

УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ ГОРНЫЙ УНИВЕРСИТЕТ
ИНСТИТУТ ГЕОЛОГИИ И ГЕОФИЗИКИ
Кафедра геоинформатики

А.В. Давыдов

СИГНАЛЫ и ЛИНЕЙНЫЕ СИСТЕМЫ

ТЕМАТИЧЕСКИЕ ЛЕКЦИИ

Давыдов А.В. Сигналы и линейные системы: Тематические лекции. – Екатеринбург: УГГУ, ИГиГ, кафедра геоинформатики. Фонд электронных документов, 2005.

Тематические лекции представляют собой учебное пособие по дисциплине "Теория сигналов и систем", которая преподается для студентов УГГУ по специальности "Информационные системы в технике и технологиях", специализации "Прикладная геоинформатика в геофизике" и "Геоинформационные системы в природо- и недропользовании". Материал лекций соответствует общим требованиям Государственного образовательного стандарта высшего профессионального образования подготовки дипломированного специалиста по направлению "Информационные системы".

Тематика лекций излагается по принципу самостоятельности для изучения каждого раздела в отдельности, что приводит к повторению отдельных положений, но удобно для освоения материала. Они могут использоваться также студентами и инженерами радиотехнических специальностей для пополнения знаний в области сигналов и линейных систем их цифровой обработки. Лекции сопровождаются практикумом на персональных компьютерах в среде Mathcad по моделированию сигналов, их преобразованиям и обработке.

Темы лекций

1. Введение в теорию сигналов и систем
2. Пространство и метрология сигналов
3. Динамическое представление сигналов
4. Спектральное представление сигналов
5. Дискретизация сигналов
6. Дискретные преобразования сигналов
7. Энергетические спектры сигналов
8. Корреляция сигналов
9. Модулированные сигналы
10. Преобразование Гильберта
- 10b. Аналитические сигналы
11. Линейные системы
12. Передача сигналов по кабелям
13. Каротажные кабели
14. Импульсные параметры кабеля
15. Повышение скорости передачи сигналов по кабелю
16. Синхронизация кода Манчестер-II
17. Случайные сигналы
18. Многомерные сигналы и системы

Литература

ВВЕДЕНИЕ

Курсы теории сигналов и систем в настоящее время занимают одно из центральных мест среди дисциплин профессиональной подготовки не только радиоинженеров - разработчиков радиотехнических систем самого различного назначения, но и всех специальностей, в той или иной мере связанных с регистрацией, обращением, обработкой и использованием информационных данных самой различной природы – пользователей систем. Это определяется тем, что информация, наряду с материей и энергией, принадлежит к фундаментальным философским категориям естествознания и является одной из движущих сил современного развития науки, техники и человеческой цивилизации в целом. Но информация не относится к числу материальных объектов и не существует в явной физической форме. Носителями информации являются сигналы в любой форме их материального представления в пределах систем, вне которых понятия сигналов также не имеют смысла. Все это и приводит к тому, что профессионально грамотная и эффективная регистрация информации, ее обработка, интерпретация и использование возможны только при хороших знаниях теории сигналов и систем.

Настоящий курс предназначен для студентов и специалистов, занимающихся как использованием информационных систем в геофизике, в природо- и недропользовании, так и геоинформатикой – регистрацией, обработкой, интерпретацией и представлением геолого-геофизических данных. По образному выражению Козьмы Пруткова, "природа не злонамеренна, но коварна". И непредсказуема, даже в сходных геологических условиях. Обработка и интерпретация геолого-геофизических данных в этих условиях – высокое искусство моделирования на основе, как правило, достаточно ограниченных данных

объективно однозначной модели среды, математический отклик которой по физическим параметрам должен быть максимально приближен к результатам исследований. Это не значит, что для каждого геологического объекта должна создаваться индивидуальная система обработки результатов его исследований. Инструменты и технологии обработки геолого-геофизических данных хорошо известны и в определенной мере стандартизованы. Тем не менее, результаты интерпретации во многом определяются опытом и квалификацией специалистов, выбирающих и применяющих эти инструменты (системы) и технологии с учетом особенностей и специфики исходных информационных данных (сигналов).

Темы настоящего курса функционально замкнуты и содержат теоретический материал, необходимый и достаточный для их изучения без ссылок на предыдущие темы. В отдельных случаях это создает некоторое дублирование материала, но компенсируется удобством работы. С этой же целью каждая тема содержит начальное оглавление по разделам и основное содержание разделов. Каждый раздел обычно содержит ссылку на литературу, которая использовалась при подготовке данных разделов, и в которой читатель, при необходимости, может найти более подробные сведения по данным вопросам. Предполагается знакомство читателей с курсами "Высшая математика", "Физика" и "Информатика" в объеме технических вузов, по крайней мере, на элементарном уровне, а фундаментальные понятия, как правило, подробно поясняются иллюстрациями. Большинство приведенных формул пригодны для непосредственного использования в системе Mathcad, в которой выполнены и все иллюстрации пособия. Самостоятельное изучение материала рекомендуется сопровождать практическими занятиями в системе Mathcad, программные документы – инструкции по которым имеются на сайте автора.

Главный сайт автора ~ Лекции по сигналам ~ Практикум

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru.

Copyright ©2005 Davydov A.V.



Тема 1: ВВЕДЕНИЕ В ТЕОРИЮ СИГНАЛОВ И СИСТЕМ

Одна из основных задач теории в любой области знаний – найти позицию, с которой объект виден в предельной простоте.

Джосайя Уиллард Гиббс. Американский физик, XIX в.

Для понимания истин, предельно простых для теоретиков, нормальному инженеру требуется специальная подготовка.

Роберт Тимофеевич Шарло. Уральский геофизик, XX в.

Содержание:

1.1. Общие сведения и понятия. Понятие сигнала. Шумы и помехи. Размерность сигналов. Математическое описание сигналов. Математические модели сигналов. Виды моделей. Классификация сигналов.

1.2. Типы сигналов. Аналоговый сигнал. Дискретный сигнал. Цифровой сигнал. Преобразования типа сигналов. Спектральное представление сигналов. Графическое отображение сигналов. Тестовые сигналы.

1.3. Системы преобразования сигналов. Общее понятие систем. Основные системные операции. Линейные системы.

1.4. Информационная емкость сигналов. Понятие информации. Количественная мера информации. Энтропия источника информации. Основные свойства энтропии Энтропия непрерывного источника информации. Информационная емкость сигналов.

Литература.

1.1. ОБЩИЕ СВЕДЕНИЯ И ПОНЯТИЯ [1,10, 15, 25]

Понятие сигнала. В XVIII веке в теорию математики вошло понятие функции, как определенной зависимости какой-либо величины y от другой величины – независимой переменной x , с математической записью такой зависимости в виде $y(x)$. Довольно скоро математика функций стала базовой основой теории всех естественных и технических наук. Особое значение функциональная математика приобрела в технике связи, где временные функции вида $s(t)$, $v(f)$ и т.п., используемые для передачи информации, стали называть сигналами.

В технических отраслях знаний термин "сигнал" (signal, от латинского signum – знак) очень часто используется в широком смысловом диапазоне, без соблюдения строгой терминологии. Под ним понимают и техническое средство для передачи, обращения и использования информации - электрический, магнитный, оптический сигнал; и физический процесс, представляющий собой материальное воплощение информационного сообщения - изменение какого-либо параметра носителя информации (напряжения, частоты, мощности электромагнитных колебаний, интенсивности светового потока и т.п.) во времени, в пространстве или в зависимости от изменения значений каких-либо других аргументов (независимых переменных); и смысловое содержание определенного физического состояния или процесса, как, например, сигналы светофора, звуковые предупреждающие сигналы и т.п. Все эти понятия объединяет конечное назначение сигналов. Это определенные сведения, сообщения, информация о каких-либо процессах, состояниях или физических величинах объектов материального мира, выраженные в форме, удобной для передачи, обработки, хранения и использования этих сведений.

Термин "сигнал" очень часто отождествляют с понятиями "данные" (data) и "информация" (information). Действительно, эти понятия взаимосвязаны и не существуют одно без другого, но относятся к разным категориям.

Понятие информации имеет много определений, от наиболее широкого (информация есть формализованное отражение реального мира) до практического (сведения и данные, являющиеся объектом хранения, передачи, преобразования, восприятия и управления). В настоящее время мировая наука все больше склоняется к точке зрения, что информация, наряду с материей и энергией, принадлежит к фундаментальным философским категориям естествознания и отнести к одному из свойств объективного мира, хотя и несколько специфичному. Что касается "данных" (от латинского datum – факт), то это совокупность фактов, ре-

зультатов наблюдений, измерений о каких-либо объектах, явлениях или процессах материального мира, представленных в формализованном виде, количественном или качественном. Это не информация, а только атрибут информации - сырье для получения информации путем соответствующей обработки и интерпретации (истолкования).

Наука и техника интернациональны, и используют, в основном, общепринятые термины, большинство из которых англоязычны. Термин "signal" в мировой практике является общепринятым для характеристики формы представления данных, при которой данные рассматриваются как результат некоторых измерений объекта исследований в виде последовательности значений скалярных величин (аналоговых, числовых, графических и пр.) в зависимости от изменения каких-либо переменных значений (времени, энергии, температуры, пространственных координат, и пр.). С учетом этого, в дальнейшем под термином "сигнал" в узком смысле этого слова будем понимать каким-либо образом упорядоченное *отображение* в изменении физического состояния какого-либо объекта – материального носителя сигнала, определенных данных о характере изменения в пространстве, во времени или по любой другой переменной физических величин, физических свойств или физического состояния объекта исследований. А так как данные содержат информацию, как об основных целевых параметрах объекта исследований, так и о различных сопутствующих и мешающих факторах измерений, то в широком смысле этого слова можно считать, что сигнал является носителем общей измерительной информации. При этом материальная форма носителей сигналов (механическая, электрическая, магнитная, акустическая, оптическая и любая другая), равно как и форма отображения в каких-либо физических параметрах или процессах носителей, значения не имеет. Информативным параметром сигнала может являться любой параметр носителя сигнала, функционально связанный со значениями информационных данных.

Сигнал, в самом общем смысле, это зависимость одной величины от другой, и с математической точки зрения представляет собой функцию. Наиболее распространенное представление сигналов - в электрической форме в виде зависимости напряжения от времени $U(t)$. Так, например, сигнал изменения напряженности магнитного поля по профилю аэро съемки – это и временная последовательность изменения электрического напряжения на выходе датчика аэромагнитометра, и запись этого напряжения на ленте регистратора, и последовательные значения цифровых отсчетов при обработке лент регистратора и вводе сигнала в ЭВМ.

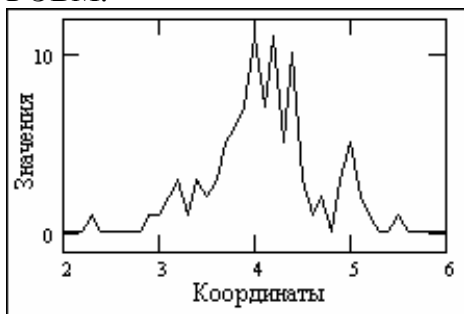


Рис. 1.1.1. Сигнал.

Сигнал - это информационная функция, несущая сообщение о физических свойствах, состоянии или поведении какой-либо физической системы, объекта или среды, а целью обработки сигналов можно считать извлечение определенных информационных сведений, которые отображены в этих сигналах (кратко - полезная или целевая информация) и преобразование этих сведений в форму, удобную для восприятия и дальнейшего использования.

Под "анализом" сигналов (analysis) имеется в виду не только их чисто математические преобразования, но и получение на основе этих преобразований выводов о специфических особенностях соответствующих процессов и объектов. Целями анализа сигналов обычно являются:

- Определение или оценка числовых параметров сигналов (энергия, средняя мощность, среднее квадратическое значение и пр.).
- Разложение сигналов на элементарные составляющие для сравнения свойств различных сигналов.
- Сравнение степени близости, "похожести", "родственности" различных сигналов, в том числе с определенными количественными оценками.

Математический аппарат анализа сигналов весьма обширен, и широко применяется на практике во всех без исключения областях науки и техники.

С понятием сигнала неразрывно связан термин *регистрации* сигналов, использование которого также широко и неоднозначно, как и самого термина сигнал. В наиболее общем смысле под этим термином можно понимать операцию выделения сигнала и его преобразования в форму, удобную для дальнейшего использования. Так, при получении информации о физических свойствах каких-либо объектов, под регистрацией сигнала понимают процесс измерения физических свойств объекта и перенос результатов измерения на материальный носитель сигнала или непосредственное энергетическое преобразование каких-либо свойств объекта в информационные параметры материального носителя сигнала (как правило - электрического). Но так же широко термин регистрации сигналов используют и для процессов выделения уже сформированных сигналов, несущих определенную информацию, из суммы других сигналов (радиосвязь, телеметрия и пр.), и для процессов фиксирования сигналов на носителях долговременной памяти, и для многих других процессов, связанных с обработкой сигналов.

Применительно к настоящему курсу под термином регистрации будем понимать *регистрацию данных* (data logging) которые проходят через конкретную систему или точку системы и определенным образом фиксируются на каком-либо материальном носителе или в памяти системы. Что касается процесса получения информации при помощи технических средств, обеспечивающих опытным путем нахождение соотношения измеряемой величины с принятой по определению образцовой единицей этой величины, и представление измеренного соотношения в какой-либо физической или числовой форме информационного сигнала, то для этого процесса будем применять, в основном, термин *детектирования*.

Шумы и помехи (noise). При детектировании сигналов, несущих целевую для данного вида измерений информацию, в сумме с основным сигналом одновременно регистрируются и мешающие сигналы - шумы и помехи самой различной природы (рис. 1.1.2). К помехам относят также искажения полезных сигналов при влиянии различных дестабилизирующих факторов на процессы измерений, как, например, влияние микрокаверн в стенках скважины на измерения в рентгенорадиометрических методах каротажа, грозовых разрядов на электро-разведочные методы измерений и т.п. Выделение полезных составляющих из общей суммы зарегистрированных сигналов или максимальное подавление шумов и помех в информационном сигнале при сохранении его полезных составляющих является одной из основных задач первичной обработки сигналов (результатов наблюдений).



Типы помех разделяют по источникам их возникновения, по энергетическому спектру, по характеру воздействия на сигнал, по вероятностным характеристикам и другим признакам.

Источники помех бывают внутренние и внешние.

Внутренние шумы могут быть присущи физической природе источников сигналов, как, например, тепловые шумы электронных потоков в электрических цепях или дробовые эффекты в электронных приборах, или возникают в измерительных устройствах и системах передачи и обработки сигналов от влияния различных дестабилизирующих факторов - температуры, повышенной влажности, нестабильности источников питания, влияния механических вибраций на гальванические соединения, и т.п.

Внешние источники шумов бывают искусственного и естественного происхождения. К искусственным источникам помех относятся промышленные помехи - двигатели, переключатели, генераторы сигналов различной формы и т.д. Естественными источниками помех являются молнии, флуктуации магнитных полей, всплески солнечной энергии, и т.д.

Электрические и магнитные поля различных источников помех вследствие наличия индуктивных, емкостных и резистивных связей создают на различных участках и цепях сигнальных систем паразитные разности потенциалов и токи, накладывающиеся на полезные сигналы.

Помехи подразделяются на флуктуационные, импульсные и периодические. Флук-

туационные или шумовые помехи представляют хаотический и беспорядочный во времени процесс в виде нерегулярных случайных всплесков различной амплитуды. Как правило, флюктуационные помехи распределены по нормальному закону с нулевым средним и оказывают существенное влияние только на сигналы низкого уровня.

Импульсные помехи во многом похожи на шумовые помехи и проявляются как в виде отдельных импульсов, так и в виде последовательности импульсов, форма и параметры которых имеют случайный характер. Причинами импульсных помех являются резкие броски тока и напряжения в промышленных установках, транспортных средствах, а также природные электрические явления. Распределение импульсных помех симметричное с произвольной плотностью распределения.

Периодические помехи вызываются периодическими низкочастотными или высокочастотными полями линий электропередач, силовых электроустановок и др. Если основная мощность помех сосредоточена на отдельных участках диапазона частот, например, на частоте напряжения промышленной сети или кратна этой частоте, то такие помехи называют сосредоточенными.

В зависимости от характера воздействия на сигнал помехи разделяют на аддитивные и мультипликативные. Аддитивные (налагающиеся) помехи суммируются с сигналом, не зависят от его значений и формы и не изменяют информативной составляющей самого сигнала. Мультипликативные или деформирующие помехи могут изменять форму информационной части сигнала, иметь зависимость от его значений и от определенных особенностей в сигнале и т.п. При известном характере мультипликативных помех возможна коррекция сигнала на их влияние.

Следует заметить, что деление сигналов на полезные и мешающие (шумовые) является достаточно условным. Источниками мешающих сигналов также являются определенные физические процессы, явления или объекты. При выяснении природы мешающих сигналов они могут переводиться в разряд информационных. Так, например, вариации диаметра скважин является мешающим фактором практически для всех ядерно-физических методов каротажа. Вместе с тем этот же фактор, при соответствующем методическом и аппаратном обеспечении, может дать возможность бесконтактного определения диаметра скважин в качестве дополнительного информационного параметра.

Размерность сигналов. Простейшими сигналами геофизической практики являются одномерные сигналы, как, например, сейсмические импульсы $s(t)$, измерения каких-либо параметров геофизических полей (электрических, магнитных, и пр.) по профилям на поверхности земли $s(x)$ или по стволу скважины $s(h)$, и т.п. Значения одномерных сигналов зависят только от одной независимой переменной, как, например, на рис. 1.1.1 и 1.1.2.

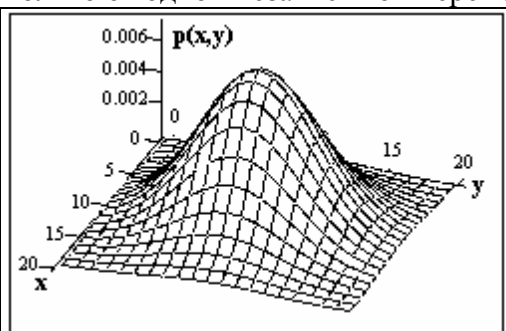


Рис. 1.1.3. Двумерный сигнал.

В общем случае сигналы являются многомерными функциями пространственных, временных и прочих независимых переменных - сейсмическая волна вдоль линии профиля $s(x,t)$, аномалия гравитационного поля на поверхности наблюдений $s(x,y)$, пространственно - энергетическое распределение потока ионизирующих частиц или квантов от источника излучения $s(x,y,z,E)$ и т.п. Все большее применение находят также многомерные сигналы, образованные некоторым множеством одномерных сигналов, как, на-

пример, комплексные каротажные измерения нескольких физических параметров горных пород по стволу скважины одновременно.

Многомерные сигналы могут иметь различное представление по своим аргументам. Так, полный акустический сигнал сейсмического профиля дискретен по пространству (точкам расположения приемников) и непрерывен по времени.

Многомерный сигнал может рассматриваться, как упорядоченная совокупность одномерных сигналов. С учетом этого при анализе и обработке сигналов многие принципы и

практические методы обработки одномерных сигналов, математический аппарат которых развит достаточно глубоко, распространяются и на многомерные сигналы. Физическая природа сигналов для математического аппарата их обработки значения не имеет.

Вместе с тем обработка многомерных сигналов имеет свои особенности, и может существенно отличаться от одномерных сигналов в силу большего числа степеней свободы. Так, при дискретизации многомерных сигналов имеет значение не только частотный спектр сигналов, но и форма раstra дискретизации. Пример не очень полезной особенности - многомерные полиномы сигнальных функций, в отличие от одномерных, не разлагаются на простые множители. Что касается порядка размерности многомерных сигналов, то ее увеличение выше двух практически не изменяет принципы и методы анализа данных, и сказывается, в основном, только на степени громоздкости формул и чисто техническом усложнении вычислений.

Учитывая эти факторы, при рассмотрении общей теории анализа, преобразований и обработки сигналов ограничимся, в основном, одно- и двумерными сигнальными функциями, а в качестве универсальных независимых переменных (аргументов функций) будем использовать, как правило, переменную "t" для одномерных сигналов и переменные "x,t" или "x,y" для двумерных сигналов, безотносительно к их физическому содержанию (пространство, время, энергия и пр.).

Математическое описание сигналов. Сигналы могут быть объектами теоретических исследований и практического анализа только в том случае, если указан способ их математического описания - математическая модель сигнала. Математическое описание позволяет абстрагироваться от физической природы сигнала и материальной формы его носителя, проводить классификацию сигналов, выполнять их сравнение, устанавливать степень тождества, моделировать системы обработки сигналов. Как правило, описание сигнала задается функциональной зависимостью определенного информационного параметра сигнала от независимой переменной (аргумента) – $s(x)$, $y(t)$ и т.п. Функции математического описания сигналов могут быть как вещественными, так и комплексными. Выбор математического аппарата описания определяется простотой и удобством его использования при анализе и обработке сигналов.

Отметим двойственность применения описания сигналов функциями типа $s(t)$ и т.п. С одной стороны $s(t)$ – это величина, равная значению функции в момент времени t . С другой стороны мы обозначаем также через $s(t)$ и саму функцию, т.е. то правило, по которому каждому значению t ставится в соответствие определенная величина s . В большинстве аналитических выражений это не вызывает недоразумений и при однозначном соответствии значений сигналов их аналитическим выражениям принимается по умолчанию.

Сделаем также одно замечание по терминологии описания сигналов. В теоретических работах по анализу сигналов конкретные значения величины сигнала (отсчеты значений по аргументу) часто именуют координатами сигнала. В отраслях знаний, связанных с геологией и горным делом, и в геофизической практике в том числе, этот термин используется по своему прямому смысловому назначению – пространственных координат результатов измерений, и является неизменным атрибутом всех геолого-геофизических данных. С учетом последнего фактора условимся применять термин “координата” по своему традиционному смысловому назначению в качестве обобщающего термина для независимых переменных сигнальных функций. При этом под понятием *координат* значений сигнала будем понимать не только какие-либо пространственные координаты, как это непосредственно имеет место для результатов измерений при геолого-геофизических съемках, но и любые другие аргументы, на числовой оси которых отложены значения или отсчеты сигнала и рассматривается динамика его изменения (пример на рис. 1.1.1).

Математические модели сигналов. Теория анализа и обработки физических данных базируется на математических моделях соответствующих физических полей и физических процессов, на основе которых создаются математические модели сигналов. Математические модели сигналов дают возможность обобщенно, абстрагируясь от физической природы, су-

дить о свойствах сигналов, предсказывать изменения сигналов в изменяющихся условиях, заменять физическое моделирование изучаемых процессов математическим. С помощью математических моделей имеется возможность описывать свойства сигналов, которые являются главными, определяющими в изучаемых процессах, и игнорировать большое число второстепенных признаков. Знание математических моделей сигналов дает возможность классифицировать их по различным признакам, характерным для того или иного типа моделей. Так, сигналы разделяются на неслучайные и случайные в зависимости от возможности точного предсказания их значений в любые моменты времени. Сигнал является неслучайным и называется *детерминированным*, если математическая модель позволяет осуществлять такое предсказание. Детерминированный сигнал задается, как правило, математической функцией или вычислительным алгоритмом, а математическая модель сигнала может быть представлена в виде

$$s = F(t, z, \omega, \dots; A, B, C, \dots),$$

где s – информативный параметр сигнала; t, z, ω, \dots – независимые аргументы (время, пространственная координата, частота и др.); A, B, C, \dots – параметры сигналов.

Модель должна быть, по возможности, проще и минимизирована по количеству независимых аргументов и адекватна изучаемому процессу, что во многом предопределяет результаты измерений. Рассмотрим этот вопрос на примере геофизических данных.

Под геофизическим полем понимают собственное или индуцированное определенным внешним воздействием распределение какой-либо физической величины, создаваемое геологическим объектом или геологической структурой в пространстве, во времени или по любому другому аргументу (независимой переменной). В простейшем случае геофизический сигнал – это изменение какой-либо составляющей геофизического поля, т.е. сечение поля по одному из аргументов. В пределе все геофизическое поле в целом может рассматриваться как первичный многомерный сигнал в прямом физическом отображении, с которого путем измерений могут сниматься формализованные копии определенных составляющих (сечений) сигнала на материальные носители информации.

Геофизическим полям в определенных условиях их регистрации соответствуют определенные математические модели сигналов, т.е. их описание на каком-либо формальном языке. Математическое описание не может быть всеобъемлющим и идеально точным и, по существу, всегда отображает не реальные объекты, а их упрощенные (гомоморфные) модели. Модели могут задаваться таблицами, графиками, функциональными зависимостями, уравнениями состояний и переходов из одного состояния в другое и т.п. Формализованное описание может считаться математической моделью оригинала, если оно позволяет с определенной точностью прогнозировать состояние и поведение изучаемых объектов путем формальных процедур над их описанием.

Неотъемлемой частью любой математической модели сигнала является также область определения сигнала, которая устанавливается интервалом задания независимой переменной. Примеры задания интервала для переменных:

$$a \leq x \leq b, \quad x \in [a, b].$$

$$a < y \leq b, \quad y \in (a, b].$$

$$a < z < b, \quad z \in (a, b).$$

Пространство значений независимой переменной обычно обозначается через индекс R . Так, например, $R := (-\infty, +\infty)$, $x \in R$.

Кроме задания области определения сигнала могут быть также заданы виды численных значений переменных (целые, рациональные, вещественные, комплексные).

Математические модели полей и сигналов на первом этапе обработки и анализа результатов наблюдений должны позволять в какой-то мере игнорировать их физическую природу и возвращать ее в модель только на заключительном этапе интерпретации данных.

