистантными[5] выборками.

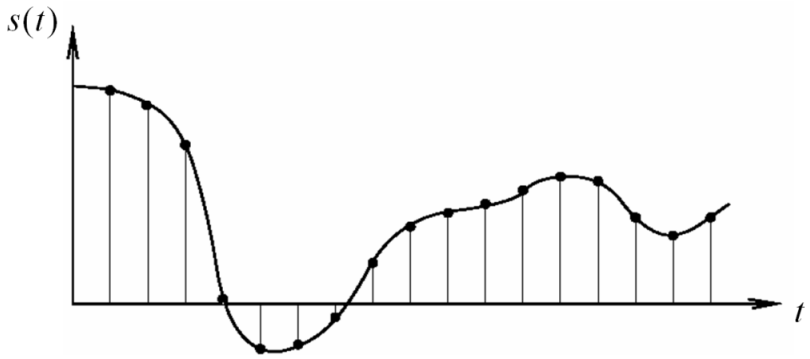


Рис. 1.1. Непрерывный и дискретный сигнал

В общем случае сигнал может быть как вещественным, так и комплексным, тогда его представление осуществляется через вещественную и мнимую части:

$$s(t) = s_1(t) + js_2(t); \quad j = \sqrt{-1}, \tag{1.1}$$

[5] Эквидистантный – равноудаленный; здесь подразумеваются равные отсчеты по времени.

где t принимает дискретные значения для дискретного сигнала и любые значения в области определения для аналогового сигнала.

Поскольку независимые переменные могут быть непрерывными или дискретными, амплитуда сигнала также может быть как непрерывной, так и дискретной. *Цифровые сигналы* – это сигналы, у которых дискретны и время, и амплитуда. Сигналы, которые дискретны по значению (амплитуде) и недискретны по переменной (времени), часто называют *квантованными* (по амплитуде).

Чтобы облегчить извлечение информации и представление ее в удобном для ЭВМ виде, сигналы должны подвергаться обработке. Техника обработки сигналов заключается в преобразовании сигнала в другой сигнал, являющийся более предпочтительным. Кроме того, может понадобиться разделение двух или большего числа сигналов, которые ранее были объединены некоторым образом; выделение некоторой компоненты или параметра сигнала либо оценка одного или нескольких параметров сигнала.

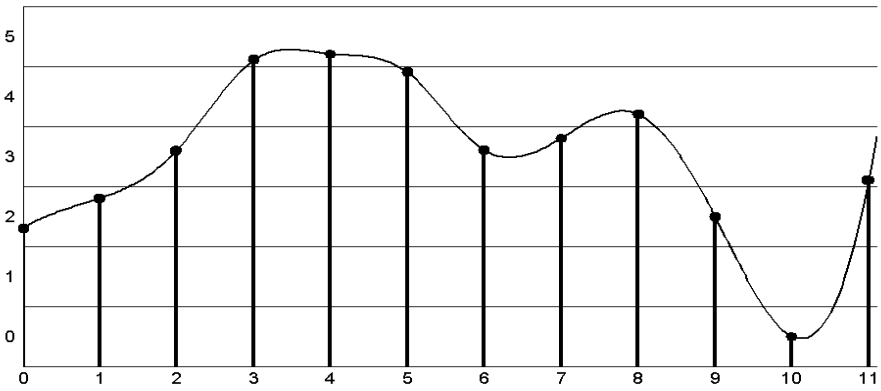
Системы обработки сигналов могут классифицироваться точно так же, как и сами сигналы. Так, системы в *непрерывном времени* – это системы, у которых на входе и выходе имеются сигналы в непрерывном времени, а *дискретные системы* (системы в дискретном времени) – это системы, у которых на входе и выходе дискретные сигналы. Точно так же *аналоговые системы* – это системы с аналоговыми сигналами на входе и выходе, а *цифровые системы* – системы с цифровыми сигналами на входе и выходе. В таком случае цифровая обработка сигналов имеет дело с преобразованиями сигналов, являющимися дискретными как по амплитуде, так и по времени.

Дискретные сигналы могут появляться при получении выборок из аналоговых сигналов, или же они могут порождаться непосредственно некоторым дискретным во времени процессом. Вне зависимости от происхождения дискретных сигналов цифровые системы обработки таких сигналов обладают рядом полезных качеств. Они могут быть реализованы с большой гибкостью на универсальных цифровых вычислительных машинах или с помощью цифровой аппаратуры. При необходимости их можно использовать для моделирования аналоговых систем или, что более важно, для преобразований сигнала, которые невозможно осуществить на аналоговой аппаратуре. Поэтому, когда требуется сложная обработка сигналов, часто желательно представить их в цифровом виде.

На рис. 1.2, *а* графически отображена процедура преобразования аналогового сигнала сначала в дискретный (дискретизация по времени),

затем в цифровой (квантование по уровню), образованный последовательностью десятичных цифр (рис. 1.2, *б*) и последовательностью двоичных кодовых групп (рис. 1.2, *в*). На рис. 1.2, *г* отображен график ошибки квантования, образованной разностью дискретного и квантованного сигналов.

а)



б) 2 2 3 5 5 4 3 3 4 2 0 3

в) 10 10 11 01 01 00 11 11 00 10 00 11

г)

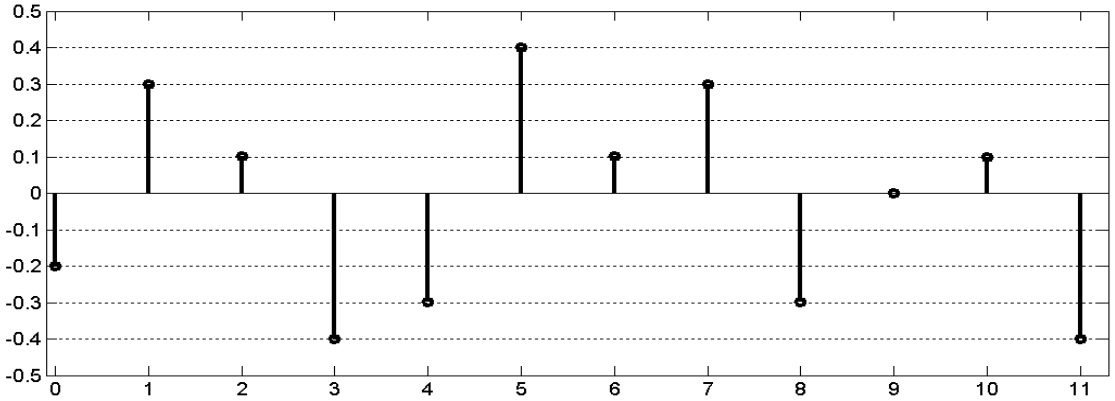


Рис. 1.2. Процедура преобразования аналогового сигнала

1.3. Представление дискретных сигналов в виде цифровых последовательностей

При цифровой обработке сигналы представляются в виде последовательности чисел – выборок или отсчетов. Последовательность чисел x, в которой n-й член последовательности обозначается как $x(n)$, может быть формально записана в виде

$$x = \{x(n)\}, \quad -\infty < n < \infty. \tag{1.2}$$

Хотя последовательности не всегда получаются путем выборки из аналоговых колебаний, для удобства мы будем называть $x(n)$ «n-й выборкой» последовательности. Хотя, строго говоря, $x(n)$ обозначает n-й член последовательности, запись (1.2) часто слишком громоздка и бо-

лее удобно говорить о «последовательности $x(n)$». Дискретные сигналы (последовательности) часто изображаются графически так, как это показано на рис. 1.3. Хотя абсцисса изображена в виде непрерывной линии, важно сознавать, что $x(n)$ определена только для целых значений n. Неправильно полагать, что $x(n)$ равна нулю для нецелых n, просто $x(n)$ не определена для нецелых значений n. Некоторые примеры последовательностей, играющих важную роль при дискретной обработке, показаны на рис. 1.4.

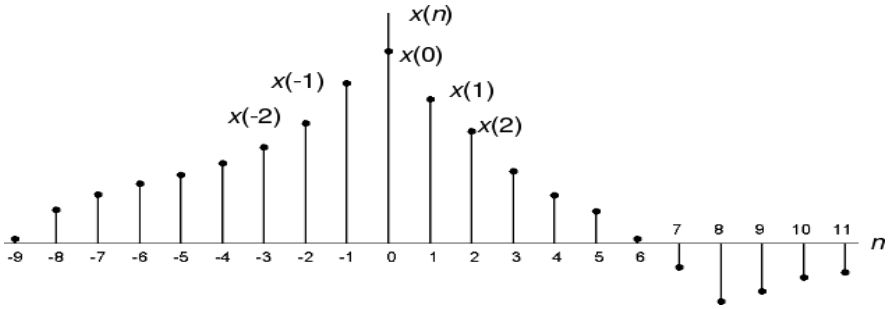


Рис. 1.3. Графическое представление дискретного сигнала

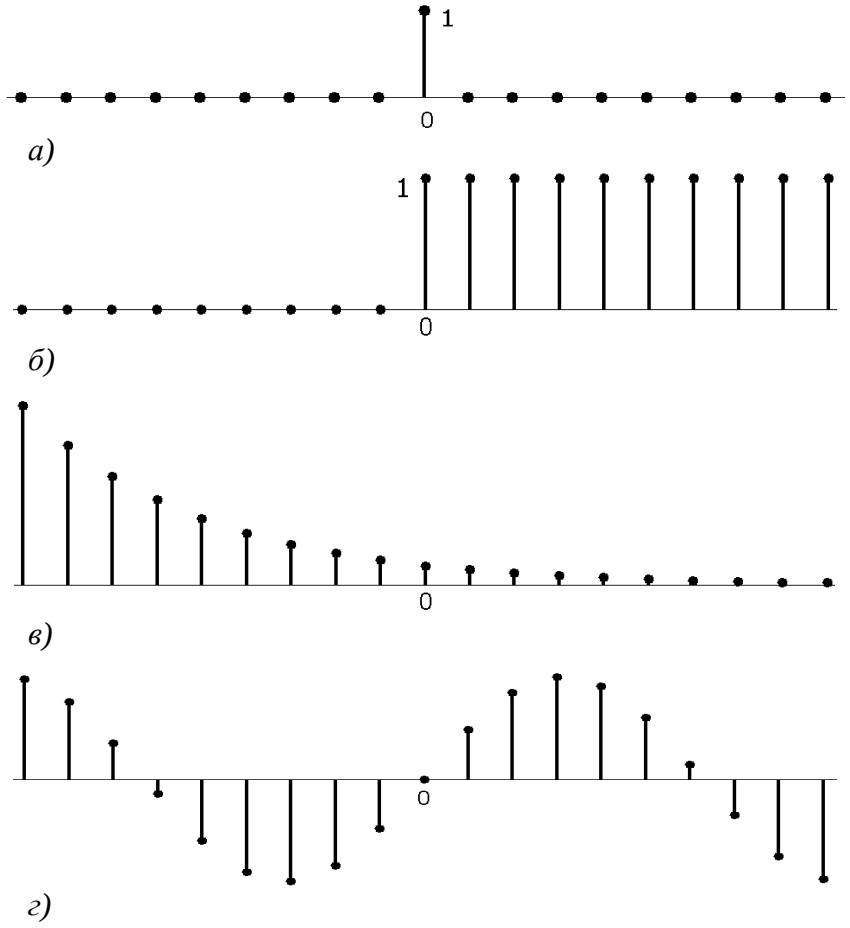


Рис. 1.4. Примеры последовательностей: *а*) единичный импульс; *б*) единичная ступенчатая последовательность; *в*) действительная экспоненциальная последовательность; *г*) синусоидальная последовательность

Действительная экспоненциальная последовательность – это последовательность со значениями вида a^n, где a – действительное число. Эту последовательность можно получить периодическим (с периодом T) взятием отсчетов (выборок) экспоненты непрерывного времени

$$e^{-\alpha t}\big|_{t=nT} = e^{-\alpha nT} = a^n,$$

где $a = e^{-\alpha T}$.

Синусоидальная последовательность имеет вид $A\cos(\omega_0 n + \varphi)$, и ее также можно получить периодическим (с периодом T) взятием отсчетов (выборок) синусоиды непрерывного времени

$$A\cos(\Omega_0 t + \varphi)_{|t=nT} = A\cos(\Omega_0 nT + \varphi) = A\cos(\omega_0 n + \varphi),$$

где $\omega_0 = \Omega_0 T$.

Комплексная экспоненциальная последовательность имеет вид

$$e^{(\sigma + j\omega_0)n} = e^{\sigma n}(\cos(\omega_0 n) + j\sin(\omega_0 n)).$$

Последовательность $x(n)$ по определению называется периодической с периодом N, если $x(n) = x(n+N)$ для всех n. Комплексная экспонента с $\sigma = 0$ и синусоидальная последовательность имеют период только тогда, когда это действительное число является целым. Если $2\pi/\omega_0$ не целое, но рациональное число, то синусоидальная последовательность будет периодической, однако с периодом, большим $2\pi/\omega_0$. Если $2\pi/\omega_0$ не рационально, то синусоидальная и комплексная экспоненциальная последовательности вовсе не будут периодическими. Параметр ω_0 будет называться цифровой частотой синусоиды или комплексной экспоненты вне зависимости от того, периодичны они или нет. Частота может быть выбрана в любом непрерывном диапазоне значений. Однако без потери общности можно ограничить этот диапазон, полагая $0 \le \omega_0 \le 2\pi$ (или $-\pi \le \omega_0 \le \pi$), так как синусоидальные и комплексные экспоненциальные последовательности, получаемые при изменении ω_0 в диапазоне $2\pi k \le \omega_0 \le 2\pi(k+1)$ в точности совпадают при любых k с последовательностями, получаемыми при изменении ω_0 в диапазоне $0 \le \omega_0 \le 2\pi$.

Иногда удобно пользоваться термином энергии последовательности. Энергия E последовательности $x(n)$ определяется как

$$E = \sum_{n=-\infty}^{\infty} |x(n)|^2. \qquad (1.3)$$

При анализе систем обработки дискретных сигналов приходится производить некоторые преобразования последовательностей.

Произведение и сумма двух последовательностей *x* и *y* определяются как произведение и сумма выборок соответственно: $xy = \{x(n), y(n)\}$. Умножение последовательности *x* на число *a* определяется как $xa = \{ax(n)\}$.

Последовательность *y* является задержанной или сдвинутой последовательностью *x*, если *y*(*n*) имеет значения $y(n) = x(n - n_0)$, где n_0 – целое число.

Произвольная последовательность может быть представлена как сумма взвешенных и задержанных единичных импульсов. Например, последовательность *p*(*n*), изображенную на рис. 1.5, можно записать как

$$p(n) = a_{-3}\delta(n+3) + a_1\delta(n-1) + a_2\delta(n-2) + a_7\delta(n-7).$$

В общем случае произвольная последовательность записывается в виде

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k). \qquad (1.4)$$

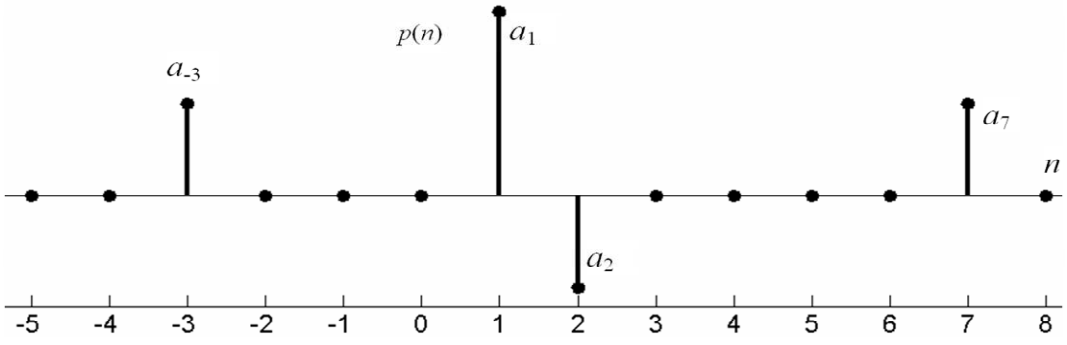


Рис. 1.5. Пример последовательности, представляющей сумму взвешенных и задержанных единичных импульсов

1.4. Представление дискретных сигналов в частотной области

Дискретные сигналы могут быть представлены различными способами, причем для дискретных сигналов важную роль играют синусоидальные и комплексные экспоненциальные последовательности. Это объясняется основным свойством линейных инвариантных к сдвигу систем, которое состоит в том, что в установившемся состоянии отклик на синусоидальный входной сигнал является синусоидой той же частоты с амплитудой и фазой, определяемыми системой. Именно это свойство линейных инвариантных к сдвигу систем делает представление сигналов через синусоиды и комплексные экспоненты таким полезным.

Чтобы убедиться в справедливости этого утверждения для дискретных систем, предположим, что входная последовательность является комплексной экспонентой круговой частоты ω, $x(n) = e^{j\omega n}$ для $-\infty < n < \infty$.

Тогда получим выходной сигнал системы с импульсной характеристикой $h(n)$:

$$y(n) = \sum_{k=-\infty}^{\infty} h(k)e^{j\omega(n-k)} = e^{j\omega n} \sum_{k=-\infty}^{\infty} h(k)e^{-j\omega k}. \qquad (1.5)$$

Если ввести

$$H(e^{j\omega}) = \sum_{k=-\infty}^{\infty} h(k)e^{-j\omega k}, \qquad (1.6)$$

то можно записать

$$y(n) = H(e^{j\omega})e^{j\omega n}. \qquad (1.7)$$

Отсюда видно, что $H(e^{j\omega})$ описывает изменение комплексной амплитуды комплексной экспоненты как функции частоты ω. Величина $H(e^{j\omega})$ называется *частотной характеристикой* системы, у которой импульсная характеристика равна $h(n)$. В общем случае $H(e^{j\omega})$ – комплексная функция и может быть выражена через свои действительную и мнимую части

$$H(e^{j\omega}) = H_{\mathrm{Re}}(e^{j\omega}) + H_{\mathrm{Im}}(e^{j\omega})$$

или через модуль и фазу

$$H(e^{j\omega}) = \left|H(e^{j\omega})\right| e^{j\cdot\arg(H(e^{j\omega}))}.$$

Иногда будет удобнее говорить о *групповой задержке*, а не о фазе. Групповая задержка определяется как взятая со знаком «минус» первая производная фазы по ω.

Поскольку синусоиду можно представить как линейную комбинацию комплексных экспонент, то частотная характеристика также выражает отклик на синусоидальный сигнал. Например, рассмотрим

$$x(n) = A\cos(\omega_0 n + \Phi) = (A/2)e^{j\Phi}e^{j\omega_0 n} + (A/2)e^{-j\Phi}e^{-j\omega_0 n}. \qquad (1.8)$$

Если $h(n)$ – действительная функция, то отклик на $(A/2)e^{-j\Phi}e^{-j\omega_0 n}$ является комплексно-сопряженным с откликом на $(A/2)e^{j\Phi}e^{j\omega_0 n}$.

Поэтому результирующий отклик равен

$$y(n) = (A/2)[H(e^{j\omega_0})e^{j\Phi}e^{j\omega_0 n} + H(e^{-j\omega_0})e^{-j\Phi}e^{-j\omega_0 n}] =$$
$$= (A/2)\left|H(e^{j\omega_0})\right|[e^{j[\omega_0 n+\Phi+\theta]} + e^{-j[\omega_0 n+\Phi+\theta]}] =$$
$$= \mathrm{Re}\{H(e^{j\omega_0})Ae^{j\Phi}e^{j\omega_0 n}\} = A\left|H(e^{j\omega_0})\right|\cos(\omega_0 n + \Phi + \theta),$$

где $\theta = \arg[H(e^{j\omega_0})]$ – значение фазочастотной характеристики системы на частоте ω_0.

Частотная характеристика $H(e^{j\omega})$ является непрерывной функцией частоты. Кроме того, это периодическая функция частоты ω с периодом 2π. Данное свойство следует непосредственно из (1.6), так как

$$e^{j(\omega_0 + 2\pi)k} = e^{j\omega k}.$$

Частотная характеристика имеет одинаковые значения на частотах ω и $\omega + 2\pi$, что означает следующее: система реагирует одинаково на комплексные экспоненты этих двух частот. Такое поведение понятно, так как эти две экспоненциальные последовательности совпадают друг с другом.

В качестве примера расчета частотной характеристики рассмотрим систему с импульсной характеристикой (рис. 1.6):

$$h(n) = \begin{cases} 1, & 0 \le n \le N-1; \\ 0 - & \text{в остальных случаях.} \end{cases} \qquad (1.9)$$

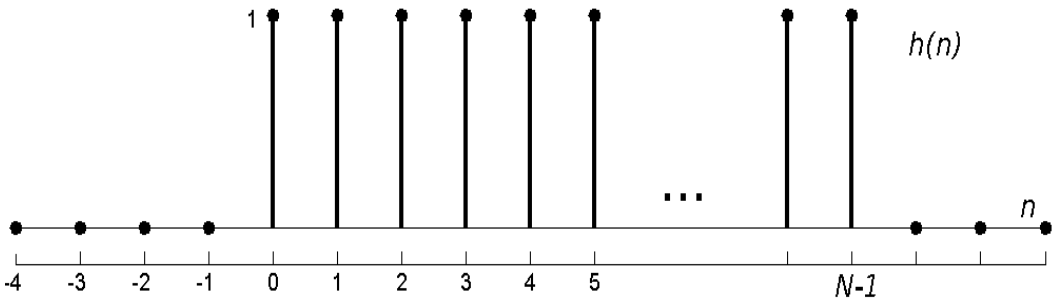


Рис. 1.6. Импульсная характеристика системы, для которой рассчитывается частотная характеристика

Частотная характеристика равна

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} e^{-j\omega n} = \frac{1 - e^{-j\omega N}}{1 - e^{-j\omega}} = \frac{\sin(\omega N/2)}{\sin(\omega/2)} e^{-j(N-1)\omega/2}. \qquad (1.10)$$

Модуль и фаза $H(e^{j\omega})$ изображены на рис. 1.7 для случая $N = 5$.

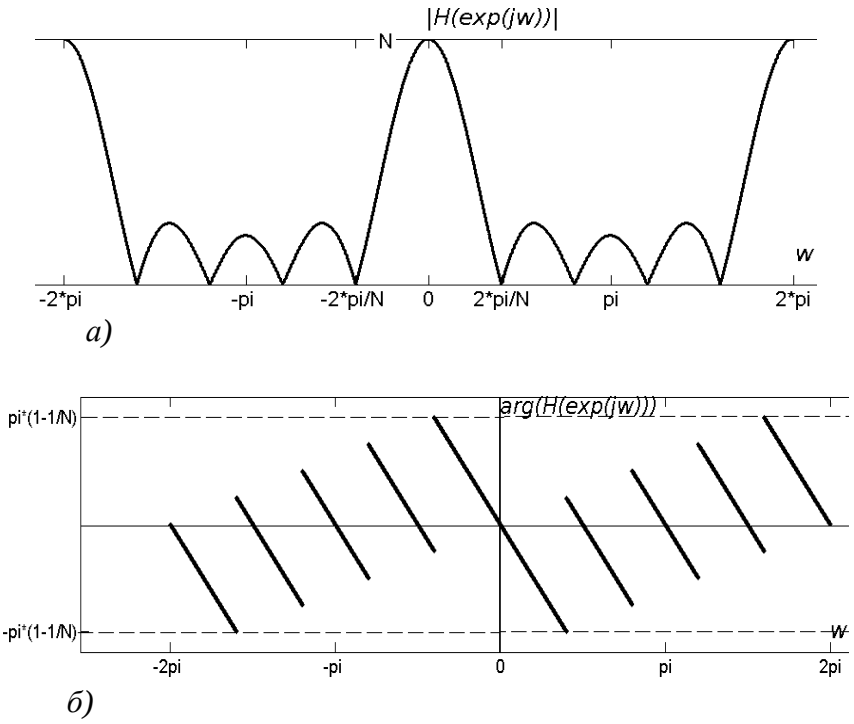


Рис. 1.7. Частотная характеристика системы с импульсной характеристикой, изображенной на рис. 1.6: модуль *(а)* и фаза *(б)*

Поскольку $H(e^{j\omega})$ – периодическая функция частоты, она может быть представлена в виде ряда Фурье. Фактически (1.6) и представляет $H(e^{j\omega})$ в виде ряда Фурье, в котором коэффициентами Фурье являются значения импульсной характеристики $h(n)$. Отсюда следует, что $h(n)$ могут быть определены через $H(e^{j\omega})$ как коэффициенты Фурье периодической функции, т. е.

$$h(n) = (1/2\pi)\int_{-\pi}^{\pi} H(e^{j\omega})e^{j\omega n}d\omega, \qquad (1.11)$$

где

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n)e^{-j\omega n}. \qquad (1.12)$$

Эти равенства можно также трактовать как представление последовательности $h(n)$, а именно, полезно рассматривать (1.11) как представление последовательности $h(n)$ в виде суперпозиции (интеграла) экспоненциальных сигналов, комплексные амплитуды которых определяются выражением (1.12). Таким образом, (1.11) и (1.12) являются парой преобразований Фурье для последовательности $h(n)$, где (1.12) играет роль прямого, а (1.11) обратного преобразования Фурье. Такое представление существует только тогда, когда ряд в (1.12) сходится.

Представление последовательности преобразованием (1.12) не ограничивается только импульсной характеристикой системы и будет справедливо для любой последовательности при условии, что ряд в (1.12) сходится. Поэтому для произвольной последовательности $x(n)$ мы определим преобразование Фурье дискретного времени (ДВПФ) соотношением

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n}, \tag{1.13}$$

а обратное преобразование Фурье – соотношением

$$x(n) = (1/2\pi)\int_{-\pi}^{\pi} X(e^{j\omega})e^{j\omega n}d\omega. \tag{1.14}$$

Ряды (1.13) не всегда сходятся, как, например, в случаях, когда $x(n)$ – единичная ступенчатая последовательность либо действительная или комплексная экспоненциальная последовательность для всех n. Имеются различные определения и интерпретации сходимости преобразования Фурье. Если $x(n)$ абсолютно суммируема, т. е. если $\sum_{n=-\infty}^{\infty} |x(n)| < \infty$, то ряд называется абсолютно сходящимся и сходится равномерно к непрерывной функции ω. Поэтому частотная характеристика устойчивой системы будет всегда сходиться. Если последовательность абсолютно суммируема, то она будет также иметь конечную энергию, т. е. $\sum_{n=-\infty}^{\infty} |x(n)|^2$ будет конечной. Это прямо следует из неравенства

$$\sum_{n=-\infty}^{\infty} |x(n)|^2 \le \left[\sum_{n=-\infty}^{\infty} |x(n)|\right]^2.$$

Однако нельзя считать, что последовательность с конечной энергией абсолютно суммируема. Примером последовательности, имеющей конечную энергию, но не абсолютно суммируемой, является последовательность

$$x(n) = \sin(\omega_0 n/\pi n).$$

Если последовательность не является абсолютно суммируемой, но имеет конечную энергию, то можно использовать тип сходимости, при которой ряд сходится так, что среднеквадратическая ошибка равна нулю.

Возможность представления последовательности как суперпозиции комплексных экспонент является очень важным качеством при анализе линейных систем с постоянными параметрами. Именно вследствие этого факта и принципа суперпозиции реакция такой системы на комплексную экспоненту полностью определяется частотной характеристикой $H(e^{j\omega})$. Если рассматривать (1.14) как суперпозицию комплексных экспонент бес-

конечно малой амплитуды, то отклик линейной системы с постоянными параметрами на $x(n)$ является суперпозицией откликов на каждую экспоненту, входящую в представление сигнала $x(n)$. Так как отклик на каждую комплексную экспоненту получается умножением на $H(e^{j\omega})$, то

$$y(n) = T[x(n)] = T\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})e^{j\omega n}d\omega\right] =$$

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

Поэтому преобразование Фурье выходного сигнала равно

$$Y(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega}). \qquad (1.15)$$

Этот результат имеет свой аналог в теории линейных систем с непрерывным временем и может быть, конечно, получен более строгим образом путем применения преобразования Фурье к свертке

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k).$$

Некоторые наиболее важные свойства ДВПФ приведены в табл. 1.1.

Таблица 1.1

Некоторые важные свойства ДВПФ

Последовательность	ДВПФ
$x(n)$	$X(e^{j\omega})$
$x(n-m)$	$X(e^{j\omega})e^{-j\omega m}$
$x(n)\cdot e^{j\omega_0 n}$	$X(e^{j(\omega-\omega_0)})$
$\sum_{k=-\infty}^{\infty} x(k)h(n-k)$	$X(e^{j\omega})H(e^{j\omega})$
$x(n)y(n)$	$\frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\theta})Y(e^{j(\omega-\theta)})d\theta$

1.5. Частотно-временные деформации дискретного сигнала

Для сигнала непрерывного времени свойство частотно-временной деформации состоит в следующем. Если спектр (преобразование Фурье) сигнала $x(t)$ равен $X(j\Omega)$, то для сигнала $x(Mt)$ спектр изменяется к виду

$$\frac{1}{M}X\left[j\frac{\Omega}{M}\right].$$

Другими словами, растяжение сигнала во времени ($M < 1$) приводит к сжатию его спектра по оси частот и наоборот, сжатие сигнала во времени

($M > 1$) растягивает его спектр по оси частот, что также сопровождается соответствующим масштабированием спектров. Аналогичные свойства имеют место и для дискретного сигнала $x(n)$, полученного в результате периодической (с периодом T) дискретизации сигнала $x(t)$. При этом сжатие и растяжение (удлинение) дискретного сигнала эквивалентны соответственно уменьшению либо увеличению частоты дискретизации $F_д = 1/T$. Рассмотрим эти процедуры подробнее.

1. Уменьшение частоты дискретизации в целое число M раз.

Эта процедура называется также *прореживанием или децимацией* дискретного сигнала $x(n)$. При этом новый дискретный сигнал $x_d(n)$, получается прореживанием исходного, т. е. сохранением лишь каждого M-го отсчета сигнала $x(n)$: $x_d(n) = x(Mn)$. Такой сигнал может быть получен в результате периодической (с периодом $T_1 = MT$) дискретизации сигнала $x(t)$, т. е. с частотой дискретизации $F_{д1} = F_д / M$.

Соотношение, связывающее спектры дискретного сигнала $x(n)$ и сигнала непрерывного времени $x(t)$, запишем в виде

$$X(e^{j\omega}) = \frac{1}{T}\sum_{n=-\infty}^{\infty} X\left[j\left(\frac{\omega}{T} - \frac{2\pi n}{T}\right)\right]. \qquad (1.16)$$

На основе (1.16) графически представлены спектры дискретных сигналов $x(n)$ и $x_d(n)$ для $M = 2$ (рис. 1.8). Чтобы гарантировать отсутствие эффекта наложения в связи с расширением спектра в M раз при прореживании дискретного сигнала, необходимо перед прореживанием пропустить его через дискретный фильтр нижних частот (ДФНЧ) с полосой пропускания $|\omega| \le \pi / M$. Таким образом, совокупность каскадно включенных ДФНЧ с полосой пропускания $|\omega| \le \pi / M$ и прореживателя в M раз выполнит операцию уменьшения частоты дискретизации в M раз без эффекта наложения. Такую систему называют *компрессором* частоты дискретизации (рис. 1.9).

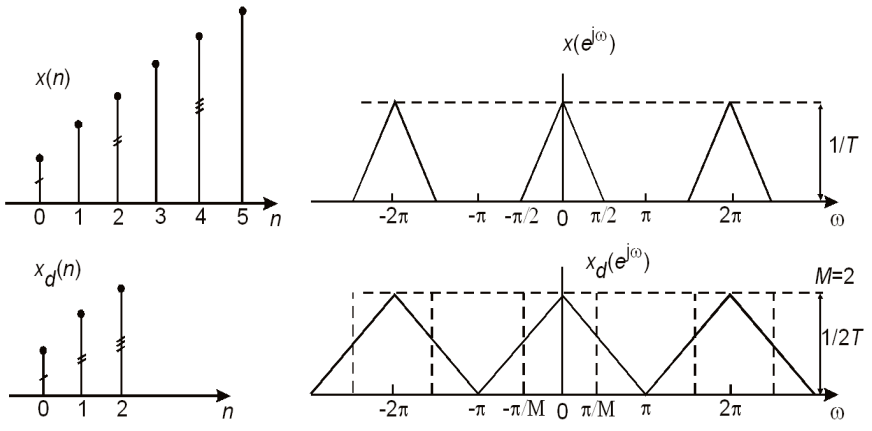


Рис. 1.8. Спектр дискретного сигнала

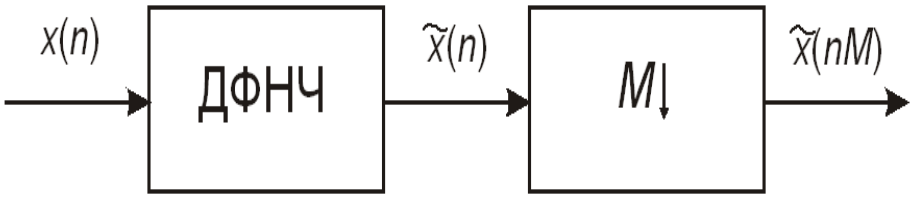


Рис. 1.9. Компрессор частоты дискретизации

2. Увеличение частоты дискретизации в целое число L раз.

Эту процедуру называют *интерполяцией* или восстановлением отсутствующих отсчетов дискретного сигнала. Для интерполяции дискретного сигнала $x(n)$ его сначала удлиняют в L раз путем вставления между каждыми двумя соседними отсчетами $L-1$ нулевых отсчетов. Такой сигнал $x_{уд}(n)$ можно записать в виде

$$x_{уд}(n)=\begin{cases} x(n/L), & n=kL, \\ 0, & n\neq kL. \end{cases}$$

Следовательно, спектр сигнала $x_{уд}(n)$ примет форму:

$$\begin{aligned} X_{уд}(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} x_{уд}(n)e^{-j\omega n} = \\ &= \sum_{n=kL} x(n/L)e^{-j\omega n} = \sum_{k=-\infty}^{\infty} x(k)e^{-j\omega Lk} = X(e^{j\omega L}). \end{aligned} \qquad (1.17)$$

Из (1.17) следует вывод о том, что при описанной выше операции удлинения дискретного сигнала, спектр сигнала $x_{уд}(n)$ сжимается по оси частот в L раз в сравнении со спектром исходного сигнала, а период повторения будет равен $2\pi/L$. Если затем пропустить $x_{уд}(n)$ через идеальный ДФНЧ с полосой пропускания $|\omega|\le\pi/L$ и усилением L, то лишние спектральные полосы будут удалены, а выходной сигнал фильтра будет точно соответствовать дискретному сигналу $x_{ин}(n)$ с периодом дискретизации T/L, т. е. частота его дискретизации увеличится в L раз, а нужные отсчеты будут восстановлены. На рис. 1.10 представлены графически: а) исходный сигнал и его спектр; б) сигнал, удлиненный добавлением нулевых отсчетов и его спектр; в) интерполированный сигнал на выходе ДФНЧ и его спектр. На рис. 1.11 представлена структура интерполятора (*экспандера* частоты дискретизации), увеличивающего частоту дискретизации в L раз.

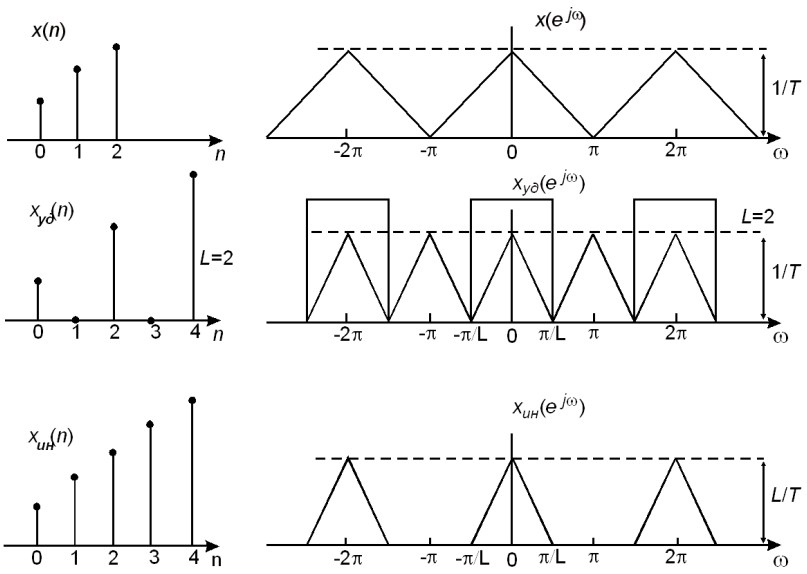


Рис. 1.10. Преобразование сигнала

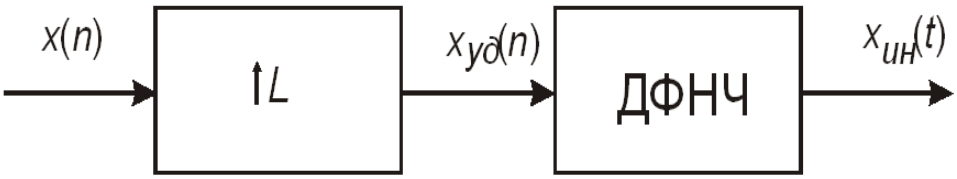


Рис. 1.11. Структура интерполятора

Если необходимо изменить частоту дискретизации исходного дискретного сигнала в L/M (рациональное число) раз, то такая операция может быть выполнена каскадным соединением интерполятора в L раз и дециматора в M раз, как это представлено на рис. 1.12. Два последовательно включенных дискретных фильтра нижних частот ДФНЧ1 и ДФНЧ2 могут быть заменены одним с меньшей частотой среза.

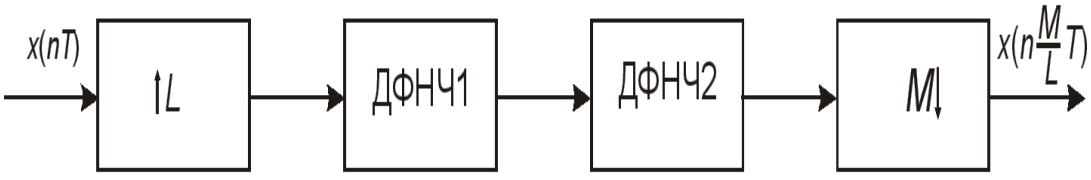


Рис. 1.12. Интерполятор и дециматор

2. ОСНОВЫ ДИСКРЕТНОЙ АРИФМЕТИКИ

2.1. Представление числа

Наиболее распространенной является позиционная система представления чисел – десятичная, двоичная и вообще m-ичная. Здесь m – основание системы счисления (целое положительное число). При применении позиционной системы любое целое число x на интервале $N = m^n$ может быть записано с помощью n разрядов:

$$x = \left(x_1 x_2 \ldots x_n\right)_m,$$

где каждый разряд представлен своим коэффициентом, $0 \le x_1 < m$ и крайний слева разряд является старшим. Такая условная запись соответствует представлению числа в виде выражения

$$x = x_1 \cdot m^{n-1} + x_2 \cdot m^{n-2} + \ldots + x_{n-1} \cdot m + x_n = \sum_{i=1}^{n} x_i \cdot m^{n-i}. \qquad (2.1)$$

Возможна также многоосновная система счисления, в которой

$$x = \sum_{i=1}^{n} x_i \cdot m_1 \cdot m_2 \cdot \ldots \cdot m_{n-i},$$

где $m_1, m_2, \ldots, m_{n-i}$ – произвольные целые положительные числа (основания системы). Такое представление чисел может быть положено в основу более общей теории дискретных сигналов на конечных интервалах.

Перепишем выражение в виде

$$x = m^n \cdot \sum_{i=1}^{n} x_i \cdot m^{-i} = N \cdot \sum_{i=1}^{n} x_i \cdot m^{-i}$$

и введем нормированную переменную $\tau_N = x/N$, которая лежит на интервале $0 \le \tau_N < 1$. Тогда выражение

$$\tau_N = \sum_{i=1}^{n} x_i \cdot m^{-i}$$

является m-ичным представлением рационального числа τ_N, принимающего N равноотстоящих значений на континуальном интервале $[0,1)$. Положив $n \to \infty$, можно получить m-ичное представление переменной τ, принимающей всевозможные (рациональные) значения на интервале $[0,1)$:

$$\tau = \sum_{i=1}^{\infty} x_i \cdot m^{-i}. \qquad (2.2)$$

По существу, эта запись есть представление рационального числа в виде дроби по отрицательным степеням модуля m (в частности, при $m = 10$

это десятичная дробь, конечная или бесконечная периодическая). Как известно, любое иррациональное число может быть также выражено в виде бесконечной (но уже непериодической) дроби, поэтому выражение может рассматриваться как представление любого действительного числа на непрерывном интервале $[0,1)$.

Важное значение имеет операция, которая носит название *m*-ичной инверсии числа и сводится к записи разрядов этого числа в обратном порядке. Например, если имеется число

$$p = (p_1 p_2 \dots p_n)_m,$$

то инверсным ему будет число

$$\langle p \rangle_m = (p_n p_{n-1} \dots p_1)_m$$

Пусть это инверсное число равно *h*:

$$\langle p \rangle_m = h = (h_1 h_2 \dots h_n)_m.$$

Это значит, что разряды чисел p и h связаны соотношениями $h_i = p_{n+1-i}$ и $p_i = h_{n+1-i}$. Для примера на рис. 2.1 показаны *m*-ично инверсные числа на интервале $N = 16$ при $m = 4$.

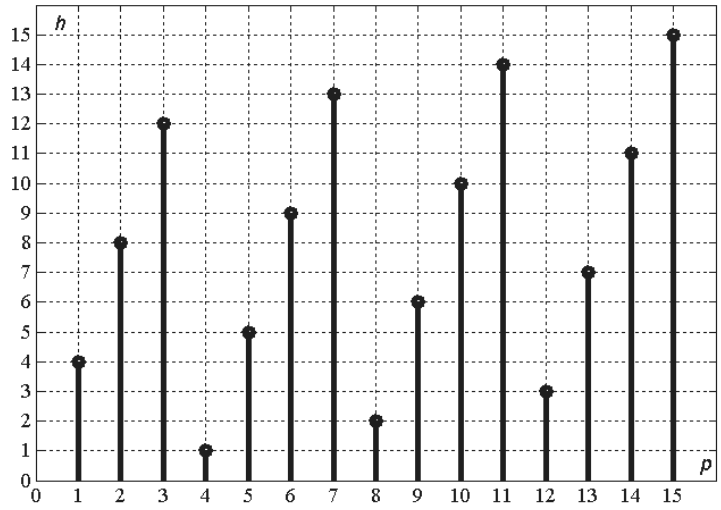


Рис. 2.1. *m*-ично инверсные числа

Инверсия числа обладает рядом замечательных свойств, которые ясно видны из рис. 2.1. Перечислим основные из них.

1) *m*-ичная инверсия взаимна, т. е. если $h = \langle p \rangle_m$, то $p = \langle h \rangle_m$. Вследствие этого переход от p к h имеет те же свойства, что и обратный переход от h к p.

2) Пусть имеются два множества *P* и *H*, содержащие одни и те же элементы, но пронумерованные по-разному, в порядке следования *m*-ично инверсных чисел p и h. Это свойство можно сформулировать так: *m*-

кратное повторение признаков в P приводит к периодическому продолжению признаков в H с периодом m.

3) Пусть в P признаки $A_2, A_3, \ldots, A_m$ одинаковы и означают, что обладающие ими элементы являются нулевыми. Тогда в H нулевые элементы заполнят промежутки между элементами с номерами $h = 0, m, 2m, \ldots, m^n$. Это свойство является частным случаем предыдущего. Коротко его можно сформулировать так: усечение в P приводит к прореживанию в H.

2.2. Сравнения

Пусть задано целое положительное число m, которое назовем модулем. Любое целое число a можно рассматривать в связи с остатком от деления его на модуль

$$a = \lceil a/m \rceil \cdot m + \left((a)\right)_m, \tag{2.3}$$

где $\lceil\ \rceil$ означает целую часть, а $\left((a)\right)_m$ – остаток; например $\left((63)\right)_3 = 0$, $\left((7)\right)_6 = 1$ и т. д.

Если два целых числа a и b имеют равные остатки, $\left((a)\right)_m = \left((b)\right)_m$, то они называются сравнимыми по модулю m, что записывается в виде сравнения

$$a \equiv b \left(\mathrm{mod}\ m\right). \tag{2.4}$$

Такая запись означает просто, что $a - b$ делится нацело на m, или что $a = b + m \cdot t$, где t – целое. Очевидно, что всегда число и его остаток сравнимы по модулю m.

Основные свойства сравнений таковы:

1. Если $a_1 \equiv b_1 \left(\mathrm{mod}\ m\right)$ и $a_2 \equiv b_2 \left(\mathrm{mod}\ m\right)$, то

$$a_1 + a_2 \equiv b_1 + b_2 \left(\mathrm{mod}\ m\right),$$
$$a_1 - a_2 \equiv b_1 - b_2 \left(\mathrm{mod}\ m\right),$$
$$a_1 \cdot a_2 \equiv b_1 \cdot b_2 \left(\mathrm{mod}\ m\right).$$

2. Сравнение не нарушается, если обе его части умножить на одно и то же число.

3. Обе части сравнения можно разделить на один и тот же множитель только в том случае, когда этот множитель взаимно прост с модулем. Например, из сравнений $42 = 2 \left(\mathrm{mod}\ 10\right)$ следует, что $14 = 4 \left(\mathrm{mod}\ 10\right)$, но $7 \neq 2 \left(\mathrm{mod}\ 10\right)$.

4. Если сравнение имеет место по модулю m, то оно справедливо и по модулю d, где d – любой делитель числа m.

В некоторых задачах сравнение содержит неизвестное число. Решить сравнение – значит найти все значения этого числа, удовлетворяющие сравнению. Например, сравнению $x^5 + x + 1 = 0 \pmod 7$ удовлетворяют два решения: $x = 2 \pmod 7$ и $x = 4 \pmod 7$.

Суммой чисел a_1 и a_2 по модулю m называется величина $\left((a_1 + a_2)\right)_m$. Число $\hat{a}$ называется противоположным числу a, если $\left((a + \hat{a})\right)_m = 0$. Хотя противоположное число и является положительным, оно выполняет ту же роль, что и отрицательное число в обычной арифметике.

Вычесть из a_1 число a_2 по модулю m означает прибавить к a_1 число, противоположное a_2, поэтому разностью чисел a_1 и a_2 по модулю m называется величина

$$\left((a_1 - a_2)\right)_m = \left((a_1 + \hat{a}_2)\right)_m = 0.$$

Например, $\left((1-2)\right)_3 = \left((1+1)\right)_3 = 2$, $\left((7-3)\right)_2 = \left((7+1)\right)_2 = 0$.

Поскольку операция вычитания сводится к сложению с противоположным числом, то ниже для краткости в основном будет идти речь только об операции сложения.

Нетрудно убедиться, что при сложении по модулю m сохраняются коммутативный и ассоциативный законы:

$$\left((x + y)\right)_m = \left((y + x)\right)_m.$$

В простейшем случае, когда $m = 2$ и числа задаются на интервале $N = 2$, операции сложения и вычитания по модулю 2 совпадают и выполняются по правилам:

$$\left((0+0)\right)_2 = \left((0-0)\right)_2 = 0;$$
$$\left((0+1)\right)_2 = \left((0-1)\right)_2 = 1;$$
$$\left((1+0)\right)_2 = \left((1-0)\right)_2 = 1;$$
$$\left((1+1)\right)_2 = \left((1-1)\right)_2 = 0.$$

При $m \to \infty$ операция сложения по модулю переходит в обычную арифметическую операцию сложения.

Пусть имеется периодическая дискретная функция с периодом N, т. е. $f(x) = f(x + N)$, на интервале, равном также N. При обычном смещении этой функции вдоль оси x на величину y, ее часть $F(x)$, попадающая в интервал N, может рассматриваться как результат сдвига исходной функции по модулю N, т. е. как круговая перестановка

$$F(x) = \left[f\left((x + N)\right)_N \right].$$

2.3. *m*-сдвиг

Рассмотрим еще более общую, чем сложение чисел по модулю, операцию, которая также подчиняется коммутативному и ассоциативному законам арифметики. Эта операция – поразрядное сложение по модулю.

Пусть a и b – это целые положительные числа, заданные на интервале $N = m^n$, где m и n – также целые положительные числа, причем $2 \le m \le N$. Числа a и b можно представить в m-ичной системе счисления в виде:

$$a = (a_1 a_2 \ldots a_n)_m = \sum_{i=1}^{n} a_i \cdot m^{n-i},$$

$$b = (b_1 b_2 \ldots b_n)_m = \sum_{i=1}^{n} b_i \cdot m^{n-i}.$$

Поразрядной суммой чисел a и b по модулю m, которую обозначим $a \underset{m}{\oplus} b$, называется число

$$c = a \underset{m}{\oplus} b = (c_1 c_2 \ldots c_n)_m = \sum_{i=1}^{n} c_i \cdot m^{n-i}, \qquad (2.5)$$

где

$$c = \left((a_i + b_i) \right)_m.$$

Введем m-ично противоположное число b^*, удовлетворяющее условию

$$b \underset{m}{\oplus} b^* = 0. \qquad (2.6)$$

Очевидно, что это условие выполняется, если для каждого разряда чисел b и b^* справедливо одно из следующих равносильных соотношений:

$$\left(\left(b_i + b_i^* \right) \right)_m = 0,\ b_i^* = \left((m - b_i) \right)_m,\ b_i^* = \begin{cases} m - b_i, b_i \neq 0 \\ 0, b_i = 0 \end{cases}. \qquad (2.7)$$

Тогда поразрядной разностью чисел a и b по модулю m будет положительное число

$$d = a \underset{m}{-} b = a \underset{m}{\oplus} b^*. \qquad (2.8)$$

Другими словами, поразрядное вычитание по модулю m – это то же самое, что и поразрядное сложение с m-ично противоположным числом по тому же модулю.

При $m = 2$ каждая точка числовой оси m-ично противоположна самой себе, а при $m = N$ m-ично противоположные точки равноудалены от концов интервала $x^* = N - x$. Рисунок 2.2 дает представление о расположении m-ично противоположных чисел.

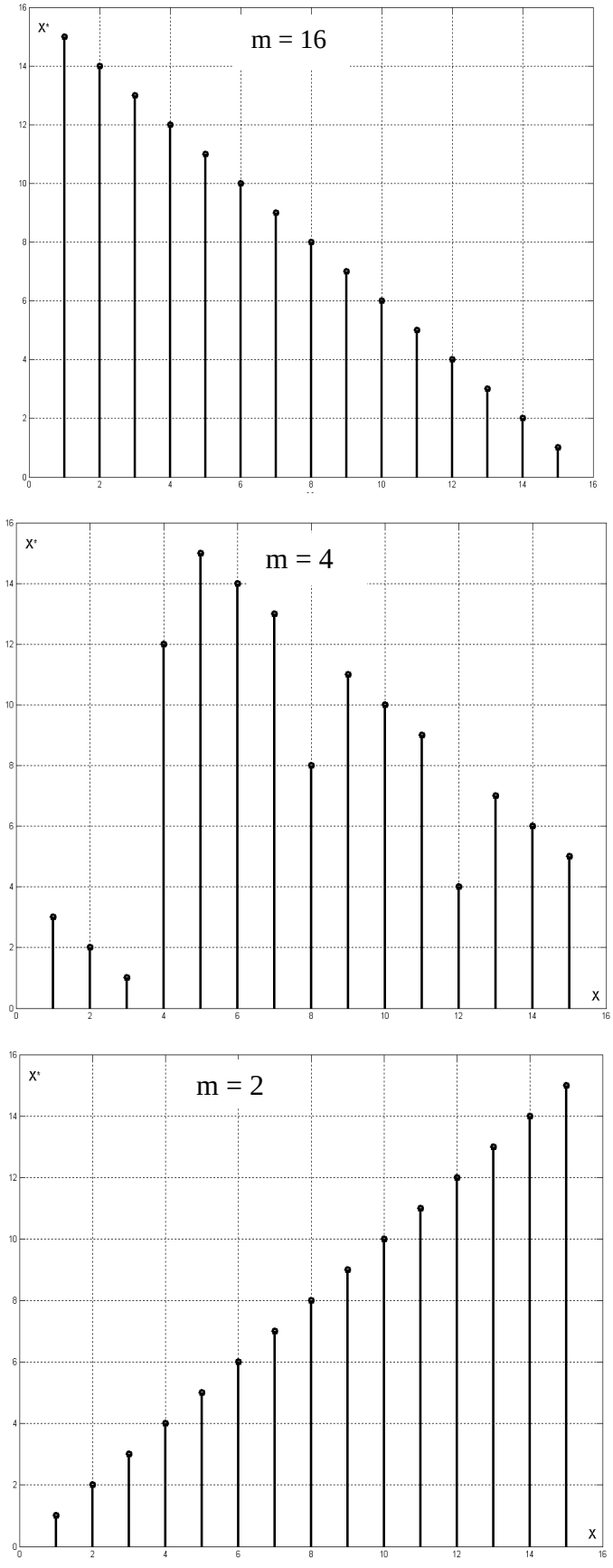


Рис. 2.2. *m*-ично противоположные числа для разных *m*

Поразрядная сумма по модулю m образуется путем поразрядного суммирования без переноса единицы в старший разряд. Понятно, что при таком способе образования сумма c или разность d всегда будут принадлежать тому же интервалу N, что и слагаемые a и b. В дальнейшем для краткости операции, связанные с таким способом сложения или вычитания, будем называть m-сдвигом.