Виды моделей сигналов. При анализе физических данных используются два основных подхода к созданию математических моделей сигналов.

Первый подход оперирует с *детерминированными* сигналами, значения которых в любой момент времени или в произвольной точке пространства (а равно и в зависимости от любых других аргументов) являются априорно известными или могут быть достаточно точно определены (вычислены). Такой подход удобен в прямых задачах геофизики (расчеты полей для заданных моделей сред), в задачах активных воздействий на среду при заранее известных параметрах и форме сигнала воздействия (вибрационная сейсморазведка, электромагнитные методы каротажа и пр.), а также при использовании хорошо известных и достоверных геолого-геофизических данных. Для описания неслучайных сигналов используются также *квазидетерминированные* модели, в которых значения одного или нескольких параметров априорно неизвестны и считаются случайными величинами с малой случайной компонентой, влиянием которой можно пренебречь.

Второй подход предполагает случайный характер сигналов, закон изменения которых во времени (или в пространстве) носит случайный характер и которые принимают конкретные значения с некоторой вероятностью. Модель такого сигнала представляет собой описание статистических характеристик случайного процесса путем задания законов распределения вероятностей, корреляционной функции, спектральной плотности энергии и др.

Случайность может быть обусловлена как собственной физической природой сигналов, что характерно, например, для методов ядерной геофизики, так и вероятностным характером регистрируемых сигналов как по времени или месту их появления, так и по содержанию. С этих позиций случайный сигнал может рассматриваться как отображение случайного по своей природе процесса или физических свойств объекта (процесса), которые определяются случайными параметрами или сложным строением геологической среды, результаты измерений в которой трудно предсказуемы.

Между этими двумя видами сигналов нет резкой границы. Строго говоря, детерминированных процессов и отвечающих им детерминированных сигналов в природе не существует. Даже сигналы, хорошо известные на входе в среду (при внешнем воздействии на нее), по месту их регистрации всегда осложнены случайными помехами, влиянием дестабилизирующих факторов и априорно неизвестными параметрами и строением самой среды. С другой стороны, модель случайного поля часто аппроксимируется методом суперпозиции (сложения) сигналов известной формы. Детерминированные модели могут использоваться и для изучения чисто случайных процессов, если уровень полезного сигнала в этом процессе значительно выше уровня статистических флуктуаций, что имеет место, например, при регистрации ионизирующих излучений от горных пород.

На выбор математической модели поля в том или ином методе геофизики в немалой степени влияет также сложность математического аппарата обработки сигналов и сложившиеся традиции геологической интерпретации результатов наблюдений. Не исключается и изменение модели, как правило, с переводом из вероятностной в детерминированную, в процессе накопления информации об изучаемом явлении или объекте.



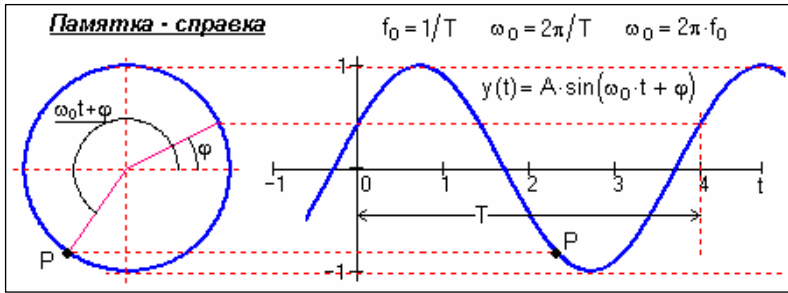
Рис. 1.1.4. Классификация сигналов.

Классификация сигналов осуществляется на основании существенных признаков соответствующих математических моделей сигналов. Все сигналы разделяют на две крупных группы: детерминированные и случайные (рис. 1.1.4).

Классификация детерминированных сигналов. Обычно выделяют два класса детерминированных сигналов: периодические и непериодические.

К **периодическим** относят гармонические и полигармонические сигналы. Для периодических сигналов выполняется общее условие $s(t) = s(t + kT)$, где $k = 1, 2, 3, \dots$ - любое целое

число, T - период, являющийся конечным отрезком времени.



Гармонические сигналы (или синусоидальные), описываются следующими формулами:

$$s(t) = A \cdot \sin(2\pi f_0 t + \varphi) = A \cdot \sin(\omega_0 t + \varphi), \quad (1.1.1)$$

или:

$$s(t) = A \cdot \cos(\omega_0 t + \varphi),$$

где A , f_0 , ω_0 , φ , ϕ - постоянные

величины, которые могут исполнять роль информационных параметров сигнала: A - амплитуда сигнала, f_0 - циклическая частота в герцах, $\omega_0 = 2\pi f_0$ - угловая частота в радианах, φ и ϕ - начальные фазовые углы в радианах. Период одного колебания $T = 1/f_0 = 2\pi/\omega_0$. При $\varphi = \phi - \pi/2$ синусные и косинусные функции описывают один и тот же сигнал. Частотный спектр сигнала представлен амплитудным и начальным фазовым значением частоты f_0 (при $t = 0$).

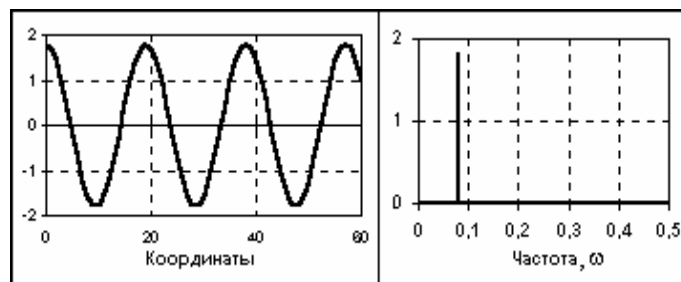


Рис. 1.1.5. Гармонический сигнал и спектр его амплитуд.

Полигармонические сигналы составляют наиболее широко распространенную группу периодических сигналов и описываются суммой гармонических колебаний:

$$s(t) = \sum_{n=0}^N A_n \sin(2\pi f_n t + \varphi_n), \quad (1.1.2)$$

или непосредственно функцией $s(t) = y(t \pm kT_p)$, $k = 1, 2, 3, \dots$, где T_p - период одного полного колебания сигнала $y(t)$, заданного на одном периоде. Значение $f_p = 1/T_p$ называют фундаментальной частотой колебаний. Полигармонические сигналы представляют собой сумму определенной постоянной составляющей ($f_0 = 0$) и произвольного (в пределе - бесконечного) числа гармонических составляющих с произвольными значениями амплитуд A_n и фаз φ_n , и с периодами, кратными периоду фундаментальной частоты f_p . Другими словами, на периоде фундаментальной частоты f_p , которая равна или кратно меньше минимальной частоты гармоник, укладывается кратное число периодов всех гармоник, что и создает периодичность повторения сигнала. Частотный спектр полигармонических сигналов дискретен, в связи с чем второе распространенное математическое представление сигналов - в виде спектров (рядов Фурье).

В качестве примера на рис. 1.1.6 приведен отрезок периодической сигнальной функции, которая получена суммированием постоянной составляющей (частота постоянной составляющей равна 0) и трех гармонических колебаний с разными значениями частоты и начальной фазы колебаний. Математическое описание сигнала задается формулой:

$$s(t) = \sum_{k=0}^3 A_k \cdot \cos(2 \cdot \pi \cdot f_k \cdot t + \varphi_k),$$

где: $A_k = \{5, 3, 4, 7\}$ - амплитуда гармоник; $f_k = \{0, 40, 80, 120\}$ - частота в герцах; $\varphi_k = \{0, -0.4, -0.6, -0.8\}$ - начальный фазовый угол колебаний в радианах; $k = 0, 1, 2, 3$. Фундаментальная частота сигнала 40 Гц.



Рис. 1.1.6. Временная модель сигнала.

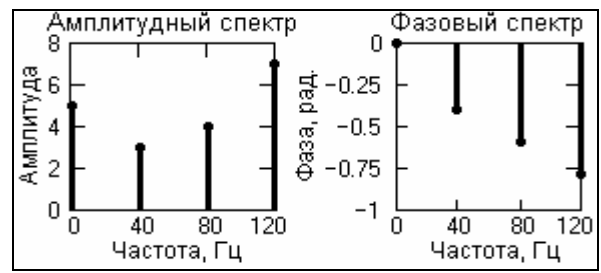


Рис. 1.1.7. Спектр сигнальной функции.

Частотное представление данного сигнала (спектр сигнала) приведено на рис. 1.1.7. Обратим внимание, что частотное представление периодического сигнала $s(t)$, ограниченного по числу гармоник спектра, составляет всего восемь отсчетов и весьма компактно по сравнению с множеством точек непрерывного временного представления.

Периодический сигнал любой произвольной формы может быть представлен в виде суммы гармонических колебаний с частотами, кратными фундаментальной частоте колебаний $f_p = 1/T_p$. Для этого достаточно разложить один период сигнала в ряд Фурье по тригонометрическим функциям синуса и косинуса с шагом по частоте, равным фундаментальной частоте колебаний $\Delta f = f_p$:

$$s(t) = \sum_{k=0}^K (a_k \cos 2\pi k \Delta f t + b_k \sin 2\pi k \Delta f t), \quad (1.1.3)$$

$$a_0 = (1/T) \int_0^T s(t) dt, \quad a_k = (2/T) \int_0^T s(t) \cos 2\pi k \Delta f t dt, \quad (1.1.4)$$

$$b_k = (2/T) \int_0^T s(t) \sin 2\pi k \Delta f t dt. \quad (1.1.5)$$

Количество членов ряда Фурье $K = k_{\max}$ обычно ограничивается максимальными частотами f_{\max} гармонических составляющих в сигналах так, чтобы $f_{\max} < K \cdot f_p$. Однако для сигналов с разрывами и скачками имеет место $f_{\max} \rightarrow \infty$, при этом количество членов ряда ограничивается по допустимой погрешности аппроксимации функции $s(t)$.

Одночастотные косинусные и синусные гармоники можно объединить и представить разложение в более компактной форме:

$$s(t) = \sum_{k=0}^K S_k \cos (2\pi k \Delta f t - \varphi_k), \quad (1.1.3')$$

$$S_k = \sqrt{a_k^2 + b_k^2}, \quad \varphi_k = \operatorname{arctg} (b_k/a_k). \quad (1.1.6)$$

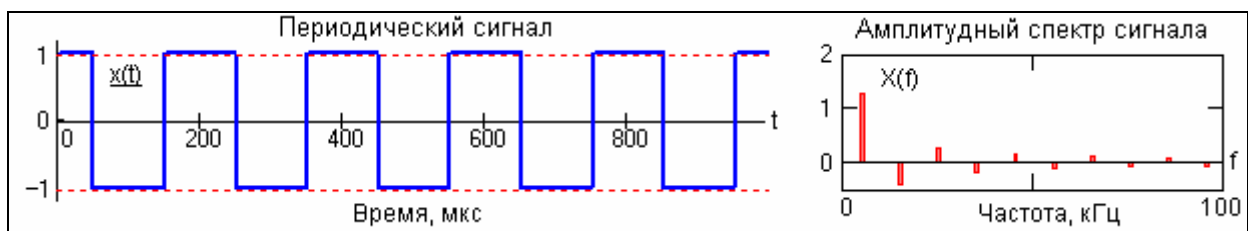


Рис. 1.1.8. Прямоугольный периодический сигнал (меандр).

Пример представления прямоугольного периодического сигнала (меандра) в виде амплитудного ряда Фурье в частотной области приведен на рис. 1.1.8 (сигнал четный относительно $t=0$, не имеет синусных гармоник, все значения φ_k для данной модели сигнала равны нулю).

Информационными параметрами полигармонического сигнала могут быть как определенные особенности формы сигнала (размах от минимума до максимума, экстремальное отклонение от среднего значения, и т.п.), так и параметры определенных гармоник в этом сигнале. Так, например, для прямоугольных импульсов информационными параметрами могут быть период повторения импульсов, длительность импульсов, скважность импульсов (от-

ношение периода к длительности). При анализе сложных периодических сигналов информационными параметрами могут также быть:

- Текущее среднее значение за определенное время, например, за время периода:

$$(1/T) \int_t^{t+T} s(t) dt.$$

- Постоянная составляющая одного периода:

$$(1/T) \int_0^T s(t) dt.$$

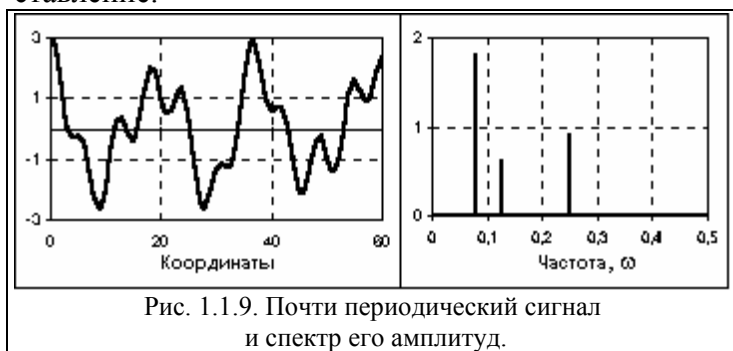
- Среднее выпрямленное значение:

$$(1/T) \int_0^T |s(t)| dt.$$

- Среднее квадратичное значение:

$$\sqrt{\frac{1}{T} \int_0^T x(t)^2 dt}.$$

К **непериодическим** сигналам относят почти периодические и аperiodические или переходные сигналы. Основным инструментом их анализа также является частотное представление.



Почти периодические сигналы

близки по своей форме к полигармоническим. Они также представляют собой сумму двух и более гармонических сигналов (в пределе – до бесконечности), но не с кратными, а с произвольными частотами, отношения которых (хотя бы двух частот минимум) не относятся к рациональным числам, вследствие чего фундаментальный период суммарных колебаний бесконечно велик. Так, например, сумма двух гармоник с частотами $2f_0$ и $3.5f_0$ дает периодический сигнал ($2/3.5$ – рациональное число) с фундаментальной частотой $0.5f_0$, на одном периоде которой будут укладываться 4 периода первой гармоники и 7 периодов второй. Но если значение частоты второй гармоники заменить близким значением $\sqrt{12} f_0$, то сигнал перейдет в разряд непериодических, поскольку отношение $2/\sqrt{12}$ не относится к числу рациональных чисел. Как правило, почти периодические сигналы порождаются физическими процессами, не связанными между собой. Математическое отображение сигналов тождественно полигармоническим сигналам (сумма гармоник), а частотный спектр также дискретен.

Апериодические сигналы составляют основную группу непериодических сигналов и задаются произвольными функциями времени. На рис. 1.1.10 показан пример аperiodического сигнала, заданного формулой на интервале $(0, \infty)$:

$$s(t) = \exp(-a \cdot t) - \exp(-b \cdot t),$$

где a и b – константы, в данном случае $a = 0.15$, $b = 0.17$.

К аperiodическим сигналам относятся также импульсные сигналы, которые в радиотехнике и в отраслях, широко ее использующих, часто рассматривают в виде отдельного класса сигналов. Импульсы представляют собой сигналы, как правило, определенной и достаточно простой формы, существующие в пределах конечных временных интервалов. Сигнал, приведенный на рис. 1.1.11, относится к числу импульсных.

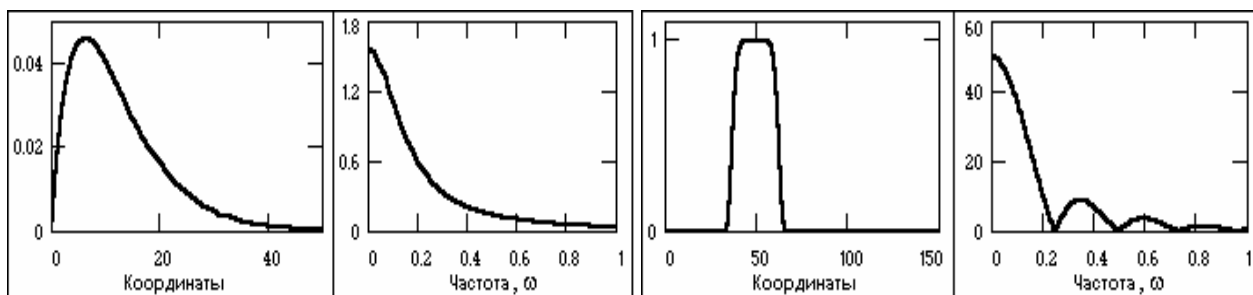


Рис. 1.1.10. Аперiodический сигнал и модуль спектра. Рис. 1.1.11. Импульсный сигнал и модуль спектра.

Частотный спектр аперiodических сигналов непрерывен и может содержать любые гармоники в частотном интервале $(0-\infty)$. Для его вычисления используется интегральное преобразование Фурье, которое можно получить переходом в формулах (1.1.3) от суммирования к интегрированию при $\Delta f \rightarrow 0$ и $k\Delta f \rightarrow f$.

$$s(t) = \int_0^{\infty} (a(f) \cos 2\pi ft + b(f) \sin 2\pi ft) df = \int_0^{\infty} S(f) \cos(2\pi ft - \varphi(f)) df. \quad (1.1.7)$$

$$a(f) = \int_0^T s(t) \cos 2\pi ft dt, \quad b(f) = \int_0^T s(t) \sin 2\pi ft dt, \quad (1.1.8)$$

$$S(f) = \sqrt{a(f)^2 + b(f)^2}, \quad \varphi(f) = \operatorname{arctg}(b(f)/a(f)). \quad (1.1.9)$$

Частотные функции $a(f)$, $b(f)$ и $S(f)$ представляют собой не амплитудные значения соответствующих гармоник на определенных частотах, а распределения спектральной плотности амплитуд этих гармоник по частотной шкале. Формулы (1.1.8-1.1.9) обычно называют формулами прямого преобразования Фурье, формулы (1.1.7) – обратного преобразования.

Если нас не интересует поведение сигнала за пределами области его задания, то эта область может восприниматься, как один период периодического сигнала и, соответственно, применяться для его представления разложение в ряды Фурье по (1.1.3-1.1.6).

В классе импульсных сигналов выделяют подкласс радиоимпульсов. Пример радиоимпульса приведен на рис. 1.1.12.

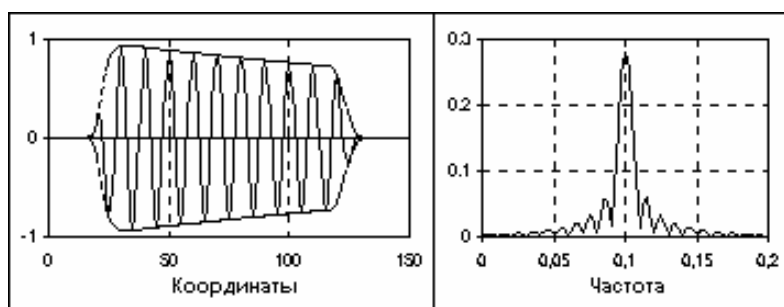


Рис. 1.1.12. Радиоимпульс и модуль его спектра.

Уравнение радиоимпульса имеет вид:

$$s(t) = u(t) \cos(2\pi f_0 t + \varphi_0).$$

где $\cos(2\pi f_0 t + \varphi_0)$ – гармоническое колебание заполнения радиоимпульса, $u(t)$ – огибающая радиоимпульса.

С энергетических позиций сигналы разделяют на два класса: с ограниченной (конечной) энергией и с бесконечной энергией.

Для сигналов с ограниченной энергией (иначе – *сигналов с интегрируемым квадратом*) должно выполняться соотношение:

$$\int_{-\infty}^{\infty} |s(t)|^2 dt < \infty.$$

Как правило, к этому классу сигналов относятся аперiodические и импульсные сигналы, не имеющие разрывов 2-го рода при ограниченном количестве разрывов 1-го рода, и особых точек, уходящих в бесконечность. Любые периодические, полигармонические и поч-

ти периодические сигналы, а также сигналы с разрывами и особыми точками 2-го рода, относятся к сигналам с бесконечной энергией. Для их анализа применяются специальные методы.

Иногда в отдельный класс выделяют сигналы конечной длительности, отличные от нуля только на ограниченном интервале аргументов (независимых переменных). Такие сигналы обычно называют *финитными*.

Классификация случайных сигналов. Случайным сигналом называют функцию времени, значения которой заранее неизвестны, и могут быть предсказаны лишь с некоторой вероятностью. Случайный сигнал отображает случайное физическое явление или физический процесс, причем зарегистрированный в единичном наблюдении сигнал не воспроизводится при повторных наблюдениях и не может быть описан явной математической зависимостью. При регистрации случайного сигнала реализуется только один из возможных вариантов (исходов) случайного процесса, а достаточно полное и точное описание процесса в целом можно произвести только после многократного повторения наблюдений и вычисления определенных статистических характеристик ансамбля реализаций сигнала. В качестве основных статистических характеристик случайных сигналов принимают:

- а) закон распределения вероятности нахождения величины сигнала в определенном интервале значений;
- б) спектральное распределение мощности сигнала.

Случайные сигналы подразделяют на стационарные и нестационарные. Случайные стационарные сигналы сохраняют свои статистические характеристики в последовательных реализациях случайного процесса. Что касается случайных нестационарных сигналов, то их общепринятой классификации не существует. Как правило, из них выделяют различные группы сигналов по особенностям их нестационарности.

1.2. ТИПЫ СИГНАЛОВ [1,10,15]

Выделяют следующие типы сигналов, которым соответствуют определенные формы их математического описания.

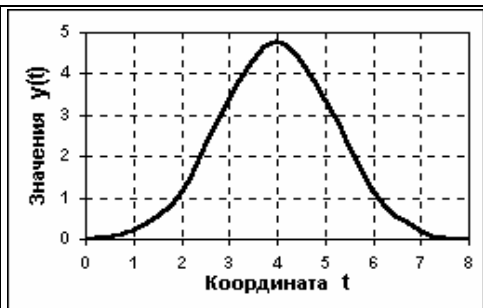


Рис. 1.2.1. Аналоговый сигнал.

Аналоговый сигнал (analog signal) является непрерывной функцией непрерывного аргумента, т.е. определен для любого значения аргументов. Источниками аналоговых сигналов, как правило, являются физические процессы и явления, непрерывные в динамике своего развития во времени, в пространстве или по любой другой независимой переменной, при этом регистрируемый сигнал подобен (“аналогичен”) порождающему его процессу. Пример математической записи сигнала: $y(t) = 4.8 \exp[-(t-4)^2/2.8]$. Пример графического отображения данного сигнала приведен на рис. 1.2.1, при этом как сама функция, так и ее аргументы, могут принимать любые значения в пределах некоторых интервалов $y_1 \leq y \leq y_2$, $t_1 \leq t \leq t_2$.

Если интервалы значений сигнала или его независимых переменных не ограничиваются, то по умолчанию они принимаются равными от $-\infty$ до $+\infty$. Множество возможных значений сигнала образует континуум - непрерывное пространство, в котором любая сигнальная точка может быть определена с точностью до бесконечности. Примеры сигналов, аналоговых по своей природе - изменение напряженности электрического, магнитного, электромагнитного поля во времени и в пространстве.



Рис. 1.2.2. Дискретный сигнал

Дискретный сигнал (discrete signal) по своим значениям также является непрерывной функцией, но определенной только по дискретным значениям аргумента. По множеству своих значений он является конечным (счетным) и описывается дискретной последовательностью отсчетов (samples) $y(n\Delta t)$, где $y_1 \leq y \leq y_2$, Δt - интервал между отсчетами (интервал или шаг дискретизации, sample time), $n = 0, 1, 2, \dots, N$. Величина, обратная шагу дискретизации: $f = 1/\Delta t$, называется частотой дискретизации (sampling frequency). Если дискретный сигнал получен дискретизацией (sampling) аналогового сигнала, то он представляет собой последовательность отсчетов, значения которых в точности равны значениям исходного сигнала по координатам $n\Delta t$.

Пример дискретизации аналогового сигнала, приведенного на рис. 1.2.1, представлен на рис. 1.2.2. При $\Delta t = \text{const}$ (равномерная дискретизация данных) дискретный сигнал можно описывать сокращенным обозначением $y(n)$. В технической литературе в обозначениях дискретизированных функций иногда оставляют прежние индексы аргументов аналоговых функций, заключая последние в квадратные скобки - $y[t]$. При неравномерной дискретизации сигнала обозначения дискретных последовательностей (в текстовых описаниях) обычно заключаются в фигурные скобки - $\{s(t_i)\}$, а значения отсчетов приводятся в виде таблиц с указанием значений координат t_i . Для числовых последовательностей (равномерных и неравномерных) применяется и следующее числовое описание: $s(t_i) = \{a_1, a_2, \dots, a_N\}$, $t = t_1, t_2, \dots, t_N$. Примеры дискретных геофизических сигналов - результаты вертикального электрического зондирования, профили геохимического опробования и т.п.



Рис. 1.2.3. Цифровой сигнал

Цифровой сигнал (digital signal) квантован по своим значениям и дискретен по аргументу. Он описывается квантованной решетчатой функцией $y_n = Q_k[y(n\Delta t)]$, где Q_k - функция квантования с числом уровней квантования k , при этом интервалы квантования могут быть как с равномерным распределением, так и с неравномерным, например - логарифмическим. Задается цифровой сигнал, как правило, в виде дискретного ряда (discrete series) числовых данных - числового массива по последовательным значениям аргумента при $\Delta t = \text{const}$,

но в общем случае сигнал может задаваться и в виде таблицы для произвольных значений аргумента.