Для выяснения сущности m-сдвига рассмотрим конкретный пример. Пусть $m=5$, $n=2$, $N=m^n=25$, $a=0,1,2,\ldots,24$ и $b=13$. Имея в виду, что число 13 в пятеричной системе записывается как $(b_1b_2)_5=(23)_5$, получим для $d=a\underset{m}{\oplus}b$ результат, изображенный на рис. 2.3. Из него видно, что m-сдвиг массива чисел $\{a\}$ на b единиц сводится к перестановке чисел.

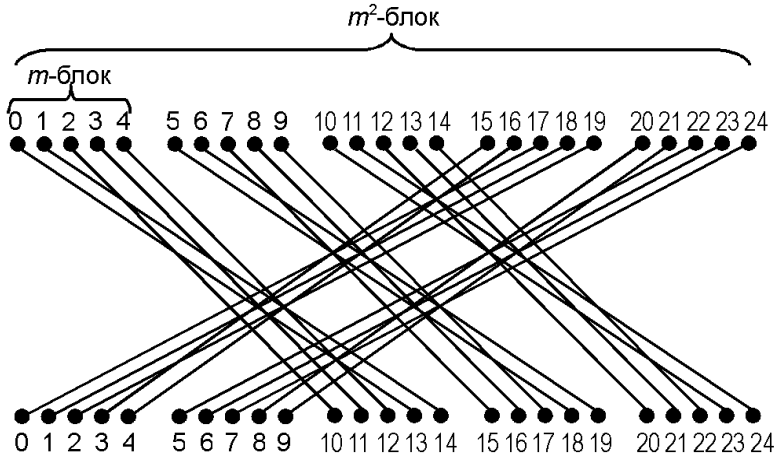


Рис. 2.3. m-сдвиг последовательности чисел

Для того чтобы пояснить механизм такой перестановки, разобьем интервал N определения массива $\{a\}$ сначала на самые мелкие подынтервалы по m чисел в каждом, затем разобьем тот же интервал N на более крупные подынтервалы по m^2 чисел в каждом, на еще более крупные подынтервалы по m^3 чисел в каждом и т. д. Для простоты назовем все эти подынтервалы соответственно m-блоками, m^2-блоками и т. д. Очевидно, что их число на интервале N соответственно будет составлять $m^{n-1},m^{n-2},\ldots,m,1$.

Тогда перестановка чисел при m-сдвиге $d=a\underset{m}{\oplus}b$ сводится к следующему.

1) Внутри каждого m-блока производится круговая перестановка чисел путем сдвига на b_n единиц.

2) Внутри каждого m^2-блока производится круговая перестановка m-блоков путем сдвига на b_{n-1} m-блоков.

3) Внутри каждого m^3-блока производится круговая перестановка m^2-блоков путем сдвига на b_{n-2} m^2-блоков и т. д.

Этот процесс продолжается n раз, пока не будут использованы все разрядные коэффициенты m-ичного представления числа b. В общем случае будет произведено n круговых перестановок в блоках, содержащих m, m^2, ..., $m^N = N$ чисел так, что m-сдвиг – это как бы многопетлевая круговая перестановка на интервале из N чисел.

Отметим, что m-сдвиг массива из N-чисел вырождается при $m = N$ в однократную круговую перестановку. В другом частном случае, когда $m = 2$ и $N = 2^n$, для m-сдвига справедливо выражение

$$c = a \underset{2}{\oplus} b = a + b - 2\sum_{i=1}^{n} a_i \cdot b_i \cdot 2^{n-i}.$$

Введенное понятие m-сдвига массива чисел требует изменения некоторых привычных представлений. Запаздывание дискретной функции $f(x)$ на τ единиц необходимо теперь рассматривать как ее m-сдвиг на τ-точек, т. е. как образование функции $f\left(x \underset{m}{\oplus} \tau\right)$ путем указанной выше перестановки отсчетов исходной функции.

Симметричными относительно точки x будут отсчеты $f\left(x \underset{m}{\oplus} \tau\right)$ и $f\left(x \underset{m}{\oplus} \tau\right)$, а соседними с отсчетом $f(x)$ – отсчеты $f\left(x \underset{m}{\oplus} 1\right)$ и $f\left(x \underset{m}{\oplus} 2\right)$ и т. д.

Определим четную и нечетную дискретные функции в смысле m -сдвига как функции, обладающие соответственно свойствами

$$f_ч\left(x^*\right) = f_ч(x),\ f_н\left(x^*\right) = -f_н(x).$$

Такие функции для краткости будем называть m-четной и m-нечетной. Следовательно, если $m = 2$, то любая дискретная функция является m-четной, а в случае $m = N$, m-четная и m-нечетная функции на интервале N имеют обычный смысл.

2.4. Линейное векторное пространство

Дискретные функции на конечном интервале N – это некоторый набор чисел или числовой вектор. Будем рассматривать такие векторы принадлежащими N - мерному линейному пространству и определим свойства такого пространства. Обозначим через $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ некоторые векторы и через λ и μ действительные или комплексные числа.

Векторное пространство является линейным. Это значит, что в нем определены две операции над векторами – сложение векторов и умноже-

ние вектора на скаляр, в результате которых образуется новый вектор в том же пространстве, причем эти операции должны удовлетворять следующим аксиомам:

1) $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ (коммутативность).

2) $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$ (ассоциативность).

3) $\mathbf{a}+0=\mathbf{a}$ (существование нуля).

4) $\mathbf{a}+(-\mathbf{a})=0$ (существование противоположного элемента).

5) $\lambda(\mu\mathbf{a})=\lambda\mu\mathbf{a}$.

6) $\mathbf{a}\cdot 1=\mathbf{a}$.

7,8) $\begin{cases}(\lambda+\mu)\mathbf{a}=\lambda\mathbf{a}+\mu\mathbf{a}\\ \lambda\cdot(\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b}\end{cases}$ (дистрибутивность).

Если числа, определяющие координаты вектора и скаляры действительные, то векторное пространство называется *действительным*. Оно может быть и комплексным, если все используемые числа – *комплексные*.

Указанные свойства линейного пространства еще не определяют такие геометрические свойства, как длину вектора, расстояние между векторами, угол между векторами и др., называемые метрическими свойствами. Далее будет показано, что метрика векторного пространства автоматически определяется, если в нем введена еще одна линейная операция $(\mathbf{a},\mathbf{b})$ – скалярное произведение. Это линейная операция над двумя векторами, в результате которой образуется не вектор, а скаляр. Она должна удовлетворять следующим трем аксиомам:

1) $(\mathbf{a},\mathbf{b})=\overline{(\mathbf{b},\mathbf{a})}$.

2) $(\mathbf{a},\mathbf{b}+1)=(\mathbf{a},\mathbf{b})+(\mathbf{a},\mathbf{c})$.

3) $(\lambda\mathbf{a},\mathbf{b})=\lambda(\mathbf{a},\mathbf{b}),(\mathbf{a},\lambda\mathbf{b})=\overline{\lambda}(\mathbf{a},\mathbf{b})$.

Если скалярное произведение удовлетворяет еще одному требованию $(\mathbf{a},\mathbf{a})\geq 0$, причем $(\mathbf{a},\mathbf{a})=0$ только при $\mathbf{a}=0$, то векторное пространство называется унитарным. Действительное конечномерное унитарное пространство называется *евклидовым*. Два вектора, для которых $(\mathbf{a},\mathbf{b})=0$, называются *ортогональными*.

Свойства векторного пространства (линейность, метричность и т. д.) сформулированы без привлечения какой-либо системы координат и, следовательно, не зависят от нее. Тем не менее, их можно выразить на языке некоторой системы координат.

Система координат в двумерном действительном пространстве может быть задана с помощью двух непараллельных векторов $\mathbf{f}_0$ и $\mathbf{f}_1$ (рис. 2.4). Эти векторы линейно-независимы, так как $\mathbf{f}_0 \neq \lambda \mathbf{f}_1$. Любой вектор $\mathbf{a}$ на плоскости может быть разложен по системе этих векторов, т. е. представлен в виде линейной комбинации

$$\mathbf{a} = \lambda_0 \mathbf{f}_0 + \lambda_1 \mathbf{f}_1,$$

поэтому три непараллельных вектора на плоскости будут уже линейно зависимы. Аналогичные рассуждения можно провести для трехмерного, четырехмерного и вообще многомерного пространства.

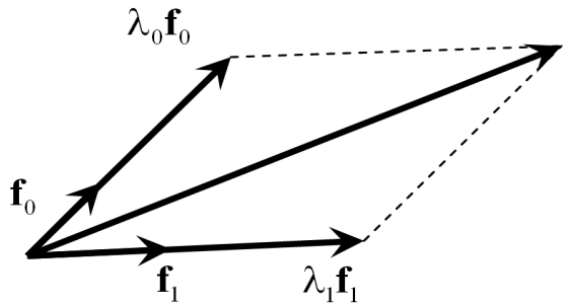


Рис. 2.4. Разложение вектора

Обобщив этот результат, получим, что в N-мерном комплексном пространстве система координат может быть задана с помощью N непараллельных векторов $\mathbf{f}_0$, $\mathbf{f}_1,\ldots,\mathbf{f}_{N-1}$, а любой вектор $\mathbf{a}$ может быть разложен в этой системе:

$$\mathbf{a} = \lambda_0 \mathbf{f}_0 + \lambda_1 \mathbf{f}_1 + \ldots + \lambda_{N-1} \mathbf{f}_{N-1} = \sum_{k=0}^{N-1} \lambda_k \mathbf{f}_k \,. \qquad (2.9)$$

Система координат $\{\mathbf{f}_k\}$, а также выражаемая ею система линейно независимых дискретных функций носят название базиса пространства.

Из (2.9) следует, что при заданной системе координат любой вектор полностью определяется с помощью N чисел $\{\lambda_k\}$, называемых координатами вектора. Совокупность координат будем называть спектром данного вектора. Очевидно, что если выбрать другой базис, то тот же вектор будет иметь другой спектр.

Особенно удобно производить разложение векторов, если система координат является ортогональной и нормированной к единице, т. е. если

$$(\mathbf{f}_k, \mathbf{f}_l) = \begin{cases} 0, k \neq l; \\ 1, k = l. \end{cases} \qquad (2.10)$$

Рассмотрим векторы $\mathbf{a}$ и $\mathbf{b}$ в двумерном пространстве с декартовым базисом $(\mathbf{f}_k, \mathbf{f}_l)$:

$$\mathbf{a} = \lambda_0 \mathbf{f}_0 + \lambda_1 \mathbf{f}_1;$$
$$\mathbf{b} = \mu_0 \mathbf{f}_0 + \mu_1 \mathbf{f}_1.$$

Скалярное произведение этих векторов можно представить в виде

$$(\mathbf{a},\mathbf{b}) = \lambda_0 \overline{\mu_0} (\mathbf{f}_0, \mathbf{f}_0) + \lambda_0 \overline{\mu_1} (\mathbf{f}_0, \mathbf{f}_1) +$$
$$+\lambda_1 \overline{\mu_0} (\mathbf{f}_1, \mathbf{f}_0) + \lambda_1 \overline{\mu_1} (\mathbf{f}_1, \mathbf{f}_1).$$

Используя свойства базиса, получаем

$$(\mathbf{a},\mathbf{b}) = \lambda_0 \overline{\mu_0} + \lambda_1 \overline{\mu_1}.$$

Аналогично в случае *N*-мерного пространства будем иметь

$$(\mathbf{a},\mathbf{b}) = \sum_{k=0}^{N-1} \lambda_k \overline{\mu_k}. \qquad (2.11)$$

Если перейти к новому базису, то координаты λ_k и μ_k изменятся, однако скалярное произведение останется тем же.

Теперь рассмотрим, каким образом определяется метрика векторного пространства, если в нем задано скалярное произведение. Длина вектора, или его норма, определяется как $\|\mathbf{a}\| = \sqrt{(\mathbf{a},\mathbf{a})}$. Скалярное произведение вектора на себя будет равно

$$(\mathbf{a},\mathbf{a}) = \sum_{k=0}^{N-1} |\lambda_k|^2. \qquad (2.12)$$

Это есть не что иное, как обобщенная теорема Пифагора. Отсюда длина вектора будет равна

$$\|\mathbf{a}\| = \sqrt{\sum_{k=0}^{N-1} |\lambda_k|^2}. \qquad (2.13)$$

Назовем расстоянием между двумя векторами длину разностного вектора. Тогда для нее будем иметь

$$\Delta = \|\mathbf{a} - \mathbf{b}\| = \sqrt{\sum_{k=0}^{N-1} |\lambda_k - \mu_k|^2}.$$

Эти выражения говорят о том, что метрика унитарного векторного пространства является среднеквадратической в том смысле, что приближение двух функций в нем понимается как минимизация среднеквадратической ошибки. Из последнего выражения следует:

$$\Delta^2 = \sum_{x=0}^{N-1} |\lambda_k - \mu_k| |\overline{\lambda_k} - \overline{\mu_k}| = \sum_{k=0}^{N-1} |\lambda_k|^2 + \sum_{k=0}^{N-1} |\mu_k|^2 - \sum_{x=0}^{N-1} |\lambda_k \overline{\mu_k} + \overline{\lambda_k} \mu_k|.$$

Здесь последняя сумма представляет собой удвоенную действительную часть скалярного произведения $(\mathbf{a},\mathbf{b})$. Учитывая это и (2.13), перейдем к векторным обозначениям и введем угол ψ следующим образом:

$$\Delta^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\cdot \mathrm{Re}(\mathbf{a},\mathbf{b}) = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\cdot\|\mathbf{a}\|\cdot\|\mathbf{b}\|\cdot\cos(\psi).$$

Это выражение напоминает теорему косинусов из элементарной тригонометрии (рис. 2.5). Из него видно, что скалярное произведение позволяет определить не только длину векторов и расстояние между векторами, но и угол между ними:

$$\cos(\psi) = \mathrm{Re}(\mathbf{a},\mathbf{b})/(\|\mathbf{a}\|\cdot\|\mathbf{b}\|).$$

Следовательно, ортогональные векторы, для которых $(\mathbf{a},\mathbf{b}) = 0$, взаимно перпендикулярны.

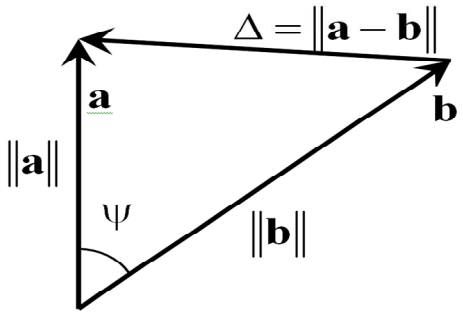


Рис. 2.5. Определение угла между векторами

Скалярное произведение позволяет также найти проекции вектора на ось координат. Для этого рассмотрим $(\mathbf{a},\mathbf{f}_k)$. Так как у базисного вектора $\mathbf{f}_k$ все координаты равны нулю, за исключением $\mu_k = 1$, то в соответствии с (2.11) получим

$$(\mathbf{a},\mathbf{f}_k) = \lambda_k . \qquad (2.14)$$

Для простоты было принято, что базис $\{\mathbf{f}_k\}$ является нормированным. Если это не выполняется, т. е. если

$$(\mathbf{f}_k,\mathbf{f}_l) = \begin{cases} 0, k \neq l \\ E, k = l, \end{cases}$$

то вместо выражений (2.11), (2.12) и (2.14) получим

$$(\mathbf{a},\mathbf{b}) = E\sum_{k=0}^{N-1} \lambda_k \overline{\mu_k}, \qquad (2.15)$$

$$(\mathbf{a},\mathbf{a}) = E\sum_{k=0}^{N-1} |\lambda_k|^2, \qquad (2.16)$$

$$\lambda_k = \frac{1}{E}(\mathbf{a},\mathbf{f}_k). \qquad (2.17)$$

3. ВВЕДЕНИЕ В ТЕОРИЮ ГРУПП, ПОЛЕЙ, КОЛЕЦ

Понятие «множество» эквивалентно понятию «совокупность». Само понятие четко не определяется, но может быть пояснено примерами. Можно говорить о множестве книг, о множестве художественной или технической литературы, о множестве точек кривой или вершин какой-либо фигуры. Таким образом, чтобы определить множество, достаточно указать общий признак, которым обладают все элементы этого множества и только они. Если данным свойством не обладает ни один из элементов множества, то говорят, что это свойство объединяет пустое множество (обозначают обычно как Ø). Принадлежность элемента *a* множеству А обозначают $a \in \text{A}$.

Определение понятия поля. Множество элементов называется полем, если заданы операции «сложения» и «умножения» (названия «сложение» и «умножение» условны, далее будем употреблять их без кавычек), удовлетворяющие следующим законам (табл. 3.1).

Таблица 3.1

Законы поля

Законы	Операция	
	сложение	умножение
Замкнутость: для каждой пары элементов $a,b \in M$ существует, и притом единственный, элемент $c \in M : c = a * b$	**A1.** $a + b = c$	**M1.** $a \cdot b = c$
Ассоциативность: $(a * b) * c = a * (b * c)$	**A2.** $(a + b) + c = a + (b + c)$	**M2.** $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
Коммутативность: $a * b = b * a$	**A3.** $a + b = b + a$	**M3.** $a \cdot b = b \cdot a$
Наличие единичного элемента: существует элемент $e \in M$, такой, что $a * e = e * a = a$, где $a \in M$	**A4.** $a + e = e + a = a$ $(e = 0)$	**M4.** $a \cdot e = e \cdot a = a$ $(e = 1)$
Наличие обратных элементов: для любого $a \in M$ существует элемент $\overline{a} \in M$ такой, что $a * \overline{a} = \overline{a} * a = e$	**A5.** $a + \overline{a} = \overline{a} + a = 0$ $(a = -\overline{a})$	**M5.** $a \cdot \overline{a} = \overline{a} \cdot a = 1$ $\left(\overline{a} = a^{-1}, a \neq 0\right)$
Дистрибутивность:	D1. $a \cdot (b + c) = a \cdot b + a \cdot c$ D2. $(b + c) \cdot a = b \cdot a + c \cdot a$	

Таким образом, в поле выполняются обычные законы замкнутости, ассоциативности, коммутативности и дистрибутивности.

Поле всегда содержит *единичный* элемент (е). При операции сложения единичный элемент называется *нулем* (0), а при операции умножения –

единицей (1). Сумма любого элемента поля и нуля, а также произведение любого элемента поля a и единицы равны a (см. табл. 3.1, А4, М4). Для всякого элемента a существует *обратный* (по сложению) элемент $a = -\overline{a}$, единственный элемент поля, удовлетворяющий уравнению $a + (-a) = 0$ (см. табл. 3.1, А5). Далее для всякого элемента a, не равного нулю, существует *обратный* (по умножению) элемент $\overline{a} = a^{-1}$ – единственный элемент поля, удовлетворяющий уравнению $a \cdot \overline{a} = 1$ (см. табл. 3.1, М5).

Поле, содержащее конечное число элементов q, называется *конечным полем* и обозначается $GF(q)$ (GF – означает Galois Field – поле Галуа, по фамилии французского математика Эвариста Галуа). Порядком конечного поля называется число элементов поля.

Определение понятия группы. Множество, в котором задана только основная операция сложения и выполняются законы А1 – А5 (см. табл. 3.1), называется *коммутативной аддитивной группой,* а множество, в котором задана только основная операция умножения и выполняются законы М1 – М5 (см. табл. 3.1) – *коммутативной мультипликативной группой.* Часто вместо термина «коммутативная группа» используется термин «абелева группа». Так как далее рассматриваются только коммутативные группы, то будем писать просто группа.

Группа, имеющая конечное число элементов, называется конечной группой. Конечная группа обозначается через $G = \{a, b, c\}$. *Порядком конечной группы* называется число элементов группы.

Все элементы любого конечного поля образуют аддитивную группу, поэтому порядок аддитивной группы поля совпадает с порядком поля. Мультипликативная группа поля включает все элементы поля, кроме нулевого, поэтому мультипликативная группа поля порядка q имеет порядок $q - 1$.

Рассмотрим некоторые примеры конечных полей. Для этого вспомним определение простого числа. Натуральное число P называется простым, если $P > 1$ и не имеет положительных делителей, отличных от 1 и P. Натуральное число N называется составным, если $N > 1$ и имеет по крайней мере один положительный делитель, отличный от 1 и N.

Простое поле. Элементарным примером конечного поля является *простое поле,* элементами которого являются целые числа по модулю простого числа p.

Сравнение целых чисел x, y по модулю m согласно (2.4) эквивалентно равенству

$$x - y = k \cdot m \qquad (3.1)$$

для некоторого целого k.

Все целые числа x, такие, что $x \equiv r \pmod m$ при фиксированном r, т. е. числа, которые дают при делении на m остаток r, образуют *класс чисел по модулю m*, который обозначают через $\{r\}$ или $r \pmod m$. Из определения следует, что всем числам класса отвечает один и тот же остаток r и можно получить все числа класса, если в форме $m \cdot k + r$ заставить k пробегать все целые числа. Соответственно m различным значениям r имеем m классов чисел по модулю m. Любое число класса $\{r\}$ называется *вычетом по модулю m* по отношению ко всем числам этого класса.

Из свойств сравнения известно, что если $x \equiv a \pmod m$ и $y \equiv b \pmod m$, то $x + y \equiv a + b \pmod m$ и $x \cdot y \equiv a \cdot b \pmod m$. Тем самым определены *сложение и умножение классов вычетов по модулю m*:

$$\{a\} + \{b\} = \{a + b\}, \tag{3.2}$$

$$\{a\} \cdot \{b\} = \{a \cdot b\}. \tag{3.3}$$

Легко проверить, что все законы поля, исключая М5, удовлетворяются для сложения и умножения классов вычетов, заданных посредством (3.2), (3.3) при произвольном m. Однако М5 выполняется только при условии, что $m = p$, где p – простое число.

Взяв от каждого класса по одному вычету, получим *полную систему вычетов*. Чаще всего в качестве полной системы вычетов употребляют наименьшие неотрицательные вычеты $0,1,\ldots, p-1$. Полная система вычетов по модулю простого числа p удовлетворяет всем законам поля.

Пример. Элементами поля $GF(5)$ являются числа 0, 1, 2, 3, 4. Аддитивная группа поля $GF(5)$ состоит из чисел 0, 1, 2, 3, 4; а мультипликативная группа – из чисел 1, 2, 3, 4. Правила сложения и умножения элементов поля $GF(5)$ определяются, соответственно, следующим образом (табл. 3.2).

Расширенное поле $GF\left(p^n\right)$. Аналогично сравнению целых чисел по модулю m можно определить сравнение для полиномов над полем $GF(p)$

Таблица 3.2

Сложение и умножение в поле *GF*(5)

+	0	1	2	3	4	·	0	1	2	3	4
0	0	1	2	3	4	0	0	0	0	0	0
1	1	2	3	4	0	1	0	1	2	3	4
2	2	3	4	0	1	2	0	2	4	1	3

Окончание табл. 3.2

3	3	4	0	1	2	3	0	3	1	4	2
4	4	0	1	2	3	4	0	4	3	2	1

Полином $A(x)=\sum_{i=0}^{n} a_i x_i$ называется полиномом над полем $GF(p)$, если его коэффициенты a_i принадлежат полю $GF(p)$. Степенью полинома $A(x)$ называется наибольшее число n такое, что $a_n \neq 0(\bmod m)$.

Сравнение полиномов $A(x)$, $B(x)$ по модулю полинома $F(x)$

$$A(x) \equiv B(x)(\bmod F(x)) \tag{3.4}$$

по определению эквивалентно равенству

$$A(x) - B(x) = K(x) \cdot F(x) \tag{3.5}$$

для некоторого полинома $K(x)$.

Все операции в сравнении выполняются по модулю p; этот факт обычно отмечают записью

$$A(x) \equiv B(x)(\operatorname{modd} F(x), p), \tag{3.6}$$

которая читается: полином $A(x)$ сравним с полиномом $B(x)$ по двойному модулю $\{F(x), p\}$.

Все полиномы $A(x)$ над полем $GF(p)$ такие, что $A(x) \equiv R(x)(\operatorname{modd} F(x), p)$ при фиксированном $R(x)$, т. е. полиномы, которые дают при делении на $F(x)$ остаток $R(x)$, образуют *класс полиномов по двойному модулю* $(F(x), p)$, который обозначают через $\{R(x)\}$ или $R(x)(\operatorname{modd} F(x), p)$.

Из определения следует, что всем полиномам класса отвечает один и тот же остаток $R(x)$ и можно получить все полиномы класса, если в форме $F(x)K(x)+R(x)$ заставить $K(x)$ пробегать все полиномы с коэффициентами из $GF(p)$.

Если $F(x)$ – полином степени n над полем $GF(p)$, то всевозможные остатки $R(x)$ – полиномы степени не выше $n-1$:

$$R(x) = a_{n-1} \cdot x^{n-1} + a_{n-2} \cdot x^{n-2} + \ldots + a_1 \cdot x + a_0, \tag{3.7}$$

где каждое из $a_{n-1}, a_{n-2}, \ldots, a_1, a_0$ может быть любым из p элементов поля $GF(p)$.

Из (3.7) следует, что существует точно p^n различных полиномов $R(x)$. Соответственно имеется p^n классов полиномов по двойному модулю $\left(F(x), p\right)$. Любой полином класса называется вычетом по двойному модулю $\left(F(x), p\right)$ по отношению ко всем полиномам этого класса.

В соответствии со свойством сравнений, если

$$A(x) \equiv R(x)\left(\operatorname{modd} F(x), p\right)$$

и

$$B(x) \equiv P(x)\left(\operatorname{modd} F(x), p\right),$$

то

$$A(x) + B(x) \equiv \left(R(x) + P(x)\right)\left(\operatorname{modd} F(x), p\right)$$

и

$$A(x) \cdot B(x) \equiv \left(R(x) \cdot P(x)\right)\left(\operatorname{modd} F(x), p\right).$$

Тем самым определены сложение и умножение классов вычетов по двойному модулю $\left(F(x), p\right)$:

$$\begin{aligned} \{R(x)\} + \{P(x)\} &= \{R(x) + P(x)\}; \\ \{R(x)\} \cdot \{P(x)\} &= \{R(x) \cdot P(x)\}. \end{aligned} \tag{3.8}$$

Легко проверить, что все законы поля, исключая М5, удовлетворяются для сложения и умножения классов вычетов, заданных посредством (3.8) при произвольном $F(x)$ над полем $GF(p)$. Однако, так же, как это имело место при сравнении целых чисел, М5 выполняется только в том случае, если $F(x) = f(x)$ – полином, неприводимый над $GF(p)$.

Полином $f(x)$ степени $n \geq 1$ с коэффициентами из поля $GF(p)$ неприводим над полем $GF(p)$, если он не может быть представлен в форме

$$f(x) = A(x) \cdot B(x),$$

где $A(x)$ и $B(x)$ – полиномы над $GF(p)$.

Взяв от каждого класса по одному вычету, получим полную систему вычетов. Обычно в качестве полной системы вычетов употребляют полиномы степени не выше $n - 1$, т. е. полиномы вида (3.7).

Полная система вычетов по двойному модулю

$$\left(f(x), p\right),$$

где $f(x)$ – полином, неприводимый над $GF(p)$, и p – простое число, удовлетворяет всем законам поля.

Таким образом, полная система вычетов по двойному модулю $\left(f(x), p\right)$ образует конечное поле, содержащее p^n элементов, которые обо-

значают через $GF(p^n)$ и называют *расширенным полем* или расширением степени n простого поля $GF(p)$.

Подчеркнем, что в отличие от простого поля элементами расширенного поля $GF(p^n)$ являются уже не числа, а полиномы степени не выше $n-1$ с коэффициентами из поля $GF(p)$.

Пример. Найдем элементы расширенного поля $GF(3^2)$. Полагая в (3.7) $n = 2$ и придавая коэффициентам a_i независимо значения элементов поля $GF(3)$ (элементы поля $GF(3)$ – это числа 0, 1, 2), получаем

$$GF\left(3^2\right)=\left\{0,1,2,x,2\cdot x,x+1,x+2,2\cdot x+1,2\cdot x+2\right\}.$$

Легко увидеть, что элементы поля $GF(3^2)$ – это все полиномы с коэффициентами из $GF(3)$ степени не больше $n-1=2-1=1$. Учитывая, что сложение элементов поля $GF(3^2)$ приводится по модулю $p = 3$, получаем табл. 3.3.

Таблица 3.3

Сложение в поле $GF(3^2)$

+	1	2	x	$2x$	$x+1$	$2x+1$	$x+2$	$2x+2$
1	2	0	$x+1$	$2x+1$	$x+2$	$2x+2$	x	$2x$
2	0	1	$x+2$	$2x+2$	x	$2x$	$x+1$	$2x+1$
x	$x+1$	$x+2$	$2x$	0	$2x+1$	1	$2x+2$	2
$2x$	$2x+1$	$2x+2$	0	x	1	$x+1$	2	$x+2$
$x+1$	$x+2$	x	$2x+1$	1	$2x+2$	2	$2x$	0
$2x+1$	$2x+2$	$2x$	1	$x+1$	2	$x+2$	0	x
$x+2$	x	$x+1$	$2x+2$	2	$2x$	0	$2x+1$	1
$2x+2$	$2x$	$2x+1$	2	$x+2$	0	x	1	$x+1$

Составление таблицы умножения элементов поля $GF(p^n)$ требует конкретизации неприводимого полинома $f(x)$, так как умножение элементов осуществляется по двойному модулю $\left(f\left(x\right),p\right)$.

Выберем неприводимый над полем $GF(3)$ полином степени $n=2$, например $f\left(x\right)=x^2-2$. В том, что полином $f\left(x\right)=x^2-2$ неприводим над полем $GF(3)$, легко убедиться, проверив, что он не делится без остатка на полиномы степени $n-1=2-1=1$ с коэффициентами из $GF(3)$, т. е. на полиномы x, $x-1$, $x-2$.

Найдем, для примера, произведение элементов $2x$ и $2x+1$. Выполняя умножение, получаем

$$\left(2x+1\right)2x=4x^2+2x\equiv x^2+2x\left(\mathrm{mod}\,3\right).$$

Разделив полином x^2+2x на полином x^2-2x и учитывая, что операция сложения проводится по модулю 3, найдем

$$\begin{array}{r|l} x^2+2x & x^2-2 \\ \underline{+\;\;2x^2-4} & \underline{\;\;2\;\;} \\ 2x-4=2x-1 & \end{array}.$$

Итак:

$$(2x+1)2x=4x^2+2x\equiv 2x-1\left(\mathrm{modd}\left(x^2-2,3\right)\right).$$

Аналогичным образом можно найти произведение всех возможных пар элементов поля $GF(3^2)$; результаты расчета приведены в табл. 3.4.

Таблица 3.4.

Умножение в поле $GF(3^2)$

×	1	2	x	$2x$	$x+1$	$2x+1$	$x+2$	$2x+2$
1	1	2	x	$2x$	$x+1$	$2x+1$	$x+2$	$2x+2$
2	2	1	$2x$	x	$2x+2$	$x+2$	$2x+1$	$x+1$
x	x	$2x$	2	1	$x+2$	$x+1$	$2x+2$	$2x+1$
$2x$	$2x$	x	1	2	$2x+1$	$2x+2$	$x+1$	$x+2$
$x+1$	$x+1$	$2x+2$	$x+2$	$2x+1$	$2x$	2	1	x
$2x+1$	$2x+1$	$x+2$	$x+1$	$2x+2$	2	x	$2x$	1
$x+2$	$x+2$	$2x+1$	$2x+2$	$x+1$	1	$2x$	x	2
$2x+2$	$2x+2$	$x+1$	$2x+1$	$x+2$	x	1	2	$2x$

Расширенное поле $GF\left(\left(p^n\right)^s\right)$. Поле порядка p^{ns} можно получить не только расширением степени ns поля $GF(p)$, но также расширением степени s поля $GF(p^n)$.

Рассуждения, аналогичные проведенным для поля $GF(p^n)$, показывают, что *полная система вычетов по тройному модулю*

$$\left(f_s(x), f_n(x), p\right),$$

где $f_s(x)$ – неприводимый над полем $GF(p^n)$ полином степени s и $f_n(x)$ – неприводимый над полем $GF(p)$ полином степени n, удовлетворяет всем законам поля. Поэтому полиномы степени не выше $n-1$

$$R(x)=a_{s-1}x^{s-1}+a_{s-2}x^{s-2}+\ldots+a_1x+a_0, \qquad (3.9)$$

где каждое из $a_{s-1}, a_{s-2}, \ldots, a_1, a_0$ может быть любым из p^n элементов поля $GF(p^n)$, образуют конечное поле $GF\left(\left(p^n\right)^s\right)$.

Из (3.9) следует, что существует точно p^{ns} различных полиномов $R(x)$. Поэтому поле $GF\left(\left(p^n\right)^s\right)$ содержит p^{ns} элементов. Поле $GF\left(\left(p^n\right)^s\right)$ называют *расширением степени s поля* $GF(p^n)$.

Подчеркнем, что в отличие от расширенного поля $GF(p^n)$, где элементами были полиномы с коэффициентами из $GF(p)$, элементами расширенного поля $GF\left(\left(p^n\right)^s\right)$ суть полиномы с коэффициентами из поля $GF(p^n)$.

Пример. Найдем элементы расширенного поля $GF\left(\left(p^n\right)^s\right)$ для $p = 2$, $n = 2$ и $s = 2$.

Полагая в (3.9) $s = 2$ и придавая коэффициентам a_i независимо значения элементов поля $GF(2^2)$, найдем элементы поля $GF\left(\left(2^2\right)^2\right)$.

Элементы поля $GF(2^2)$ – это 0, 1, a, $a + 1$; поэтому

$$GF\left(\left(2^2\right)^2\right)=\left\{\begin{array}{c} 0,1,a,a+1,x,x+1,x+a,x+a+1,ax,ax+1,ax+a, \\ ax+a+1,(a+1)x,(a+1)x+1,(a+1)x+a,(a+1)x+a+1 \end{array}\right\}.$$

Легко увидеть, что элементы поля $GF\left(\left(2^2\right)^2\right)$ – это все полиномы степени не больше $s-1=2-1=1$ с коэффициентами из $GF(2^2)$.

3.1. Мультипликативная структура полей Галуа

Период элемента поля $GF(q^s)$. Согласно свойству М1 (см. табл. 3.1), если поле $GF(q^s)$ содержит элемент a, то оно должно содержать также степени $a^2 = aa$, $a^3 = aa^2$. Естественно, что в конечном поле не все степени будут различными. Поэтому должен существовать такой наименьший положительный показатель степени ε, что $a^{\varepsilon} \equiv 1 \ \left(\mathrm{mod}\, f(x), p\right)$; ε в этом случае называется *периодом элемента a*. Если период элемента a равен ε, то элементы $a^0, a^1, ..., a^{\varepsilon-1}$ все различны. Так как порядок мультипликативной группы поля $GF(q^s)$ равен $q^s - 1$, то максимально возможный период элементов поля $GF(q^s)$ равен

$$\varepsilon_{\max} = q^s - 1. \qquad (3.10)$$

Подчеркнем, что в поле характеристики p не может быть элементов, период которых кратен p. Действительно, если допустить обратное, что $\varepsilon = kp$, то по определению периода элемента $a^{kp} \equiv 1$ и $a^{kp} - 1 \equiv 0$. Но в поле характеристики p всегда справедливо

$$a^p \pm b^p \equiv (a \pm b)^p . \tag{3.11}$$

Поэтому $0 \equiv a^{kp} - 1 \equiv \left(a^k - 1\right)^p$; отсюда следует, что период ε элемента a есть k, что противоречит сделанному допущению.

Теорема о периоде элемента поля $GF(q^s)$. Если известен период произвольного элемента a поля $GF(q^s)$, то можно установить период элемента a^k, где k – произвольное целое число, если воспользоваться теоремой о периоде элемента поля.

$$\left.\begin{array}{c} \text{Если период элемента } a \text{ равен } \varepsilon, \\ \text{то период элемента } a^k \text{ есть } \varepsilon/(\varepsilon, k), \\ \text{где} (\varepsilon, k) - \text{наибольший общий делитель чисел } \varepsilon \text{ и } k. \end{array}\right\} \tag{3.12}$$

В частности, из (3.12) следует, что порядок мультипликативной группы поля $GF(q^s)$ всегда кратен периоду любого его ненулевого элемента или, другими словами, что период ε любого ненулевого элемента поля $GF(q^s)$ всегда делит порядок мультипликативной группы поля $q^s - 1$; последнее записывается как $\varepsilon \mid q^s - 1$.

$$\left.\begin{array}{c} \text{Также из теоремы о периоде элемента следует,} \\ \text{что если } \varepsilon \text{ и } k \text{ взаимно простые числа,} \\ \text{то период элемента } a^k \text{ равен периоду элемента } a, \\ \text{т. е., если } (\varepsilon, k) = 1, \text{ и } a^{\varepsilon} \equiv 1, \text{ то } a^{\varepsilon_1} \equiv 1, \ \varepsilon_1 = k\varepsilon. \end{array}\right\} \tag{3.13}$$

Первообразный элемент поля $GF(q^s)$. Элемент a, имеющий максимально возможный период (3.10), называется *первообразным элементом поля* $GF(q^s)$. Если обозначить первообразный элемент буквой θ, то степени $\theta^0, \theta^1, ..., \theta^{q^s - 2}$ различны и пробегают все ненулевые элементы поля $GF(q^s)$. Поэтому первообразный элемент поля является образующим элементом мультипликативной группы поля G:

$$\left.\begin{array}{c} G=\left\{\theta^0,\theta^1,\theta^2,...,\theta^{q^s-2}\right\} \\ u \\ GF\left(q^s\right)=\left\{0,\theta^0,\theta^1,\theta^2,...,\theta^{q^s-2}\right\}. \end{array}\right\} \qquad (3.14)$$

Так как $\theta^{q^s-1}\equiv 1$, то $\theta^{q^s}\equiv\theta,\theta^{q^s+1}\equiv\theta^2,...,$ и мультипликативная группа поля $GF(q^s)$ *циклична.*

Если θ – первообразный элемент поля $GF(q^s)$, то все степени θ^k, где k и q^s-1 – взаимно простые числа, также являются первообразными элементами этого поля. Таких чисел k имеется $\varphi\left(q^s-1\right)$, где φ – функция Эйлера. Функция Эйлера $\varphi(k)$ определена для всех. Целых положительных k и представляет собой число чисел ряда 0, 1, ..., $k-1$ взаимно простых с k.

Следовательно, в поле $GF(q^s)$ имеется

$$\varphi\left(q^s-1\right) \qquad (3.15)$$

первообразных элементов.

Далее, если θ – первообразный элемент поля $GF(q^s)$, то мультипликативно обратный элемент θ^{-1} тоже является первообразным, так как $-1\equiv q^s-2\left(\mathrm{mod}\, q^s-1\right)$, a и q^s-1 всегда взаимно просты. Поэтому в поле $GF(q^s)$ имеется

$$\frac{1}{2}\varphi\left(q^s-1\right) \qquad (3.16)$$

первообразных элементов $\theta_1,\theta_2,...,\theta_{\frac{1}{2}\varphi\left(q^s-1\right)}$ и столько же мультипликативно обратных им.

Если элемент a поля $GF(q^s)$ имеет период ε, где $\varepsilon\left|q^s-1\right.$, то все степени a^k, где k – числа взаимно простые с ε, также имеют период ε. Поэтому число элементов поля $GF(q^s)$ имеющих период ε, равно

$$\varphi(\varepsilon). \qquad (3.17)$$

Так как порядок мультипликативной группы поля $GF(q^s)$ равен q^s-1, то очевидно что

$$\sum_{\varepsilon\left|q^s-1\right.}\varphi(\varepsilon)=q^s-1$$

Пример. Найти степени ненулевых элементов поля $GF(3^2)$. Запишем степени ненулевых элементов $a^k\ (k=1,2,...,8)$ в табл. 3.5.

Таблица 3.5

Степени ненулевых элементов поля $GF(3^2)$

a	a^2	a^3	a^4	a^5	a^6	a^7	a^8
1	1	1	1	1	1	1	1
2	1	2	1	2	2	2	1
x	2	$2x$	1	x	x	$2x$	1
$2x$	2	x	1	$2x$	$2x$	x	1
$x+1$	$2x$	$x+2$	1	$x+1$	$x+1$	$x+2$	1
$2x+1$	x	$x+1$	2	$2x+1$	$x+2$	$2x+2$	1
$x+2$	x	$2x+2$	2	$x+2$	$2x+1$	$x+1$	1
$2x+2$	$2x$	$x+2$	2	$2x+2$	$x+1$	$2x+1$	1

Из табл. 3.5 следует, что элемент 1 имеет период $\varepsilon=1$, элемент 2 – период $\varepsilon=2$; элементы x, $2x$ – период $\varepsilon=4$; элементы $x+1$, $2x+1$, $x+2$, $2x+2$ – период $\varepsilon_{\max}=8$.

Элементы с периодом $\varepsilon_{\max}=8$ являются первообразными элементами поля $GF(3^2)$. Степени первообразных элементов поля пробегают все ненулевые элементы поля $GF(3^2)$. Первообразные элементы $x+1$ и $x+2$, $2x+1$ и $2x+2$ являются соответственно мультипликативно обратными, так как

$$(x+1)(x+2)\equiv 1\left(\operatorname{modd} x^2-2,3\right)$$

и

$$(2x+1)(2x+2)\equiv 1\left(\operatorname{modd} x^2-2,3\right).$$

Подгруппа. По определению элемент a, имеющий период $\varepsilon_{\max}=q^s-1$, является образующим элементом мультипликативной группы для $GF(q^s)$.

Если элемент a имеет период ε, где $\varepsilon\,|\,q^s-1$, то он является образующим элементом мультипликативной группы порядка ε; при этом, если $\varepsilon<q^s-1$, то такая группа называется подгруппой мультипликативной группы поля $GF(q^s)$, так как состоит из части элементов группы. Подгруппу обозначают через $H=\{h_1,h_2,...\}$. Так как порядок мультипликативной группы поля $GF(q^s)$ всегда кратен периоду любого ненулевого элемента этого поля, а порядок подгруппы равен периоду образующего элемента, то сразу же следует, что порядок мультипликативной группы поля всегда кратен порядку любой его подгруппы.

Пример. Найдем все подгруппы мультипликативной группы поля $GF(3^2)$.

Элемент 2, имеющий период $\varepsilon = 2$, является образующим для подгруппы порядка 2:

$$H = \left\{2^0, 2^1\right\} \equiv \{1, 2\} (\mathrm{mod}\, 3).$$

Элементы *x*, 2*x*, имеющие период $\varepsilon = 4$, являются образующими для подгруппы порядка 4. Если выбрать, например, в качестве образующего элемент *x*, то получим

$$H = \left\{x^0, x^1, x^2, x^3\right\} \equiv \{1, x, 2, 2x\} \left(\mathrm{modd}\, x^2 - 2, 3\right).$$

Смежные классы. Пусть $G = \{a_1, a_2, ...\}$ – мультипликативная группа и $H = \{h_1, h_2, ...\}$ – ее подгруппа. Тогда множество элементов вида *ah*, где *h* принимает значения из *H* и *a* – фиксированный элемент *G*, называется смежным классом по *H* и обозначается *aH*.

Из определения следует, что смежный класс по *H* и подгруппа группы *G* содержат одинаковое число элементов. Кроме того, можно доказать, что если у двух смежных классов оказался общий элемент, то они совпадают.

Элемент a_i, принадлежащий смежному классу $\{a_i, H\}$, называется образующим элементом этого смежного класса (не путать с понятием образующего элемента мультипликативной группы). В качестве образующего может быть выбран любой элемент $a \in \{a_i H\}$, так как $\{aH\} = \{a_1 H\}$.

Группу *G* можно представить как объединение непересекающихся смежных классов, которые исчерпывают всю группу:

$$G = \{a_1 H\} + \{a_2 H\} + ... + \{a_k H\}. \qquad (3.18)$$

Здесь символ сложения следует рассматривать как символ объединения. Представление (3.18) называется разложением группы *G* на смежные классы.

Пример. Найти разложение мультипликативной группы поля $GF(3^2)$ на смежные классы. Выберем подгруппу $H = \{1, x, 2, 2x\}$ и образующий элемент $a_2 = x + 1$.

$$\{a_2 H\} = \{(x+1)1, (x+1)x, (x+1)2, (x+1)2x\} \equiv$$
$$\equiv \{x+1, x+2, 2x+2, 2x+1\} \left(\mathrm{modd}\, x^2 - 2, 3\right).$$

Таким образом,

$$G = \{1, x, 2, 2x\} + \{x+1, x+2, 2x+2, 2x+1\}.$$

Подполе. Если часть элементов поля порядка q^s характеристики *p* сама образует поле, то это поле называется подполем этого поля.

Из определения расширения степени n поля $GF(p)$ следует, что все подполя $GF(p^n)$ – это в точности поля $GF(p^k)$, где k делит n. Аналогично, все подполя $GF(q^s)$ $q = p^{n1}$ – это в точности поля $GF(p^k)$, где k делит $n_1 s$.

Отсюда непосредственно следует, что порядок поля всегда кратен порядку его подполя. Можно показать, что ненулевой элемент a поля $GF(q^s)$ принадлежит его подполю $GF(p^k)$, если $p^k - 1$ кратно ε и ε – период элемента a в мультипликативной группе поля $GF(q^s)$.

Элементы поля $GF(p^k)$ удовлетворяют сравнению

$$x^{pk} - x \equiv 0 \left(\operatorname{modd} f(x), p\right), \tag{3.19}$$

и обратно: если элемент поля $GF(q^s)$ удовлетворяет сравнению (3.19), то он является элементом его подполя $GF(p^k)$.

Всякое расширенное поле $GF(q^s)$ характеристики p имеет своим подполем простое поле $GF(p)$. Числа поля $GF(q^s)$ (но не элементы поля!) удовлетворяют сравнению

$$x^p - x \equiv 0 \left(\operatorname{mod} p\right). \tag{3.20}$$

Обратно, если элемент поля $GF(q^s)$ удовлетворяет сравнению (3.20), то он является элементом поля $GF(p)$. Сравнения (3.19) и (3.20) получаются как следствия из теоремы Ферма.

Пример. Найдем все подполя поля $GF(3^2)$. Для поля $GF(3^2)$ делителями показателя степени $n = 2$ являются числа 1, 2. Поэтому поле $GF(3^2)$ содержит единственное подполе $GF(3)$. Легко убедиться, что числа поля $GF(3^2)$ (и только они!) – 0, 1, 2 удовлетворяют сравнению $x^3 - x \equiv 0 \left(\operatorname{mod} 3\right)$ (см. (3.20)).

Характер мультипликативной группы. Если в двух множествах M, $\bar{M}$ определены некоторые соотношения между элементами и если элементу a из M поставлен в соответствие один и только один элемент $\bar{a}$ из $\bar{M}$ так, что

– каждый элемент $\bar{a}$ из $\bar{M}$ является образом, по меньшей мере, одного элемента из M;

– все соотношения между элементами из M выполняются и для соответствующих элементов из $\bar{M}$, то $\bar{M}$ называется гомоморфным образом множества M; в этом случае пишут $M \sim \bar{M}$ *или* $a \sim \bar{a}$.

В частности, если множество представляет собой мультипликативную группу поля $GF\left(p^n\right) - G$, то каждому ее элементу можно поставить в соответствие элемент $\psi(a)$, так чтобы выполнилось условие:

$$\begin{gathered} \text{если } a \sim \psi(a),\ b \sim \psi(b),\ c = ab, \\ \text{то } c \sim \psi(c) = \psi(ab) = \psi(a)\psi(b);\ a,b,c \in G. \end{gathered} \tag{3.21}$$

Функция ψ называется *характером* мультипликативной группы.

Из (3.21) следует, что характер мультипликативной группы определяется на ненулевых элементах поля $GF(p^n)$ и заранее исключается нулевое представление $\psi(a)=0$ для всех $a\in G$. Из (3.21) также непосредственно вытекает, что $\psi\left(a^2\right)=\left[\psi(a)\right]^2$ и следовательно,

$$\psi(1)=1. \tag{3.22}$$

Пусть порядок мультипликативной группы G есть p^n-1. Тогда, согласно (3.21)

$$\left[\psi(a)\right]^{p^n-1}=\overbrace{\psi(a)\psi(a)...\psi(a)}^{p^n-1}=\psi\left(a^{p^n-1}\right)=\psi(1)=1, \tag{3.23}$$

так как $a^{p^n-1}\equiv 1$.

Характер называется *нетривиальным*, если он не равен тождественно единице. Характер является *K-значным*, если функция ψ принимает ровно K различных значений.

K-значный характер мультипликативной группы G существует лишь в том случае, если K делит порядок мультипликативной группы, $K\,|\,p^n-1$. Это объясняется тем, что различные характеры образуют мультипликативную подгруппу порядка K, а порядок группы всегда кратен порядку подгруппы.

Так как существуют, вообще говоря, несколько различных K-значимых характеров, то определенный характер фиксируется следующей комплекснозначной функцией:

$$\psi(a)=\exp(j2\pi u/K), \tag{3.24}$$

где u определяется из условия $a\equiv\theta^u\left(\mathrm{modd}\, f(x), p\right)$;

θ – парообразный элемент поля $GF(p^n)$ (образующий элемент мультипликативной группы порядка p^n-1).

Чтобы распространить понятие характера на все элементы поля $GF(p^n)$, доопределим функцию ψ, полагая

$$\psi(0)=0. \tag{3.25}$$

Известно, что для любого нетривиального характера справедливо

$$\sum_{x\in GF\left(p^n\right)}\psi(ax+b)=0,\ a\neq 0(\mathrm{mod}\, p),\ \ a,d\in GF\left(p^n\right). \tag{3.26}$$

Из соотношения (3.26) непосредственно следует

$$\sum_{x\in GF\left(p^n\right)}\psi(ax+b)=\sum_{\substack{x\in GF\left(p^n\right)\\ x\neq 0(\mathrm{mod}\, p)}}\psi(ax+b)+\psi(b)=0, \tag{3.27}$$

откуда

$$\sum_{\substack{x\in GF\left(p^n\right)\\ x\neq 0(\mathrm{mod}\ p)}} \psi(ax+b) = -\psi(b). \tag{3.28}$$

В частном случае, когда $K=2$, характер называется *двузначным*, при этом функция ψ согласно (3.24) принимает значения ± 1:

$$\psi(a) = e^{j\pi u} = \begin{cases} 1, & \text{если } u \equiv 0(\mathrm{mod}\,2); \\ -1, & \text{если } u \equiv 1(\mathrm{mod}\,2). \end{cases} \tag{3.29}$$

Отметим некоторые свойства двузначного характера:

$$\psi(-1) = \begin{cases} 1, & \text{если } \dfrac{p^n-1}{2} \equiv 0(\mathrm{mod}\,2); \\ -1, & \text{если } \dfrac{p^n-1}{2} \equiv 1(\mathrm{mod}\,2), \end{cases} \tag{3.30}$$

$$\psi(a) = \psi\left(a^{-1}\right),\ a \neq 0(\mathrm{mod}\ p); \tag{3.31}$$

$$\psi\left(a^2\right) = 1. \tag{3.32}$$

Известно, что обе части сравнения и модуль можно умножить на одно и то же целое число. Поэтому следствием (3.30) является

$$\psi(-1) = \begin{cases} -1, & \text{если } p^n - 1 \equiv 2(\mathrm{mod}\,4); \\ 1, & \text{если } p^n - 1 \equiv 0(\mathrm{mod}\,4). \end{cases} \tag{3.33}$$

3.2. Алгебраическая структура полей Галуа

Представление полинома $x^{q^s} - x$ в виде произведения множителей $x - a_i$, где a_i – элементы поля $GF(q^s)$. Согласно (3.20), любой элемент поля $GF(q^s)$ удовлетворяет уравнению $x^{q^s} - x = 0$. Поэтому если $a_0, a_1, ..., a_{q^s-1}$ – все элементы поля $GF(q^s)$, то полином $x^{q^s} - x$ должен делиться без остатка на полином

$$f(x) = \prod_{i=0}^{q^s-1} (x - a_i).$$

Но степени и старшие коэффициенты этих полиномов равны; поэтому, как легко доказать,

$$x^{q^s} - x = \prod_{i=0}^{q^s-1} (x - a_i). \qquad (3.34)$$

Таким образом, элементами поля $GF\left(q^s\right)$ являются корни уравнения $x^{q^s} - x = 0$ или, другими словами, уравнение (3.20) имеет в поле $GF(q^s)$ ровно q^s различных корней.

Представление полинома $x^{q^s} - x$ в виде произведения неприводимых полиномов. Напомним, что полином $f_k(x)$ степени $k \geq 1$ неприводим над полем $GF(p)$, если не существует таких двух полиномов $A(x)$ и $B(x)$ с коэффициентами из $GF(p)$, каждый степени меньше k, что

$$f_k(x) = A(x)B(x).$$

Из определения следует, что полином $f_k(x)$, неприводимый над полем $GF(p)$, не может иметь в этом поле корней. Однако в расширенном поле $GF(q)$, $q = q^n$ полином $f_k(x)$ уже имеет корни. Точнее, полином $f_k(x)$ степени k, неприводимый над полем $GF(p)$, имеет в поле

$$GF(q),\ q = q^n,$$

где n кратно k, ровно k корней.