По существу, цифровой сигнал по своим значениям (отсчетам) является формализованной разновидностью дискретного сигнала при округлении отсчетов последнего до определенного количества цифр, как это показано на рис. 1.2.3. Цифровой сигнал конечен по множеству своих значений. Процесс преобразования бесконечных по значениям аналоговых отсчетов в конечное число цифровых значений называется квантованием по уровню, а возникающие при квантовании ошибки округления отсчетов (отбрасываемые значения) - шумами (noise) или ошибками (error) квантования (quantization).

В дискретных системах и в ЭВМ сигнал всегда представлен с точностью до определенного количества разрядов, а, следовательно, всегда является цифровым. С учетом этих факторов при описании цифровых сигналов функция квантования обычно опускается (подразумевается равномерной по умолчанию), а для описания сигналов используются правила описания дискретных сигналов. Что касается формы обращения цифровых сигналов в системах хранения, передачи и обработки, то, как правило, они представляют собой комбинации коротких одно- или двуполярных импульсов одинаковой амплитуды, которыми в двоичном коде с определенным количеством числовых разрядов кодируются числовые последователь-

ности сигналов (массивов данных).

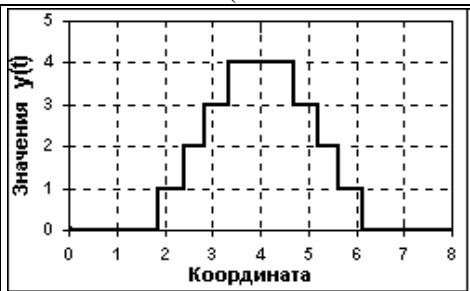


Рис. 1.2.4. Дискретно-аналоговый сиг-

В принципе, квантованными по своим значениям могут быть и аналоговые сигналы, зарегистрированные соответствующей аппаратурой (рис. 1.2.4), которые принято называть дискретно-аналоговыми. Но выделять эти сигналы в отдельный тип не имеет смысла - они остаются аналоговыми кусочно-непрерывными сигналами с шагом квантования, который определяется допустимой погрешностью измерений.

Большинство дискретных и цифровых сигналов, с которыми приходится иметь дело при обработке геофизических данных, являются аналоговыми по своей природе, дискретизированными в силу методических особенностей измерений или технических особенностей регистрации. Но существуют и сигналы, которые изначально относятся к классу цифровых, как, например отсчеты количества гамма-квантов, зарегистрированных по последовательным интервалам времени.

Преобразования типа сигналов. Формы математического отображения сигналов, особенно на этапах их первичной регистрации (детектирования) и в прямых задачах описания геофизических полей и физических процессов, как правило, отражают их физическую природу. Однако последнее не является обязательным и зависит от методики измерений и технических средств детектирования, преобразования, передачи, хранения и обработки сигналов. На разных этапах процессов получения и обработки информации как материальное представление сигналов в устройствах регистрации и обработки, так и формы их математического описания при анализе данных, могут изменяться путем соответствующих операций преобразования типа сигналов.

Операция дискретизации (discretization) осуществляет преобразование аналоговых сигналов (функций), непрерывных по аргументу, в функции мгновенных значений сигналов по дискретному аргументу, как, например $s(t) \Rightarrow s(n\Delta t)$, где значения $s(n\Delta t)$ представляют собой отсчеты функции $s(t)$ в моменты времени $t = n\Delta t$, $n = 0, 1, 2, \dots, N$.

Операция восстановления аналогового сигнала из его дискретного представления обратна операции дискретизации и представляет, по существу, интерполяцию данных.

В общем случае, дискретизация сигналов может приводить к определенной потере информации о поведении сигналов в промежутках между отсчетами. Однако существуют условия, определенные теоремой Котельникова-Шеннона, согласно которым аналоговый сигнал с ограниченным частотным спектром может быть без потерь информации преобразован в дискретный сигнал, и затем абсолютно точно восстановлен по значениям своих дискретных отсчетов.

Операция квантования или аналого-цифрового преобразования (АЦП; английский термин Analog-to-Digital Converter, ADC) заключается в преобразовании дискретного сигнала $s(n\Delta t)$ в цифровой сигнал $s(n) = s_n \approx s(n\Delta t)$, $n = 0, 1, 2, \dots, N$, как правило, кодированный в двоичной системе счисления. Процесс преобразования отсчетов сигнала в числа называется квантованием по уровню (quantization), а возникающие при этом потери информации за счет округления – ошибками или шумами квантования (quantization error, quantization noise).

При преобразовании аналогового сигнала непосредственно в цифровой сигнал операции дискретизации и квантования совмещаются.

Операция цифро-аналогового преобразования (ЦАП; Digital-to-Analog Converter, DAC) обратна операции квантования, при этом на выходе регистрируется либо дискретно-аналоговый сигнал $s(n\Delta t)$, который имеет ступенчатую форму (рис. 1.2.4), либо непосредственно аналоговый сигнал $s(t)$, который восстанавливается из $s(n\Delta t)$, например, путем сглаживания.

Так как квантование сигналов всегда выполняется с определенной и неустранимой погрешностью (максимум - до половины интервала квантования), то операции АЦП и ЦАП

не являются взаимно обратными с абсолютной точностью.

Спектральное представление сигналов. Кроме привычного динамического представления сигналов и функций в виде зависимости их значений от определенных аргументов (времени, линейной или пространственной координаты и т.п.) при анализе и обработке данных широко используется математическое описание сигналов по аргументам, обратным аргументам динамического представления. Так, например, для времени обратным аргументом является частота. Возможность такого описания определяется тем, что любой сколь угодно сложный по своей форме сигнал, не имеющий разрывов первого рода (бесконечных значений на интервале своего задания), можно представить в виде суммы более простых сигналов, и, в частности, в виде суммы простейших гармонических колебаний, что выполняется при помощи преобразования Фурье. Соответственно, математическое разложение сигнала на гармонические составляющие описывается функциями значений амплитуд и начальных фаз колебаний по непрерывному или дискретному аргументу – частоте изменения функций на определенных интервалах аргументов их динамического представления. Совокупность амплитуд гармонических колебаний разложения называют *амплитудным спектром* сигнала, а совокупность начальных фаз – *фазовым спектром*. Оба спектра вместе образуют полный частотный спектр сигнала, который по точности математического представления тождественен динамической форме описания сигнала.

Линейные системы преобразования сигналов описываются дифференциальными уравнениями, причем для них верен принцип суперпозиции, согласно которому реакция систем на сложный сигнал, состоящий из суммы простых сигналов, равна сумме реакций от каждого составляющего сигнала в отдельности. Это позволяет при известной реакции системы на гармоническое колебание с определенной частотой определить реакцию системы на любой сложный сигнал, разложив его в ряд гармоник по частотному спектру сигнала.

Широкое использование гармонических функций при анализе сигналов объясняется тем, что они являются достаточно простыми ортогональными функциями и определены при всех значениях t . Кроме того, они являются собственными функциями времени, сохраняющими свою форму при прохождении колебаний через любые линейные системы и системы обработки данных с постоянными параметрами (изменяются только амплитуда и фаза колебаний). Немаловажное значение имеет и то обстоятельство, что для гармонических функций и их комплексного анализа разработан мощный математический аппарат.

Примеры частотного представления сигналов уже приводились выше в разделе классификации сигналов (рис. 1.1.5 – 1.1.12)

Кроме гармонического ряда Фурье применяются и другие типы разложения: по функциям Уолша, Бесселя, Хаара, полиномам Чебышева, Лаггера, Лежандра и др.

Графическое отображение сигналов общеизвестно и особых пояснений не требует. Для одномерных сигналов график – это совокупность пар значений $\{t, s(t)\}$ в прямоугольной системе координат (рис. 1.2.1 – 1.2.4). При графическом отображении дискретных и цифровых сигналов используется либо способ непосредственных дискретных отрезков соответствующей масштабной длины над осью аргумента, либо способ огибающей (плавной или ломанной) по значениям отсчетов. В силу непрерывности геофизических полей и, как правило, вторичности цифровых данных, получаемых дискретизацией и квантованием аналоговых сигналов, второй способ графического отображения будем считать основным.

Тестовые сигналы (test signal). В качестве тестовых сигналов, которые применяются при моделировании и исследовании систем обработки данных, обычно используются сигналы простейшего типа: гармонические синус-косинусные функции, дельта-функция и функция единичного скачка.

Дельта-функция или функция Дирака. По определению, дельта-функция описывается следующими математическими выражениями (в совокупности):

$$\delta(t-\tau) = 0 \quad \text{при } t \neq \tau,$$

$$\int_{-\infty}^{\infty} \delta(t-\tau) dt = 1.$$

Функция $\delta(t-\tau)$ не является дифференцируемой, и имеет размерность, обратную размерности ее аргумента, что непосредственно следует из безразмерности результата интегрирования. Значение дельта-функции равно нулю везде за исключением точки τ , где она представляет собой бесконечно узкий импульс с бесконечно большой амплитудой, при этом площадь импульса равна 1.

Дельта-функция является полезной математической абстракцией. На практике такие функции не могут быть реализованы с абсолютной точностью, так как невозможно реализовать значение, равное бесконечности, в точке $t = \tau$ на аналоговой временной шкале, т.е. определенной по времени также с бесконечной точностью. Но во всех случаях, когда площадь импульса равна 1, длительность импульса достаточно мала, а за время его действия на входе какой-либо системы сигнал на ее выходе практически не изменяется (реакция системы на импульс во много раз больше длительности самого импульса), входной сигнал можно считать *единичной импульсной функцией* со свойствами дельта - функции.

При всей своей абстрактности дельта - функция имеет вполне определенный физический смысл. Представим себе импульсный сигнал прямоугольной формы $\Pi(t-\tau)$ длительностью θ , амплитуда которого равна $1/\theta$, а площадь соответственно равна 1. При уменьшении значения длительности θ импульс, сокращаясь по длительности, сохраняет свою площадь, равную 1, и возрастает по амплитуде. Предел такой операции при $\theta \Rightarrow 0$ и носит название дельта - импульса. Этот сигнал $\delta(t-\tau)$ сосредоточен в одной координатной точке $t = \tau$, конкретное амплитудное значение сигнала не определено, но площадь (интеграл) остается равной 1. Это не мгновенное значение функции в точке $t = \tau$, а именно импульс (импульс силы в механике, импульс тока в электротехнике и т.п.) – математическая модель короткого действия, значение которого равно 1.

Дельта-функция обладает *фильтрующим свойством*. Суть его заключается в том, что если дельта-функция $\delta(t-\tau)$ входит под интеграл какой-либо функции в качестве множителя, то результат интегрирования равен значению подынтегральной функции в точке τ расположения дельта-импульса, т.е.:

$$\int_{-\infty}^{\infty} f(t) \delta(t-\tau) dt = f(\tau).$$

Интегрирование в этом выражении может ограничиваться ближайшими окрестностями точки τ .

Функция единичного скачка или функция Хевисайда иногда называется также функцией включения. Полное математическое выражение функции:

$$\sigma(t) = \begin{cases} 0 & \text{если } t < 0, \\ 1/2 & \text{если } t = 0, \\ 1 & \text{если } t > 0. \end{cases}$$

При моделировании сигналов и систем значение функции скачка в точке $t=0$ очень часто принимают равным 1, если это не имеет принципиального значения.

Функция единичного скачка используется также при создании математических моделей сигналов конечной длительности. При умножении любой произвольной функции, в том числе периодической, на прямоугольный импульс, сформированный из двух последовательных функций единичного скачка

$$s(t) = \sigma(t) - \sigma(t-T),$$

из нее вырезается участок на интервале $0-T$, и обнуляются значения функции за пределами этого интервала.

Функция Кронекера. Для дискретных и цифровых систем разрешающая способность по

аргументу сигнала определяется интервалом его дискретизации Δt . Это позволяет в качестве единичного импульса использовать дискретный интегральный аналог дельта-функции - функцию единичного отсчета $\delta(k\Delta t - n\Delta t)$, которая равна 1 в координатной точке $k = n$, и нулю во всех остальных точках. Функция $\delta(k\Delta t - n\Delta t)$ может быть определена для любых значений $\Delta t = \text{const}$, но только для целых значений координат k и n , поскольку других номеров отсчетов в дискретных функциях не существует.

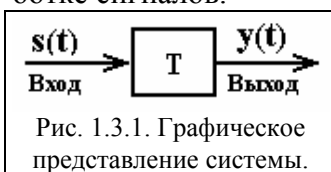
Математические выражения $\delta(t - \tau)$ и $\delta(k\Delta t - n\Delta t)$ называют также импульсами Дирака и Кронекера. Однако, применяя такую терминологию, не будем забывать, что это не просто единичные импульсы в координатных точках τ и $n\Delta t$, а импульсные функции, определяющие как значения импульсов в определенных координатных точках, так и нулевые значения по всем остальным координатам, в пределе от $-\infty$ до ∞ .

1.3. СИСТЕМЫ ПРЕОБРАЗОВАНИЯ СИГНАЛОВ [1,9,14,18]

Сигналы, в любой форме материального представления, содержат определенную полезную информацию. Если при преобразованиях сигналов происходит нарушение заключенной в них информации (частичная утрата, количественное изменение содержания и т.п.), то такие изменения называются *искажениями* сигнала. Если полезная информация остается постоянной, то такие изменения называются *преобразованиями* сигнала.

Любые изменения сигналов сопровождаются изменением их спектра, и по характеру этого изменения разделяются на два вида: линейные и нелинейные. К *нелинейным* относят изменения, при которых в составе спектра сигналов появляются новые гармонические составляющие. При *линейных* изменениях сигналов изменяются амплитуды и/или начальные фазы гармонических составляющих спектра. И линейные, и нелинейные изменения сигналов могут происходить как с сохранением полезной информации, так и с ее искажением. Это зависит не только от характера изменения спектра сигналов, но и от спектрального состава самой полезной информации.

Общее понятие систем. Преобразование и обработка сигналов осуществляется в системах. Понятия сигнала и системы неразрывны, так как любой сигнал существует в пределах какой-либо системы. Система обработки сигналов может быть реализована как в материальной форме (специальное устройство, измерительный прибор, совокупность физических объектов с определенной структурой взаимодействия и т.п.), так и программно на ЭВМ или другом специализированном вычислительном устройстве. Форма реализации системы существенного значения не имеет, и определяет только ее возможности при анализе и обработке сигналов.



Безотносительно к назначению система всегда имеет *вход*, на который подается внешний входной сигнал, в общем случае многомерный, и *выход*, с которого снимается обработанный выходной сигнал. Если устройство системы и внутренние операции сигнальных преобразований принципиального значения не имеют, то система в целом может восприниматься в формализованной форме “черного ящика”.

Формализованная система представляет собой *системный оператор* (алгоритм) преобразования входного сигнала $s(t)$ – *воздействия* или *возбуждения*, в сигнал на выходе системы $y(t)$ – *отклик* или *выходную реакцию* системы. Символическое обозначение операции преобразования (трансформации сигнала): $y(t) = T[s(t)]$.

Системный оператор T - это набор правил преобразования (transformation) сигнала $s(t)$ в сигнал $y(t)$. Так, например, в самом простейшем случае таким правилом может быть таблица перекодировки входных сигналов в выходные. Для общеизвестных операций преобразования сигналов применяются двухсимвольные индексы операторов трансформации, где вторым символом обозначается конкретный вид операции (например, TF- преобразование Фурье).

Входной сигнал системы может представлять собой m - мерный вектор, а выходной сигнал n - мерный вектор, при этом система будет иметь m входов и n выходов. Пример та-

кой системы в геофизике: трехканальный гамма-спектрометр, на три входа решающего блока которого поступают потоки сигналов от калиевого, радиевого и ториевого каналов амплитудного анализатора спектрометра, а на три выхода решающего блока подаются результаты количественной интерпретации входной информации - сигналы количественных содержаний калия, урана и тория, при этом системный оператор реализует алгоритм решения системы трех линейных уравнений с тремя неизвестными.

Для детерминированных входных сигналов соотношение между выходными и входными сигналами всегда однозначно задается системным оператором. В случае реализации на входе системы случайного входного процесса аналогично существует однозначное соответствие процессов на выходе и входе системы, однако при этом одновременно происходит изменение статистических характеристик сигнала (математического ожидания, дисперсии, корреляционной функции и пр.), которое также определяется системным оператором.

Для полного определения системы необходимо задание характера, типа и области допустимых величин входных и выходных сигналов. Как правило, системы выполняются на сигналы одного типа по входу/выходу, и подразделяются на системы непрерывного времени (аналоговые или дискретные сигналы на входе и выходе) и цифровые системы. Совокупность системного оператора T и областей входных/выходных сигналов образует математическую модель системы.

Линейные и нелинейные системы составляют два основных класса систем обработки сигналов. Термин линейности (linear) означает, что система преобразования сигналов должна иметь произвольную, но в обязательном порядке линейную связь между входным сигналом (возбуждением) и выходным сигналом (откликом) с определенным изменением спектрального состава входного сигнала (усиление или подавление определенных частотных составляющих сигнала). В нелинейных (nonlinear) системах связь между входным и выходным сигналом определяется произвольным нелинейным законом с дополнением частотного состава входного сигнала частотными составляющими, отсутствующими во входном сигнале.

Математически связь между аналоговыми сигналами входа $s(t)$ и выхода $y(t)$ в линейной системе обычно задается линейным дифференциальным уравнением:

$$\sum_{m=0}^M a_m \frac{d^m [y(t)]}{dt^m} = \sum_{n=0}^N b_n \frac{d^n [s(t)]}{dt^n}. \quad (1.3.1)$$

где a_m и b_n – параметрические коэффициенты системы. Максимальный порядок производной входного сигнала в уравнении (1.3.1) не превышает порядка производной выходного сигнала, т.е. $m \leq n$. Значение m называется *порядком* системы. При нормировке уравнения к $a_0 = 1$ получаем неоднородное линейное дифференциальное уравнение для произвольного входного сигнала $s(t)$, решение которого дает выходной сигнал $y(t)$:

$$y(t) = \sum_{n=0}^N b_n \frac{d^n [s(t)]}{dt^n} - \sum_{m=1}^M a_m \frac{d^m [y(t)]}{dt^m}. \quad (1.3.1')$$

Аналогичная связь выхода с входом в дискретной (цифровой) системе описывается разностными уравнениями:

$$\sum_{m=0}^M a_m y((k-m)\Delta t) = \sum_{n=0}^N b_n s((k-n)\Delta t). \quad (1.3.2)$$

$$y(k\Delta t) = \sum_{n=0}^N b_n s((k-n)\Delta t) - \sum_{m=1}^M a_m y((k-m)\Delta t). \quad (1.3.2')$$

Уравнение (1.3.2') можно рассматривать как алгоритм последовательного вычисления значений $y(k\Delta t)$, $k = 0, 1, 2, \dots$, по значениям входного сигнала $s(k\Delta t)$ при известных значениях коэффициентов a_m , b_n и задании определенных начальных условий $s(-k\Delta t)$, $y(-k\Delta t)$.

Стационарные и нестационарные системы. Система считается стационарной и имеет *постоянные параметры*, если ее свойства (математический алгоритм оператора преобра-

зования) в пределах заданной точности не зависят от входного и выходного сигналов и не изменяются ни во времени, ни от каких-либо других внешних факторов. Математически это означает задание системы уравнениями типа (1.3.1-2) с постоянными значениями коэффициентов a_m и b_n . В противном случае система является нестационарной и называется *параметрической* или *системой с переменными параметрами*. Среди последних большое значение имеют так называемые адаптивные системы обработки данных. В этих системах производится, например, оценивание определенных параметров входных и выходных сигналов, по результатам сравнения которых осуществляется подстройка параметров преобразования (переходной характеристики системы) таким образом, чтобы обеспечить оптимальные по производительности условия обработки сигналов или минимизировать погрешность обработки.

Основные системные операции. К базовым линейным операциям, из которых могут быть сформированы любые линейные операторы преобразования, относятся операции скалярного умножения, сдвига и сложения сигналов:

$$y(t) = c \times s(t), \quad y(t) = s(t-\Delta t), \quad y(t) = a(t)+b(t).$$

Отметим, что строго корректно операции сложения и умножения являются линейными только для аналоговых и дискретных сигналов. В случае цифровых сигналов они линейны относительно самих цифровых сигналов, но если последние - результат операции АЦП, то сложение и умножение не может считаться линейным абсолютно точно по отношению к исходным сигналам.

Для нелинейных систем выделим важный тип безинерционных операций нелинейной трансформации сигнала, результаты которой зависят только от его входных значений. К ним относятся, например, операции квадратирования и логарифмирования сигнала:

$$y(t) = [s(t)]^2, \quad y(t) = \log[s(t)].$$

Линейные системы. Система считается линейной, если ее реакция на входные сигналы аддитивна (выполняется принцип суперпозиции сигналов) и однородна (выполняется принцип пропорционального подобия).

Принцип *аддитивности* требует, чтобы реакция системы на сумму входных сигналов была равна сумме реакций на каждый сигнал в отдельности. Так, для двух сигналов должно иметь место:

$$T[a(t)+b(t)] = T[a(t)]+T[b(t)].$$

Принцип *однородности* или пропорционального подобия требует сохранения однозначности масштаба преобразования при любой амплитуде входного сигнала:

$$T[c \times s(t)] = c \times T[s(t)].$$

Другими словами, *отклик линейной системы на взвешенную сумму входных сигналов должен быть равен взвешенной сумме откликов на отдельные входные сигналы независимо от их количества и для любых весовых коэффициентов*, в том числе комплексных.

Примеры.

1. Система $y(t) = a^2t$.

$$y(t_1) = a^2t_1, \quad y(t_2) = a^2t_2, \quad y(ct) = a^2ct.$$

$$y(t_1+t_2) = a^2(t_1+t_2) = a^2t_1+a^2t_2 = y(t_1)+y(t_2). \quad \text{Система аддитивна.}$$

$$cy(t) = ca^2t = a^2ct = y(ct). \quad \text{Система однородна. Следовательно, система линейна.}$$

2. Система $y(t) = a \cdot t^2$.

$$y(t_1) = at_1^2, \quad y(t_2) = at_2^2, \quad y(ct) = a(ct)^2 = ac^2t^2.$$

$$y(t_1+t_2) = a(t_1+t_2)^2 \neq y(t_1)+y(t_2) = at_1^2+at_2^2. \quad \text{Система не аддитивна.}$$

$$y(t) = c at^2 \neq y(ct) = ac^2t^2. \quad \text{Система неоднородна. Следовательно, система нелинейна.}$$

При программной реализации линейных систем на ЭВМ особых затруднений с обеспечением линейности в разумных пределах значений входных и выходных сигналов, как правило, не возникает. При физической (аппаратной) реализации систем обработки данных диапазон входных и/или выходных сигналов, в котором обеспечивается линейность преобразования сигналов, всегда ограничен и должен быть специально оговорен в технической документации или методической инструкции.

Инвариантность систем к сдвигу. Система называется инвариантной к сдвигу, если сдвиг входного сигнала по аргументам (времени, координатам пространства и т.п.) вызывает соответствующий сдвиг выходного сигнала:

$$y(x,t) = T[s(x,t)], \quad T[s(x-\Delta x, t-\Delta t)] = y(x-\Delta x, t-\Delta t).$$

Инвариантность системы к сдвигу является одним из подтверждений постоянства ее параметров.

Линейные системы, инвариантные к сдвигу. Линейность и инвариантность к сдвигу являются независимыми свойствами систем и не определяют друг друга. Так, например, операция квадратирования сигнала инвариантна к сдвигу, но нелинейна.

В теории анализа и обработки данных основное место занимают системы, линейные и инвариантные к сдвигу (ЛИС - системы). Они обладают достаточно широкими практически возможными возможностями при относительной простоте математического аппарата. В дальнейшем, если это специально не оговаривается, будем иметь в виду именно такие системы.

Преимущество, которое отдается ЛИС - системам в методах обработки информации, базируется на возможности разложения входного сигнала любой, сколь угодно сложной формы, на составляющие простейших форм, отклик системы на которые известен и хорошо изучен, с последующим вычислением выходного сигнала в виде суммы откликов на все составляющие входного сигнала. В качестве простейших форм разложения сигналов используются, как правило, единичные импульсы и гармонические составляющие. Разложение по единичным импульсам применяется при динамическом представлении сигнала в зависимости от реальных физических аргументов (времени, координат и пр.) и использует операцию свертки. Разложение на гармонические составляющие использует спектральное (частотное) представление сигнала и преобразование Фурье.



Рис. 1.3.2 Соединения систем.

Соединения ЛИС - систем. При последовательном (каскадном) соединении систем выходной сигнал одной системы служит входным сигналом для второй и т.д. в зависимости от количества составляющих систем каскада. По отношению к общей системной операции преобразования порядок соединения входящих в нее систем значения не имеет. Так, для двух последовательно соединенных систем на рис. 1.3.2:

$$y(t) = T_2[T_1[s(t)]] = T_1[T_2[s(t)]].$$

При параллельном соединении входной сигнал поступает одновременно на входы всех составляющих систем, а выходные сигналы систем суммируются:

$$y(t) = T_1[s(t)] + T_2[s(t)] + \dots + T_N[s(t)].$$

Образуемые в результате соединений системы в целом также являются ЛИС - системами, если линейны и инвариантны к сдвигу системы, в них входящие.

Обобщенная схема системы цифровой обработки сигналов на рис. 1.3.3 приведена в качестве примера.