Пусть некоторый элемент поля $GF(q) - a$, имеющий период ε, является корнем полинома $f_k(x)$ и

$$q^k \equiv 1 (\mathrm{mod}\, \varepsilon),\ k \geq 1. \qquad (3.35)$$

Тогда единственными корнями полинома $f_k(x)$ являются элементы поля $GF(q)$

$$a_0, a_1, ..., a^{q^k-1}, \qquad (3.36)$$

т. е.

$$f_k(x) = \prod_{i=0}^{k-1} \left(x - a^{p^i}\right). \qquad (3.37)$$

Полином (3.37) называется *минимальным полиномом элемента a,* так как любой полином $A(x)$, для которого $A(a) = 0$, удовлетворяет также условиям $A\left(a^{p^i}\right) = 0$ и, следовательно, делится на $f_k(x)$.

Степень k полинома $f_k(x)$ называется степенью элемента a. Так как элементы $a^p, ..., a^{p^k-1}$ имеют тот же минимальный полином, что и элемент a, то они называются *p -сопряженными* с a.

Аналогично в расширенном поле $GF\left(q^s\right)$ полином $f_k(x)$ степени k, s кратно k, неприводимый над полем $GF(q)$, имеет в поле $GF\left(q^s\right)$ ровно k корней.

Если некоторый элемент поля $GF\left(q^s\right) - a$, имеющий период ε, является корнем полинома $f_k(x)$ и

$$q^{s^k} \equiv 1(\mathrm{mod}\,\varepsilon),\ k \geq 1, \tag{3.38}$$

то единственными корнями полинома $f_k\ (x)$ являются элементы поля $GF\left(q^s\right)$:

$$a, a^q, ..., a^{q^{k-1}}, \tag{3.39}$$

т. е.

$$f_k(x) = \prod_{i=0}^{k-1}\left(x - a^{q^i}\right). \tag{3.40}$$

Полином (3.40) называется *минимальным полиномом элемента a поля* $GF\left(q^s\right)$, *а степень* k *полинома* $f_k(x)$ *называется степенью элемента* a.

Так как элементы $a^q, ..., a^{q^{k-1}}$ имеют тот же минимальный полином, что и элемент a, то они называются q *– сопряженными* с a.

Заметим, что все q*-сопряженные* (p-сопряженные) элементы имеют один и тот же период ε. Последнее непосредственно вытекает из теоремы о периоде, если учесть, что числа q^i, $\left(p^i\right)$ в поле характеристики p всегда взаимно простые с делителями числа $p^s - 1,\ q - 1$

Пример. Найдем все минимальные полиномы ненулевых элементов полей $GF\left(2^2\right)$, $GF\left(3^2\right)$ и $GF\left[\left(2^2\right)^1\right]$.

Для нахождения минимальных полиномов элементов поля $GF\left(2^2\right)$ воспользуемся формулами (3.35) и (3.37), тогда получим табл. 3.6.

Таблица 3.6

Минимальные полиномы элементов поля

$a \in GF\left(2^2\right)$	ε	k	$f_k(x)$
1	1	1	$\prod_{i=0}^{0}(x-1)=x+1$
a	3	2	$a+1\prod_{i=0}^{1}\left(x-a^{2i}\right)=(x-a)\left(x-a^2\right)\equiv x^2+x+1$
	3	2	$\prod_{i=0}^{1}\left[x-(a+1)^{2^i}\right]=\left[x-(a+1)\right]\left[x-(a+1)^2\right]\equiv x^2+x+1$

Теперь перейдем к решению задачи представления полинома $x^{q^s}-x$ в виде произведения неприводимых полиномов. Действительно, из проведенного анализа следует, что элементы поля $GF(q^s)$ являются, с одной стороны, корнями полинома $x^{q^s}-x$, а с другой стороны, – корнями минимальных полиномов (3.37) или (3.40). Отсюда следует что $x^{q^s}-x$ делится на все минимальные полиномы элементов поля $GF(q)$. Более того, произведение всех различных минимальных полиномов элементов поля $GF(q^s)$ равно $x^{q^s}-x$. Тем самым решена задача представления полинома $x^{q^s}-x$ в виде произведения неприводимых полиномов.

Из предыдущего примера следует, что может существовать несколько различных неприводимых полиномов, корни которых имеют в заданном поле $GF(q^s)$ одинаковый период. Так, например, имеется два неприводимых полинома степени $k=2$, корни которых имеют в поле $GF(q^s)$ период ε, равный 8, четыре неприводимых полинома степени $k=2$, корни которых в поле $GF\left[(2^2)^2\right]$ имеют период $\varepsilon=15$ и т. д.

В общем случае число неприводимых полиномов степени k (k делит n при $s=1$ или k делит s при $s>1$), корни которых имеют в поле $GF(q^s)$ одинаковый период, равно

$$\varphi(\varepsilon)/k,\ \varepsilon\left|\left(q^s-1\right)\right.. \qquad (3.41)$$

Первообразные полиномы. Полином $f(x)$ степени s, неприводимый над полем $GF(q)$, называется *первообразным*, если его корень θ в поле $GF(q^s)$ имеет период $\varepsilon_{\max} = q^s - 1$, то есть если корень полинома является первообразным элементом поля $GF(q^s)$. Число первообразных полиномов степени s равно

$$\varphi(q^s - 1)/s. \qquad (3.42)$$

В частном случае, если $s = 1$ и $q = p^n$, первообразным называется такой неприводимый над полем $GF(p^n)$ полином степени n, корень которого в поле $GF(p^n)$ имеет период $\varepsilon = p^n - 1$. Число первообразных полиномов степени n равно

$$\varphi(q-1)/n, \quad q = p^n. \qquad (3.43)$$

Пример. Подсчитаем по формулам (3.42) и (3.43) число первообразных полиномов для $q^s < 10^4$.

Для заданных p и n число первообразных неприводимых над $GF(p)$ полиномов степени n указано в табл. 3.7. Для заданных q и s число первообразных неприводимых над полем $GF(q)$ полиномов степени s указано в табл. 3.8.

Таблица 3.7

Число первообразных неприводимых полиномов степени n над $GF(p)$

n	p				
	2	3	5	7	11
2	1	2	4	8	16
3	2	4	20	36	176
4	2	8	48	160	
5	6	22	280		

Таблица 3.8

Число первообразных неприводимых полиномов степени s над $GF(q)$

s	q				
	2^2	2^3	2^4	2^5	2^6
2	4	18	16	64	96
3	12	144	96	576	
4	32	432	640		
5	120				
6	288				

Взаимные полиномы. Если θ – первообразный элемент поля $GF(q^s)$, то, согласно (3.40), первообразный полином степени s:

$$f(x)=\prod_{i=0}^{s-1}\left(x-\theta^{q^i}\right). \tag{3.44}$$

Далее, если θ – первообразный элемент, то θ^{-1} тоже первообразный элемент того же поля, поэтому полином $f_1(x)=\prod_{i=0}^{s-1}\left(x-\theta^{-q^i}\right)$ – также первообразный полином.

Первообразные полиномы $f(x)$ и $f_1(x)$ называются взаимными. Очевидно, что каждому первообразному полиному соответствует взаимный, поэтому число невзаимных первообразных полиномов степени s равно

$$\varphi(q^s-1)/2s. \tag{3.45}$$

Сопряженные полиномы. Пусть θ – первообразный элемент поля $GF(q^s)$, $q=p^n, s>1$. Полиномы

$$f^k(x)=\prod_{i=0}^{s-1}\left(x-\left(\theta^{q^i}\right)^{p^k}\right),\ k=1,2,\ldots,n\text{-}1, \tag{3.46}$$

у которых корни являются сопряженными с корнями полинома (3.44), тоже будут первообразными. Полиномы (3.46) называются *p-сопряженными полиномами.*

Из определения следует, что для поля $GF(q^s)$, $q=p^n, s>1$, число p-сопряженных полиномов равно n. Таким образом, в поле $GF(q^s)$, $q=p^n, s>1$, число невзаимных и несопряженных первообразных полиномов степени s равно

$$\varphi(q^s-1)/2ns. \tag{3.47}$$

Автоморфизм. Взаимно однозначное отображение элементов поля на элементы того же поля, при котором удовлетворяются все законы поля, называется автоморфизмом поля.

Отображение $a\to a^q$ для всякого $a\in GF(q^s)$ есть автоморфизм поля $GF(q^s)$. Действительно, такое отображение однозначно обратимо, так как

для каждого элемента a в поле $GF(q^s)$ содержится элемент a^q, и обратно, для каждого элемента содержится элемент a (3.39). При этом

$$(a \pm b)^q \equiv a^q \pm b^q (\text{modd } f(x), p),$$

и, следовательно, различные элементы имеют и различные q-е степени.

Далее, если $a \to a^q$ и $b \to b^q$, то

$$(a \pm b) \to (a \pm b)^q = a^q \pm b^q, \tag{3.48}$$

$$(ab) \to (ab)^q = a^q b^q. \tag{3.49}$$

Так как отображение $a \to a^{q^i}$ переводит элементы поля $GF(q^s)$ в q-сопряженные элементы, а отображение $a \to a^{p^i}$ переводит элементы поля $GF(q^s)$ в p-сопряженные элементы, то автоморфизм – это такое взаимно-однозначное соответствие, при котором элементы поля переводятся в сопряженные элементы или, другими словами, такое взаимно однозначное соответствие, при котором корни данного неприводимого полинома переводятся в другие корни того же полинома.

Пример. Установим взаимно однозначное соответствие между элементами полей $GF(3^2)$, построенными по различным корням первообразного полинома $f(x) = x^2 + x + 2$.

Так как полином $x^2 + x + 2$ является первообразным, то его корни имеют период $q - 1 = 3^2 - 1 = 8$ и, следовательно, являются первообразными элементами поля $GF(3^2)$. Обозначим через θ_1 корень полинома $f(x)$, то есть $f(\theta_1) = 0$. Тогда, согласно (3.39), вторым (сопряженным) корнем является θ_1^3. Тогда

$$GF_1(3^2) = \left\{0, \theta_1^0, \theta_1^1, \ldots, \theta_1^7\right\} \equiv$$

$$\equiv \left\{0,1,x,2x+1,2x+2,2,2x,x+2,x+1\right\}\left(\text{modd } x^2 + x + 2, 3\right);$$

$$GF_2(3^2) = \left\{0, \left(\theta_1^3\right)^0, \left(\theta_1^3\right)^1, \ldots, \left(\theta_1^3\right)^7\right\} \equiv$$

$$\equiv \left\{0,1,2x+2,x+2,x,2,x+1,2x+1,2x\right\}\left(\text{modd } x^2 + x + 2, 3\right);$$

$GF_1(3^2)$ для θ_1, $GF_2(3^2)$ для θ_1^3.

Таким образом, между элементами полей $GF_1(3^2)$ и $GF_2(3^2)$ имеется следующее взаимно однозначное соответствие, определяемое автоморфизмом поля:

$$0 \to 0, 1 \to 1, x \to 2x+2, 2x+1 \to x+2,$$
$$2x+2 \to x, 2 \to 2, 2x \to x+1, x+1 \to 2x.$$

Изоморфизм. Взаимно однозначное отображение элементов полей заданного порядка, построенных по различным неприводимым полиномам, при котором удовлетворяются все законы поля, называются изоморфизмом.

Поле, изоморфное полю $GF(q^s)$, будем обозначать через $\widetilde{GF}(q^s)$ и писать $GF(q^s) \approx \widetilde{GF}(q^s)$. Элементы поля $GF(q^s)$ будем по-прежнему обозначать через $a, b, c, \ldots$, а элементы поля $\widetilde{GF}(q^s)$ – через $\tilde{a}, \tilde{b}, \tilde{c}, \ldots$.

Если преобразование $x \to z$ переводит неприводимый над полем $GF(q)$ полином степени s – $f(x)$ в неприводимый над полем $GF(q)$ полином степени s – $\tilde{f}(x)$, т. е.

$$f(z) \equiv 0 (\text{modd } \tilde{f}(x), p), \tag{3.50}$$

то отображение

$$a(\text{modd } f(x), p) \approx \tilde{a}(\text{modd } \tilde{f}(x), p), \tag{3.51}$$

где

$$a = a_{s-1}x^{s-1} + a_{s-2}x^{s-2} + \ldots + a_1 x + a_0,$$
$$\tilde{a} = a_{s-1}z^{s-1} + a_{s-2}z^{s-2} + \ldots + a_1 z + a_0,$$
$$a_i \in GF(q)$$

есть изоморфизм поля $GF(q^s)$.

Действительно, такое отображение однозначно обратимо, так как для каждого элемента $a \in GF(q^s)$ существует элемент $\tilde{a} \in \widetilde{GF}(q^s)$ и обратно.

Далее, если $a, b \in GF(q^s)$, $\tilde{a}, \tilde{b} \in \widetilde{GF}(q^s)$ и $a \approx \tilde{a}$, $b \approx \tilde{b}$, то

$$a + b \approx \tilde{a} + \tilde{b},\ ab \approx \tilde{a}\tilde{b}. \tag{3.52}$$

Пример. Установим взаимно однозначное соответствие между элементами полей $GF(3^2)$, построенными по различным первообразным полиномам $f(x) = x^2 + x + 2$ и $\tilde{f}(x) = x^2 + 2x + 2$.

Легко установить, что преобразование $x = 2z$ $(z = 2x)$ переводит полином $f(x)$ в полином $\widetilde{f}(x)$. Действительно,

$$f(z) = z^2 + z + 2 = 4x^2 + 2x + 2 \ (\mathrm{modd}\, x^2 + 2x + 2).$$

Далее обозначим через x_1 корень полинома $f(x)$, тогда элементы поля $GF(3^2)$ суть 0 и степени x_1:

$$\begin{gathered} GF(3^2) = \left\{0, x_1^0, x_1^1, \ldots, x_1^7\right\} \equiv \\ \equiv \left\{0,1, x_1, 2x_1 + 1, 2x_1 + 2, 2, 2x_1, x_1 + 2, x_1 + 1\right\}\left(\mathrm{modd}\, {x_1}^2 + x + 2, 3\right). \end{gathered} \tag{3.53}$$

Корень полинома $\widetilde{f}(x)$ обозначим через x_2, тогда элементы поля $\widetilde{GF}(3^2)$ суть 0 и степени x_2:

$$\widetilde{GF}(3^2) = \left\{0, x_2^0, x_2^1, \ldots, x_2^7\right\} \equiv$$

$$\equiv \left\{0,1, x_2, x_2 + 1, 2x_2 + 1, 2, 2x_2, 2x_2 + 2, x_2 + 2\right\}\left(\mathrm{modd}\, {x_2}^2 + 2x_2 + 2, 3\right).$$

Согласно (3.51), находим

$$1 \approx 1, 2 \approx 2, x_1 \approx 2x_2, 2x_1 \approx x_2, x_1 + 1 \approx 2x_2 + 1,$$
$$2x_1 + 1 \approx x_2 + 1, 2x_1 + 2 \approx x_2 + 2, x_1 + 2 \approx 2x_2 + 2.$$

3.3. Разностные множества

Множество $\mathrm{D}(N, K, \lambda)$, состоящее из K вычетов $\{a_1, a_2, \ldots, a_k\}$ по модулю N, называется *разностным множеством* (РМ), сбалансированным на один уровень, если для каждого $d \neq 0 \bmod N$ существует точно λ упорядоченных пар $a_i, a_j \in \mathrm{D}$ таких, что $a_i - a_j \equiv d \bmod N$.

Пример. Множество $\mathrm{D}(7,3,1) = \{1,2,4\}$ является РМ, сбалансированным на один уровень $\lambda = 1$, так как для каждого $d \neq 0$ существует точно одна пара $a_i - a_j \equiv d \bmod N$:

$$\begin{array}{lll} 1 = a_2 - a_1 = 2 - 1; & 2 = a_3 - a_2 = 4 - 2; & 3 = a_3 - a_1 = 4 - 1; \\ 4 = a_1 - a_3 = 1 - 4; & 5 = a_2 - a_3 = 2 - 4; & 6 = a_1 - a_2 = 1 - 2; \end{array} \bmod 7.$$

Разностные множества $\mathrm{D}(N, K, 1)$ с параметром $\lambda = 1$ называются совершенными разностными множествами.

Известны разностные множества со следующими параметрами (ниже приняты следующие обозначения: P, P_1, P_2 – простые числа, t, u и m – натуральные числа):

1.) Зингера:

$$N = \frac{q^n - 1}{q - 1};\ K = \frac{q^{n-1} - 1}{q - 1};\ \lambda = \frac{q^{n-2} - 1}{q - 1}; q = P^m; \quad (3.54)$$

2.) квадратичных вычетов:

$$N = P^m = 4t - 1;\ K = 2t - 1;\ \lambda = t - 1; \quad (3.55)$$

3.) Холла:

$$N = P = 4u^2 + 27 = 4t - 1;\ K = 2t - 1;\ \lambda = t - 1; \quad (3.56)$$

4.) Якоби:

$$N = P_1 \cdot P_2 = 4t - 1;\ K = \frac{N - 1}{2} = 2t - 1;\ \lambda = \frac{N - 3}{4} = t - 1;\ P_2 = P_1 + 2; \quad (3.57)$$

5.) биквадратичных вычетов:

$$N = P = 4(2u + 1)^2 + 1;\ K = (2u + 1)^2;\ \lambda = u(u + 1); \quad (3.58)$$

6.) биквадратичных вычетов и нуля:

$$N = P = 4(2u + 1)^2 + 9;\ K = (2u + 1)^2 + 3;\ \lambda = u(u + 1) + 1; \quad (3.59)$$

7.) восьмеричных вычетов:

$$N = P = 8u^2 + 1 = 64t^2 + 9;\ K = u^2;\ \lambda = t^2; \quad (3.60)$$

8.) восьмеричных вычетов и нуля:

$$N = P = 8(2u + 1)^2 + 49 = 128t^2 + 9;\ K = (2u^2 + 1)^2;\ \lambda = 4t^2. \quad (3.61)$$

Заметим, что это не полный перечень разностных множеств.

Разностные множества применятся в том числе и для формирования дискретно-кодовых последовательностей, определяющие закон изменения манипулируемого параметра (амплитуда, фаза, частота) в дискретно-кодовых сигналах (ДКС). При этом для разностного множества со значением параметра $\lambda = 1$ можно сопоставить закон формирования отдельных символов. В зависимости от значения разностное множество формирует дискретно-кодовую последовательность.

4. СПЕКТРАЛЬНЫЙ АНАЛИЗ В БАЗИСЕ ВКФ

4.1. Дискретные преобразования Фурье

Дискретные функции, заданные на интервале N, могут рассматриваться как векторы в N-мерном линейном евклидовом пространстве. Простейшая система базисных векторов в таком пространстве может быть задана единичной матрицей:

$$\begin{array}{c} x\rightarrow \\ k\downarrow \end{array} \quad \begin{array}{ccccc} 0 & 1 & 2 & \dots & N-1 \end{array}$$

$$\{u(k,x)\}\rightarrow \begin{array}{c} 0 \\ 1 \\ 2 \\ \dots \\ N-1 \end{array} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = 1$$

Здесь каждая строка выражает единичный импульс, смещенный на k позиций:

$$u(k,x)=\begin{cases} 0, k\neq x \\ 1, k=x \end{cases},$$

и любые две строки ортогональны:

$$\sum_{x=0}^{N-1} u(k,x)\cdot u(l,x)=\begin{cases} 0, k\neq l \\ 1, k=l \end{cases}.$$

При этом энергия любой базисной функции $E_u=1$.

Дискретный сигнал $s(x)$ можно в соответствии с векторным выражением (2.9) рассматривать как результат разложения по системе базисных функций $\{u(k,x)\}$:

$$s(x)=\sum_{k=0}^{N-1} S(k)\cdot u(k,x) \qquad (4.1)$$

или в матричной форме

$$\mathbf{s}_x=\mathbf{1}\mathbf{s}_k,$$

где $\mathbf{s}_x$, $\mathbf{s}_k$ – матрицы-столбцы сигнала и его спектра.

Выражение (4.1) означает, что отсчеты сигнала $s(0),s(1),\dots,s(N-1)$ являются его координатами в базисной системе $\{u(k,x)\}$. Оно выражает

фильтрующее свойство единичного импульса и показывает, что этот импульс играет в дискретном анализе роль, аналогичную дельта-функции в непрерывном анализе.

Теперь рассмотрим тот же сигнал $s(x)$ (в общем случае комплексный) в произвольном ортогональном и полном базисе.

Разложение сигнала в соответствии с (2.9) представляется рядом

$$s(x)=\sum_{k=0}^{N-1}S(k)\cdot f(k,x). \qquad (4.2)$$

Обратная формула для вычисления спектра $S(k)$ в соответствии с выражениями (2.17) запишется следующим образом:

$$S(k)=\frac{1}{E_f}\big(s(x),f(k,x)\big),$$

где E_f – энергия базисной функции $f(k,x)$.

Фигурирующее здесь скалярное произведение можно вычислить, воспользовавшись тем, что оно не меняется при изменении системы координат. Поэтому, рассматривая это скалярное произведение в базисе $\{u(k,x)\}$, где координатами функций $s(x)$ и $f(k,x)$ являются их отсчеты при $x=0,1,\ldots,N-1$, и применяя векторную формулу (2.11), получаем

$$S(k)=\frac{1}{E_f}\sum_{k=0}^{N-1}s(x)\cdot\overline{f}(k,x). \qquad (4.3)$$

Совокупность формул (4.2) и (4.3) является парой дискретных преобразований Фурье, причем (4.2) – это формула обращения (обратное преобразование), а (4.3) – формула разложения (прямое преобразование).

Базисную систему функций можно представить в виде матрицы:

$$\{f(k,x)\}\rightarrow$$

$$\begin{array}{c} x\rightarrow \\ k\downarrow \end{array}\quad \begin{array}{ccccc} 0 & 1 & 2 & \ldots & N-1 \end{array}$$

$$\begin{array}{c} 0 \\ \rightarrow\ 1 \\ 2 \\ \ldots \\ N-1 \end{array}\quad \begin{bmatrix} f(0,0) & f(0,1) & f(0,2) & \ldots & f(0,N-1) \\ f(1,0) & f(1,1) & f(1,2) & \ldots & f(1,N-1) \\ f(2,0) & f(2,1) & f(2,2) & \ldots & f(2,N-1) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ f(N-1,0) & f(N-1,1) & f(N-1,2) & \ldots & f(N-1,N-1) \end{bmatrix}=\mathbf{F}_{kx}$$

Тогда преобразования Фурье можно записать и в матричной форме:

$$\mathbf{s}_x = \mathbf{F}_{xk}\mathbf{s}_k = \mathbf{F}_{kx}^T\mathbf{s}_k,$$
$$\mathbf{s}_k = \frac{1}{E_f}\overline{\mathbf{F}}_{xk}\mathbf{s}_x \qquad (4.4)$$

Умножим обе части равенства слева на матрицу $\mathbf{F}_{xk}$:

$$\mathbf{F}_{xk}\mathbf{s}_k = \frac{1}{E_f}\cdot\mathbf{F}_{xk}\cdot\overline{\mathbf{F}}_{xk}\cdot\mathbf{s}_k.$$

Сравнив это выражение с первым из выражений (4.4), получим

$$\frac{1}{E_f}\cdot\mathbf{F}_{xk}\cdot\overline{\mathbf{F}}_{xk}^T = 1.$$

Последнее выражение совпадает с определением унитарной матрицы. Таким образом, представление дискретного сигнала в виде линейной комбинации (4.2) или в матричной форме в виде (4.4) возможно всегда, если базисная система функций $\{f(k,x)\}$ является полной и линейно-независимой, а выражение для спектра (4.3) и (4.4) возможно только в том случае, когда система $\{f(k,x)\}$ еще и ортогональна (матрица $\mathbf{F}_{xk}$ – унитарна). Заметим, что если базисная матрица является симметрической, т. е. $\mathbf{F}_{kx} = \mathbf{F}_{xk}$, то прямое и обратное преобразования Фурье вычисляются умножением на сопряженные матрицы:

$$\mathbf{s}_x = \mathbf{F}_{xk}\mathbf{s}_k, \quad \mathbf{s}_k = \frac{1}{E_f}\overline{\mathbf{F}}_{kx}\mathbf{s}_x. \qquad (4.5)$$

Перейдем к энергетическим соотношениям при дискретном преобразовании Фурье. Векторной формуле (2.16) обобщенная теорема Пифагора соответствует следующее выражение:

$$\big(s(x), s(x)\big) = E_f\sum_{k=0}^{N-1}\big|S(k)\big|^2. \qquad (4.6)$$

Скалярное произведение в левой части равенства вычислим по формуле (2.16) в базисе функций $\{u(k,x)\}$, для которых $E_u = 1$. Тогда получим

$$\sum_{x=0}^{N-1}\big|s(x)\big|^2 = E_f\sum_{k=0}^{N-1}\big|S(k)\big|^2. \qquad (4.7)$$

Выражение является равенством Парсеваля для дискретных сигналов. Оно показывает, что энергия сигнала определяется самим сигналом и не зависит от выбора базиса, в котором он описывается. Энергию спектральных составляющих

$$e(k) = E_f \left| S(k) \right|^2$$

будем называть энергетическим спектром сигнала (в теории континуальных сигналов энергетический спектр имеет другой смысл, а именно спектральной плотности энергии – энергии, приходящейся на полосу в 1Гц).

Пусть при $n \le N-1$

$$s_1(x) = \sum_{k=0}^{n} S(k) \cdot f(k,x),$$

причем $S(k)$ вычисляется по формуле (4.3). Метрика евклидова пространства такова, что приближение $s(x) \approx s_1(x)$ имеет минимальную среднеквадратическую ошибку по сравнению с любым другим выбором коэффициентов *S(k)*):

$$\Delta = \sqrt{\frac{1}{E_f} \sum_{x=0}^{N-1} \left| s(x) - s_1(x) \right|^2} \quad .$$

При $n = N-1$ эта ошибка становится равной нулю.

Хотя переменная x может иметь любой физический смысл (время, пространственная координата и т. д.), будем в дальнейшем условно называть область изменения этой переменной временной областью в отличие от области изменения переменной k, которую будем называть спектральной областью.

4.2. Преобразование Фурье в базисе ВКФ

Любая из рассмотренных ранее систем ВКФ представляет собой ортогональный и полный базис, причем энергия базисной функции есть $E_f = N$. Формулы преобразования Фурье, и основные свойства этого преобразования в любом симметрическом базисе ВКФ будут одни и те же. Поэтому ниже для определенности рассмотрение ведется в базисе ВКФ–Пэли, а в тех случаях, когда свойства в разных базисах различны, делаются

необходимые оговорки. Пара дискретных преобразований Фурье в базисе ВКФ–Пэли имеет вид

$$s(x)=\sum_{p=0}^{N-1} S(p)\cdot \mathrm{Pal}(p,x);\; S(p)=\sum_{x=0}^{N-1} s(x)\cdot \overline{\mathrm{Pal}}(p,x). \qquad (4.8)$$

В общем случае сигнал $s(x)$ и его спектр $S(p)$ – это комплексные функции. Они могут быть представлены в виде суммы действительной и мнимой частей или в показательной форме:

$$s(x)=\mathrm{Re}\left[s(x)\right]+ j\,\mathrm{Im}\left[s(x)\right]=\left|s(x)\right|\cdot e^{j\psi(x)};$$
$$S(p)=\mathrm{Re}\left[S(p)\right]- j\,\mathrm{Im}\left[S(p)\right]=\left|S(p)\right|\cdot e^{-j\varphi(p)}.$$

Здесь $\left|s(x)\right|$ и $\left|S(p)\right|$ – огибающие, а $\psi(x)$ и $\varphi(p)$ – фазы сигнала и его спектра. Огибающую $\left|S(p)\right|$, кроме того, называют амплитудным спектром, а фазу $\varphi(p)$ – фазовым спектром. Очевидно, что

$$\left|S(p)\right|=\sqrt{\left\{\mathrm{Re}\left[S(p)\right]\right\}^2+\left\{\mathrm{Im}\left[S(p)\right]\right\}^2};$$
$$\varphi(p)=\mathrm{arctg}\left(\frac{\mathrm{Im}\left[S(p)\right]}{\mathrm{Re}\left[S(p)\right]}\right).$$

Дискретные преобразования Фурье можно записать в матричной форме:

$$\mathbf{s}_x=\mathbf{P}\cdot\mathbf{s}_p;\;\; \mathbf{s}_p=\frac{1}{N}\overline{\mathbf{P}}\cdot\mathbf{s}_x. \qquad (4.9)$$

Энергетические соотношения в спектре сигнала в базисе ВКФ определяются равенством Парсеваля , которое в данном случае принимает вид

$$\frac{1}{N}\sum_{x=0}^{N-1}\left|s(x)\right|^2=\sum_{p=0}^{N-1}\left|S(p)\right|^2. \qquad (4.10)$$

Таким образом, функции

$$\rho(p)=\left|S(p)\right|^2=S(p)\overline{S}(p),$$
$$e(p)=N\rho(p)=N\left|S(p)\right|^2 \qquad (4.11)$$

являются соответственно спектром мощности и энергетическим спектром дискретного сигнала в базисе ВКФ. Базисная функция в общем случае – комплексная, т. е.

$$\mathrm{Pal}(p,x)=\mathrm{Re}\left[\mathrm{Pal}(p,x)\right]+ j\,\mathrm{Im}\left[\mathrm{Pal}(p,x)\right].$$

Действительная и мнимая составляющие этой функции равны

$$\begin{aligned}\mathrm{Re}\big[\mathrm{Pal}(p,x)\big] &= \frac{1}{2}\Big[\,\mathrm{Pal}(p,x) + \mathrm{Pal}(p^*,x)\,\Big],\\ \mathrm{Im}\big[\mathrm{Pal}(p,x)\big] &= \frac{1}{2j}\Big[\,\mathrm{Pal}(p,x) - \mathrm{Pal}(p^*,x)\,\Big],\end{aligned} \tag{4.12}$$

где p и p^* – m-ично противоположные числа, соответствующие номерам комплексно-сопряженных функций.

Выражения (4.12), по существу, являются обобщенными формулами Эйлера. При $m = N$ они превращаются в обычные формулы для экспоненциальных функций.

Характерной особенностью ВКФ является присущее им специфическое понятие сдвига аргумента как m-сдвига. В соответствии с этим было сформулировано свойство мультипликативности базисных функций:

$$\mathrm{Pal}(a,x)\,\mathrm{Pal}(b,x) = \mathrm{Pal}(c,x),\ c = \Big(a \underset{m}{\oplus} b\Big). \tag{4.13}$$

Для сигналов, рассматриваемых в базисе ВКФ, сдвиг как во временной, так и в спектральной области также должен пониматься в смысле m-сдвига. Обычные представления о четности сигнала в базисе ВКФ заменяются новыми представлениями о m-четности. В соответствии с этими представлениями любой дискретный сигнал можно разделить на m-четную и m-нечетную части:

$$\begin{aligned}s_{ч}(x) &= \frac{1}{2}\Big[\,s(x) + s(x^*)\,\Big] = s_{ч}(x^*),\\ s_{н}(x) &= \frac{1}{2}\Big[\,s(x) - s(x^*)\,\Big] = -s_{н}(x^*).\end{aligned} \tag{4.14}$$

При $m = N$, когда базис ВКФ переходит в базис ДЭФ, формулы (4.14) дают четную и нечетную части сигнала на интервале N в обычном смысле слова. В другом крайнем случае, при $m = 2$, любой сигнал является четным по модулю 2. Четная и нечетная по модулю m части действительного сигнала ортогональны. Среднее значение нечетной части на интервале $[0, N-1]$ равно нулю.

Действительная составляющая ВКФ – это m-четная функция, а мнимая составляющая – m-нечетная функция. В самом деле, учитывая свойство ВКФ

$$\mathrm{Pal}(p,x) = \overline{\mathrm{Pal}}(p^*,x) = \overline{\mathrm{Pal}}(p,x^*),$$

получаем из (4.12)

$$\mathrm{Re}\left[\mathrm{Pal}(p,x)\right]=\mathrm{Re}\left[\mathrm{Pal}(p^*,x)\right];\ \mathrm{Im}\left[\mathrm{Pal}(p,x)\right]=-\mathrm{Im}\left[\mathrm{Pal}(p^*,x)\right]. \quad (4.15)$$

Рассмотрим особенности спектров некоторых типов сигналов. Как уже было отмечено, система ВКФ содержит действительные и попарно комплексно-сопряженные функции. Следовательно, спектральные компоненты действительного сигнала в базисе ВКФ также будут либо действительными, либо попарно комплексно-сопряженными, т. е.

$$S(p)=\overline{S}(p^*).$$

Спектр комплексного сигнала таким свойством не обладает.

Пусть $s(x)$ – действительная m-четная функция. Перепишем второе выражение в виде

$$S(p)=\frac{1}{N}\sum_{x=0}^{N-1}s(x)\cdot\mathrm{Re}\left[\mathrm{Pal}(p,x)\right]-j\frac{1}{N}\sum_{x=0}^{N-1}s(x)\cdot\mathrm{Im}\left[\mathrm{Pal}(p,x)\right]. \quad (4.16)$$

Здесь учтено, что

$$\mathrm{Re}\left[\overline{\mathrm{Pal}}(p,x)\right]=\mathrm{Re}\left[\mathrm{Pal}(p,x)\right],$$

а

$$\mathrm{Im}\left[\overline{\mathrm{Pal}}(p,x)\right]=-\mathrm{Im}\left[\mathrm{Pal}(p,x)\right].$$

Вторая сумма в (4.16) равна нулю, так как она представляет собой среднее значение m-нечетной функции. Таким образом, спектр действительной m-четной функции $s(x)$ есть также действительная m-четная функция. Если же $s(x)$ – действительная m-нечетная функция, то первая сумма в выражении (4.16) равна нулю, и спектр $S(p)$ будет мнимой m-нечетной функцией.

Наконец, если сигнал $s(x)$ – произвольная действительная функция, то его спектр $S(p)$ – комплексная функция

$$S(p)=S_ч(p)-jS_н(p),$$

где

$$S_ч(p)=\frac{1}{N}\sum_{x=0}^{N-1}s(x)\cdot\mathrm{Re}\left[\mathrm{Pal}(p,x)\right],$$

$$S_н(p)=\frac{1}{N}\sum_{x=0}^{N-1}s(x)\cdot\mathrm{Im}\left[\mathrm{Pal}(p,x)\right],$$

причем действительная часть $S_ч(p)$ – это m-четная, а мнимая часть $S_н(p)$ – m-нечетная функция. Следовательно, у такого сигнала амплитудный спектр есть m-четная функция, а фазовый спектр – m-нечетная функция:

$$\left|S(p)\right|=\left|S(p^*)\right|,\ \ \varphi(p)=-\varphi(p^*).$$

В общем случае рассмотрим комплексный сигнал

$$s(x) = \mathrm{Re}\big[s(x)\big] + j\,\mathrm{Im}\big[s(x)\big] = s_1(x) + js_2(x) \qquad (4.17)$$

со спектром в базисе ВКФ $S(p) = S_1(p) + jS_2(p)$, где $S_1(p)$ и $S_2(p)$, есть комплексные спектры действительных сигналов $s_1(x)$ и $s_2(x)$ соответственно. Эти слагаемые комплексного сигнала $s(x)$ благодаря множителю j линейно независимы. В свою очередь, каждый из них можно представить как сумму m-четной и m-нечетной частей, которые также линейно-независимы и, кроме того, ортогональны:

$$s_1(x) = s_{1ч}(x) + s_{1н}(x), \;\; s_2(x) = s_{2ч}(x) + s_{2н}(x).$$

В результате получим, что комплексный сигнал $s(x)$ является линейной комбинацией четырех компонентов:

$$s(x) = \big[s_{1ч}(x) + s_{1н}(x)\big] + j\big[s_{2ч}(x) + s_{2н}(x)\big].$$

Его спектр также может быть выражен в виде аналогичной линейной комбинации

$$S(p) = a(p) - jb(p) = \big[S_{1ч}(p) + S_{2н}(p)\big] - j\big[S_{1н}(p) - S_{2ч}(p)\big],$$

где компоненты $S_{1ч}(p)$ и $S_{1н}(p)$ определяются только действительным сигналом $s_1(x)$, а компоненты $S_{2н}(p)$ и $S_{2ч}(p)$ – только действительным сигналом $s_2(x)$. Отсюда следует, что спектр комплексного сигнала можно всегда разделить и выделить из него спектры сигналов $s_1(x)$ и $s_2(x)$.

4.3. Свертка и корреляция

В соответствии с введенным понятием m-сдвига, определим m-свертку как

$$s_1(x) \underset{m}{*} s_2(x) = \sum_{\tau=0}^{N-1} s_1(\tau) \cdot \bar{s}_2(x \underset{m}{\ominus} \tau). \qquad (4.18)$$

m-свертке присущи те же свойства, что и обычной свертке, а именно:

– дистрибутивность:

$$s_1(x) \underset{m}{*} \big(s_2(x) + s_3(x)\big) = s_1(x) \underset{m}{*} s_2(x) + s_1(x) \underset{m}{*} s_3(x);$$

– ассоциативность:

$$s_1(x) \underset{m}{*} s_2(x) \underset{m}{*} s_3(x) = \Big[s_1(x) \underset{m}{*} s_2(x) \Big] \underset{m}{*} s_3(x);$$

– коммутативность:

$$s_1(x) \underset{m}{*} s_2(x) = \overline{s_2(x) \underset{m}{*} s_1(x)};$$

– свертка с единичным импульсом:

$$s_1(x) \underset{m}{*} u(x \underset{m}{\ominus} \tau) = s(x \underset{m}{\ominus} \tau).$$

Поскольку m-сдвиг не выводит сигнал за пределы интервала N, то и m-свертка дает новый сигнал, располагающийся на том же интервале N. В частном случае, при $m = N$ (базис ДЭФ), m-сдвиг представляет собой круговую перестановку отсчетов, а m-свертка – круговую или периодическую свертку. В другом частном случае, при $m = 2$ (базис функций Уолша), m-свертка представляет собой диадную свертку.

Теорема о спектре свертки. Представим в (4.18) сигнал $s_2(x \underset{m}{\ominus} \tau)$ в виде дискретного преобразования (4.8). Тогда получим

$$s_1(x) \underset{m}{*} s_2(x) = \sum_{\tau=0}^{N-1} s_1(\tau) \sum_{p=0}^{N-1} \overline{S}_2(p) \cdot \overline{\mathrm{Pal}}(p, x \underset{m}{\ominus} \tau) =$$
$$= N \sum_{p=0}^{N-1} \overline{S}_2(p) \left[\frac{1}{N} \sum_{\tau=0}^{N-1} s_1(\tau) \mathrm{Pal}(p, \tau) \right] \overline{\mathrm{Pal}}(p, x).$$

Выражение в прямых скобках есть в соответствии с (4.8) комплексно-сопряженный спектр $\overline{S}_1(p)$, поэтому

$$s_1(x) \underset{m}{*} s_2(x) = N \sum_{p=0}^{N-1} \overline{S}_2(p) \overline{S}_1(p) \overline{\mathrm{Pal}}(p, x).$$

Перейдем в обеих частях равенства к комплексно-сопряженным функциям:

$$\frac{1}{N} \overline{s_1(x) \underset{m}{*} s_2(x)} = \sum_{p=0}^{N-1} S_1(p) S_2(p) \mathrm{Pal}(p, x)$$

или сокращенно

$$\overline{s_1(x) \underset{m}{*} s_2(x)} \leftrightarrow N S_1(p) S_2(p).$$

Таким образом, спектр комплексно-сопряженной свертки двух сигналов есть произведение спектров этих сигналов (с коэффициентом N).

Теорема о спектре произведения двух сигналов. Этот спектр определяется выражением

$$S_{\Pi}(p)=\frac{1}{N}\sum_{x=0}^{N-1}s_1(x)s_2(x)\overline{\mathrm{Pal}}(p,x).$$

Применив к $s_1(x)$ формулу обращения из (4.8) и изменив порядок суммирования, получим

$$S_{\Pi}(p)=\sum_{q=0}^{N-1}S_1(q)\left[\frac{1}{N}\sum_{x=0}^{N-1}s_2(x)\overline{\mathrm{Pal}}(p\underset{m}{\ominus}q,x)\right]=\sum_{q=0}^{N-1}S_1(q)S_2(p\underset{m}{\ominus}q).$$

Правая часть этого равенства есть свертка функций $S_1(p)$ и $\overline{S}_2(p)$, поэтому можно записать

$$s_1(x)\underset{m}{*}s_2(x)\leftrightarrow S_1(p)\underset{m}{*}\overline{S}_2(p). \qquad (4.19)$$

Следовательно, при перемножении сигналов происходит свертывание спектра одного из них с комплексно-сопряженным спектром другого.

Взаимокорреляционная функция двух сигналов $B_{12}(\tau)$ и автокорреляционная функция сигнала $B(\tau)$, если их связывать с энергией сигналов, определяются следующим образом:

$$B_{12}(\tau)=\sum_{x=0}^{N-1}s_1(x)\cdot\bar{s}_2\left(x\underset{m}{\ominus}\tau\right), \qquad (4.20)$$

$$B(\tau)=\sum_{x=0}^{N-1}s(x)\cdot\bar{s}\left(x\underset{m}{\ominus}\tau\right). \qquad (4.21)$$

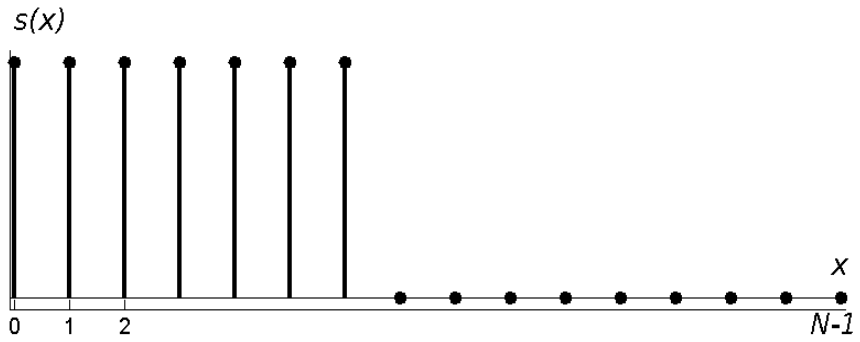
Заметим, что в отличие от свертки (4.18) здесь суммирование ведется по переменной x, поэтому значения автокорреляционной функции зависят только от *с-сдвига* комплексно-сопряженной копии относительно самого сигнала.

На рис. 4.1 приведена в качестве примера автокорреляционная функция дискретного прямоугольного импульса, вычисленная для различных m. Видно, что в различных базисах ВКФ автокорреляционная функция одного и того же сигнала имеет различный вид. Это открывает возможности выбора оптимального базиса для решения задач такого типа, как обнаружение и различение сигналов в шумах.

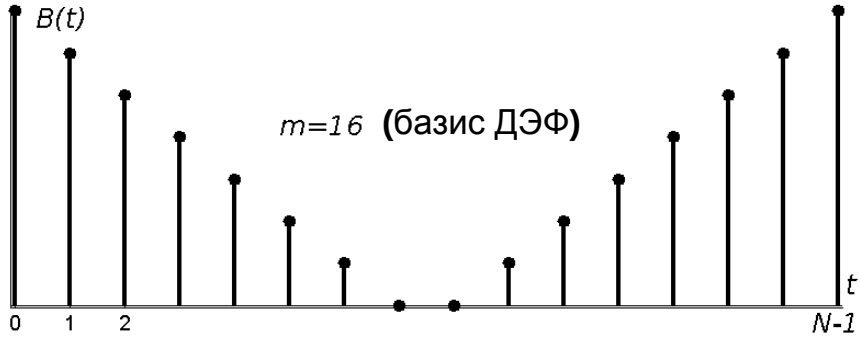
Можно убедиться в том, что m-автокорреляционная функция обладает свойствами, аналогичными свойствам обычной автокорреляционной функции, в частности:

– при $\tau=0$ она имеет действительное значение, равное энергии сигнала $B(0)=E$;

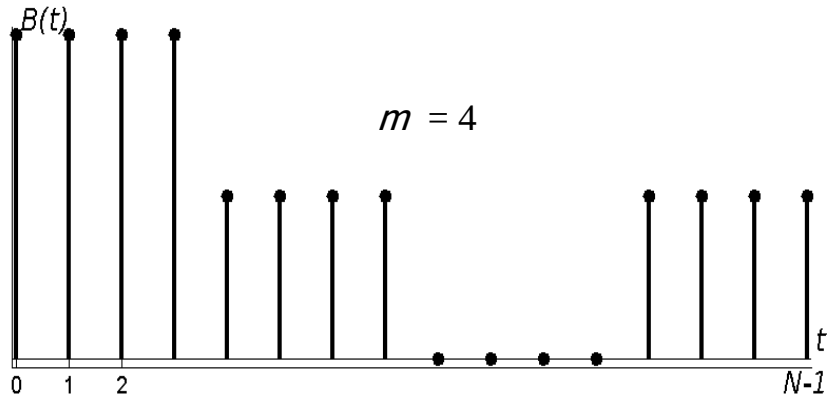
– ее действительная часть является m-четной функцией, а мнимая – m-нечетной функцией:



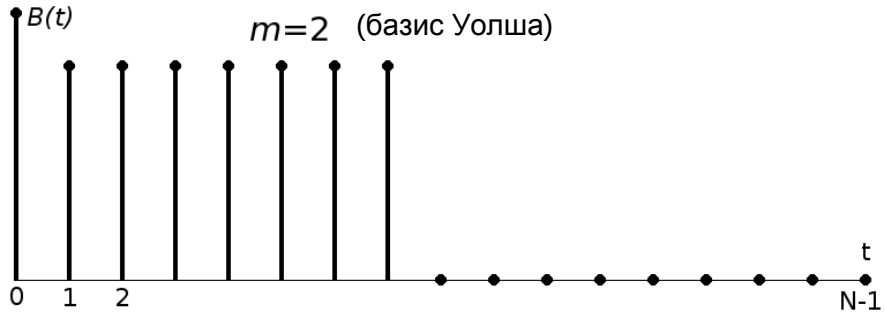
а)



б)



в)



г)

Рис. 4.1. m-автокорреляционная функция прямоугольного импульса

$$\mathrm{Re}\left[B(\tau)\right]=\mathrm{Re}\left[B(\tau^{*})\right];$$

$$\mathrm{Im}\left[B(\tau)\right]=-\mathrm{Im}\left[B(\tau^{*})\right],$$

где τ и τ^{*} – m-ично противоположные числа;

– любое ее значение не превосходит по модулю энергию сигнала $\left|B(\tau)\right|\le E$.

Теорема Винера - Хинчина. Покажем, что в базисе ВКФ автокорреляционная функция и энергетический спектр

$$e(p)=NS(p)\overline{S}(p)$$

связаны между собой ДПФ. Действительно,

$$B(\tau)=\sum_{x=0}^{N-1}s(x)\cdot\overline{s}\left(x\underset{m}{\ominus}\tau\right)=\sum_{x=0}^{N-1}s(x)\sum_{p=0}^{N-1}\overline{S}(p)\cdot\overline{\mathrm{Pal}}\left(p,x\underset{m}{\ominus}\tau\right)=$$

$$=N\sum_{p=0}^{N-1}S(p)\overline{S}(p)\cdot\mathrm{Pal}\left(p,\tau\right)=,$$

или сокращенно

$$B(\tau)\leftrightarrow e(p).\qquad(4.22)$$

Это выражение можно трактовать как теорему Винера - Хинчина применительно к базису ВКФ.

Теорема об инвариантности энергетического спектра, относительно *m*-сдвига. Важным свойством энергетического спектра в базисе ВКФ является то, что он не изменяется при *m-сдвиге* сигнала:

$$\begin{aligned}e_{\tau}(p)=NS_{\tau}(p)\overline{S}_{\tau}(p)=NS(p)\overline{\mathrm{Pal}}\left(p,x\underset{m}{\ominus}\tau\right)\overline{S}(p)\cdot\mathrm{Pal}\left(p,\tau\right)=\\=NS(p)\overline{S}(p)=e(p).\end{aligned}\qquad(4.23)$$

4.4. Спектры некоторых сигналов

Для иллюстрации особенностей дискретного преобразования Фурье в базисе ВКФ рассмотрим спектры некоторых часто употребляемых дискретных сигналов: единичного импульса, прямоугольного импульса и действительной или мнимой части ВКФ.

Спектр единичного импульса. Единичный импульс

$$u(x)=\begin{cases}1,x=0\\0,0<x\le N-1\end{cases}$$

имеет равномерный спектр с нулевой фазой:

$$U(p)=\frac{1}{N}\sum_{x=0}^{N-1}u(x)\cdot\overline{\mathrm{Pal}}(p,x)=\frac{1}{N}\overline{\mathrm{Pal}}(p,0)=\frac{1}{N}. \qquad (4.24)$$

Спектр постоянного сигнала $s(x)=1$. Учитывая, что

$$\sum_{x=0}^{N-1}\mathrm{Pal}(p,x)=0,$$

имеем

$$S(p)=\frac{1}{N}\sum_{x=0}^{N-1}\overline{\mathrm{Pal}}(p,x)=U(p), \qquad (4.25)$$

где $U(p)$ – единичный «импульс» по частоте:

$$U(p)=\begin{cases}1, p=0;\\ 0, 0<p\le N-1.\end{cases}$$

Спектр дискретного прямоугольного импульса. Такой импульс определяется выражением

$$s(x)=\begin{cases}1, 0\le x\le r;\\ 0, r+1\le x\le N-1,\end{cases}$$

и его спектр находится следующим образом:

$$\begin{aligned}S(p)&=\frac{1}{N}\sum_{x=0}^{N-1}s(x)\cdot\overline{\mathrm{Pal}}(p,x)=\frac{1}{N}\sum_{x=0}^{N-1}\exp\left(-j\frac{2\pi}{m}\sum_{i=1}^{n}p_{n+1-i}\cdot x_i\right)=\\ &=\frac{1}{N}\sum_{x_1=0}^{r_1}\exp\left(-j\frac{2\pi}{m}p_n\cdot x_i\right)\ldots\sum_{x_n=0}^{r_n}\exp\left(-j\frac{2\pi}{m}p_1\cdot x_n\right)\end{aligned}, \qquad (4.26)$$

где $r_1, r_2, \ldots, r_n$ – разряды m-ичного представления числа r. В правой части этого выражения i-я сумма представляет собой сумму геометрической прогрессии

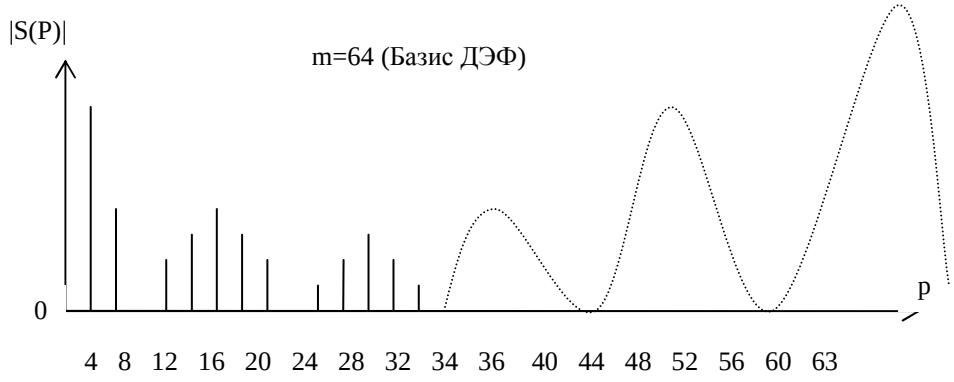
$$\sum_{x_i=0}^{r_i-1}\exp\left(-j\frac{2\pi}{m}p_{n+1-i}\cdot x_i\right)=\frac{\sin\left(\frac{\pi}{m}p_{n+1-i}\cdot(r_i+1)\right)}{\sin\left(\frac{\pi}{m}p_{n+1-i}\right)}\exp\left(-j\frac{\pi}{m}p_{n+1-i}\cdot r_i\right).$$

Следовательно, окончательно получим

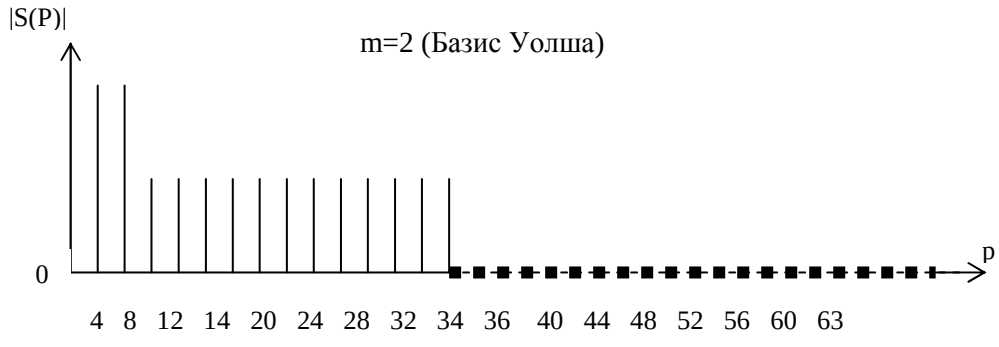
$$S(p)=\frac{1}{N}\prod_{i=1}^{n}\frac{\sin\left(\frac{\pi}{m}p_{n+1-i}\cdot(r_i+1)\right)}{\sin\left(\frac{\pi}{m}p_{n+1-i}\right)}\exp\left(-j\frac{\pi}{m}p_{n+1-i}\cdot r_i\right). \qquad (4.27)$$

На рис. 4.2 приведены амплитудные спектры $|S(p)|$ прямоугольного импульса в разных системах ВКФ. Сопряженные спектральные компоненты не показаны и обозначены лишь огибающей.