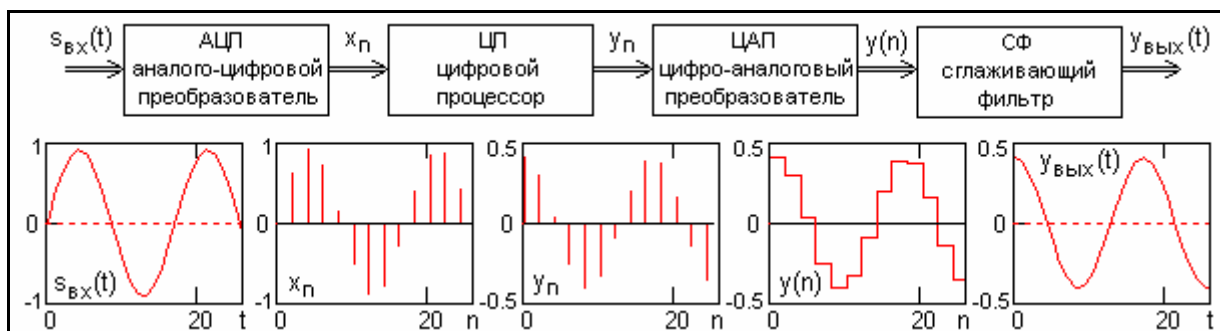


Рис. 1.3.3. Структурная схема системы дифференцирования сигналов.

1.4. ИНФОРМАЦИОННАЯ ЕМКОСТЬ СИГНАЛОВ [10,12,28]

Объем информации, находящейся в обращении и необходимой для функционирования

ния и развития современного общества, нарастает примерно пропорционально квадрату развития производительных сил. В передовых по научно-техническому развитию странах мира доля рабочей силы, занятой вопросами сбора, обработки и обеспечения информацией, превышает долю рабочей силы в сфере материального производства. Применение методов и средств автоматизации на всех этапах обращения информации, эффективная организация ее хранения, обработки и обмена, приобретают все большее значение в качестве основного условия успешного функционирования экономики стран.

Понятие информации. В настоящее время нет общепринятого и однозначного понимания термина "Информация". Спектр бытующих понятий весьма широк, от общего философского - информация есть отражение реального мира, до узко практического - информация есть сведения, являющиеся объектом хранения, передачи и преобразования. Расхождения существуют и по вопросу места информации в материальном мире. Это свойство индивидуальных объектов или результат их взаимодействия? Присуща ли информация всем видам материи или лишь определенным образом организованной материи?

В информатике под информацией понимается, как правило, совокупность сведений смыслового содержания, которые можно собирать, обрабатывать, передавать и т.п. Причем именно сведений в изначальном смысле латинского слова *informatio*, а не данных или сигналов, которые являются носителями этих сведений. В таком понимании процессы извлечения сведений из данных и их интерпретации неразрывно связаны с разумом, а конечным результатом обработки и восприятия информации с помощью разума является раскрытие неопределенности знаний о каком-либо объекте, явлении или процессе. Но при таком подходе размывается само понятие разума.

С одной стороны, существование любого живого существа поддерживается до тех пор, пока действуют его органы чувств (датчики), преобразующие физические воздействия окружающего мира в определенным образом организованные сигналы, которые в материальной форме отображают данные об этих воздействиях. Данные собираются и интерпретируются определенной системой, из общей суммы данных извлекаются определенные сведения, степень неопределенности сведений об окружающей обстановке снижается, и ... лиса распутывает заячий след. Живое существо существует до тех пор, пока способно воспринимать и обрабатывать внешние и внутренние воздействия. Нет сомнений и в том, что в коллективных сообществах его члены не только способны собирать и обрабатывать информацию, но и передавать ее другим членам сообщества, как, например, в пчелиной семье точный путь до продуктивного цветочного массива. Информационный танец пчелы в этом отношении по компактности содержания ничем не уступает телеграфному сообщению. Естественно, в принятой у пчел символической форме.

С другой стороны, если информация неразрывно связана с разумом, то в этом случае нельзя отказать в разуме и электронной вычислительной машине, обыгрывающей в шахматы чемпиона мира, а равно и любым устройствам технической кибернетики любого уровня, так как все они имеют определенные системы сбора, передачи, накопления, хранения и обработки информации той или иной степени сложности, и на основе этой информации способны формировать сигналы обратной связи для управления определенными процессами.

В технических отраслях знаний, где вопросы соотношения информации с разумом не стоят на первом месте, преобладает понимание информации в виде отображения такого всеобщего свойства материи, как разнообразие, как характеристики внутренней организованности материальных систем, процессов или явлений по множеству состояний, которые для них возможны. В такой трактовке информация существует независимо от того, воспринимается она каким-либо "разумом" или нет, и является одним из свойств материальных объектов. "Информация есть информация, а не материя и не энергия" (Норберт Винер). Это свойство в какой-то мере имеет потенциальный характер. Информация может проявлять себя только при взаимодействии объектов или процессов, может возникать (создаваться) и исчезать (уничтожаться).

Но и в такой трактовке возникает очень много вопросов, на которые трудно дать од-

нозначные ответы. Насекомое третичного периода, неизвестное в настоящее время ученым, прилипло к капле смолы хвойного дерева. Новый слой смолы закрыл насекомое. Дерево упало и его занесло песком. Смола превратилась в янтарь. Янтарь в потенциале содержит полную информацию о насекомом, потому как в нем десятки тысяч фрагментов ДНК - информация, достаточная для восстановления ДНК и воспроизводства насекомого, если не в настоящее время, то в ближайшем будущем. Но когда она возникла? В момент появления насекомого с его ДНК? В момент прилипания к смоле? В момент окаменения? Можно ли говорить о появлении информации, если еще не существовал субъект, способный извлечь и использовать эту информацию? Наконец, янтарь с насекомым найден и попал на глаза палеонтолога. Определен новый вид насекомого. Появилась первая частичная информация? Так может быть информация появляется только при активном и целенаправленном воздействии на объект исследований? А если янтарь оказался непрозрачным, и его переплавили? Исчезла ли информация? И можно ли считать, что она вообще была?

Ответы на эти и подобные им вопросы тяготеют к двум полюсам, а по существу, к двум диаметрально противоположным философским позициям.

Сторонники первой позиции считают, что информация присуща только определенным образом организованным объектам или процессам, и понимают под информацией только то, что может восприниматься, обрабатываться, осмысливаться и использоваться, т.е. является продуктом целенаправленного процесса сбора, организации, систематизации и использования сведений о материальных объектах и процессах.

Противоположная позиция, это понятие информации как свойства объектов и процессов воспринимать и перерабатывать внутреннее состояние и внешнее воздействие окружающей среды, сохранять его результаты и передавать их другим объектам. С этой позиции все материальные объекты и процессы являются источниками, носителями и потребителями информации, на основе которой и идет развитие реального мира. По существу, это соответствует принятию материальности информации и информационной основы мироздания.

При неопределенности самого понятия информации можно достаточно обоснованно считать, что информация проявляется, хранится и передается от одного объекта к другому в материально - энергетической форме в виде сигналов. Сигналом, как материальным носителем информации, может быть любой физический процесс (электрический, магнитный, оптический, акустический и пр.), определенные параметры которого (амплитуда, частота, энергия, интенсивность и др.) однозначно отображают информационные данные (сообщения).

Количественная мера информации. Теория любого явления начинается с появления количественных взаимоотношений между объектами исследований, т.е. при установлении принципов измеряемости каких-либо свойств объектов. Единицу количественной меры информации - БИТ (сокращение binary digit - двоичная цифра), впервые предложил Р. Хартли в 1928 году. 1 бит - это информация о двух возможных равновероятных состояниях объекта, неопределенность выбора из двух равновероятных событий. Математически это отображается состоянием 1 или 0 одного разряда двоичной системы счисления. Количество информации H (в битах), необходимое и достаточное для полного снятия неопределенности состояния объекта, который имеет N равновозможных состояний, измеряется как логарифм по основанию 2 из числа возможных состояний:

$$H = \log_2 N. \quad (1.4.1)$$

Соответственно, двоичный числовой информационный код одного из N возможных состояний объекта занимает H двоичных разрядов.

Пример. Необходимо поднять груз на определенный этаж 16 -ти этажного здания (нумерация этажей 0-15, $N = 16$). Сколько бит информации полностью определяют задание?

$$H = \log_2 N = \log_2 16 = 4.$$

Следовательно, 4 бита информации должны быть необходимы и достаточны для полного снятия неопределенности выбора. В этом можно убедиться применением логики исчисления с последовательным делением пополам интервалов состояний. Например, для 9-го этажа:

1. Выше 7-го этажа? Да = 1.
2. Выше 11-го этажа? Нет = 0.
3. Выше 9-го этажа? Нет = 0.
4. Выше 8-го этажа? Да = 1.

Итог: этаж номер 9 или 1001 в двоичном исчислении, четыре двоичных разряда.

Если в приведенном примере на этажах имеется по 4 квартиры с нумерацией на каждом этаже 0-3 ($M=4$), то при адресации груза в квартиру потребуется еще 2 бита информации. Такой же результат получим, если вместо независимой нумерации этажей и квартир на этажах (два источника неопределенности) мы будем иметь только сквозную нумерацию квартир (один обобщенный источник):

$$H = \log_2 N + \log_2 M = \log_2 16 + \log_2 4 = 6 \equiv \log_2 (N \times M) = \log_2 64 = 6,$$

т.е. количество информации отвечает требованию аддитивности: неопределенность объединенного источника равна сумме неопределенностей исходных источников, что соответствует интуитивному требованию к информации: она должна быть однозначной, а ее количество должно быть одним и тем же независимо от способа задания.

Основание логарифма не имеет принципиального значения и определяет только масштаб или единицу неопределенности. Так, если за единицу неопределенности принять три равновероятных состояния, то для определения, например, одной фальшивой золотой монеты (более легкой) из 27 внешне неотличимых монет потребуется только $H = \log_3 27 = 3$, т.е. три взвешивания на равноплечных весах. Логику исчисления взвешиваний предлагается определить самостоятельно.

Двоичная мера информации получила общее признание в связи с простотой реализации информационной техники на элементах с двумя устойчивыми состояниями. В десятичном исчислении единицей информации является один десятичный разряд - ДИТ.

Энтропия источника информации. Степень неопределенности состояния объекта (или так называемого источника информации) зависит не только от числа его возможных состояний, но и от вероятности этих состояний. При неравновероятных состояниях свобода выбора для источника ограничивается. Так, если из двух возможных состояний вероятность одного из них равна 0.999, то вероятность другого состояния соответственно равна $1-0.999 = 0.001$ и при взаимодействии с таким источником результат практически предрешен.

В общем случае, в соответствии с теорией вероятностей, источник информации однозначно и полно характеризуется ансамблем состояний $U = \{u_1, u_2, \dots, u_N\}$ с вероятностями состояний соответственно $\{p(u_1), p(u_2), \dots, p(u_N)\}$ при условии, что сумма вероятностей всех состояний равна 1. Мера количества информации, как неопределенности выбора дискретным источником состояния из ансамбля U , предложена К. Шенноном в 1946 году и получила название *энтропии дискретного источника информации* или энтропии конечного ансамбля:

$$H(U) = - \sum_{n=1}^N p_n \log_2 p_n. \quad (1.4.2)$$

Выражение Шеннона совпадает с выражением Больцмана для энтропии физических систем при оценке степени разнообразия их состояний. Мера энтропии Шеннона является обобщением меры Хартли на случай ансамблей с неравновероятными состояниями, в чем нетрудно убедиться, если в выражении (1.4.2) значение p_n заменить значением $p=1/N$ для ансамбля равновероятных состояний. *Энтропия конечного ансамбля $H(U)$ характеризует неопределенность, приходящуюся в среднем на одно состояние ансамбля.*

Учитывая, что в дальнейшем во всех математических выражениях, касающихся энтропии, мы будем использовать только двоичное основание логарифма, индекс 2 основания логарифма в формулах будем подразумевать по умолчанию.

Пример. Вычислить энтропию ансамбля 32 букв русского алфавита. Вероятности использования букв приведены в таблице. Сравнить энтропию с неопределенностью, которая была бы у алфавита при равновероятном их использовании.

Неопределенность на одну букву при равновероятности использования:

$$H(u) = \log 32 = 5$$

Энтропия алфавита по ансамблю таблицы:

$$H(u) = -0.064 \log 0.064 - 0.015 \log 0.015 - \dots - 0.143 \log 0.143 \approx 4.42.$$

Таким образом, неравновероятность состояний снижает энтропию источника.

u_i	p_i	u_i	p_i	u_i	p_i	u_i	p_i	u_i	p_i
а	.064	з	.015	о	.096	х	.009	э	.003
б	.015	и	.064	п	.024	ц	.004	ю	.007
в	.039	й	.010	р	.041	ч	.013	я	.019
г	.014	к	.029	с	.047	ш	.006	-	.124
д	.026	л	.036	т	.056	щ	.003		
е,ё	.074	м	.026	у	.021	ъ,ь	.015		
ж	.008	н	.056	ф	.020	ы	.016		

Основные свойства энтропии:

1. Энтропия является величиной вещественной и неотрицательной, т.к. значения вероятностей p_n находятся в интервале 0-1, значения $\log p_n$ всегда отрицательны, а значения $-p_n \log p_n$ в (1.4.2) соответственно положительны.
2. Энтропия - величина ограниченная, т.к. при $p_n \Rightarrow 0$ значение $-p_n \log p_n$ также стремится к нулю, а при $0 < p_n \leq 1$ ограниченность суммы всех слагаемых очевидна.
3. Энтропия равна 0, если вероятность одного из состояний источника информации равна 1 и тем самым состояние источника полностью определено (вероятности остальных состояний источника равны нулю, т.к. сумма вероятностей должна быть равна 1).
4. Энтропия максимальна при равной вероятности всех состояний источника информации:

$$H_{\max}(U) = - \sum_{n=1}^N (1/N) \log (1/N) = \log N.$$

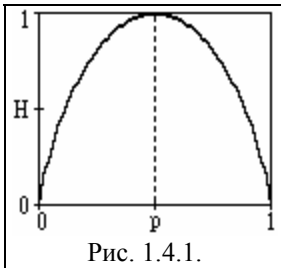


Рис. 1.4.1.

5. Энтропия источника с двумя состояниями u_1 и u_2 при изменении соотношения их вероятностей $p(u_1)=p$ и $p(u_2)=1-p$ определяется выражением:

$$H(U) = -[p \log p + (1-p) \log (1-p)],$$

и изменяется от 0 до 1, достигая максимума при равенстве вероятностей. График изменения энтропии приведен на рис. 1.4.1.

6. Энтропия объединенных статистически независимых источников информации равна сумме их энтропий.

Рассмотрим это свойство на двух источниках информации u и v . При объединении источников получаем обобщенный источник информации (u,v) , который описывается вероятностями $p(u_n v_m)$ всех возможных комбинаций состояний u_n источника u и v_m источника v . Энтропия объединенного источника при N возможных состояниях источника u и M возможных состояниях источника v :

$$H(UV) = - \sum_{n=1}^N \sum_{m=1}^M p(u_n v_m) \log p(u_n v_m),$$

Источники статистически независимыми друг от друга, если выполняется условие:

$$p(u_n v_m) = p(u_n) \cdot p(v_m).$$

С использованием этого условия соответственно имеем:

$$\begin{aligned} H(UV) &= - \sum_{n=1}^N \sum_{m=1}^M p(u_n) p(v_m) \log [p(u_n) p(v_m)] = \\ &= - \sum_{n=1}^N p(u_n) \log p(u_n) \sum_{m=1}^M p(v_m) - \sum_{m=1}^M p(v_m) \log p(v_m) \sum_{n=1}^N p(u_n). \end{aligned}$$

С учетом того, что $\sum_{n=1}^N p(u_n) = 1$ и $\sum_{m=1}^M p(v_m) = 1$, получаем:

$$H(UV) = H(U) + H(V). \quad (1.4.3)$$

7. Энтропия характеризует среднюю неопределенность выбора одного состояния из ансамбля, полностью игнорируя содержательную сторону ансамбля. С одной стороны, это расширяет возможности использования энтропии при анализе самых различных явлений, но, с другой стороны, требует определенной дополнительной оценки возникающих ситуаций. Как это следует из рис. 1.4.1, энтропия состояний может быть неоднозначной, и если в каком-либо экономическом начинании действие u с вероятностью $p_u=r$ приводит к успеху, а действие v с вероятностью $p_v=1-r$ к банкротству, то выбор действий по оценке энтропии может оказаться и прямо противоположным, т.к. энтропия при $p_v=r$ равна энтропии при $p_u=r$.

Энтропия непрерывного источника информации должна быть бесконечна, т. к. неопределенность выбора из бесконечно большого числа возможных состояний бесконечно велика.

Разобьем диапазон изменения непрерывной случайной величины U на конечное число n малых интервалов Δu . При реализации значений u в интервале $(u_n, u_n+\Delta u)$ будем считать, что реализовалось значение u_n дискретной случайной величины U' , вероятность реализации которой:

$$p(u_n < u < u_n + \Delta u) = \int_{u_n}^{u_n + \Delta u} p(u) du \approx p(u_n) \Delta u.$$

Энтропия дискретной величины U' :

$$H(U') = - \sum_{n=1}^N p(u_n) \Delta u \log (p(u_n) \Delta u).$$

Заменяем $\log (p(u_n) \Delta u) = \log p(u_n) + \log \Delta u$, принимаем во внимание, что сумма $p(u_n) \Delta u$ по всем возможным значениям u_n равна 1, и получаем:

$$H(U') = - \sum_{n=1}^N p(u_n) \Delta u \log p(u_n) - \log \Delta u. \quad (1.4.4)$$

В пределе, при $\Delta u \rightarrow 0$, получаем выражение энтропии для непрерывного источника:

$$H(U) = - \int_{-\infty}^{\infty} p(u) \log p(u) du - \lim_{\Delta u \rightarrow 0} \log \Delta u. \quad (1.4.5)$$

Значение энтропии в (1.4.5), как и ожидалось, стремится к бесконечности за счет второго члена выражения. Для получения конечной характеристики информационных свойств непрерывных сигналов используют только первый член выражения (1.4.5), получивший название дифференциальной энтропии. Ее можно трактовать, как среднюю неопределенность выбора произвольной случайной величины по сравнению со средней неопределенностью выбора случайной величины U' , имеющей равномерное распределение в диапазоне (0-1). Действительно, для такого распределения $p(u_n) = 1/N$, $\Delta u = 1/N$, и при $N \rightarrow \infty$ из (1.4.4) следует:

$$H(U') = - (\log N)/N - \log \Delta u \rightarrow - \lim_{\Delta u \rightarrow 0} \log \Delta u.$$

Соответственно, разность энтропий дает дифференциальную энтропию:

$$h(U) = H(U) - H(U') = - \int_{-\infty}^{\infty} p(u) \log p(u) du. \quad (1.4.6)$$

Дифференциальная энтропия не зависит от конкретных значений величины U :

$$h(U+a) = h(U), \quad a = \text{const},$$

но зависит от масштаба ее представления:

$$h(kU) = h(U) + \log k.$$

Практика анализа и обработки сигналов обычно имеет дело с сигналами в определенном интервале $[a, b]$ их значений, при этом максимальной дифференциальной энтропией обладает равномерное распределение значений сигналов:

$$h(U) = - \int_a^b p(u) \log p(u) du = \log (b-a).$$

По мере сужения плотности распределения значение $h(U)$ уменьшается, и в пределе при $p(u) \rightarrow \delta(u-c)$, $a < c < b$ стремится к нулю.

Информационная емкость сигналов существенно зависит от типа сигналов и определяет требования к каналам передачи данных (каналам связи), равно как и технические характеристики каналов связи определяют требования к информационной емкости сигналов, передаваемых по этим каналам.

Для каналов передачи дискретных сигналов (дискретные канала связи) используют понятия технической и информационной скорости передачи данных.

Под технической скоростью передачи подразумевают число элементарных сигналов (символов), передаваемых по каналу в единицу времени. Простейший элементарный символ – однополярный электрический импульс длительностью τ на тактовом интервале T . В дискретных каналах используют, как правило, двуполярные импульсы, положительные на первой половине интервала T и отрицательные на второй половине. Это позволяет поддерживать нулевой потенциал кабеля и выполнять тактовую синхронизацию приема-передачи сигналов. Единицей измерения технической скорости $V_t = 1/T$ служит БОД – один символ в секунду. Полоса пропускания канала связи обычно ограничивается определенной предельной частотой $F_{\text{пред}}$ по уровню затухания сигнала до уровня статистических помех, при этом значение технической скорости передачи данных, естественно, не может быть выше $F_{\text{пред}}$ без каких-либо специальных устройств выделения информационных сигналов.

При известной технической скорости V_t скорость передачи информации измеряется в битах в секунду и при уровне помех меньше амплитудных значений символьных импульсов задается соотношением:

$$V_h = V_t H(s),$$

где $H(s)$ – энтропия символа. Для двоичных дискретных символов с возможными состояниями $[0, 1]$ (однополярные – есть/нет импульс на такте, для двуполярных – порядок полярности импульсов на такте, например, 0: плюс/минус, 1: минус/плюс) при постоянной амплитуде импульсов значение $H(s)$ равно 1. При числе L возможных равновероятных уровней амплитуды импульсов (уровень помех меньше разности уровней амплитуд импульсов) значение $H(s)$ равно $\log L$.

Информационная емкость сигнала или полное количество информации в сигнале S (сообщении, кодовой последовательности/слове) определяется полным количеством $N = t/T$ энтропии символов в битах на интервале задания сигнала t :

$$I_t(S) = N \log L = (t/T) \log L. \quad (1.4.7)$$

Увеличение числа уровней L увеличивает пропускную способность каналов связи, но усложняет аппаратуру кодирования данных и снижает помехоустойчивость связи.

Для непрерывных сигналов передача по каналам связи возможна только при условии, что максимальная информационная частота в сигнале F_{max} не превышает предельной частоты $F_{\text{пред}}$ передачи сигналов каналом связи. Для оценки информационной емкости непрерывного сигнала выполним его дискретизацию с интервалом $\Delta t = 1/2F_{\text{max}}$. Как установлено Котельниковым и Шенноном, по мгновенным отсчетам непрерывного сигнала с таким интервалом дискретизации аналоговый сигнал может быть восстановлен без потери информации. При полной длительности сигнала T_s число отсчетов:

$$N = T_s/\Delta t = 2F_{\text{max}} T_s.$$

Определим максимально возможное число выборок в каждом отсчете при наличии шума в канале со средней мощностью $P_{\text{ш}} = \delta^2$. При средней мощности сигнала $P_s = s^2$:

$$L = \sqrt{\frac{s^2 + \delta^2}{\delta^2}} = \sqrt{1 + \frac{s^2}{\delta^2}}.$$

Информационная емкость сигнала:

$$I(S) = 2F_{\text{max}} T_s \log L. \quad (1.4.8)$$

Информационные возможности сигнала возрастают с расширением его спектра и превышением его уровня над уровнем помех.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
9. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. - М.: Мир, 1988. - 488 с.
10. Дмитриев В.И. Прикладная теория информации: Учебник для вузов. - М.: Высшая школа, 1989.
12. Игнатов В.А. Теория информации и передачи сигналов. - М.: Советское радио, 1979.
14. Купер Дж., Макгиллем А. Вероятностные методы анализа сигналов и систем. - М.: Мир, 1989.
15. Лосев А.К. Линейные радиотехнические цепи: Учебник для вузов. - М.: Высшая школа, 1971.
18. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. - М.: Связь, 1979. - 416 с.
25. Сергиенко А.Б. Цифровая обработка сигналов. / Учебник для вузов. - СПб.: Питер, 203. - 608 с.
28. Колесник В.Д., Полтырев Г.Ш. Курс теории информации. - М.: Наука, 1982. - 416 с.

Главный сайт автора ~ Лекции по сигналам ~ Практикум

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru.

Copyright ©2005 Davydov A.V.

Тема 2: МЕТРОЛОГИЯ СИГНАЛОВ

Физическая величина более точно определяется уравнением, чем измерением.
Макс Планк. Немецкий физик – теоретик, XVIII-XIX в.

Между тем, уравнение только математическая модель физической величины. А без измерений понятия точности вообще не существует.
Борис Старцев. Уральский геофизик – практик, XX-XXI в.

Содержание: 2.1. Пространство сигналов. Линейное пространство сигналов. Норма сигналов. Метрика сигналов. Скалярное произведение сигналов. Коэффициент корреляции сигналов. Координатный базис пространства. 2.2. Мощность и энергия сигналов. Понятия мощности и энергии сигналов. 2.3. Пространства функций. Нормирование метрических параметров. Ортогональные сигналы. Ортонормированный базис пространства. Разложение сигнала в ряд. Ортонормированные системы функций. Разложение энергии сигнала. 2.4. Функции корреляции сигналов. Корреляционные функции сигналов. Взаимная корреляционная функция. 2.5. Математическое описание шумов и помех. Шумы и помехи. Природа помех. Характеристики помех. Литература.

ВВЕДЕНИЕ

В данной теме метрология сигналов рассматривается, в основном, на уровне понятий и базовых определений, предваряя их более подробное изучение в дальнейших темах курса. Это объясняется тем, что при детальном изучении каких-либо характеристик или свойств сигналов их рассмотрение не может выполняться в отрыве от других метрологических характеристик рассматриваемых типов сигналов и требует определенной ориентировки в общей метрологии сигналов, хотя бы на уровне понятий.

2.1. ПРОСТРАНСТВО СИГНАЛОВ [1,3,16,29].

Важнейшее свойство аналоговых и дискретных сигналов заключается в том, что их линейные комбинации также являются аналоговыми или дискретными сигналами. Линейные комбинации цифровых сигналов, в силу их ограничения по разрядности, в принципе относятся к разряду нелинейных операций, однако последним фактором можно пренебречь, если ошибки, которые вносятся в результаты наблюдений при квантовании отсчетов, достаточно малы по сравнению с шумами зарегистрированной информации. При дискретизации и квантовании данных непосредственно на входах в ЭВМ последнее выполняется практически всегда, поскольку ошибки определяются разрядностью ЭВМ и программными системами обработки данных, которые обычно не ниже 6-12 десятичных разрядов.