а)



б)



в)

Рис. 4.2. Амплитудные спектры прямоугольного импульса

На рис. 4.3 показан амплитудный спектр дискретного прямоугольного импульса той же длительности, что и на рис. 4.2, но на меньшем интервале и при $m = N = 8$ (базис ДЭФ).

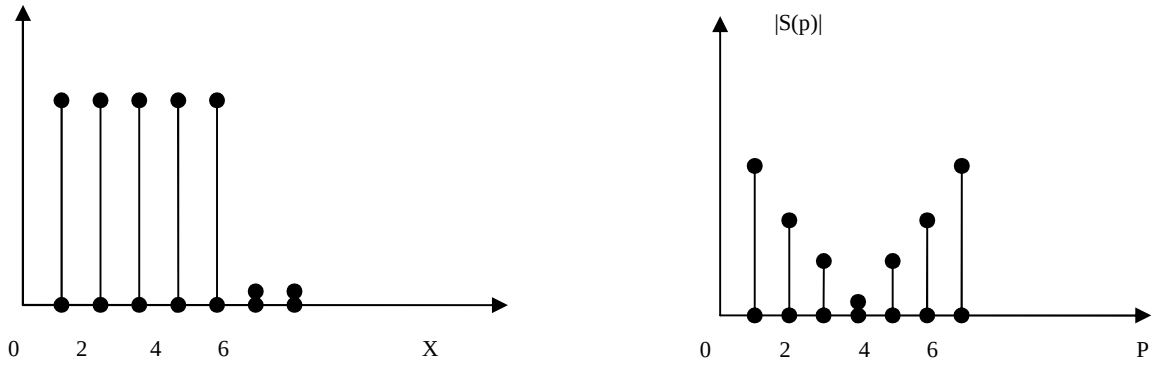


Рис. 4.3. Амплитудный спектр прямоугольного импульса в базисе ДЭФ

Переход от импульса на рис. 4.3 к импульсу на рис. 4.2 можно трактовать как увеличение интервала N в восемь раз за счет добавления 56 нулевых точек.

Спектр действительной и мнимой частей ВКФ. Использовав выражения (4.12) для действительной и мнимой частей ВКФ $\mathrm{Re}\left[\mathrm{Pal}(q,x)\right]$ и $\mathrm{Im}\left[\mathrm{Pal}(q,x)\right]$, нетрудно получить спектр этих функций. Так, действительная часть ВКФ имеет спектр

$$S_R(p) = \frac{1}{N}\sum_{x=0}^{N-1}\frac{1}{2}\left[\mathrm{Pal}(q,x) + \mathrm{Pal}(q^*,x)\right]\overline{\mathrm{Pal}}(p,x) =$$
$$= \frac{1}{2N}\sum_{x=0}^{N-1}\mathrm{Pal}(p \underset{m}{\ominus} q,x) + \frac{1}{2N}\sum_{x=0}^{N-1}\mathrm{Pal}(p \underset{m}{\ominus} q^*,x).$$

Учитывая, что ВКФ $\mathrm{Pal}(p,x)$ при $p \neq 0$ имеет нулевое среднее значение, получаем, что суммы в правой части этого равенства равны нулю при всех p, за исключением $p = q$ и $p = q^*$.

Если ввести единичный «импульс» по частоте $U(p)$, то спектр действительной части ВКФ можно записать в виде

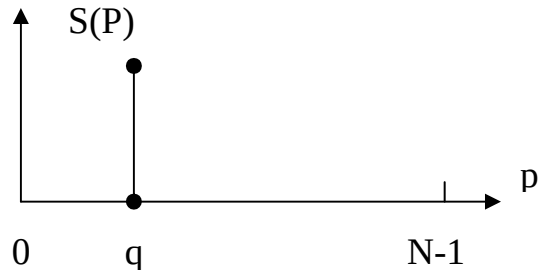
$$S_R(p) = \frac{1}{2}\left[U\left(p \underset{m}{\ominus} q\right) + U\left(p \underset{m}{\ominus} q^*\right)\right];$$

$\mathrm{Re}\left[\mathrm{Pal}(q,x)\right]$; в) $\mathrm{Im}\left[\mathrm{Pal}(q,x)\right]$.

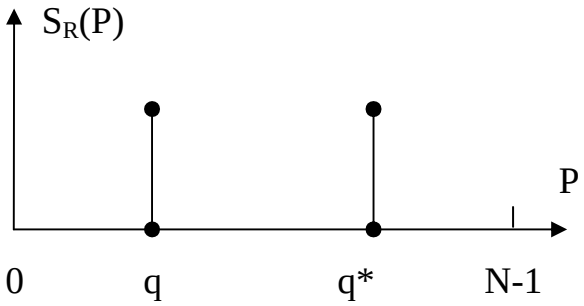
Аналогично спектр мнимой части ВКФ равен

$$S_R(p) = \frac{1}{2j}\left[U\left(p \underset{m}{\ominus} q\right) - U\left(p \underset{m}{\ominus} q^*\right)\right].$$

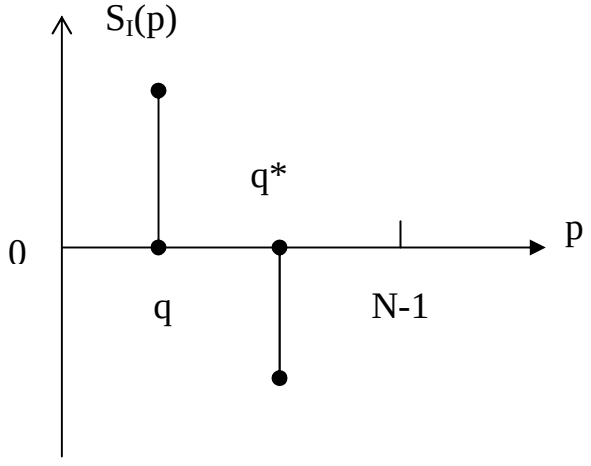
Эти спектры показаны на рис. 4.4, где для сравнения приведен также спектр ВКФ.



а) Pal(q,x);



б) Re[Pal(q,x)];



в) Im[Pal(q,x);

Рис. 4.4. Спектр в базисе ВКФ-Пэли:
а) Pal(q,x);
б) Re[Pal(q,x)];
в) Im[Pal(q,x);

5. ВВЕДЕНИЕ В ЦИФРОВОЙ СПЕКТРАЛЬНЫЙ АНАЛИЗ

Спектральный анализ – это один из методов обработки сигналов, который позволяет охарактеризовать частотный состав измеряемого сигнала. Преобразование Фурье является математической основой, которая связывает временной или пространственный сигнал (или же некоторую модель этого сигнала) с его представлением в частотной области. Существенный вклад в развитие цифровых методов спектрального анализа внесли эффективные алгоритмы, предназначенные для вычисления дискретного преобразования Фурье, предложенные Д. Кули и Д. Тьюки в 1965 году. 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 в некоторых встречающихся на практике случаях, позволяющее ускорить вычисление ДПФ в 100 и более раз 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. Возможно, именно алгоритмы БПФ более, чем какие-либо другие методы существенно расширили область применения методов спектрального анализа как средства обработки сигналов. Поэтому начнем рассмотрение вопросов ЦСА с алгоритмов БПФ, включающих алгоритмы с основанием 2 и прореживанием по времени и по частоте.

5.1. Введение в алгоритмы БПФ с основанием 2

Напомним, что прямое ДПФ конечной последовательности $x(n)$, $0 \le n \le N-1$ определяется выражением

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}kn},\ k = \overline{0, N-1} \qquad (5.1)$$

или в более удобном виде как

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N{}^{kn},\ k = \overline{0, N-1}, \qquad (5.2)$$

где

$$W_N = e^{-j(2\pi/N)},$$

а обратное ДПФ имеет вид

$$x(n)=\frac{1}{N}\sum_{k=0}^{N-1}X(k)W_N^{-kn},\ n=\overline{0,N-1}. \qquad (5.3)$$

Выражения (5.2) и (5.3) различаются только знаком экспоненты от W_N и коэффициентом $1/N$, поэтому рассуждения, касающиеся вычислительных процедур для (5.2), применимы с очевидными изменениями к (5.3).

Из соотношения (5.2) следует, что в случае, когда последовательность $x(n)$ является комплексной, при прямом вычислении N-точечного ДПФ нужно выполнить $(N-1)^2$ комплексных умножений и $N\cdot(N-1)$ комплексных сложений. Таким образом, для достаточно больших N (порядка 1000) прямое вычисление ДПФ требует выполнения чрезмерного количества вычислительных операций.

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 получилось ДПФ исходной N-точечной последовательности. Так, например, если N четное, а исходная N-точечная последовательность разбита на две $(N/2)$-точечные последовательности, то для вычисления искомого N-точечного ДПФ потребуется порядка $\left(\frac{N}{2}\right)^2 2=\frac{N^2}{2}$ комплексных умножений, т. е. вдвое меньше по сравнению с прямым вычислением. Здесь множитель $\left(\frac{N}{2}\right)^2$ 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 (предполагая, что $N/2$ – четное) и сокращая тем самым объем вычислений еще в два раза. Выигрыш 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 две $(N/2)$-точечные последовательности $\{x_1(n)\}$ и $\{x_2(n)\}$ из четных и нечетных членов $x(n)$ соответственно, т. е.

$$\begin{aligned} x_1(n)&=x(2n),\ n=0,1,\ldots,N/2-1;\\ x_2(n)&=x(2n+1),\ n=0,1,\ldots,N/2-1. \end{aligned} \qquad (5.4)$$

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

$$X(k)=\sum_{n=0}^{N-1}x(n)W_N^{kn}=\sum_{n=0}^{N/2-1}x(2n)W_N^{2nk}+\sum_{n=0}^{N/2-1}x(2n+1)W_N^{(2n+1)k}=$$
$$=\sum_{n=0}^{N/2-1}x_1(n)W_{N/2}^{kn}+W_N^k\sum_{n=0}^{N/2-1}x_2(n)W_{N/2}^{kn}=X_1(k)+W_N^kX_2(k), \tag{5.5}$$

где учтено, что $W_N^2=[e^{-j(2\pi/N)}]^2=e^{-j[2\pi/(N/2)]}=W_{N/2}$, и введены обозначения для $(N/2)$-точечных ДПФ последовательностей $\{x_1(n)\}$ и $\{x_2(n)\}$:

$$X_1(k)=\sum_{n=0}^{N/2}x_1(n)W_{N/2}^{nk},$$
$$X_2(k)=\sum_{n=0}^{N/2}x_2(n)W_{N/2}^{nk}.$$

Из формулы (5.5) следует, что N-точечное ДПФ $X(k)$ может быть разложено на два $(N/2)$-точечных ДПФ, результаты которых объединяются согласно (5.5). Если бы $(N/2)$-точечные ДПФ вычислялись обычным способом, то для вычисления N-точечного ДПФ потребовалось бы $N^2/2+N$ комплексных умножений. При больших N (когда $N^2/2\gg N$) это позволяет сократить время вычисления на 50%.

Поскольку $X(k)$ определено при $0\le k\le N-1$, а $X_1(k)$ и $X_2(k)$ определены при $0\le k\le N/2-1$, необходимо доопределить формулу (5.5) для $k>N/2$. Это определение достаточно очевидно и может быть записано следующим образом:

$$X(k+\frac{N}{2})=X_1(k+\frac{N}{2})+W_N^{k+N/2}X_2(k+\frac{N}{2})=$$
$$=X_1(k)-W_N^kX_2(k),\quad k=\overline{0,(N/2)-1}, \tag{5.6}$$

где учтено, что $W_N^{k+N/2}=e^{-j\frac{2\pi N}{2N}}W_N^k=e^{-j\pi}W_N^k=-W_N^k$.

Таким образом, вычисление $X(k)$ по $X_1(k)$ и $X_2(k)$ можно представить в виде

$$X(k)=X_1(k)+W_N^kX_2(k),$$
$$X\left(k+\frac{N}{2}\right)=X_1(k)-W_N^kX_2(k),\quad k=\overline{0,(N/2)-1}. \tag{5.7}$$

На рисунке 5.1 с помощью направленного графа представлена последовательность операций при вычислении восьмиточечного ДПФ с использованием двух четырехточечных ДПФ. Незачерненный кружок графа означает операцию сложения / вычитания, причем верхний выход соответствует сумме, а нижний – разности. Стрелка обозначает операцию умножения на значение множителя a, указанного над стрелкой. 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 рассчитываются их преобразования $X_1(k)$ и $X_2(k)$. Затем в соответствии с формулой (5.7) получают $X(k)$.

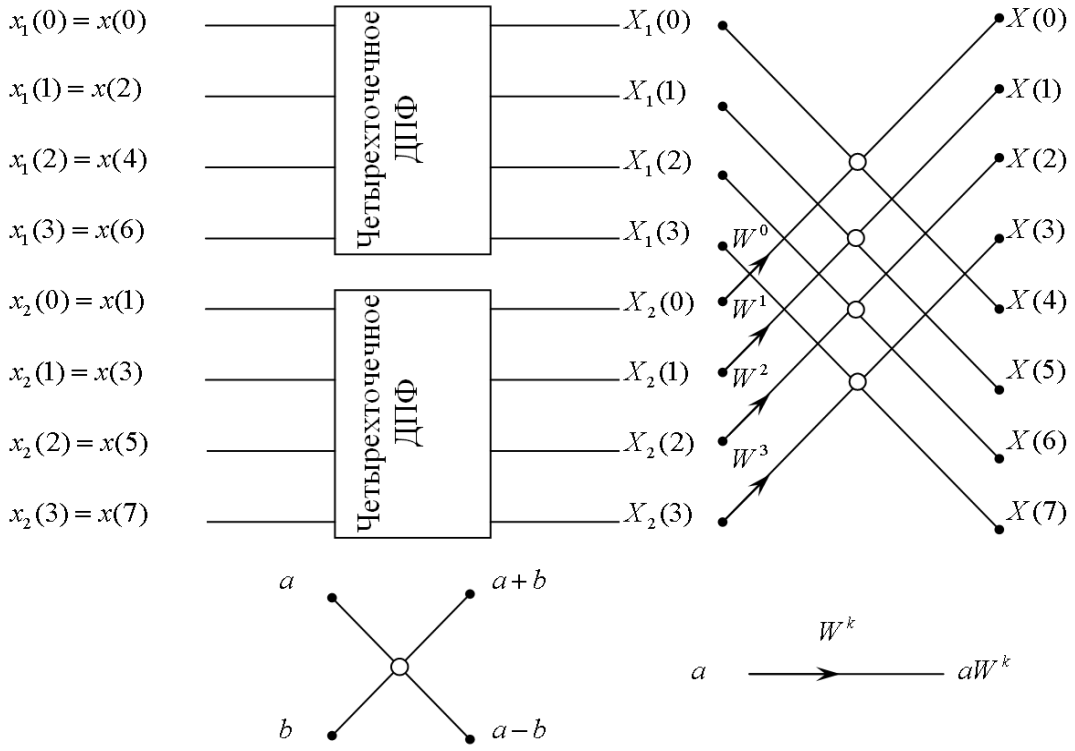


Рис. 5.1. Вычисление восьмиточечного ДПФ через два четырехточечных ДПФ

Выражение (5.7) соответствует разбиению исходного N-точечного вычисления ДПФ на два $(N/2)$-точечных вычислений. Если $N/2$ – четное число, что имеет место всегда, когда N равно степени 2, то можно вычислять каждое $(N/2)$-точечное ДПФ в (5.7) путем разбиения сумм на два $(N/4)$-точечных ДПФ, которые затем объединяются, давая $(N/2)$-точечное ДПФ. Каждая из последовательностей $x_1(n)$ и $x_2(n)$ разбивается на две последовательности, состоящие из четных и нечетных членов. Ана-

логично $(N/2)$-точечные ДПФ могут быть записаны как комбинации двух $(N/4)$-точечных ДПФ, т. е.

$$\begin{aligned} X_1(k) &= A(k) + W_{N/2}^{k} B(k), \\ X_1(k+N/4) &= A(k) - W_{N/2}^{k} B(k); \\ X_2(k) &= C(k) + W_{N/2}^{k} D(k), \\ X_2(k+N/4) &= C(k) - W_{N/2}^{k} D(k), \end{aligned} \tag{5.8}$$

или

$$\begin{aligned} X_1(k) &= A(k) + W_{N}^{2k} B(k), \\ X_1(k+N/4) &= A(k) - W_{N}^{2k} B(k); \\ X_2(k) &= C(k) + W_{N}^{2k} D(k), \\ X_2(k+N/4) &= C(k) - W_{N}^{2k} D(k), \end{aligned} \tag{5.9}$$

где $0 \le k \le N/4-1$, $A(k)$ и $B(k)$ – $(N/4)$-точечные ДПФ соответственно четных и нечетных членов $x_1(n)$, $C(k)$ и $D(k)$ – $(N/4)$-точечные ДПФ соответственно четных и нечетных членов $x_2(n)$. На рис. 5.2 показан результирующий направленный граф, в котором четырехточечные ДПФ из рис. 5.1 рассчитываются согласно (5.9).

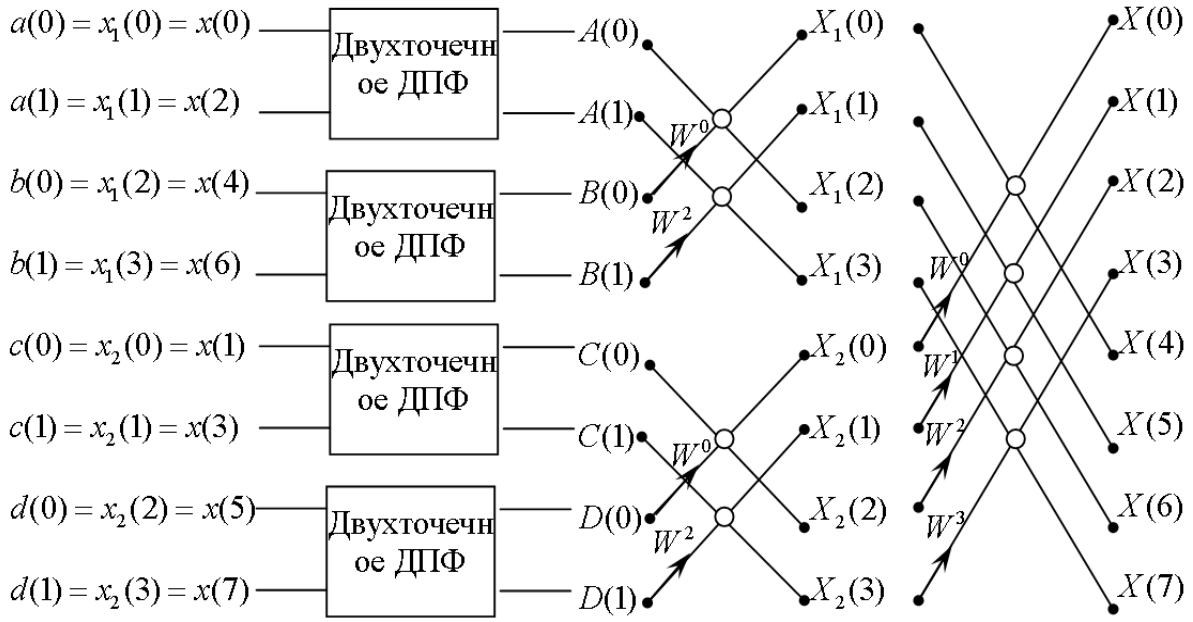


Рис.5.2. Вычисление восьмиточечного ДПФ через два четырехточечных ДПФ, которые в свою очередь вычисляются через четыре двухточечных ДПФ.

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, например, $A(k)$, $k=0,1$, может быть рассчитано без использования умножений по формулам

$$\begin{aligned} A(0) &= a(0) + W_2^0 a(1) = a(0) + W_8^0 a(1) = a(0) + a(1), \\ A(1) &= a(0) + W_2^1 a(1) = a(0) + W_8^4 a(1) = a(0) - a(1). \end{aligned} \tag{5.10}$$

Таким образом, восьмиточечное ДПФ (см. рис. 5.1 и рис. 5.2) в итоге сводится к алгоритму, описываемому направленным графом, представленным на рис. 5.3.

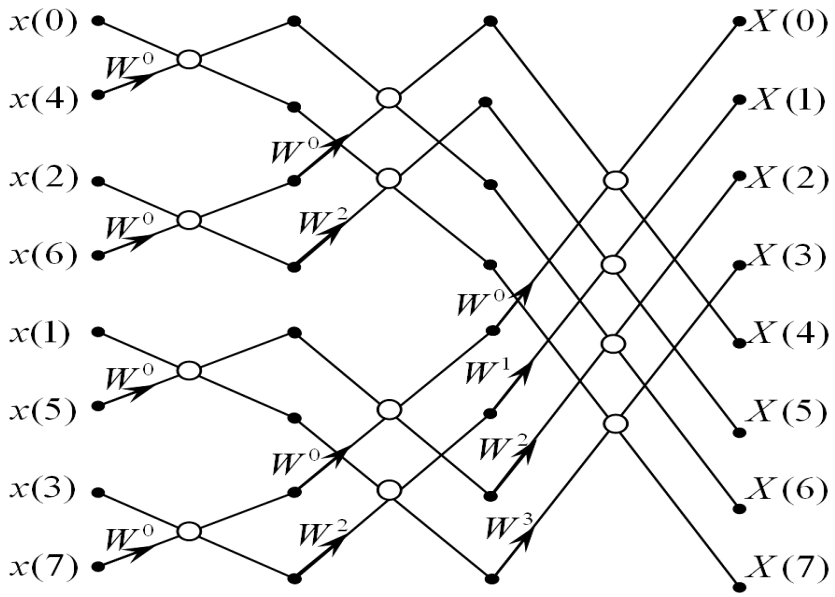


Рис. 5.3. Восьмиточечное ДПФ, полученное последовательным прореживанием в 2 раза

Анализ графа на рис. 5.3 и процедуры последовательного сокращения вдвое размеров преобразований показывает, что на каждом этапе БПФ (т. е. при каждом сокращении размеров 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 $\frac{N}{2}\log_2 N$. 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. Так, например, на рис. 5.3 первый этап БПФ содержит только сложения и вычитания комплексных чисел. Даже на

втором этапе используются только сложения и вычитания комплексных чисел. Фактически, как следует из направленного графа на рис. 5.3, вместо ожидаемых $\frac{8}{2}\log_2 8 = 12$ умножений достаточно выполнить всего два нетривиальных умножения. Однако для больших значений N фактическое число нетривиальных умножений хорошо аппроксимируется выражением $\frac{N}{2}\log_2 N$.

Описанный выше алгоритм был назван алгоритмом с прореживанием по времени, поскольку на каждом этапе входная (т. е. временная) последовательность разделяется на две обрабатываемые последовательности меньшей длины, т. е. входная последовательность прореживается на каждом этапе. Другая форма алгоритма БПФ (с прореживанием по частоте) будет описана ниже, а сейчас целесообразно рассмотреть некоторые общие свойства алгоритмов БПФ.

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:

$$\begin{aligned} X &= A + W_N^k B, \\ Y &= A - W_N^k B. \end{aligned} \qquad (5.11)$$

На рис. 5.4 изображен направленный граф базовой операции (5.11).

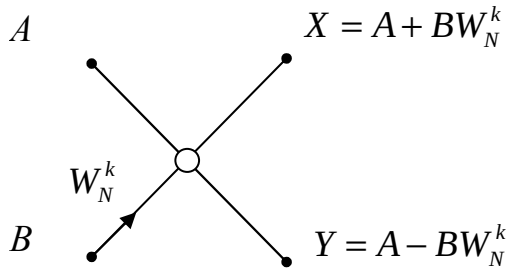


Рис. 5.4. Базовая операция алгоритма БПФ.

Внимательное рассмотрение направленного графа на рис. 5.3 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 находились 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.

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. $k = 0,1,\ldots,N-1$. В примере на рис. 5.3 для этого требовался следующий порядок размещения входной последовательности: $x(0)$, $x(4)$, $x(2)$, $x(6)$, $x(1)$, $x(5)$, $x(3)$ и $x(7)$. Характер перестановки элементов входной последовательности может быть описан сравнительно просто. Далее будет показано, что в случае, когда N является степенью 2, входная последовательность должна быть расположена в памяти в двоично-инверсном порядке для того, чтобы выходная последовательность получалась 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 $N = 2^3 = 8$ прямой порядок номеров приведен в табл. 5.1 слева, а двоично-инверсный порядок – справа. Таким образом, для двоичной инверсии входной последовательности необходим соответствующий алгоритм.

Из сказанного выше ясно, что перестановку входной последовательности можно произвести с замещением, меняя в парах местами числа с прямым и двоично-инверсным номерами и используя для этого лишь одну вспомогательную ячейку памяти. На рис. 5.5 показана схема перестановки данных, представленных в табл. 5.1.

Таблица 5.1

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

Окончание табл. 5.1

4	100	001	1
5	101	101	5
6	110	011	3
7	111	111	7

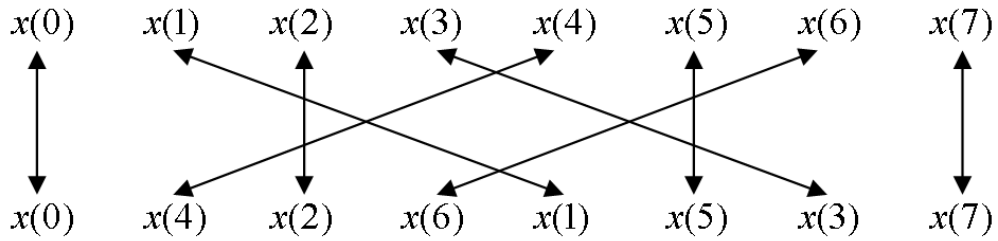


Рис. 5.5. Перестановка данных с замещением

5.2. 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 входная 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.

$$\begin{aligned} x_1(n) &= x(n), \quad n = 0,1,...,(N/2)-1, \\ x_2(n) &= x(n+N/2), \quad n = 0,1,...,(N/2)-1. \end{aligned} \tag{5.12}$$

При таком разбиении N-точечное ДПФ последовательности $x(n)$ можно записать в виде

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

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

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

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

где учтено, что $W_N^{kN/2} = e^{-j\frac{2\pi N}{N2}k} = e^{-j\pi k} = (-1)^k$.

Запишем выражения раздельно для четных и нечетных отсчетов ДПФ:

$$X(2k) = \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)](W_N^2)^{nk} = \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)]W_{N/2}^{nk}, \qquad (5.13)$$

$$X(2k+1) = \sum_{n=0}^{N/2-1} [x_1(n) - x_2(n)]W_N^{n(2k+1)} = \sum_{n=0}^{N/2-1} \{[x_1(n) - x_2(n)]W_N^n\}W_{N/2}^{nk}. \qquad (5.14)$$

Из выражений (5.13) и (5.14) видно, что четные и нечетные отсчеты ДПФ можно получить из $(N/2)$-точечных ДПФ последовательностей $f(n)$ и $g(n)$, равных

$$\begin{aligned} f(n) &= x_1(n) + x_2(n), \quad n = 0,1,\ldots,N/2-1, \\ g(n) &= [x_1(n) - x_2(n)]W_N^n, \quad n = 0,1,\ldots,N/2-1. \end{aligned} \qquad (5.15)$$

Таким образом, снова вычисление N-точечного ДПФ удалось свести к вычислению двух $(N/2)$-точечных ДПФ. На рис. (5.6) эта методика иллюстрируется для случая $N = 8$.

Описанную методику можно применить повторно и представить каждое из $(N/2)$-точечных ДПФ в виде комбинации двух $(N/4)$-точечных ДПФ. На рис. 5.7, показан переход от четырехточечных ДПФ к двухточечным ДПФ с последующим прямым вычислением двухточечных ДПФ.

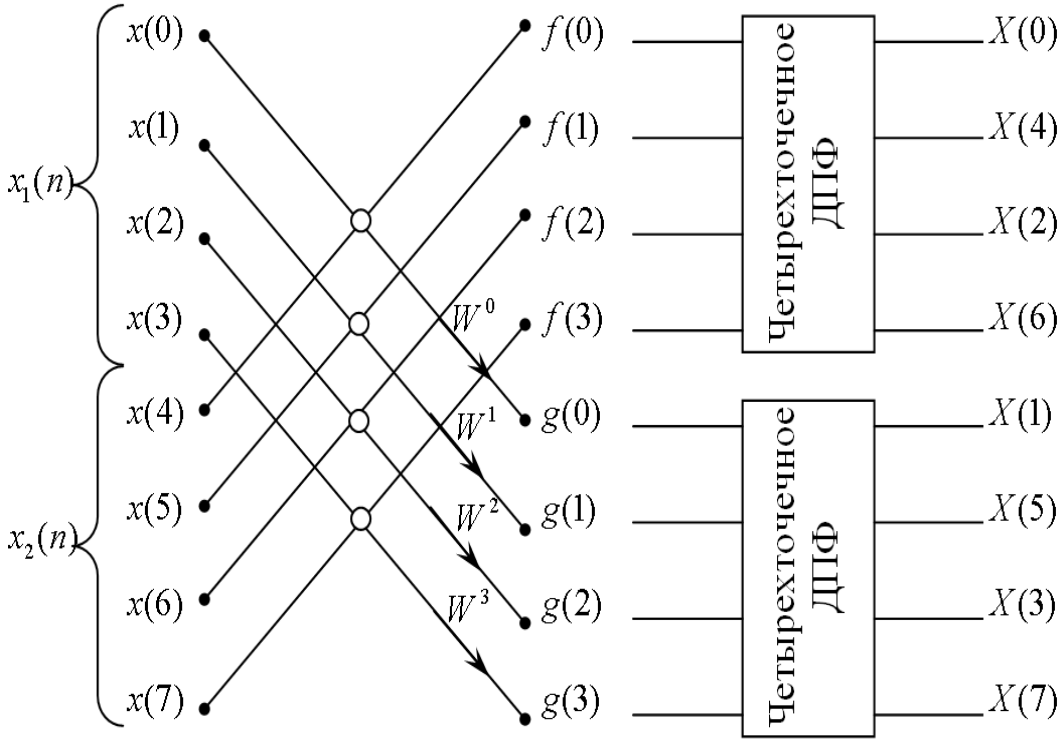


Рис. 5.6. Переход от восьмиточечного ДПФ к двум четырехточечным ДПФ при прореживании по частоте

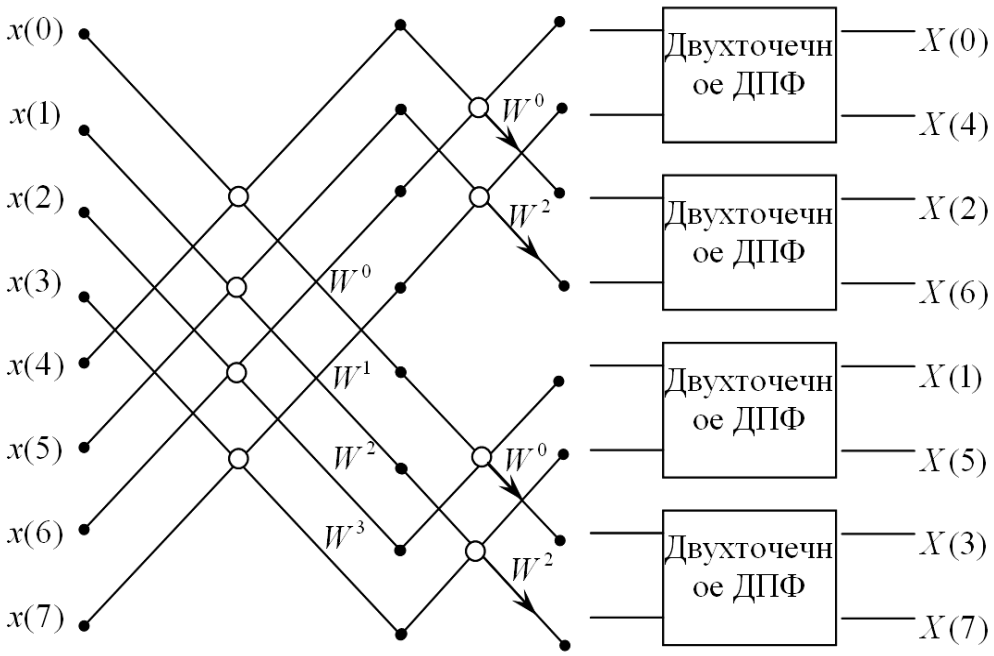


Рис. 5.7. Переход от четырехточечных ДПФ к двухточечным ДПФ

Сравнение алгоритмов, представленных на рис. 5.3 и 5.7, позволяет выявить два очевидных различия между ними. Во-первых, при прореживании по времени порядок следования входных отсчетов двоично-инверсный, а выходных – прямой и наоборот при прореживании по частоте (см. рис. 5.7). Второе отличие заключается в несколько ином выполнении базовой операции (см. рис. 5.8 и рис. 5.4): при прореживании по частоте комплексное умножение выполняется *после* сложения – вычитания.

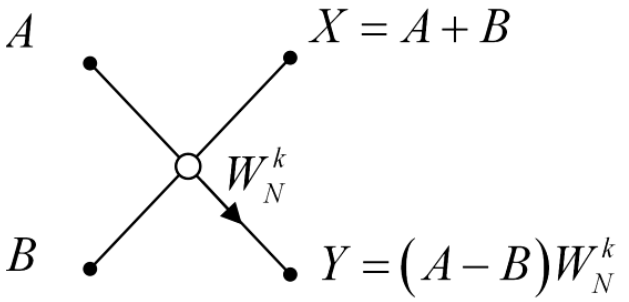


Рис. 5.8. Базовая операция алгоритма БПФ с прореживанием по частоте

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

операциям («бабочкам»), и между рис. 5.3 и рис. 5.9, т. е. то, что рис. 5.9 можно получить из рис. 5.3, изменив направление стрелок (сигналов) и поменяв вход и выход. Полный направленный граф восьмиточечного ДПФ с замещением и прореживанием по частоте представлен на рис. 5.9.

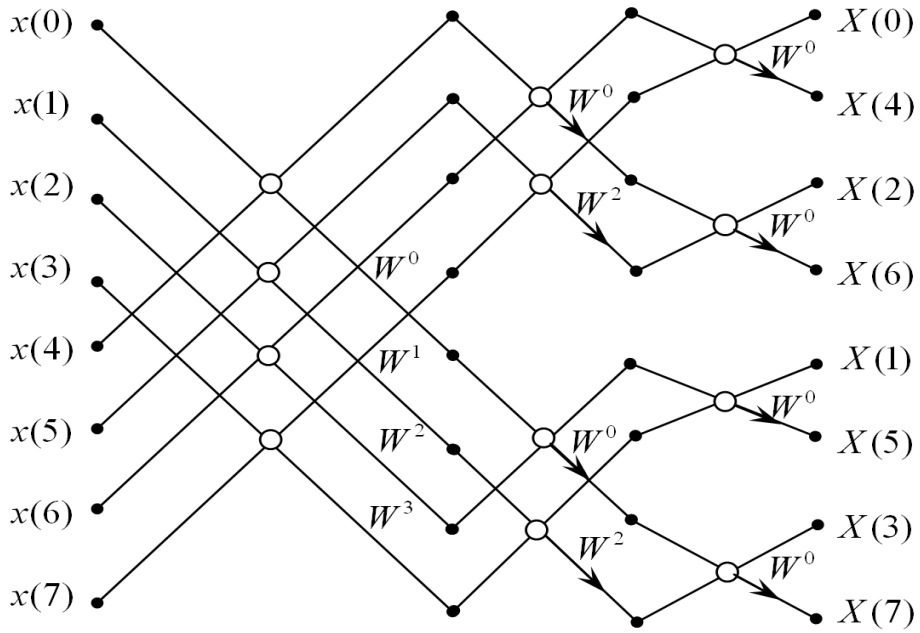


Рис. 5.9 Полный направленный граф восьмиточечного ДПФ с замещением и прореживанием по частоте

Как уже отмечалось, для вычисления обратного ДПФ (5.2) можно использовать БПФ-алгоритм, если разделить результат на N и использовать вместо степеней W_N степени W_N^{-1}.

5.3. Алгоритмы БПФ для составного значения *N*

Рассмотренные выше алгоритмы предполагали, что N является целой степенью 2, т. е. $N = 2^L$. В более общем случае эффективное вычисление ДПФ связано с представлением N в виде сомножителей:

$$N = p_1 p_2 \ldots p_L . \tag{5.16}$$

Если N представлено в виде произведения одинаковых сомножителей r, то соответствующий алгоритм называют алгоритмом БПФ по основанию r. Понятие «смешанное основание» означает, что не все сомножители N одинаковы. N-точечные алгоритмы по основанию 2 особенно просты для реализации, поэтому на практике выгодно всегда иметь дело с по-

следовательностями длины $N = 2^L$. Это можно сделать во многих случаях, просто дополняя последовательность конечной длины нулями, если это необходимо. В случаях, когда невозможно выбрать $N = 2^L$, приходится рассматривать более общую ситуацию (5.16).

Рассмотрим применение принципа прореживания по времени, когда N является произведением сомножителей, не все из которых равны 2. Пусть $q_1 = p_2 p_3 \dots p_L$, так что $N = p_1 q_1$. Если N является степенью 2, можно выбрать $p_1 = 2$, а $q_1 = N/2$. Используя прореживание по времени, можно разложить $x(n)$ на две последовательности длины $N/2$, состоящие из четных и нечетных выборок соответственно, как это мы делали выше. Если $N = p_1 q_1$, то можно разделить входную последовательность на p_1 последовательностей длины q_1, так что каждая p_1-я выборка попадает в одну последовательность. Например, если $p_1 = 3$, а $q_1 = 4$, так что $N = 12$, то можно разложить $x(n)$ на три последовательности длины 4, причем первая последовательность состоит из выборок $x(0), x(3), x(6), x(9)$, вторая – из $x(1), x(4), x(7), x(10)$, а третья – из $x(2), x(5), x(8), x(11)$. В общем случае можно записать $X(k)$ в виде

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{kn} = \sum_{r=0}^{q_1-1} x(p_1 r) W_N^{p_1 rk} + \sum_{r=0}^{q_1-1} x(p_1 r + 1) W_N^{k} W_N^{p_1 rk} + \dots +$$
$$+ \sum_{r=0}^{q_1-1} x(p_1 r + p_1 - 1) W_N^{(p_1-1)} W_N^{p_1 rk} = \sum_{l=0}^{p_1-1} W_N^{lk} \sum_{r=0}^{q_1-1} x(p_1 r + l) W_N^{p_1 rk}. \qquad (5.17)$$

Внутренние суммы можно представить как q_1-точечные ДПФ:

$$G(k) = \sum_{r=0}^{q_1-1} x(p_1 r + l) W_{q_1}^{rk},$$

потому что, как легко проверить, $W_N^{p_1 rk} = W_{q_1}^{rk}$ при $N = p_1 q_1$.

Таким образом, (5.17) представляет $X(k)$ в виде p_1 ДПФ последовательностей длины q_1. Чтобы определить число комплексных умножений и сложений для вычисления ДПФ по (5.17), будем считать, что p_1-точечные ДПФ получаются путем прямого вычисления. Из (5.17) видно, что нужно рассчитать p_1 q_1-точечных ДПФ. Поэтому общее число требуемых комплексных сложений и умножений равно. Внешняя сумма в (5.17) получа-

ется путем умножения q_1-точечных ДПФ на коэффициенты W_N^{lk} и сложения результатов. Так как двойное суммирование в (5.17) выполняется для N значений k, то для объединения p_1 q_1-точечных ДПФ требуется $N(p_1-1)$ комплексных сложений и умножений. Следовательно, общее число комплексных умножений и сложений для вычисления ДПФ по (5.17) равно $N(p_1-1)+p_1\cdot q_1^2$. Теперь q_1-точечное ДПФ может быть разложено аналогичным образом. В частности, если представить q_1 в виде $q_1=p_2q_2$, то q_1-точечные последовательности во внутренней сумме (5.17) могут быть разбиты на p_2 последовательностей, каждая из которых состоит из q_2 точек, так, что внутренняя сумма в (5.17) может быть заменена на двойную сумму тем же способом, с которого мы начали. Тогда число операций требуемых для расчета q_1-точечных ДПФ в (5.17) вместо q_1^2 станет равным $q_1(p_2-1)+p_2\cdot q_2^2$. Следовательно, общее число комплексных умножений и сложений для вычисления ДПФ станет равным

$$N(p_1-1)+N(p_2-1)+p_1\cdot p_2\cdot q_1^2.$$

Если продолжить эту процедуру, разлагая далее q_2-точечные ДПФ, то, в конце концов, общее число комплексных умножений и сложений для вычисления ДПФ станет равным

$$N(p_1+p_2+\ldots+p_L-L). \qquad (5.18)$$

Из (5.18) видно, что лучше производить разложение на максимально возможное число сомножителей и лучше выбирать простые сомножители.

5.4. Реализация ДПФ на основе цифровой фильтрации

Покажем, что ДПФ можно рассматривать как отклик цифрового фильтра и найдем структуру фильтра, определяющего один спектральный отсчет $X(k)$ в точке $\omega_k=2\pi k/N$:

$$X(k)=\sum_{n=0}^{N-1}x(n)e^{-j\omega_k n}=\sum_{n=0}^{N-1}x(n)h_k(N-1-n)=y_k(N-1). \qquad (5.19)$$

В такой форме записи (5.19) можно рассматривать $X(k)$ как выходной сигнал $y_k(N-1)$ фильтра с импульсной характеристикой $h_k(n)$ в мо-

мент $(N-1)$. Из (5.19) импульсная характеристика $h_k(n)$ определится соотношением:

$$\begin{aligned} &h_k(N-1-n) = e^{-j\omega_k n}; \\ &h_k(n) = e^{j\omega_k n} e^{-j\omega_k (N-1)} = e^{j\omega_k n} e^{j\omega_k}, \quad 0 \le n \le N-1. \end{aligned} \tag{5.20}$$

Таким образом, цифровой фильтр – это КИХ-фильтр (комплексный), передаточная функция которого имеет вид

$$\begin{aligned} H_k(z) &= \sum_{n=0}^{N-1} h_k(n) z^{-n} = e^{j\omega_k} \sum_{n=0}^{N-1} (e^{j\omega_k} z^{-1})^n = \\ &= e^{j\omega_k} \frac{1 - e^{j\omega_k N} z^{-N}}{1 - e^{j\omega_k} z^{-1}} = (1 - z^{-N}) \frac{e^{j\omega_k}}{1 - e^{j\omega_k} z^{-1}}. \end{aligned} \tag{5.21}$$

Эта структура (5.21) уже встречалась – один канал фильтра с частотной выборкой. Для вычисления всех значений $X(k)$ $(k = 0,1,\ldots,N-1)$ необходимо N таких фильтров – частотных каналов анализатора спектра. Каждый частотный канал анализатора спектра представляет собой комплексный резонатор – когерентный накопитель отсчетов комплексной гармоники $e^{j\omega_k n}$. На рисунке 5.10 представлена структурная схема анализатора спектра в виде гребенки фильтров. Достоинством такого выполнения является возможность вычисления «скользящего» $X_n(k)$ спектра по N предшествующим текущему моменту n отсчетам: $x(n)$, $x(n-1)$, $x(n-2)$, …, $x(n-N+1)$. Заметим, что при скользящем анализе на основе БПФ для каждого нового входного отсчета приходится вычислять полное БПФ, т.е. $\frac{N}{2}\log_2 N$ базовых операций на отсчет. При анализе с помощью гребенки фильтров (см. рис. 5.10) число комплексных умножений на отсчет равно N. Таким образом, в большинстве представляющих интерес случаев гребенка фильтров оказывается эффективнее скользящего БПФ.

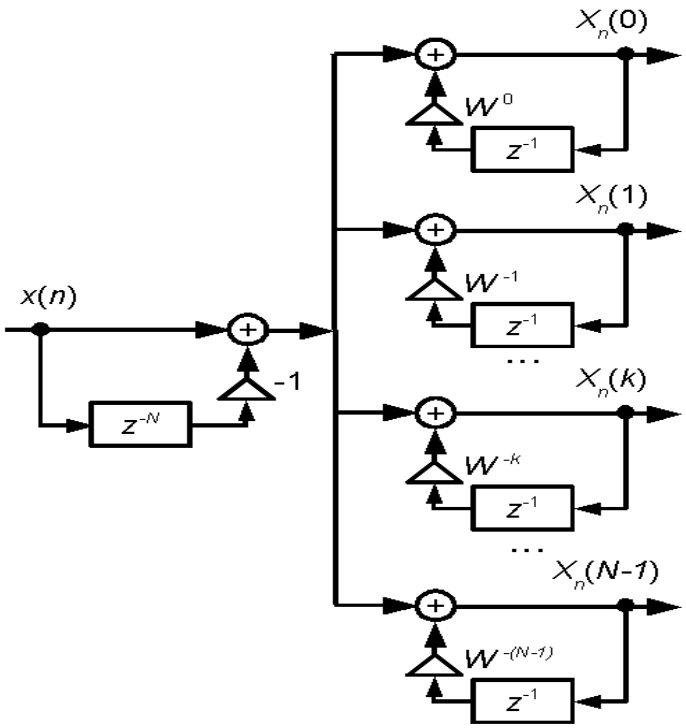


Рис. 5.10. Анализатор спектра в виде гребенки фильтров

Определим амплитудно-частотную характеристику одного канала анализатора спектра, вычислив $H_k(z)$ на единичной окружности:

$$\left|H_k(e^{j\omega})\right| = \left|\frac{1-e^{-jN\omega}}{1-e^{-j(\omega-\omega_k)}}\right| = \left|\frac{1-e^{-jN(\omega-\omega_k)}}{1-e^{-j(\omega-\omega_k)}}\right| = \left|e^{-j\frac{N-1}{2}(\omega-\omega_k)}\,\frac{\sin\frac{N(\omega-\omega_k)}{2}}{\sin\frac{(\omega-\omega_k)}{2}}\right| =$$

$$= \left|\frac{\sin\frac{N(\omega-\omega_k)}{2}}{\sin\frac{(\omega-\omega_k)}{2}}\right| = \left|D_N(\omega-\omega_k)\right|, \quad \text{где} \quad D_N(\omega) = \frac{\sin\frac{N(\omega-\omega_k)}{2}}{\sin\frac{(\omega-\omega_k)}{2}}.$$

На рис. 5.11 изображены графики функций $D_N(\omega\text{-}\omega_k)$ для 16 –точечного ДПФ, реализованного гребенкой фильтров в соответствии с рис. 5.10. Частотные характеристики фильтров с четными номерами *k* показаны на рис. 5.11 вверху, а с нечетными номерами *k* – внизу, при этом для всех фильтров, за исключением 8-го, изображены лишь главные лепестки. Видно, что скользящее ДПФ эквивалентно довольно грубому набору фильтров с АЧХ, имеющими относительно большие боковые лепестки и существенным перекрытием между соседними фильтрами. При поступлении на вход анализатора спектра комплексной гармоники с частотой $\omega = \omega_k = 2\pi k / N$, все отсчеты $X(i) \equiv 0$, $(i \neq k)$, а $X(k) = N$. Если же $\omega \neq 2\pi k/N$, то отличны от нуля все $X(i)$, т.е.

происходит «растекание» (просачивание) энергии входной гармоники по всем *N* спектральным отсчетам. Чтобы улучшить частотную характеристику канальных фильтров анализатора можно ввести весовую обработку входной последовательности *x(n)* умножением ее на весовую функцию (окно) *w(n)*. Аналогично проделанной выше процедуре можно показать, что АЧХ канальных фильтров с весовой обработкой совпадает с амплитудно-частотным спектром окна, сдвинутого на частоту $\omega_к$. Выбор функций окна представляет собой отдельную проблему.

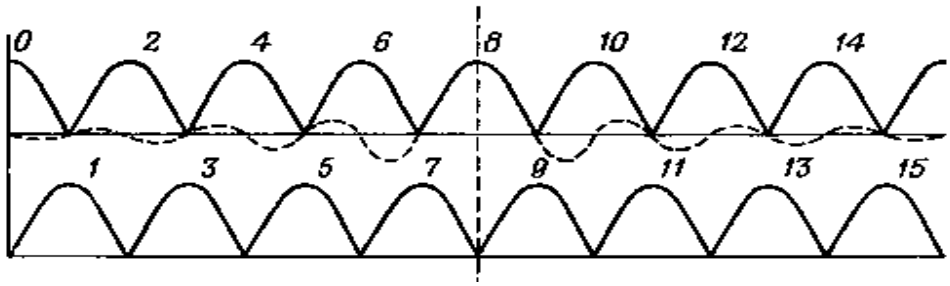


Рис. 5.11. АЧХ гребенки фильтров, эквивалентной скользящему ДПФ с размером $N = 16$

На основе цифровых фильтров предложен более экономный по числу операций, чем прямой алгоритм вычисления *X(k)*, получивший название *алгоритма Герцеля* [3]. Заметим, что в этих алгоритмах нет необходимости вычислять все *N* значений *X(k)*, т. е. в общем случае можно вычислять *X(k)* для любых *M* значений *k*. Поэтому, при $M < log_2 N$ прямой метод или метод Герцеля могут быть наиболее эффективными методами вычисления ДПФ.

СПИСОК СОКРАЩЕНИЙ

БПФ – быстрое преобразование Фурье
ВКФ – взаимно корреляционная функция
ДВПФ – преобразование Фурье дискретного времени
ДКС – дискретно-кодовый сигнал
ДФНЧ – дискретный фильтр нижних частот
ДЭФ – дискретно-экспоненциальные функции
ПФА – периодическая функция автокорреляции
РМ – разностное множество
ЦОС – цифровая обработка сигналов

ЛИТЕРАТУРА

Основная:

1. Блейхут, Р. Быстрые алгоритмы цифровой обработки сигналов / Р. Блейхут. пер. с англ. – М. : Мир, 1989. – 448 с.

2. Варакин, Л. Е. Системы связи с шумоподобными сигналами / Л. Е. Варакин. – М. : Радио и связь, 1985. – 384 с.

3. Гольденберг, Л. М. Цифровая обработка сигналов / Л. М. Гольденберг, Б. Д. Матюшкин, М. Н. Поляк. – М. : Радио и связь, 1985. – 312 с.

4. Кривошеев, В. И. Цифровая обработка сигналов: метод. указания / В. И. Кривошеев. – Нижний Новгород, 2003. – 99 с.

5. Лосев, В. В. Микропроцессорные устройства обработки информации. Алгоритмы цифровой обработки: учеб. пособие для вузов / В. В. Лосев. – Минск : Выш. шк., 1990. – 132 с.

6. Свердлик, М. Б. Оптимальные дискретные сигналы / М. Б Свердлик. – М. : Советское радио, 1975. – 200 с.

7. Трахтман, А. М., Основы теории дискретных сигналов на конечных интервалах / А. М. Трахтман, В. А Трахтман. – М. : Советское радио, 1975. – 208 с.

8. Цифровая обработка сигналов: практический подход / пер. с англ. – 2-е изд.: – М. : Вильямс, 2004 г. – 992 с.: ил.

Дополнительная:

9. Блейхут, Р. Теория и практика кодов, контролирующих ошибки / Р. Блейхут. – М. : Мир, 1986. – 514 с., ил.

10. Дэвенпорт, Г. Высшая математика. Введение в теорию чисел / Г. Дэвенпорт. – М. : Наука, 1965. – 165 с.

11. Френкс, Л. Теория сигналов / Л. Френкс; пер. с англ. – М. : Советское радио, 1974. – 344 с.

12. Айерлэнд, К. Классическое введение в современную теорию чисел / К. Айерлэнд, М. Роузен; пер. с англ. – М. : Мир, 1987. – 416 с.

СОДЕРЖАНИЕ

Учебное издание

МАЛЬЦЕВ Сергей Васильевич

ОСНОВЫ ДИСКРЕТНОЙ МАТЕМАТИКИ

Учебно-методический комплекс
для студентов специальности 1-39 01 01 «Радиотехника»

Редактор *Т. Н. Лупенько*

Дизайн обложки *В. А. Виноградовой*

Подписано в печать 09.12.09. Формат 60x84 1/16. Гарнитура Таймс. Бумага офсетная.
Ризография. Усл. печ. л. 5,56. Уч.-изд. л. 4,93. Тираж 35 экз. Заказ 2163.

Издатель и полиграфическое исполнение:
учреждение образования «Полоцкий государственный университет»

ЛИ № 02330/0548568 от 26.06.2009 ЛП № 02330/0494256 от 27.05.2009

211440 г. Новополоцк, ул. Блохина, 29