Пространство сигналов. Пусть $L\{s_1(t), s_2(t), \dots\}$ - множество сигналов, которые имеют какие-то общие свойства и определенную структуру связи между сигналами. Например, множество может состоять из сигналов вида $s_n(t) = A_n \cos(\omega_n t + \varphi_n) \cdot \exp(-at^2)$ – затухающих гармонических колебаний с определенными значениями амплитуд, частот и начальных фаз. Путем введения структурных ограничений множество сигналов может быть превращено в функциональное пространство сигналов. Так, если пространство значений независимой переменной t задано выражением $R := (-\infty, +\infty)$, то пространство сигналов $L^p[R]$ определяет множество сигналов в этом пространстве, для которых выполняется условие однозначной реализации:

$$\int_{-\infty}^{\infty} |s(t)|^p dt < \infty.$$

Для анализа сигналов наиболее часто используется гильбертово пространство, сигналы в котором должны удовлетворять условию интегрирования с квадратом:

$$\int_{-\infty}^{\infty} |s(t)|^2 dt < \infty.$$

Периодические сигналы обычно рассматриваются в пространстве $L^2[0, 2\pi]$ одного периода:

$$\int_0^{2\pi} |s(t)|^2 dt < \infty.$$

Линейное пространство сигналов. Множество сигналов L образует линейное пространство сигналов, если для него справедливы следующие аксиомы:

1. Для любых сигналов $u(t) \in L$ и $v(t) \in L$ существует их сумма $s(t) = u(t) + v(t)$, которая также содержится в L . Операция суммирования коммутативна: $u(t) + v(t) = v(t) + u(t)$, и ассоциативна: $u(t) + (v(t) + x(t)) = (u(t) + v(t)) + x(t)$.
2. Для любого сигнала $s(t) \in L$ и числа α определен сигнал $y(t) = \alpha s(t)$, $y(t) \in L$.
3. Множество L содержит такой нулевой элемент \emptyset , что для всех сигналов $u(t) \in L$ выполняется равенство $u(t) + \emptyset = u(t)$.

Пример. Множество сигналов L состоит из импульсных сигналов произвольной формы с амплитудой не более 10 вольт. Образуют ли эти сигналы линейное пространство?

Нет, не образуют, так как не выполняется, по крайней мере, первая аксиома линейного пространства (сумма двух сигналов с амплитудой более 5 вольт превышает 10 вольт). Требуются дополнительные структурные ограничения по параметрам сигналов.

Сигналы могут описываться как вещественными, так и комплексными функциями, и линейные пространства также могут быть вещественными или комплексными.

Множество L , для которого выполняются данные аксиомы, при анализе сигналов и систем может рассматриваться как специальным образом сконструированное многомерное (в пределе – бесконечномерное) геометрическое пространство. Сигналы таких линейных пространств называют *векторами* в силу аналогии их свойств со свойствами векторов. Рассмотрим это на конкретном примере.

Представим себе произвольный сигнал $s(t)$, заданный на интервале $[a, b]$. Дискретизируем сигнал с равномерным шагом дискретизации и переведем в цифровую форму (представим N последовательными выборками):

$$s = (s_1, s_2, \dots, s_N).$$

В таком представлении величина s может рассматриваться в виде N -мерного вектора в N -мерном пространстве, в котором значения s_n представляют собой проекции s -вектора на координатные оси данного пространства. Двумерный вектор в двумерном пространстве – это точка с координатами s_1 и s_2 на рис. 2.1.1. Соответственно, в трехмерном пространстве сигнал s представлен точкой в трехмерном пространстве. Представить себе N -мерное пространство при $N > 3$ можно только абстрактно, но с математических позиций такое пространство вполне реально и N -мерный сигнал s отображается вполне определенной точкой в этом пространстве с координатами s_n по осям пространства. При уменьшении интервала дискретизации сигнала до бесконечно малой величины значение N стремится к бесконечности, и пространство сигналов превращается в бесконечномерное пространство аналоговых сигналов. Следовательно, и аналоговые сигналы могут рассматриваться как предельный случай бесконечномерных векторов.

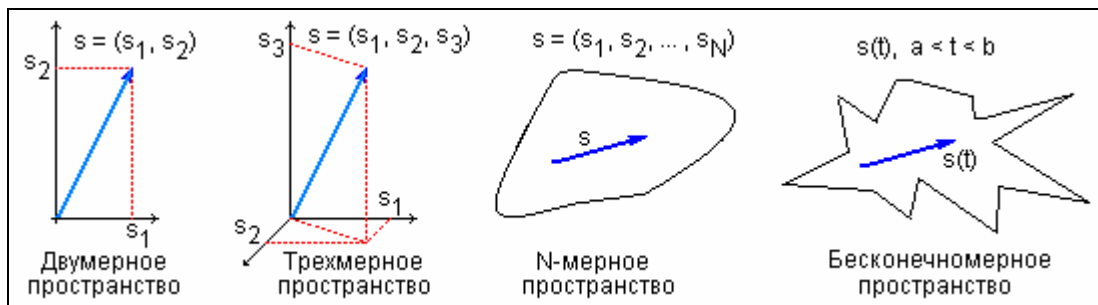


Рис. 2.1.1. Пространства сигналов и функций.

С учетом вышеизложенного, для математического анализа систем и сигналов в линейном пространстве может использоваться математика векторов. Основными метрическими параметрами векторного анализа являются норма, метрика и скалярное произведение сигналов.

Норма сигналов в линейном пространстве является аналогом длины векторов и обо-

значается индексом $\|s(t)\|$ - норма (norm). В математике существуют различные формы норм. При анализе сигналов обычно используются квадратичные нормы:

$$\|s(t)\| = \sqrt{\int_{-\infty}^{\infty} s^2(t) dt} . \quad (2.1.2)$$

Соответственно, для дискретных сигналов:

$$\|s(t)\| = \sqrt{\sum_{n=-\infty}^{\infty} s^2(n)} . \quad (2.1.2')$$

Для комплексных сигналов:

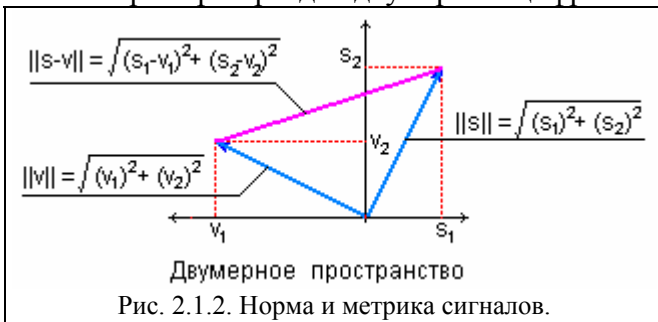
$$\|s(t)\| = \sqrt{\int_{-\infty}^{\infty} s(t) s^*(t) dt} , \quad (2.1.2'')$$

где $s^*(t)$ – величины, комплексно сопряженные с $s(t)$.

Линейное пространство сигналов L является нормированным, если каждому сигналу пространства $s(t)$ однозначно сопоставлена его числовая норма $\|s(t)\|$, и выполняются следующие аксиомы:

1. Норма неотрицательна и равна нулю тогда и только тогда, когда сигнал равен нулю ($\|s(t)\| = 0$, при $s(t) = 0$).
2. Для любого числа b должно быть справедливо равенство: $\|bs(t)\| = |b| \cdot \|s(t)\|$.
3. Если $v(t)$ и $u(t)$ – сигналы из пространства L , то должно выполняться неравенство треугольника: $\|v(t)+u(t)\| \leq \|v(t)\| + \|u(t)\|$.

Пример норм для двумерных цифровых сигналов приведен на рис. 2.1.2.



Метрика сигналов. Линейное пространство сигналов L является метрическим, если каждой паре сигналов $s(t) \in L$ и $v(t) \in L$ однозначно сопоставляется неотрицательное число $\rho(s(t), v(t))$ – метрика (metric) или расстояние между векторами. Пример метрики для двух векторов в двумерном пространстве приведен на рис. 2.1.2.

Для метрик сигналов в метрическом пространстве любой размерности должны выполняться аксиомы:

1. $\rho(s(t), v(t)) = \rho(v(t), s(t))$ – рефлексивность метрики.
2. $\rho(s(t), s(t)) = 0$ для любых $s(t) \in L$.
3. $\rho(s(t), v(t)) \leq \rho(s(t), a) + \rho(a, v(t))$ для любых $a \in L$.

Метрика определяется нормой разности двух сигналов (см. рис. 2.1.2):

$$\rho(s(t), v(t)) = \|s(t) - v(t)\| . \quad (2.1.3)$$

В свою очередь норму можно отождествлять с расстоянием от выбранного элемента пространства до нулевого: $\|s(t)\| = \rho(s(t), 0)$.

По метрике сигналов можно судить, например, о том, насколько точно один сигнал может быть аппроксимирован другим сигналом, или насколько изменяется выходной сигнал относительно входного при прохождении через какое-либо устройство.

Пример. Сигнал на интервале $(0, T)$ представляет собой половину периода синусоиды амплитудой A : $s(t) = A \cdot \sin(\pi t/T)$, $0 \leq t \leq T$. Требуется аппроксимировать сигнал прямоугольным импульсом $p(t)$ (см. рис. 2.1.3).

Если принять амплитуду импульса $p(t)$ равной B , то квадрат расстояния между сигналами: $\rho^2(s, p) = \int_0^T (A \sin(\pi t/T) - B)^2 dt = A^2 T/2 - 4ABT/\pi + B^2 T$.

Для решения задачи требуется найти минимум выражения $\rho^2(s, p)$. Дифференцируем полученное выражение по B , приравняем нулю и, решая относительно B , находим значение экстремума: $B = 2A/\pi \approx 0.64A$. Это искомое значение минимума функции $\rho^2(s, p)$, так как вторая производная функции по B положительна. При этом минимальное значение метрики: $\rho_{\min} \approx 0.31A \sqrt{T}$. Вычислим нормы сигналов при $A = 1$:

$$E_s = A^2 \int_0^T \sin^2(\pi t/T) dt = A^2 T/2 = 10. \text{ Норма: } \|s(t)\| = \sqrt{E_s} = 0.707 A \sqrt{T} \approx 3.16.$$

$$E_p = B^2 \int_0^T dt = B^2 T \approx 8.1. \text{ Норма: } \|p(t)\| = \sqrt{E_p} = B \sqrt{T} \approx 2.85.$$

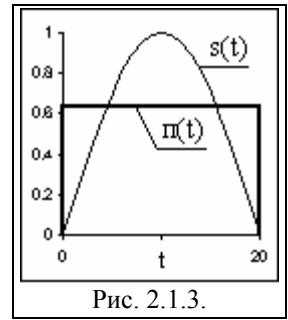


Рис. 2.1.3.

Скалярное произведение произвольных сигналов $u(t)$ и $v(t)$ отражает степень их связи (сходства) по форме и положению в пространстве сигналов, и обозначается как $\langle u(t), v(t) \rangle$.

$$\langle u(t), v(t) \rangle = \|u(t)\| \cdot \|v(t)\| \cos \varphi, \quad (2.1.4)$$

Физическую сущность скалярного произведения векторов в двумерном пространстве можно видеть достаточно наглядно (рис. 2.1.4). Это произведение "длины" (нормы) одного вектора на проекцию второго вектора по "направлению" первого вектора.

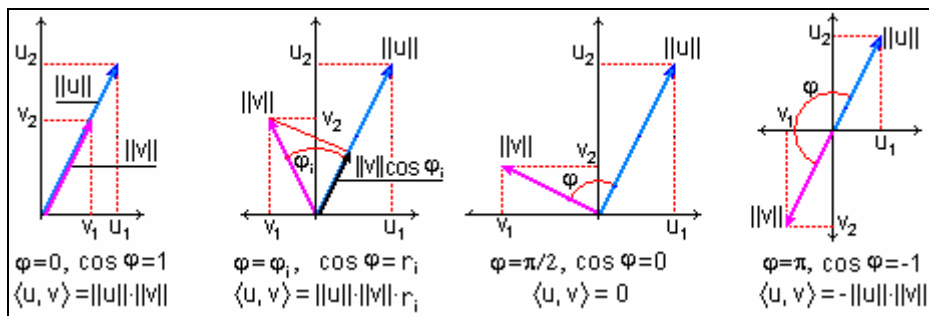


Рис. 2.1.4. Скалярное произведение сигналов в двумерном пространстве.

При кажущейся абстрактности скалярного произведения сигналов оно может приобретать вполне конкретный физический смысл для конкретных физических процессов, которые отображаются этими сигналами. Так, например, если $v = F$ – сила, приложенная к телу, а $u = s$ – перемещение тела под действием этой силы, то скалярное произведение $W = F \cdot s$ определяет выполненную работу, при условии совпадения силы с направлением перемещения. В противном случае, при наличии угла φ между векторами силы и перемещения, работа будет определяться проекцией силы в направлении перемещения, т.е. $W = s \cdot F \cdot \cos \varphi$.

Вычисление скалярного произведения обычно производится непосредственно по сигнальным функциям. Поясним это на примере двумерных сигналов с использованием рисунка 2.1.2. Для квадрата метрики сигналов s и v имеем:

$$\begin{aligned} \|s-v\|^2 &= \|s\|^2 + \|v\|^2 - 2 \|s\| \|v\| \cos \varphi = \|s\|^2 + \|v\|^2 - 2 \langle s, v \rangle. \\ 2 \langle s, v \rangle &= \|s\|^2 + \|v\|^2 - \|s-v\|^2 = (s_1^2 + s_2^2) + (v_1^2 + v_2^2) - \{(s_1 - v_1)^2 + (s_2 - v_2)^2\} = 2 (s_1 v_1 + s_2 v_2). \\ \langle s, v \rangle &= s_1 v_1 + s_2 v_2. \end{aligned}$$

Обобщая полученное выражение на аналоговые сигналы:

$$\langle s(t), v(t) \rangle = \int_{-\infty}^{\infty} s(t)v(t) dt. \quad (2.1.5)$$

Соответственно, для дискретных сигналов в N -мерном пространстве:

$$\langle s_n, v_n \rangle = \sum_{n=1}^N s_n v_n. \quad (2.1.5')$$

Скалярное произведение обладает следующими свойствами:

1. $\langle s, v \rangle \geq 0$;
2. $\langle s, v \rangle = \langle v, s \rangle$;
3. $\langle as, v \rangle = a \langle s, v \rangle$, где a – вещественное число;
4. $\langle s+v, a \rangle = \langle s, a \rangle + \langle v, a \rangle$.

Линейное пространство аналоговых сигналов с таким скалярным произведением называется *гильбертовым пространством* H (второе распространенное обозначение - L_2). Линейное пространство дискретных и цифровых сигналов - *пространством Евклида* (обозначение пространства - R_2). В этих пространствах справедливо фундаментальное неравенство Коши-Буняковского (модуль косинуса в (2.1.4) может быть только равным или меньше 1):

$$|\langle s, v \rangle| \leq \|s\| \cdot \|v\|. \quad (2.1.6)$$

Для комплексного гильбертова пространства скалярное произведение вычисляется по формуле:

$$\langle s(t), v(t) \rangle = \int_{-\infty}^{\infty} s(t)v^*(t) dt. \quad (2.1.7)$$

При определении функций в пространстве $L^2[a, b]$ вычисление скалярного произведения производится соответственно с пределами интегрирования от a до b .

Из выражения (2.1.4) следует косинус угла между сигналами:

$$\cos \varphi = \langle s(t), v(t) \rangle / (\|s\| \cdot \|v\|). \quad (2.1.8)$$

Пример. Имеется два смещенных во времени прямоугольных импульса с одинаковой амплитудой и длительностью: $s_1(t) = 2$ при $0 \leq t \leq 5$, $s_1(t) = 0$ при других t ; и $s_2(t) = 2$ при $4 \leq t \leq 9$, $s_2(t) = 0$ при других t .

Квадраты норм сигналов: $\|s_1\|^2 = \int_0^5 s_1^2(t) dt = 20$. $\|s_2\|^2 = \int_4^9 s_2^2(t) dt = 20$

Скалярное произведение: $\langle s_1, s_2 \rangle = \int_4^5 s_1(t) s_2(t) dt = 8$.

Отсюда имеем: $\cos \varphi = \langle s_1, s_2 \rangle / (\|s_1\| \cdot \|s_2\|) = 8/20 = 0.4$ и $\varphi \approx 1.16$ радиан $\approx 66^\circ$

При полном совмещении сигналов: $\langle s_1, s_2 \rangle = \int_t^{t+5} s_1(t) s_2(t) dt = 20$, $\cos \varphi = 1$, $\varphi = 0$.

При отсутствии перекрытия сигналов; $\langle s_1, s_2 \rangle = 0$, $\cos \varphi = 0$, $\varphi = 90^\circ$.

Физическое понятие "угла" между многомерными сигналами довольно абстрактно. Однако при рассмотрении выражения (2.1.8) совместно с выражением для квадрата метрики сигналов

$$\rho^2(s, v) = \int_{-\infty}^{\infty} [s(t) - v(t)]^2 dt = \|s\|^2 + \|v\|^2 - 2 \cdot \|s\| \cdot \|v\| \cos \varphi.$$

можно отметить следующие закономерности. При $\varphi = 0$ ($\cos \varphi = 1$) сигналы "совпадают по направлению" и расстояние между ними минимально. При $\varphi = \pi/2$ ($\cos \varphi = 0$) сигналы "перпендикулярны друг другу" (иначе говоря – ортогональны), и проекции сигналов друг на друга равны 0. При $\varphi = \pi$ ($\cos \varphi = -1$) сигналы "противоположны по направлению" и расстояние между сигналами максимально. Фактор расстояния между сигналами играет существенную роль при их селекции в многоканальных системах.

Коэффициент корреляции сигналов. Одновременно заметим, что значение косинуса в (2.1.8) изменяется от 1 до -1, и не зависит от нормы сигналов ("длины" векторов). Максимальное значение $\cos \varphi = 1$ соответствует полной тождественности относительной динамики сигналов, минимальное значение $\cos \varphi = -1$ наблюдается при полной противоположности значений относительной динамики сигналов. По существу, коэффициент $r = \cos \varphi$ является

интегральным коэффициентом степени сходства формы сигналов по пространству их задания. С учетом этого он и получил название *коэффициента корреляции сигналов*. На рис. 2.1.5 можно наглядно видеть значения коэффициента корреляции двух сигналов в зависимости от их формы и сдвига по независимой переменной.

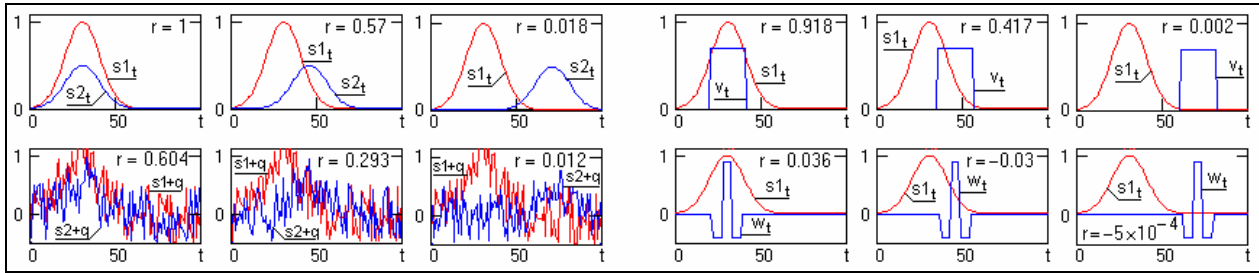


Рис. 2.1.5. Коэффициент корреляции сигналов.

Однако количественные значения коэффициентов корреляции существенно зависят от выбора нулевой точки сигнального пространства. Рассмотрим это более детально на конкретном примере.

На рис. 2.1.6 приведено изменение средней месячной температуры воздуха в трех городах земного шара в течение одного календарного года. Характер корреляции между изменениями температур в городах достаточно хорошо виден на графиках. Вычислим (см. пример ниже) значения коэффициентов корреляции для шкалы температур по Цельсию.

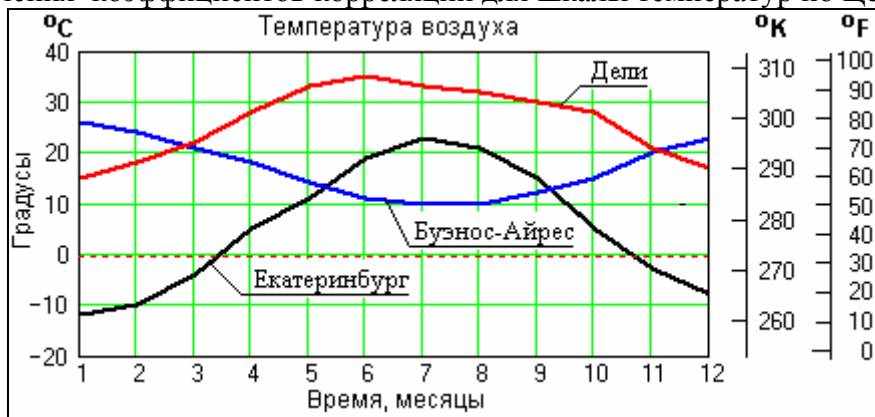


Рис. 2.4.1.

Пример. Среднемесячная температура воздуха в городах по Цельсию:

Екатеринбург: $E_k = \{-12, -10, -4, 5, 11, 19, 23, 21, 15, 5, -3, -8\}$. Дели: $D_k = \{15, 18, 22, 28, 33, 35, 33, 32, 30, 28, 21, 17\}$.

Буэнос - Айрес: $B_k = \{26, 24, 21, 18, 14, 11, 10, 10, 12, 15, 20, 23\}$. Нумерация месяцев: $k = 1, 2, 3, \dots, 12$.

Норма сигналов: $\|E\| = \sqrt{\sum_k (E_k)^2} = 45.39$, $\|D\| = \sqrt{\sum_k (D_k)^2} = 93.05$, $\|B\| = \sqrt{\sum_k (B_k)^2} = 61.9$.

Скалярные произведения: $\langle E, D \rangle = \sum_k (E_k D_k) = 2542$, $\langle E, B \rangle = 268$, $\langle B, D \rangle = 4876$.

Коэффициенты корреляции: Екатеринбург – Дели: $r_{ED} = \langle E, D \rangle / (\|E\| \|D\|) = 0.602$.

Екатеринбург – Буэнос-Айрес: $r_{EB} = 0.095$, Дели – Буэнос-Айрес: $r_{DB} = 0.847$.

Как следует из вычислений, полученные коэффициенты корреляции маловыразительны. Практически не регистрируется разнонаправленная корреляция Екатеринбург - Буэнос-Айрес, и не различаются одно- (Екатеринбург – Дели) и разнонаправленные (Дели – Буэнос-Айрес) типы корреляции.

Повторим вычисления в шкале Фаренгейта ($0^\circ\text{F} = -17,8^\circ\text{C}$, $100^\circ\text{F} = +37,8^\circ\text{C}$), и в абсолютной шкале температур Кельвина. Дополнительно вычислим значения коэффициентов корреляции в шкале Цельсия и Фаренгейта для центрированных сигналов. Центрированный сигнал вычисляется путем определения среднего значения сигнала по интервалу его задания и вычитания этого среднего значения из исходных значений сигнала, т.е. среднее значение центрированного сигнала равно нулю. Сводные результаты вычислений приведены в таблице.

Таблица 2.1.1.

Коэффициенты корреляции сигналов

Пары городов	Нецентрированные сигналы			Центрированные сигналы	
	Цельсий	Фаренгейт	Кельвин	Цельсий	Фаренгейт
Екатеринбург – Дели	0.602	0.943	1	0.954	0.954
Екатеринбург – Буэнос-Айрес	0.095	0.803	0.998	-0.988	-0.988
Дели – Буэнос-Айрес	0.847	0.953	0.999	-0.960	-0.960

Как видно из таблицы, значения коэффициента корреляции нецентрированных сигналов существенно зависят от положения сигналов относительно нулевой точки пространства. При одностороннем смещении сигналов относительно нуля (шкала Фаренгейта) значение коэффициента корреляции может быть только положительным, и тем ближе к 1, чем дальше от сигналов нулевая точка (шкала Кельвина), т.к. при больших значениях сигналов-векторов значение скалярного произведения сигналов стремится к значению произведения норм сигналов.

Для получения значений коэффициентов корреляции, независимых от нуля сигнального пространства и от масштаба единиц измерений, необходимо вычислять коэффициент по центрированным сигналам, при этом в оценках коэффициента, как это видно из результатов, приведенных в таблице, появляется знаковый параметр совпадения (или несовпадения) по "направлению" корреляции и исчезает зависимость от масштаба представления сигналов. Это позволяет вычислять коэффициенты корреляции различных сигналов вне зависимости от физической природы сигналов и их величины.

Координатный базис пространства. Для измерения и отображения одномерных величин достаточно одного нормированного параметра – стандарта величины или единицы ее измерения (для измерения длины – сантиметр, для измерения тока – ампер, и т.п.).

В пространстве сигналов роль такого метрологического стандарта выполняет координатный базис пространства - подмножество векторов $\{e_1, e_2, e_3, \dots\}$ со свойствами ортогональных координатных осей, по которым можно разложить любой произвольный сигнал, принадлежащий этому линейному пространству.

Совокупность векторов e_i пространства L является линейно независимой и образует координатный базис пространства, если равенство $\sum_i a_i e_i = \emptyset$ выполняется только в случае одновременного обращения в нуль всех числовых коэффициентов a_i . При этом произвольный сигнал $s(t)$ может быть разложен по координатному базису e_i в виде

$$s(t) = \sum_i c_i e_i, \quad (2.1.9)$$

где числа c_i – проекции сигнала $s(t)$ на координатный базис.

Число базисных векторов определяет размерность векторного пространства. Так, для двумерных векторов в качестве ортогонального базиса пространства могут быть приняты векторы $\{v_1, v_2\}$, если выполняется условие их взаимной перпендикулярности – нулевое значение скалярного произведения $\langle v_1, v_2 \rangle = 0$. При $\|v_1\| = \|v_2\| = 1$ эта пара векторов является ортонормированным базисом с единичными векторами координатных осей в качестве стандарта (единицы измерения) пространства.

Пример. Могут ли быть приняты в качестве координатного базиса двумерного пространства векторы

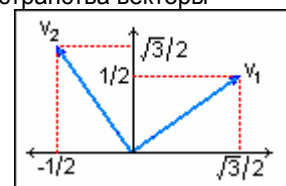
$$v_1 = (\sqrt{3}/2, 1/2), \quad v_2 = (-1/2, \sqrt{3}/2).$$

$$\langle v_1, v_2 \rangle = (\sqrt{3}/2) \cdot (-1/2) + (1/2) \cdot (\sqrt{3}/2) = 0. \quad \text{Векторы ортогональны.}$$

$$\|v_1\| = \sqrt{\left(\frac{\sqrt{3}}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = 1.$$

$$\|v_2\| = \sqrt{\left(-\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2} = 1. \quad \text{Векторы нормированы.}$$

Векторы могут быть ортонормированным базисом пространства.



Разложение произвольного двумерного вектора - сигнала s в двумерном пространстве

по координатным осям элементарно:

$$s = c_1 v_1 + c_2 v_2, \quad (2.1.10)$$

где коэффициенты c_1 и c_2 выражают значения составляющих вектора s по направлениям векторов v_1 и v_2 , т.е. являются проекциями вектора s на координатный базис пространства $\{v_1, v_2\}$. Значения проекций определяются скалярными произведениями:

$$c_1 = \langle s, v_1 \rangle, \quad c_2 = \langle s, v_2 \rangle.$$

В последнем нетрудно убедиться, если вычислить скалярные произведения левой и правой части выражения (2.1.10) сначала с вектором v_1 :

$$\langle s, v_1 \rangle = \langle (c_1 v_1 + c_2 v_2), v_1 \rangle = \langle c_1 v_1, v_1 \rangle + \langle c_2 v_2, v_1 \rangle = c_1 \langle v_1, v_1 \rangle + c_2 \langle v_2, v_1 \rangle.$$

При ортонормированности базиса $\{v_1, v_2\}$ имеем:

$$\langle v_1, v_1 \rangle = \|v_1\|^2 = 1, \quad \langle v_2, v_1 \rangle = 0.$$

Отсюда следует: $\langle s, v_1 \rangle = c_1$. Аналогичным образом можно получить и выражение для значения $c_2 = \langle s, v_2 \rangle$.

Пример. Разложить вектор $s = (\sqrt{3}/2, 5/2)$ по базису, представленному векторами

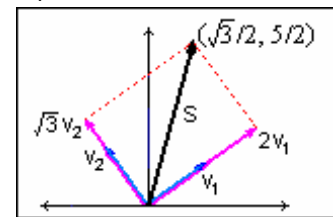
$$v_1 = (\sqrt{3}/2, 1/2) \text{ и } v_2 = (-1/2, \sqrt{3}/2) \text{ из предыдущего примера.}$$

$$s = c_1 v_1 + c_2 v_2.$$

$$c_1 = \langle s, v_1 \rangle = (\sqrt{3}/2) \cdot (\sqrt{3}/2) + (5/2) \cdot (1/2) = 2.$$

$$c_2 = \langle s, v_2 \rangle = (\sqrt{3}/2) \cdot (-1/2) + (5/2) \cdot (\sqrt{3}/2) = \sqrt{3}.$$

Результат: В пространстве с базисом $\{v_1, v_2\}$ вектор s однозначно определяется двумя векторами v_1 и $\sqrt{3}$ векторами v_2 . $s = 2v_1 + \sqrt{3}v_2$.



2.2. МОЩНОСТЬ И ЭНЕРГИЯ СИГНАЛОВ [1,3,16].

Понятия мощности и энергии в теории сигналов не относятся к характеристикам каких-либо физических величин сигналов, а являются их количественными характеристиками, отражающими определенные свойства сигналов и динамику изменения их значений (отсчетов) во времени, в пространстве или по любым другим аргументам.

Для произвольного, в общем случае комплексного, сигнала $s(t) = a(t) + jb(t)$, где $a(t)$ и $b(t)$ - вещественные функции, мгновенная мощность (instantaneous power) сигнала *по определению* задается выражением:

$$w(t) = s(t) s^*(t) = [a(t) + jb(t)] [a(t) - jb(t)] = a^2(t) + b^2(t) = |s(t)|^2, \quad (2.2.1)$$

т.е. функция распределения мгновенной мощности по аргументу сигнала равна квадрату функции его модуля, для вещественных сигналов - квадрату функции амплитуд.

Аналогично для дискретных сигналов:

$$w_n = s_n s_n^* = [a_n + jb_n] [a_n - jb_n] = a_n^2 + b_n^2 = |s_n|^2, \quad (2.2.1')$$

Энергия сигнала (также по определению) равна интегралу от мощности по всему интервалу существования или задания сигнала. В пределе:

$$E_s = \int_{-\infty}^{\infty} w(t) dt = \int_{-\infty}^{\infty} |s(t)|^2 dt. \quad (2.2.2)$$

$$E_s = \sum_{n=-\infty}^{\infty} w_n = \sum_{n=-\infty}^{\infty} |s_n|^2. \quad (2.2.2')$$

Мгновенная мощность $w(t)$ является плотностью мощности сигнала, так как измерения мощности возможны только через энергию на интервалах ненулевой длины:

$$w(\tau) = (1/\Delta t) \int_{\tau - \Delta t/2}^{\tau + \Delta t/2} |s(t)|^2 dt.$$

Энергия сигналов может быть конечной или бесконечной. Конечную энергию имеют финитные сигналы и сигналы, затухающие по своим значениям в пределах конечной длительности, которые не содержат дельта-функций и особых точек (разрывов второго рода и

ветвей, уходящих в бесконечность). В противном случае их энергия равна бесконечности. Бесконечна также энергия периодических сигналов.

Как правило, сигналы изучаются на определенном интервале T , для периодических сигналов - в пределах одного периода T , при этом средняя мощность (average power) сигнала:

$$W_T(\tau) = (1/T) \int_{\tau}^{\tau+T} w(t) dt = (1/T) \int_{\tau}^{\tau+T} |s(t)|^2 dt. \quad (2.2.3)$$

Понятие средней мощности может быть распространено и на незатухающие сигналы, энергия которых бесконечно велика. В случае неограниченного интервала T строго корректное определение средней мощности сигнала должно производиться по формуле:

$$W_s = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T w(t) dt. \quad (2.2.3')$$

Квадратный корень из значения средней мощности характеризует *действующее* (среднеквадратическое) значение сигнала (root mean square, RMS).

Применительно к электрофизическим системам, данным понятиям мощности и энергии соответствуют вполне конкретные физические величины. Допустим, что функцией $s(t)$ отображается электрическое напряжение на резисторе, сопротивление которого равно R Ом. Тогда рассеиваемая в резисторе мощность, как известно, равна (в вольт-амперах):

$$w(t) = |s(t)|^2/R,$$

а полная выделенная на резисторе тепловая энергия определяется соответствующим интегрированием мгновенной мощности $w(t)$ по интервалу задания напряжения $s(t)$ на резисторе R . Физическая размерность мощности и энергии в этом случае определяется соответствующей физической размерностью функции напряжения $s(t)$ и сопротивления резистора R . Для безразмерной величины $s(t)$ при $R=1$ это полностью соответствует выражению (2.2.1). В теории сигналов в общем случае сигнальные функции $s(t)$ не имеют физической размерности, и могут быть формализованным отображением любого процесса или распределения какой-либо физической величины, при этом понятия энергии и мощности сигналов используются в более широком смысле, чем в физике. Они представляют собой специфические метрологические характеристики сигналов.

Из сравнения выражений (2.1.2) и (2.2.2) следует, что энергия и норма сигнала связаны соотношениями:

$$E_s = \|s(t)\|^2, \quad \|s(t)\| = \sqrt{E_s} \quad (2.2.4)$$

Пример. Цифровой сигнал задан функцией $s(n) = \{0, 1, 2, 3, 4, 5, 4, 3, 2, 1, 0, 0, 0, 0, \dots\}$.

Энергия сигнала: $E_s = \sum_{n=0}^N s^2(n) = 1+4+9+16+25+16+9+4+1 = 85$. Норма: $\|s(n)\| = \sqrt{85} \approx 9.22$

Вычислим энергию суммы двух произвольных сигналов $u(t)$ и $v(t)$:

$$E = \int_{-\infty}^{\infty} [u(t)+v(t)]^2 dt = E_u + E_v + 2 \int_{-\infty}^{\infty} u(t)v(t) dt. \quad (2.2.5)$$

Как следует из этого выражения, энергии сигналов (а равно и их мощности), в отличие от самих сигналов, в общем случае не обладают свойством аддитивности. Энергия суммарного сигнала $u(t)+v(t)$, кроме суммы энергий составляющих сигналов, содержит в себе и так называемую энергию взаимодействия сигналов или взаимную энергию:

$$E_{uv} = 2 \int_{-\infty}^{\infty} u(t)v(t) dt. \quad (2.2.6)$$

Нетрудно заметить, что энергия взаимодействия сигналов равна их удвоенному скалярному произведению:

$$E_{uv} = 2 \langle u(t), v(t) \rangle. \quad (2.2.6')$$

При обработке данных используются также понятия мощности взаимодействия двух

сигналов $x(t)$ и $y(t)$:

$$\begin{aligned} w_{xy}(t) &= x(t) y^*(t), \\ w_{yx}(t) &= y(t) x^*(t), \\ w_{xy}(t) &= w_{yx}^*(t). \end{aligned} \quad (2.2.7)$$

Для вещественных сигналов:

$$w_{xy}(t) = w_{yx}(t) = x(t) y(t). \quad (2.2.8)$$

С использованием выражений (2.2.7-8) интегрированием по соответствующим интервалам вычисляются значения средней мощности взаимодействия сигналов на определенных интервалах T и энергия взаимодействия сигналов.

2.3. ПРОСТРАНСТВА ФУНКЦИЙ [1,3,11,16,29].

Пространства функций можно считать обобщением пространства N -мерных сигналов – векторов на аналоговые сигналы, как бесконечномерные векторы, с некоторыми чисто практическими уточнениями.

Нормирование метрических параметров. Норма функций в пространстве $L^2[a, b]$ определяется выражением:

$$\|s(t)\| = \sqrt{\int_a^b s^2(t) dt}.$$

Нетрудно заключить, что чем больше интервал $[a, b]$ в этой формуле, тем больше (при прочих равных условиях) будет значение нормы. При анализе и сравнении сигналов (как аналоговых, так и многомерных дискретных) такое понятие не всегда удобно, и вместо него очень часто используют понятие нормы, нормированной относительно длины интервала $[a, b]$. Для символического обозначения нормирования будем применять знак ∇ :

$$\|s(t)\|^\nabla = \sqrt{\frac{1}{b-a} \int_a^b s^2(t) dt}, \quad \|s_n\|^\nabla = \sqrt{\frac{1}{N} \sum_{n=1}^N s_n}.$$

Метрика сигналов (расстояние между сигналами) при аналогичном нормировании:

$$d^\nabla(s(t), v(t)) = \sqrt{\frac{1}{b-a} \int_a^b [s(t) - v(t)]^2 dt}, \quad d^\nabla(s_n, v_n) = \sqrt{\frac{1}{N} \sum_{n=1}^N (s_n - v_n)^2}$$

Эти выражения применяются для вычисления среднеквадратического расхождения сигналов или среднеквадратической погрешности выполнения какой-либо операции при сравнении ее результата с теоретически ожидаемым или априорно известным.

Нормированное скалярное произведение сигналов:

$$\langle s(t), v(t) \rangle^\nabla = \frac{1}{b-a} \int_a^b s(t)v(t) dt = \|s(t)\|^\nabla \|v(t)\|^\nabla \cos \varphi.$$

$$\langle s_n, v_n \rangle^\nabla = (1/N) \sum_{n=1}^N s_n v_n = \|s_n\|^\nabla \|v_n\|^\nabla \cos \varphi.$$

Косинус угла (коэффициент корреляции) между сигналами – функциями не изменяет своих значений при вычислении как по нормированным, так и по ненормированным значениям скалярного произведения и нормы сигналов (значения нормировки в числителе и знаменателе выражения (2.1.8) сокращаются). Взаимная перпендикулярность функций определяется аналогично взаимной перпендикулярности векторов условием нулевого значения скалярного произведения.

Норма, метрика и скалярное произведение периодических функций обычно нормируются на длину главного периода T .

Ортогональные сигналы. Два сигнала называются ортогональными (orthogonal), если имеют нулевое скалярное произведение:

$$\langle u(t), v(t) \rangle = \int_a^b u(t)v(t) dt = 0.$$

Соответственно, два таких сигнала в своем функциональном пространстве являются

взаимно перпендикулярными (угол между сигналами равен $\varphi = 90^\circ$), полностью независимыми друг от друга (некоррелированными, $r = \cos \varphi = 0$), и имеют нулевую энергию взаимодействия ($E_{uv} = 0$).

На рисунке 2.3.1 приведены примеры взаимно ортогональных сигналов. Нулевое скалярное произведение двух левых сигналов обеспечивается их формой (равна нулю сумма положительных и отрицательных значений произведения сигналов), а двух правых - взаимным расположением (ненулевые значения сигналов не имеют общих координат).

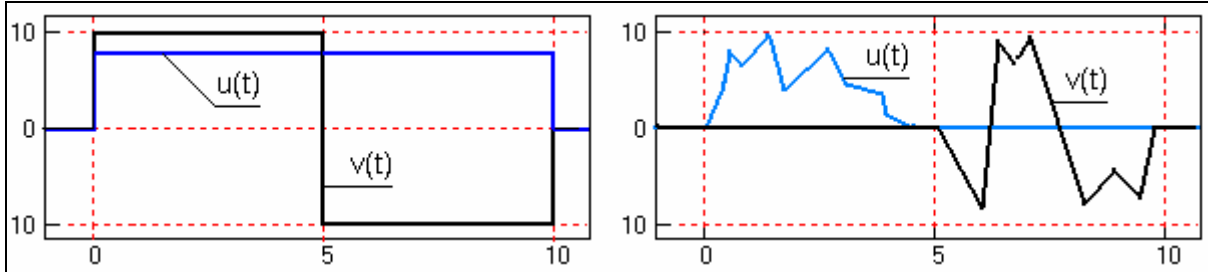


Рис. 2.3.1. Ортогональные сигналы.

Попутно заметим, что энергия и мощность суммы ортогональных сигналов обладают свойством аддитивности, т.к. имеют нулевое значение скалярного произведения и, соответственно, энергии взаимодействия.

Ортонормированный базис пространства. Множество сигналов – векторов $\{v_k, k = 1, 2, \dots, N\}$ в N -мерном декартовом пространстве при единичной норме и выполнении условий взаимной ортогональности:

$$\langle v_m, v_n \rangle = \begin{cases} 0; & m \neq n \\ 1; & m = n \end{cases} \quad (2.3.1)$$

могут быть приняты в качестве ортонормированного базиса данного пространства. Выражение (2.3.1) обычно записывается в следующей форме:

$$\langle v_m, v_n \rangle = \delta_{mn}, \quad (2.3.1')$$

где δ_{mn} – импульс Кронекера, равный правой части выражения (2.3.1).

С использованием ортонормированного базиса любой произвольный сигнал можно представить в виде линейной комбинации взвешенных базисных векторов:

$$s = c_1 v_1 + c_2 v_2 + \dots + c_N v_N,$$

где весовое значение c_k определяется проекцией вектора s на соответствующее координатное направление:

$$c_k = \langle s, v_k \rangle.$$

При распространении данных положений на функциональное пространство $L^2[a, b]$ в качестве координатного базиса пространства мы должны использовать совокупность функций $\{u_0(t), u_1(t), u_2(t), \dots\}$, в пределе – бесконечную, которая должна быть *системой ортогональных функций* $\{u_k(t), k=0, 1, 2, \dots\}$, т.е. все функции на этом отрезке должны быть взаимно ортогональны:

$$\langle u_m(t), u_n(t) \rangle = \int_a^b u_m(t) u_n(t) dt = 0, \quad m = 1, 2, \dots; \quad n = 1, 2, \dots; \quad m \neq n.$$

Система ортогональных функций на интервале $[a, b]$ будет *ортонормированной* (orthonormal functions), если все функции системы при $m=n$ имеют единичную норму, т.е. выполняются условия:

$$\langle u_m(t), u_m(t) \rangle = \|u_m(t)\|^2 = \int_a^b (u_m(t))^2 dt = 1, \quad \|u_m(t)\| = 1, \quad m = 1, 2, \dots$$

Эти условия можно записать в следующей обобщенной форме:

$$\int_a^b u_m(t) \cdot u_n^*(t) dt = \delta_{m,n}.$$

Система ортогональных функций всегда может быть превращена в ортонормирован-

ную путем нормировки, т.е. деления всех функций на их норму.

Разложение сигнала в ряд. Произвольный сигнал $s(t) \in H$ (пространство Гильберта), заданный на интервале $[a, b]$, может быть разложен в ряд по упорядоченной системе ортонормированных базисных функций $u_k(t)$:

$$s(t) = \sum_{k=0}^{\infty} c_k u_k(t). \quad (2.3.2)$$

Для нахождения значений коэффициентов c_k умножим обе части данного выражения на базисную функцию $u_m(t)$ с произвольным номером m и проинтегрируем результаты по переменной t , при этом получим:

$$\int_a^b s(t)u_m(t) dt = \sum_{k=0}^{\infty} c_k \int_a^b u_m u_k dt.$$

С учетом ортонормированности функций $u_i(t)$, в правой части этого равенства остается только один член суммы с номером $m = k$ при $\int_a^b u_k u_k dt = 1$, который, по левой части уравнения, представляет собой скалярное произведение сигнала и соответствующего $m = k$ базисного вектора, т.е. проекцию сигнала на соответствующее базисное направление:

$$c_k = \int_a^b s(t)u_k(t) dt. \quad (2.3.2)$$

Таким образом, в геометрической интерпретации коэффициенты c_k представляют собой проекции вектор - сигнала $s(t)$ на соответствующие базисные направления $u_k(t)$, т.е. координаты вектора $s(t)$ по координатному базису, образованному системой ортогональных функций $u(t)$, в пределе - бесконечномерной. При практическом использовании количество членов ряда (2.3.2) ограничивается определенным значением N , при этом для любого значения N совокупность коэффициентов c_k обеспечивают наименьшее по средней квадратической погрешности приближение к заданному сигналу.

Соответственно, энергия взаимодействия двух сигналов $x(t)$ и $y(t)$ может вычисляться по скалярному произведению их координатных проекций, которое, с учетом взаимной ортогональности всех проекций, будет равно:

$$\langle x(t), y(t) \rangle = \int_a^b x(t)y(t) dt = \int_a^b \left[\sum_{n=0}^{\infty} a_n u_n(t) \right] \left[\sum_{m=0}^{\infty} b_m u_m(t) \right] dt = \sum_{n=0}^{\infty} a_n \cdot b_n. \quad (2.3.3)$$

Косинус угла между векторами $x(t)$ и $y(t)$ с использованием выражения (2.3.3):

$$\cos \varphi = \sum_{n=0}^{\infty} a_n \cdot b_n / (||x(t)|| \cdot ||y(t)||).$$

Возможность разложения непрерывных сигналов и функций в обобщенные ряды по системам ортогональных функций имеет огромное принципиальное значение, так как позволяет вместо изучения несчетного множества точек сигнала ограничиться счетной системой коэффициентов ряда.

К системам базисных функций, которые используются при разложении сигналов, предъявляются следующие основные требования:

- для любого сигнала ряды разложения должны сходиться;
- при ограничении ряда по уровню остаточной погрешности расхождения с заданным сигналом количество членов ряда должно быть минимальным;
- функции должны иметь достаточно простую аналитическую форму;
- коэффициенты разложения должны вычисляться относительно просто.

Согласно теореме Дирехле, любой сигнал $s(t)$, имеющий конечное число точек нарушения непрерывности первого рода и конечный по энергии на интервале $[a, b]$, может быть разложен по системе ортонормальных функций, если существуют интегралы модуля сигнала и модуля его первой производной, т.е.:

$$\int_a^b |s(t)| dt < \infty, \quad \int_a^b |s'(t)| dt < \infty.$$

Ортонормированные системы функций хорошо известны в математике. Это полиномы Эрмита, Лежандра, Чебышева, функции Бесселя, Лагерра и целый ряд других. Выбор типа функций в качестве координатного базиса сигнального пространства, как и координатных осей для обычного трехмерного пространства (декартовы, цилиндрические, сферические и пр.), определяется удобством и простотой последующего использования при математической обработке сигналов. При спектральном анализе сигналов используются, в основном, два вида ортонормированных функций: гармонические функции и функции Уолша.

На интервале $[-\pi, \pi]$ рассмотрим систему следующих гармонических функций:

$$\{1, \sin t, \sin 2t, \dots, \sin kt\}, \quad k = 1, 2, 3, \dots \quad (2.3.4)$$

Вычислим нормированные на интервал скалярные произведения системы:

$$\langle 1, \sin kt \rangle = (1/2\pi) \int_{-\pi}^{\pi} \sin kt dt = (1/2k\pi) [\cos k\pi - \cos(-k\pi)] = 0, \quad k = 1, 2, 3, \dots$$

$$\begin{aligned} \langle \sin mt, \sin nt \rangle &= (1/2\pi) \int_{-\pi}^{\pi} \sin mt \sin nt dt = (1/4\pi) \int_{-\pi}^{\pi} \{\cos(m+n)t - \cos(m-n)t\} dt = \\ &= \frac{1}{4\pi(m+n)} [\sin(m+n)t]_{-\pi}^{\pi} + \frac{1}{4\pi(m-n)} [\sin(m-n)t]_{-\pi}^{\pi} = 0, \quad \text{при } m \neq n. \end{aligned}$$

Следовательно, система (2.3.4) является системой взаимно ортогональных функций.

Норма функций:

$$\|\sin kt\|^2 = (1/2\pi) \int_{-\pi}^{\pi} \sin^2 kt dt = (1/4\pi) \int_{-\pi}^{\pi} (1 - \cos 2kt) dt = \frac{1}{4\pi} [t]_{-\pi}^{\pi} + \frac{1}{8k\pi} [\sin 2kt]_{-\pi}^{\pi} = 1/2.$$

$$\|\sin kt\| = 1/\sqrt{2}, \quad k = 1, 2, 3, \dots$$

Соответственно, для превращения системы (2.3.4) в ортонормированную следует разделить все функции системы на значение нормы (рис. 2.3.2):

$$\{1, u_k(t) = \sqrt{2} \sin kt\}, \quad k = 1, 2, 3, \dots \quad (2.3.4')$$

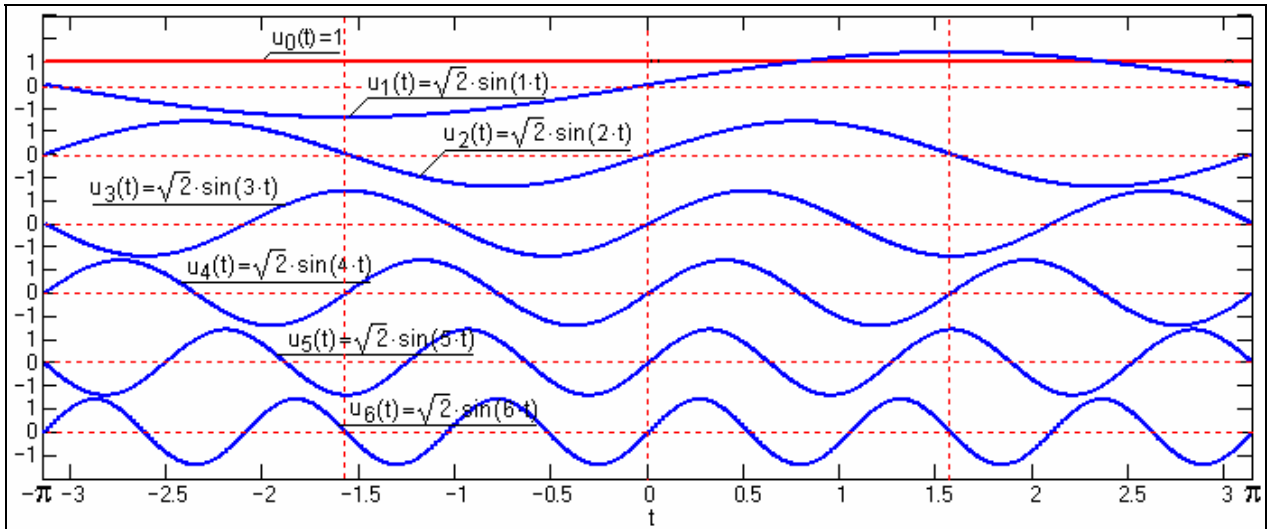


Рис. 2.3.2. Ортонормированный базис гармонических функций.

Аналогичным образом можно убедиться в ортонормированности косинусной системы гармонических функций:

$$\{1, u_k(t) = \sqrt{2} \cos kt\}, \quad k = 1, 2, 3, \dots, \quad (2.3.5)$$

и объединенной синус-косинусной системы:

$$\{1, u_k(t) = \sqrt{2} \sin kt, u_k(t) = \sqrt{2} \cos kt\}, \quad k = 1, 2, 3, \dots \quad (2.3.6)$$

Наибольшее распространение в качестве базисных функций частотного разложения нашли комплексные экспоненциальные функции $\exp(pt)$ при $p = jf$ (преобразование Фурье) и $p = s + jf$ (преобразование Лапласа), от которых с использованием формул Эйлера

$$\begin{aligned} \exp(j\omega t) &= \cos(\omega t) + j \sin(\omega t), & \exp(-j\omega t) &= \cos(\omega t) - j \sin(\omega t), \\ \cos(\omega t) &= [\exp(j\omega t) + \exp(-j\omega t)]/2, & \sin(\omega t) &= [\exp(j\omega t) - \exp(-j\omega t)]/2j \end{aligned}$$

всегда можно перейти к синус-косинусным функциям. Термин "частотного" обязан своим происхождением независимой переменной частотного представления сигналов, которая измеряется в единицах, обратных единицам времени, т.е. в единицах частоты $f = 1/|t|$. Однако понятие частотного преобразования не следует связывать только с временным представлением сигналов, т.к. математический аппарат преобразования не зависит от физического смысла переменных. Так, например, при переменной "x", как единице длины, значение f будет представлять собой пространственную частоту - число периодических изменений сигнала на единице длины с размерностью $1/|x|$.

Ортонормированная система функций Уолша, по существу, является предельной модификацией системы периодических функций с кратными частотами, при этом функции принимают значения только ± 1 . Пример четырех первых функций Уолша на интервале T от $-0,5$ до $0,5$ приведен на рис. 2.3.3. Ортогональность и нормированность функций следует из принципа их построения. Стандартное математическое обозначение функций Уолша: $wal(k, x)$, где $k = 0, 1, 2, \dots$ – порядковый номер функции, $x = t/T$ – безразмерная координата (нормированная на интервал T независимая переменная).

Наряду с функциями Уолша применяются также две связанные с ними системы: четные и нечетные функции $cal(n, x) = wal(2n, x)$, – аналогичные косинусам, и $sal(n, x) = wal(2n-1, x)$, – аналогичные синусам.

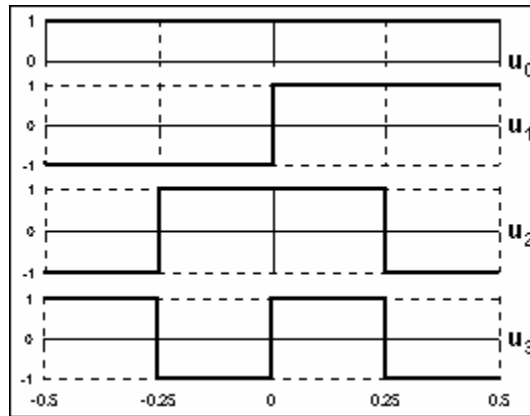


Рис. 2.3.3. Функции Уолша.

При разложении сигналов форма спектров Уолша практически тождественна спектрам гармонических функций.

Разложение энергии сигнала. Допустим, что сигнал $s(t)$ разложен в обобщенный ряд Фурье (2.3.2). Вычислим энергию сигнала непосредственной подстановкой выражения (2.3.2) в выражение (2.2.2):

$$E_s = \int_{-\infty}^{\infty} s^2(t) dt = \int_{-\infty}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_m c_n u_m u_n dt = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_m c_n \int_{-\infty}^{\infty} u_m u_n dt. \quad (2.3.7)$$

В этом выражении в силу ортонормированности базисной системы отличны от нуля только члены с номерами $m = n$. Отсюда:

$$E_s = \int_{-\infty}^{\infty} s^2(t) dt = \sum_{n=0}^{\infty} c_n^2, \quad (2.3.8)$$

т.е. при корректном разложении сигнала в обобщенный ряд Фурье энергия сигнала не изменяется, и равна сумме энергии всех составляющих ряда. Это соотношение называют равенством Парсевалья.

2.4. ФУНКЦИИ КОРРЕЛЯЦИИ СИГНАЛОВ [1, 25, 29].

Функции корреляции сигналов применяются для интегральных количественных оценок формы сигналов и степени их сходства друг с другом.

Автокорреляционные функции (АКФ) сигналов (correlation function, CF). Применительно к детерминированным сигналам с конечной энергией АКФ является количественной интегральной характеристикой формы сигнала, и представляет собой интеграл от произведения двух копий сигнала $s(t)$, сдвинутых относительно друг друга на время τ :

$$B_s(\tau) = \int_{-\infty}^{\infty} s(t) s(t+\tau) dt. \quad (2.4.1)$$

Как следует из этого выражения, АКФ является скалярным произведением сигнала и его копии в функциональной зависимости от переменной величины значения сдвига τ . Соответственно, АКФ имеет физическую размерность энергии, а при $\tau = 0$ значение АКФ непосредственно равно энергии сигнала и является максимально возможным (косинус угла взаимодействия сигнала с самим собой равен 1):

$$B_s(0) = \int_{-\infty}^{\infty} s(t)^2 dt = E_s.$$

Функция АКФ является непрерывной и четной. В последнем нетрудно убедиться заменой переменной $t = t - \tau$ в выражении (2.4.1):

$$B_s(\tau) = \int_{-\infty}^{\infty} s(t) s(t-\tau) dt = \int_{-\infty}^{\infty} s(t-\tau) s(t) dt = B_s(-\tau).$$

С учетом четности, графическое представление АКФ обычно производится только для положительных значений τ . Знак $+\tau$ в выражении (2.4.1) означает, что при увеличении значений τ от нуля копия сигнала $s(t+\tau)$ сдвигается влево по оси t . На практике сигналы обычно также задаются на интервале положительных значений аргументов от $0-T$, что дает возможность продления интервала нулевыми значениями, если это необходимо для математических операций. В этих границах вычислений более удобным является сдвиг копии сигнала влево по оси аргументов, т.е. применение в выражении (2.4.1) функции $s(t-\tau)$:

$$B_s(\tau) = \int_{-\infty}^{\infty} s(t) s(t-\tau) dt. \quad (2.4.1')$$

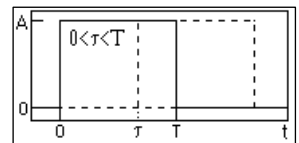
По мере увеличения значения величины сдвига τ для финитных сигналов временное перекрытие сигнала с его копией уменьшается, а, соответственно, косинус угла взаимодействия и скалярное произведение в целом стремятся к нулю:

$$\lim_{|\tau| \rightarrow \infty} B_s(\tau) = 0.$$

Пример. На интервале $(0, T)$ задан прямоугольный импульс с амплитудным значением, равным A . Вычислить автокорреляционную функцию импульса.

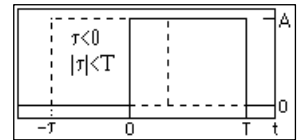
При сдвиге копии импульса по оси t вправо, при $0 \leq \tau \leq T$ сигналы перекрываются на интервале от τ до T . Скалярное произведение:

$$B_s(\tau) = \int_{\tau}^T A^2 dt = A^2(T-\tau).$$



При сдвиге копии импульса влево, при $-\tau \leq \tau < 0$ сигналы перекрываются на интервале от 0 до $T-\tau$. Скалярное произведение:

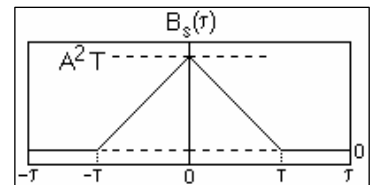
$$B_s(\tau) = \int_0^{T+\tau} A^2 dt = A^2(T+\tau).$$



При $|\tau| > T$ сигнал и его копия не имеют точек пересечения и скалярное произведение сигналов равно нулю (сигнал и его сдвинутая копия становятся ортогональными).

Обобщая вычисления, можем записать:

$$B_s(\tau) = \begin{cases} A^2(T - |\tau|), & |\tau| \leq T \\ 0, & |\tau| > T \end{cases}$$



В случае периодических сигналов АКФ вычисляется по одному периоду T , с усреднением скалярного произведения и его сдвинутой копии в пределах этого периода:

$$B_s(\tau) = (1/T) \int_0^T s(t) s(t-\tau) dt.$$

При $\tau=0$ значение АКФ в этом случае равно не энергии, а средней мощности сигналов в пределах интервала T . АКФ периодических сигналов при этом также является периодической функцией с тем же периодом T . Так, для сигнала $s(t) = A \cos(\omega_0 t + \varphi_0)$ при $T=2\pi/\omega_0$ имеем:

$$B_s(\tau) = \frac{\omega_0}{2\pi} \int_{-\pi/\omega_0}^{\pi/\omega_0} A \cos(\omega_0 t + \varphi_0) A \cos(\omega_0(t-\tau) + \varphi_0) dt = (A^2/2) \cos(\omega_0 \tau).$$

Отметим, что полученный результат не зависит от начальной фазы гармонического сигнала, что характерно для любых периодических сигналов и является одним из свойств КФ.

Для сигналов, заданных на определенном интервале $[a, b]$, вычисление АКФ также производится с нормировкой на длину интервала $[a, b]$:

$$B_s(\tau) = \frac{1}{b-a} \int_a^b s(t) s(t+\tau) dt. \quad (2.4.2)$$

В пределе, для непериодических сигналов с измерением АКФ на интервале T :

$$B_s(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T s(t) s(t+\tau) dt. \quad (2.4.2')$$

Автокорреляция сигнала может оцениваться и коэффициентом автокорреляции, вычисление которого производится по формуле (по центрированным сигналам):

$$r_s(\tau) = \cos \varphi(\tau) = \langle s(t), s(t+\tau) \rangle / \|s(t)\|^2.$$

Взаимная корреляционная функция (ВКФ) сигналов (cross-correlation function, CCF) показывает степень сходства сдвинутых экземпляров двух разных сигналов и их взаимное расположение по координате (независимой переменной), для чего используется та же формула (2.4.1), что и для АКФ, но под интегралом стоит произведение двух разных сигналов, один из которых сдвинут на время τ :

$$B_{12}(\tau) = \int_{-\infty}^{\infty} s_1(t) s_2(t+\tau) dt. \quad (2.4.3)$$

При замене переменной $t = t-\tau$ в формуле (2.4.3), получаем:

$$B_{12}(\tau) = \int_{-\infty}^{\infty} s_1(t-\tau) s_2(t) dt = \int_{-\infty}^{\infty} s_2(t) s_1(t-\tau) dt = B_{21}(-\tau)$$

Отсюда следует, что для ВКФ не выполняется условие четности, а значения ВКФ не обязаны иметь максимум при $\tau = 0$. Это можно наглядно видеть на рис. 2.4.1, где заданы два одинаковых сигнала с центрами на точках 0.5 и 1.5. Вычисление по формуле (2.4.3) с постепенным увеличением значений τ означает последовательные сдвиги сигнала $s_2(t)$ влево по оси времени (для каждого значения $s_1(t)$ для подынтегрального умножения берутся значения $s_2(t+\tau)$).

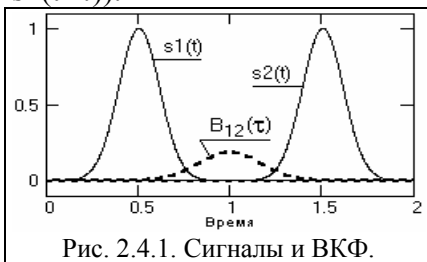


Рис. 2.4.1. Сигналы и ВКФ.

При $\tau=0$ сигналы ортогональны и значение $B_{12}(\tau)=0$. Максимум $B_{12}(\tau)$ будет наблюдаться при сдвиге сигнала $s_2(t)$ влево на значение $\tau=1$, при котором происходит полное совмещение сигналов $s_1(t)$ и $s_2(t+\tau)$. При вычислении значений $B_{21}(-\tau)$ аналогичный процесс выполняется последовательным сдвигом сигнала $s_1(t)$ вправо по временной оси с постепенным увеличением отрицательных значений τ , а соответственно значения $B_{21}(-\tau)$ являются зеркальным (относительно оси $t=0$) отображением значений $B_{12}(\tau)$, и наоборот. На рис. 2.4.2 это можно видеть наглядно.

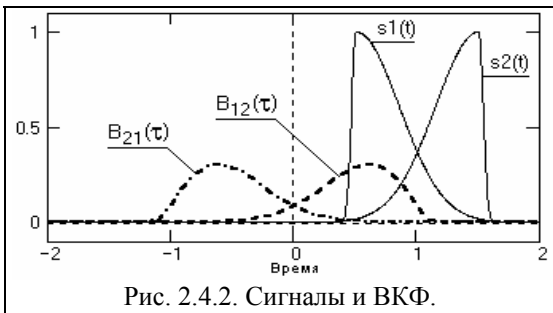


Рис. 2.4.2. Сигналы и ВКФ.

Таким образом, для вычисления полной формы ВКФ числовая ось τ должна включать отрицательные значения, а изменение знака τ в формуле (2.4.3) равносильно перестановке сигналов.

Для периодических сигналов понятие ВКФ обычно не применяется, за исключением сигналов с одинаковым периодом, например, сигналов входа и выхода систем при изучении характеристик систем.

Коэффициент взаимной корреляции двух сигналов вычисляется по формуле (по центрированным сигналам):

$$r_{sv}(\tau) = \cos \varphi(\tau) = \langle s(t), v(t+\tau) \rangle / \|s(t)\| \|v(t)\|. \quad (2.4.4)$$

Значение коэффициента взаимной корреляции может изменяться от -1 до 1.

2.5. МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ШУМОВ И ПОМЕХ [1, 30].

Шумы и помехи (noise). При детектировании сигналов в сумме с основным информационным сигналом одновременно регистрируются и мешающие сигналы - шумы и помехи самой различной природы. К помехам относят также искажения информационных сигналов при влиянии различных дестабилизирующих факторов на процессы измерений, как, например, влияние микрокаверн в стенках скважины на измерения в рентгенометрических методах каротажа, грозовых разрядов на электроразведочные методы измерений и т.п. Выделение информационных составляющих из зарегистрированных сигналов или максимальное подавление шумов и помех в информационном сигнале при сохранении его полезных составляющих является одной из основных задач первичной обработки сигналов (результатов наблюдений).

Если помехи известны и регулярны, как например, фон переменного тока, то борьба с ними особых затруднений не представляет. Наибольшие трудности представляет борьба со случайными (непредсказуемыми) помехами. В общей форме влияние помех на регистрируемый сигнал записывается в следующем виде:

$$y(t) = V(s(t), q(t)), \quad (2.5.1)$$

где $s(t)$ – информационная (полезная) часть сигнала, $q(t)$ – помеха.

Помеха называется *аддитивной*, и обычно именуется шумом, если выражение (2.5.1) представляет собой простую сумму сигнала и помехи:

$$y(t) = s(t) + q(t). \quad (2.5.2)$$

Если случайный процесс $v(t)$, оказывающий влияние на сигнал, является неотрицательным, а его влияние выражается в форме:

$$y(t) = v(t) \cdot s(t), \quad (2.5.3)$$

то помеху $v(t)$ называют *мультипликативной*.

В общем случае в сигнале могут присутствовать оба вида помех:

$$y(t) = v(t) s(t) + q(t). \quad (2.5.4)$$

Природа помех. Как правило, случайные шумовые помехи (аддитивные) порождаются различного рода физическими флуктуациями – случайными отклонениями тех или иных физических величин от своих средних значений. Природа флуктуаций обычно определяется статистической природой физических процессов. Многие физические величины представляют собой результаты усреднения определенных параметров физических процессов, дискретных и случайных по своей природе. Так, например, тепловой шум регистрируемого напряжения на резисторах электрических цепей обуславливается флуктуациями теплового движения носителей зарядов - случайностью процесса дрейфа отдельных электронов по резистору, по суммарной интенсивности движения которых и формируется падение напряжения на резисторе. Дискретной является природа электромагнитных видов излучения – дискретный квант энергии излучения (фотон) определен значением $h\nu$, где h – постоянная Планка, ν – частота. Флуктуации физических величин, дискретных и случайных по своей природе, принципиально неустраняемы, и речь может идти только о том, чтобы уменьшать их относитель-

ную величину имеющимися в нашем распоряжении средствами.

Природа мультипликативных помех обычно связана с изменениями условий измерений, параметров каналов передачи данных и систем их обработки, т.е. когда случайные помехи накладываются не на сам сигнал непосредственно, а на системы, в которых этот сигнал формируется и обращается, вызывая опосредствованные искажения сигнала, как линейные, так и нелинейные.

Характеристики помех. В математическом описании помехи представляются случайными функциями времени. Случайную функцию непрерывного времени обычно называют *случайным процессом*, ее дискретный аналог – *случайной последовательностью*. Как правило, помехи относятся к классу стационарных случайных процессов, и характеризуются своими распределениями и моментами распределений, как их числовыми параметрами. Основное распределение большинства шумовых сигналов – нормальное (гауссов процесс). Это объясняется тем, что распределение сумм независимых случайных величин, из которых складываются случайные помехи, сходится к нормальному, вне зависимости от характера распределения слагаемых (теорема Ляпунова).

Момент первого порядка выражает среднее значение (постоянную составляющую) случайного процесса. Теоретическое значение и оценка момента (по интервалу $[a, b]$):

$$M\{q\} = \bar{q} = \int_{-\infty}^{\infty} q \cdot p(q) dq. \quad (2.5.5)$$

где $p(q)$ – плотность вероятностей значений q .

Центральный момент второго порядка определяет *дисперсию* процесса:

$$D\{q\} = \sigma^2 = \int_{-\infty}^{\infty} (q - \bar{q})^2 \cdot p(q) dq = \overline{q^2} - \bar{q}^2. \quad (2.5.6)$$

Дисперсия выражает мощность переменной составляющей процесса. Корень квадратный из значения дисперсии, т.е. значение σ , является средним квадратическим значением разброса случайных значений q относительно среднего значения \bar{q} .

Смешанный момент второго порядка называется функцией автокорреляции случайного процесса $q(t)$:

$$M\{q(t)q(t+\tau)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \cdot p(x_1, x_2) dx_1 dx_2 = B(\tau). \quad (2.5.7)$$

Величина $B(\tau)$ при $\tau = 0$ равна общей мощности случайного процесса $q(t)$.

На практике большинство случайных процессов обладают свойством эргодичности. Оно заключается в том, что средние значения по множеству реализаций (математические ожидания, вычисляемые по плотностям распределений (2.5.5-7)) совпадают со средними значениями по времени T одной реализации процесса при $T \Rightarrow \infty$. Это позволяет производить оценку числовых значений параметров помех непосредственно по произвольным интервалам $[a, b]$ задания сигналов:

$$\bar{q} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T q(t) dt \approx \frac{1}{b-a} \int_a^b q(t) dt. \quad (2.5.8)$$

$$\sigma^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (q(t) - \bar{q})^2 dt \approx \frac{1}{b-a} \int_a^b (q(t) - \bar{q})^2 dt. \quad (2.5.9)$$

$$B(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T q(t)q(t+\tau) dt \approx \frac{1}{b-a} \int_a^b q(t)q(t+\tau) dt. \quad (2.5.10)$$

Спектральная плотность мощности случайного процесса (распределение мощности помех и шумов по частоте) связано с функцией автокорреляции преобразованием Фурье. В одностороннем (физическом) представлении спектра:

$$B(f) = 4 \int_0^{\infty} B(\tau) \cos 2\pi f \tau dt. \quad (2.5.11)$$

$$B(\tau) = \int_0^{\infty} B(f) \cos 2\pi f \tau dt. \quad (2.5.12)$$

Аддитивную помеху с равномерным спектром

$$B(f) = B_0 = \text{const}$$

называют *белым* шумом. Мощность белого шума в полосе частот 0-F пропорциональна ширине полосы:

$$W_F = \int_0^F B(f) df = B_0 F.$$

При белом шуме полоса частот всегда полагается конечной, т.к. в противном случае мы получим бесконечную мощность шумов.

Сигнал с аддитивной помехой обычно характеризуют не абсолютной мощностью помехи, а отношением средних мощностей сигнала и помехи, которое кратко называют *отношением сигнал/помеха*:

$$\rho = W_c / W_q.$$

Значения случайных процессов являются некоррелированными только при неограниченной полосе частот. Любое ограничение частотной полосы вносит определенную корреляцию в процесс и независимыми друг от друга можно считать только значения процесса, отстоящие друг от друга как минимум на интервал корреляции τ_0 :

$$\tau_0 = (2/W_F) \int_0^\infty B(t) dt = 1/2F.$$

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
3. Васильев Д.В. Радиотехнические цепи и сигналы: Учебное пособие для вузов. - М.: Радио и связь, 1982. - 528 с.
11. Зиновьев А.Л., Филиппов Л.И. Введение в теорию сигналов и цепей: Учебное пособие для вузов. - М.: Высшая школа, 1975. - 264 с.
16. Макс Ж. Методы и техника обработки сигналов при физических измерениях: В 2-х томах.- М.: Мир, 1983.
25. Сергиенко А.Б. Цифровая обработка сигналов. / Учебник для вузов. – СПб.: Питер, 203. – 608 с.
29. Сато Ю. Обработка сигналов. Первое знакомство. – Изд.: ДОДЭКА, 2002.
30. Харкевич А.А. Борьба с помехами. – М.: Наука, 1965.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 3: ДИНАМИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ СИГНАЛОВ

Вовсе не безразлично, каким образом следует разрезать курицу или зайца.

Децим Юний Ювенал. Сатиры. Римский поэт, I-II в.

Римлянин с высшим образованием не мог себе позволить без соответствующего философского обоснования потрошить курицу, а тем более зайца. Но лично меня всегда больше интересует, что я с этого буду иметь.

Владимир Бакаев. Комментарии. Уральский геофизик, XX-XXI в.

Содержание: Введение. 3.1. Разложение сигналов по единичным импульсам. Единичные импульсы. Разложение сигнала. Импульсный отклик линейной системы. 3.2. Свертка (конволюция). Интеграл Дюамеля. Техника свертки. Свойства свертки. Системы свертки. Начальные условия свертки. Литература.

ВВЕДЕНИЕ

Динамическая форма представления сигналов соответствует их естественной и привычной для нас форме математического описания в виде функций независимых переменных (аргументов). Моделирование и анализ линейных стационарных систем обработки сигналов произвольной формы в динамическом представлении базируется на разложении сигналов по единичным импульсам простейшей формы.

3.1. РАЗЛОЖЕНИЕ СИГНАЛОВ ПО ЕДИНИЧНЫМ ИМПУЛЬСАМ [1,11].

Единичные импульсы. В качестве математической модели единичного импульса при анализе аналоговых сигналов используют дельта-функцию.

Дельта-функция описывается следующими математическими выражениями:

$$\delta(t-\tau) = 0 \quad \text{при } t \neq \tau, \quad \int_{-\infty}^{\infty} \delta(t-\tau) dt = 1.$$

Функция $\delta(t-\tau)$ равна нулю везде за исключением точки τ , в которой она бесконечно велика и не является дифференцируемой. На практике такие функции не могут быть реализованы с абсолютной точностью. Но во всех случаях, когда площадь импульса равна 1, а реакция системы на импульс во много раз больше длительности самого импульса, входной сигнал можно считать *единичной импульсной функцией* со свойствами дельта-функции.



Функция Кронекера. Для дискретных и цифровых систем в качестве единичного импульса используется дискретный интегральный аналог дельта-функции - функция единичного отсчета $\delta(k\Delta t - n\Delta t)$, которая равна 1 в координатной точке $k = n$ и нулю во всех остальных точках, при этом функция $\delta(k\Delta t - n\Delta t)$ определена только для целых значений координат k и n .

Математические выражения $\delta(t-\tau)$ и $\delta(k\Delta t - n\Delta t)$ называют также импульсами Дирака и Кронекера. Однако, применяя такую терминологию, не следует забывать, что это не просто единичные импульсы в координатных точках τ и $n\Delta t$, а импульсные функции, определяющие как значения импульсов в определенных координатных точках, так и нулевые значения по всем остальным координатам, в пределе от $-\infty$ до ∞ .

Разложение сигнала по единичным импульсам. Импульсы Дирака и Кронекера используются для разложения, соответственно, произвольных аналоговых сигналов $s(t)$ и дискретных сигналов $s(k\Delta t)$ в непрерывную последовательность неперекрывающихся (ортогональных) импульсов:

$$s(t) = \int_{-\infty}^{\infty} s(\tau) \delta(t-\tau) dt. \quad (3.1.1)$$

$$s(k\Delta t) = \sum_{n=-\infty}^{\infty} s(n\Delta t)\delta(k\Delta t - n\Delta t). \quad (3.1.1')$$

Для аналоговых сигналов разложение (3.1.1) в физическом представлении эквивалентно сканированию значений сигнала $s(t)$ в моменты времени $t = \tau$ бесконечно узкой щелью, бегущей вдоль оси t . Для цифровых сигналов эта щель равна одному отсчету. Пример разложения дискретного сигнала приведен на рис. 3.1.1.

Единичные импульсные функции $\delta(t-\tau)$, $-\infty < \tau < \infty$, и $\delta(k\Delta t - n\Delta t)$, $-\infty < n < \infty$, образуют в бесконечномерных пространствах системы координатных базисов $\{\delta(t-\tau)\}$ и $\{\delta(k\Delta t - n\Delta t)\}$, т.к. единичные импульсные функции не перекрываются и, соответственно, взаимно ортогональны. По этим координатным системам и производится разложение сигналов $s(t)$ и $s(k\Delta t)$. Совокупности проекций сигналов на координатные базисы представляют собой векторные описания сигналов.

Импульсный отклик линейной системы. Если на вход линейной системы в момент времени $t = 0$ подать единичный импульс (Дирака или Кронекера, в зависимости от типа системы), то на выходе мы получим реакцию системы на единичный входной сигнал. Эта реакция называется функцией *импульсного отклика* системы или импульсной характеристикой. Она однозначно определяется оператором преобразования $h(\cdot)$:

$$y(t) = T[\delta(t-0)] = h(t). \quad (3.1.2)$$

$$y(k\Delta t) = T[\delta(k\Delta t-0)] = h(k\Delta t). \quad (3.1.2')$$

Импульсный отклик аналоговой системы на входную дельта-функцию также в определенной степени представляет собой математическую абстракцию идеального преобразования. С практической точки зрения под импульсным откликом можно понимать отображение реакции системы на импульсный входной сигнал произвольной формы с единичной площадью, если длительность этого сигнала пренебрежимо мала по сравнению с временной (координатной) разрешающей способностью системы. Для цифровых систем импульсный отклик однозначно определяется реакцией системы на импульс Кронекера. Функцию импульсного отклика называют также *весовой функцией* системы.

Очевидно, что в линейных и инвариантных к сдвигу системах форма импульсного отклика не зависит от времени прихода входного сигнала и определяет только его положение на временной оси. Так, если входной импульс задержан (относительно 0) на время t_0 , то соответствующий выходной сигнал будет определяться выражением:

$$y(t) = T[\delta(t-t_0)] = h(t-t_0).$$

В любой системе, работающей в реальном масштабе времени, сигнала на выходе системы не может быть, если нет сигнала на ее входе. Отсюда следует односторонность импульсного отклика физических систем:

$$h(t-\tau) = 0 \quad \text{при } t < \tau.$$

Для программных систем, работающих с зарегистрированными массивами цифровых данных, импульсный отклик может быть и двусторонним, так как при обработке сигналов в любой текущей точке $k\Delta t$ системе доступны как "прошлые" отсчеты $k\Delta t - n\Delta t$, так и "будущие" отсчеты $k\Delta t + n\Delta t$. Это резко расширяет возможности программной обработки сигналов по сравнению с физическими системами.

На рисунке 3.1.2 приведен пример импульсного отклика $h(t)$ элементарной физической системы преобразования электрических сигналов – динамической интегрирующей RC-цепи. Подобные схемы очень часто применяются в полевых геофизических приборах (например, в радиометрах) в качестве интенсиметров - измерителей средней скорости счета импульсных потоков сигналов.

При подаче на вход RC-цепи единичного и очень короткого ($\Delta t \ll RC$) импульса заряда Δq емкость C заряжается до напряжения $V_0 = \Delta q/C$, и начинает разряжаться через сопротивление R , при этом напряжение на емкости изменяется по закону $v(t) = V_0 \exp(-t/RC) = (\Delta q/C) \exp(-t/RC)$. Отсюда, импульсный отклик RC-цепи на единичный входной сигнал с еди-

ничным значением заряда $\Delta q = 1$ равен: $h(t) = (1/C)\exp(-t/RC)$, где форма отклика определяется функцией экспоненты, а множитель $(1/C)$ является масштабным преобразователем сигнала (заряда в напряжение). По существу, импульсным откликом системы определяется доля входного сигнала, которая действует на выходе системы по истечении времени t после поступления сигнала на вход (запаздывающая реакция системы).

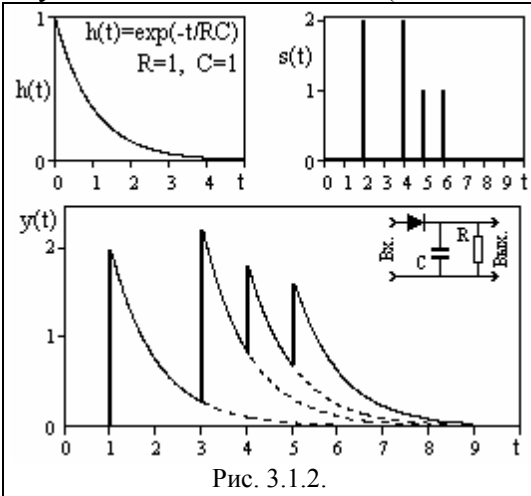


Рис. 3.1.2.

Если функция импульсного отклика системы известна, то, с учетом принципа суперпозиции сигналов в линейной системе, можно выполнить расчет реакции системы в любой произвольный момент времени на любое количество входных сигналов в любые моменты времени их прихода путем суммирования запаздывающих реакций системы на эти входные сигналы. На рис. 3.1.2 приведен пример входного сигнала $s(t)$ для RC-цепи в виде последовательности импульсов и реакция системы $y(t)$ на такой входной сигнал, образованная суммированием реакций системы на каждый импульс.

Допустим, что на вход RC-цепи в моменты времени $t_1=1$ и $t_2=2$ поступили очень короткие (по сравнению со значением RC) импульсы заряда величиной A и B . Математически это можно отобразить сигналом $s(t) = q_1(t)+q_2(t)$, где $q_1(t) = A\cdot\delta(t-t_1)$ и $q_2 = B\cdot\delta(t-t_2)$. Выходной сигнал системы при известном импульсном отклике $h(t)$ отобразится формулой:

$$y(t) = T[q_1(t)+q_2(t)] = T[A\delta(t-t_1)]+T[B\delta(t-t_2)] = A\cdot T[\delta(t-t_1)]+B\cdot T[\delta(t-t_2)] = A\cdot h(t-t_1)+B\cdot h(t-t_2).$$

При расчете значений выходного сигнала в произвольный момент времени t после прихода на вход системы сигналов q_1 и q_2 , например, для $t = 5$, для каждого из сигналов вычисляются значения их запаздывающих реакций: $y_1 = A\cdot h(5-1) = A\cdot h(4)$ и $y_2 = B\cdot h(5-2) = B\cdot h(3)$, после чего значения запаздывающих реакций суммируются $y = y_1+y_2$. Пример этой операции можно видеть на рис. 3.1.3, где для удобства графического представления приняты значения $A=1$ и $B=1$. Сущность операции не изменяется при любых значениях A и B , а в общем случае и для любого количества импульсов.

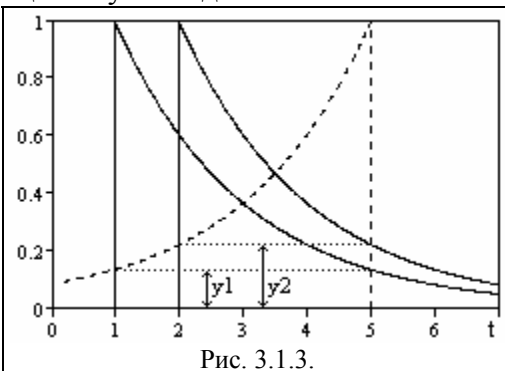


Рис. 3.1.3.

Однако эту же операцию можно рассматривать и с другой позиции. Развернем импульсный отклик $h(t)$ системы на 180° и поместим его начало $h(0)$ непосредственно в точку, для которой нужно выполнить расчет выходного сигнала, т.е. в точку $t=5$ для нашего примера. Если теперь отсчет координат для функции $h(t)$ повести назад от точки расчета по аргументу τ , т.е. перейти на вычисление $h(\tau)$, где значение τ изменяется от 0 и далее (в пределе до ∞), то нетрудно убедиться (на рисунке это наглядно видно), что функция $h(\tau)$ пересечет входные импульсы на тех же значениях y_1 и y_2 . Для этих точек пересечения первого и второго импульсов соответственно имеет место $\tau_1 = t-t_1$ и $\tau_2 = t-t_2$, как и при прямом методе расчета запаздывающих реакций при расчете значений $h(t-t_1)$ и $h(t-t_2)$. После умножения полученных значений $h(\tau_1)$ и $h(\tau_2)$ на значения входного сигнала A и B , получаем полную аналогию: $y_1 = A\cdot h(\tau_1) = A\cdot h(t-t_1)$ и $y_2 = B\cdot h(\tau_2) = B\cdot h(t-t_2)$, и соответственно суммарный сигнал $y = y_1+y_2$.

Такое, чисто математическое представление расчета более удобно для составления математических алгоритмов вычислений. Условно этот процесс для коротких входных импульсных сигналов может быть представлен в следующем виде. Для любой точки расчета t_i выходного сигнала инвертированная по направлению функция импульсного отклика $h(\tau)$ помещается в эту точку t_i и просматривается по своей координате τ с одновременным синхрон-

ным просмотром входного сигнала $s(t)$ назад от точки расчета (прошлые значения входного сигнала), т.е. по координатам $t_i - \tau$. Значения всех встреченных при просмотре импульсов $s(t_i - \tau)$ перемножаются со значениями $h(\tau)$ и суммируются, т.е. для каждой текущей точки расчета t_i в аналоговой системе выполняется операция:

$$y(t_i) = \int_0^{\infty} h(\tau) \cdot s(t_i - \tau) d\tau. \quad (3.1.3)$$

Соответственно в цифровых системах для произвольной точки k_i :

$$y(k_i \Delta t) = \sum_{n=0}^{\infty} h(n \Delta t) \cdot s(k_i \Delta t - n \Delta t). \quad (3.1.3')$$

Полученная сумма значений и будет представлять собой запаздывающую реакцию системы на все импульсы, поступившие на вход системы до текущей точки расчета выходного сигнала.

Таким образом, для линейных и стационарных систем легко определить их реакцию на любой входной сигнал, если известен импульсный отклик систем на единичный входной сигнал.

3.2. СВЕРТКА (КОНВОЛЮЦИЯ) [1,11].

Интеграл Дюамеля. Произвольный сигнал на входе системы с использованием выражений разложения сигнала может быть представлен в виде последовательной линейной комбинации взвешенных единичных импульсов:

$$y(t) = T[s(t)] = T\left[\int_{-\infty}^{\infty} s(\tau) \delta(t - \tau) d\tau\right].$$

На основании принципа суперпозиции линейный оператор T может быть внесен под знак интеграла, т.к. последний представляет собой предельное значение суммы. При этом операция преобразования действует только по переменной t . Отсюда следует:

$$y(t) = \int_{-\infty}^{\infty} s(\tau) T[\delta(t - \tau)] d\tau = \int_{-\infty}^{\infty} s(\tau) h(t - \tau) d\tau. \quad (3.2.1)$$

Это выражение представляет собой интеграл Дюамеля или свертку (конволюцию) входного сигнала с импульсной характеристикой системы. Заменой переменных $t - \tau = \tau$ можно убедиться в том, что свертка коммутативна:

$$\int_{-\infty}^{\infty} s(\tau) h(t - \tau) d\tau \equiv \int_{-\infty}^{\infty} h(\tau) s(t - \tau) d\tau. \quad (3.2.1')$$

Аналогично, для дискретных сигналов:

$$y(k \Delta t) = \sum_{n=-\infty}^{\infty} h(n \Delta t) s(k \Delta t - n \Delta t). \quad (3.2.1'')$$

Выражения (3.2.1) имеют специальную форму упрощенной математической записи в символическом виде:

$$y(t) = s(t - \tau) * h(\tau) \equiv s(t) * h(t).$$

Сравнением выражений (3.2.1' и 3.2.1'') с выражениями (3.1.3) нетрудно убедиться в их полной идентичности, за исключением нижнего предела интегрирования (суммирования). Это и понятно, так как выражения (3.1.3) были получены при рассмотрении реальной физической системы, работающей в реальном масштабе времени, импульсный отклик которых является односторонним (равен нулю при $\tau < 0$). Для таких систем интегрирование (и суммирование) от $-\infty$ до 0 не имеет смысла. Кроме того, в реальных физических системах импульсный отклик, как правило, отличен от нуля только на определенном интервале, при этом верхний интервал интегрирования (суммирования) в выражениях (3.2.1) также может устанавливаться равным значению этого интервала. В общем случае пределы интегрирования (суммирования) ограничиваются значениями τ , на которых функция $h(\tau)$ существует или имеет зна-

чимые значения.

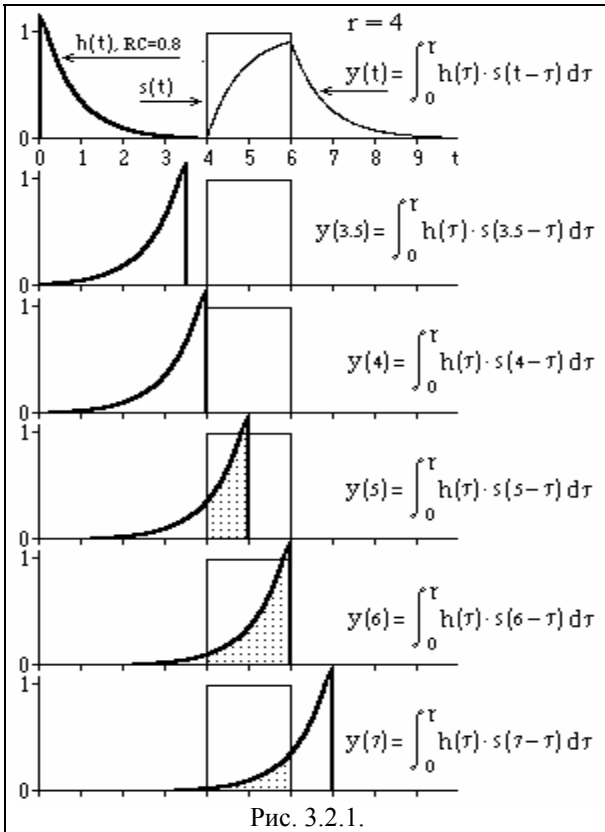


Рис. 3.2.1.

последующих графиках рисунка демонстрируется вычисление результатов свертки в ряде последовательных точек $t_i = \{3.5, 4, 5, 6, 7\}$ временной оси. В силу отрицательного знака τ в аргументах функции $s(t-\tau)$ интегрирование произведения $h(\tau)s(t-\tau)$ выполняется назад по времени и может ограничиваться только определенной длиной значимых значений импульсного отклика (которая в данном случае установлена равной $r = 4$), а результат относится к начальной точке $h(0)$ импульсного отклика. Так как входной сигнал, рассмотренный на рисунке, представляет собой прямоугольный импульс с амплитудой 1, то интеграл свертки в каждой текущей точке расчета равен площади импульсного отклика в пределах границ входного прямоугольного импульса (заполнено точками).

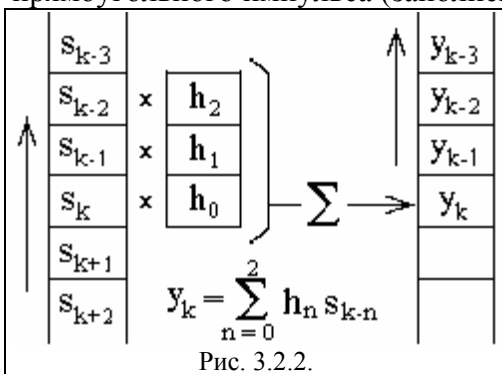


Рис. 3.2.2.

располагается против s_k , все значения s_{k-n} перемножаются с расположенными против них значениями h_n и суммируются. Результаты суммирования являются выходным значением функции y_k , после чего оператор h_n сдвигается на один номер k вперед (или функция s_k сдвигается ему навстречу) и вычисление повторяется для номера $k+1$ и т.д.

Свойства свертки. Для свертки характерны следующие свойства:

- 1. Дистрибутивность: $h(t) * [a(t)+b(t)] = h(t) * a(t)+h(t) * b(t)$.
- 2. Коммутативность: $h(t) * a(t) * b(t) = a(t) * b(t) * h(t)$.
- 3. Ассоциативность: $[a(t) * b(t)] * h(t) = h(t) * a(t) * b(t)$.

Техника свертки. Для вычисления свертки по выражению (3.2.1') функция импульсного отклика реверсируется по своей координате, т.е. строится в режиме обратного времени, и движется относительно функции входного сигнала в сторону возрастания значений t . В каждый текущий момент времени значения обеих функций перемножаются (по всем значениям своих аргументов), а произведение интегрируется. Полученный результат относится к той координатной точке, против которой находится значение импульсного отклика $h(0)$.

На рис. 3.2.1. приведен пример выполнения свертки прямоугольного импульса с импульсным откликом RC-цепи, площадь которого нормирована к 1. Если площадь импульсного отклика $h(t)$ равна 1, то площадь выходного сигнала свертки всегда должна быть равна площади входного сигнала, что можно видеть на верхнем графике рисунка, при этом одно-масштабное сравнение входного и выходного сигналов наглядно демонстрирует характер преобразования сигнала в данной системе. На

Еще более наглядна техника выполнения цифровой свертки, приведенная на рис. 3.2.2. Для вычисления свертки массив одной из функций (s_k - входного или свертываемого сигнала) располагается по ходу возрастания номеров. Массив второй функции (h_n - более короткой, которая обычно называется оператором свертки или оператором фильтра), строится параллельно первому массиву в обратном порядке (по ходу уменьшения номеров первого массива или в режиме обратного времени). Для вычисления y_k значение h_0

Преобразование свертки однозначно определяет выходной сигнал $y(t)$ для установленного значения входного сигнала $s(t)$ при известном значении функции импульсного отклика системы $h(t)$. Обратная задача деконволюции - определение функции $s(t)$ по функциям $y(t)$ и $h(t)$, относится к разряду некорректных, и имеет решение только при вполне определенных условиях. Это объясняется тем, что свертка может существенно изменить частотный спектр сигнала $y(t)$ относительно $s(t)$ и восстановление функции $s(t)$ становится невозможным, если определенные частоты ее спектра в сигнале $y(t)$ полностью утрачены.

Любая практическая система должна быть *устойчивой*, т.е. для сигналов, конечных по энергии или средней мощности, выходные сигналы также должны быть конечными по этим параметрам. Устойчивость обеспечивается при выполнении условия абсолютной интегрируемости импульсного отклика системы:

$$\int_{-\infty}^{\infty} |h(t)| dt < \infty.$$

Для систем с m входами и n выходами аналогично определяются парциальные импульсные отклики $h_{ij}(t)$, $i = \{1, 2, \dots, n\}$, $j = \{1, 2, \dots, m\}$, каждым из которых отображается сигнал на i -м выходе при поступлении сигнала $\delta(t)$ на j -й вход. Полная совокупность импульсных откликов образует матрицу:

$$\bar{\mathbf{h}}(t) = \begin{pmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,m} \\ h_{2,1} & h_{2,2} & \dots & h_{2,m} \\ \cdot & \cdot & \cdot & \cdot \\ h_{n,1} & h_{n,2} & \dots & h_{n,m} \end{pmatrix},$$

а выражение свертки приобретает вид:

$$\bar{\mathbf{Y}}(t) = \int_{-\infty}^{\infty} \bar{\mathbf{h}}(\tau) \bar{\mathbf{S}}(t-\tau) d\tau.$$

Здесь (и в дальнейшем тексте) жирным шрифтом с "крышкой" выделяются векторные величины.

Системы свертки. Свертка выполняется системой (физическим или программным устройством). Физические системы, работающие в реальном времени, вычисляют текущее значение выходного сигнала по всем прошлым значениям входного сигнала, и не могут иметь в своем распоряжении будущих значений входного сигнала. Операторы таких систем являются односторонними (каузальными). Вышеприведенная нормированная к 1 по площади функция RC-цепи $h(t) = (1/RC) \cdot \exp(-t/RC)$, принятая в качестве системного оператора на рис. 3.2.1, является именно таким односторонним каузальным оператором. При сравнении выходного сигнала такой системы с входным нетрудно заметить, что выходной сигнал сдвигается относительно входного сигнала. Для каузальных систем такой "сдвиг по фазе" существует всегда и не может быть исключен (сигнал на выходе системы не может быть раньше сигнала на ее входе).

Входным сигналом программных систем является сигнал в целом, записанный в память вычислительного устройства. При обработке таких данных в распоряжении системы при вычислении любой текущей точки выходного сигнала имеются как "прошлые" для данной точки, так и "будущие" значения входного сигнала. Это позволяет создавать системы без сдвига фазы выходного сигнала относительно входного. Для создания таких систем может использоваться два способа:

1. Первый способ иллюстрирует рис. 3.2.3. Задается система с односторонним каузальным оператором $h(\tau)$. Входной сигнал $s(t)$ пропускается через систему в обычном порядке, и выполняется свертка $g(t) = h(\tau) * s(t)$. Затем выходной сигнал $g(t)$ реверсируется ($g(t) \Rightarrow g(-t)$, конец сигнала становится его началом в порядке возрастания t) и повторно пропускается через систему, т.е. выполняется свертка $y(-t) = h(\tau) * g(-t)$. Полученный сигнал сно-

ва реверсируется ($y(-t) \Rightarrow y(t) = h(-\tau)*g(t)$) и является окончательным выходным сигналом $y(t)$ системы.

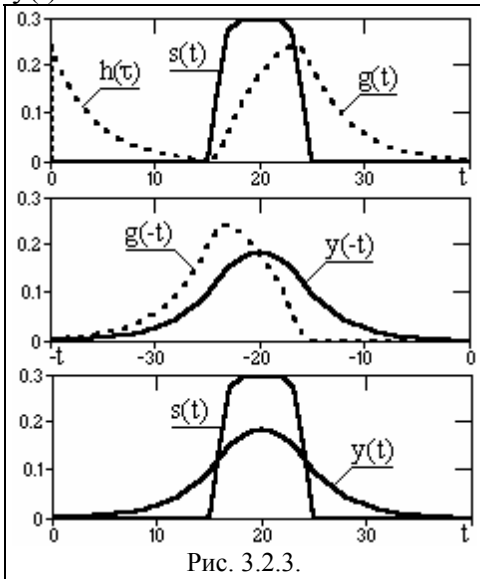


Рис. 3.2.3.

Три последние операции (реверс $g(t) \Rightarrow$ свертка с $h(\tau) \Rightarrow$ реверс выходного сигнала) эквивалентны свертке сигнала $g(t)$ с реверсированным откликом системы $h(-\tau)$, и сдвиг по фазе при свертке реверсированного сигнала компенсирует сдвиг по фазе сигнала, полученный при первой свертке. Общий результат операции $y(t) = h(\tau)*h(-\tau)*s(t)$ не имеет сдвига по фазе выходного сигнала относительно входного. Такую операцию приходится выполнять для исключения сдвига фазы при применении рекурсивных фильтров, которые всегда являются односторонними.

2. Выходной результат $y(t) = h(\tau)*h(-\tau)*s(t)$ предыдущей операции позволяет, используя свойство коммутативности свертки, сначала выполнить свертку $h(\tau)*h(-\tau) = h(\pm\tau)$ и получить один системный оператор $h(\pm\tau)$ (см. рис. 3.2.4), обеспечивающий свертку без сдвига фазы. Этот системный оператор является двусторонним и симметричным относительно $\tau = 0$. Но использование его возможно только для предварительно записанных сигналов, т.к. при выполнении свертки $y(t) = h(\pm\tau)*s(t-\tau)$ для отрицательных значений τ требуются "будущие" значения входного сигнала $s(t+\tau)$. Результат свертки с симметричным оператором полностью аналогичен первой операции

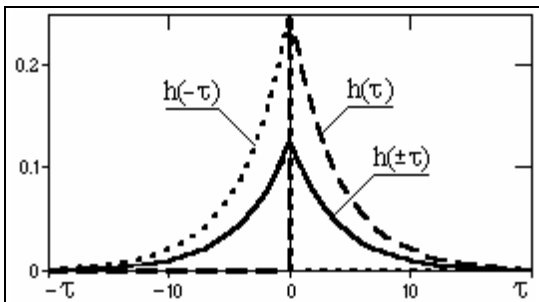


Рис. 3.2.4.

(сигнал $y(t)$ на рис. 3.2.3).

Приведенное выше формирование двустороннего симметричного оператора свертки имеет чисто познавательный характер. На практике вполне естественным является расчет непосредственно симметричных двусторонних операторов под требуемые задачи обработки числовых данных (сигналов, зарегистрированных в дискретной числовой форме).

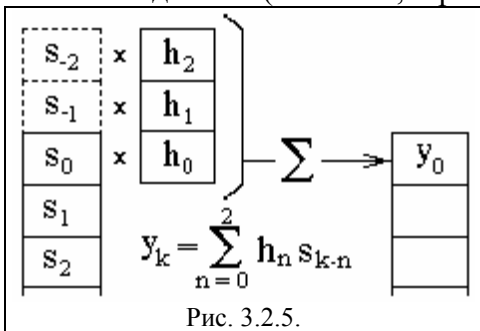


Рис. 3.2.5.

Начальные условия свертки. В начальный момент свертки, при вычислении значений $y(t_i)$ для значений $t_i < \tau_{\max}$ оператора $h(\tau)$, функция оператора, построенная в режиме обратного времени, при $\tau > t_i$ "зависает" для значений $t_i - \tau$ против отсутствующих значений входной функции. Пример такого зависания оператора дискретной свертки против несуществующих отсчетов s_{-1} и s_{-2} входного массива данных при вычислении отсчета y_0 приведен на рис. 3.2.5. Зависание исключают либо заданием начальных условий - дополнительных отсчетов, чаще всего нулевых или равных первому отсчету входной функции, либо началом свертки с отсчета входной функции $k_i = n_{\max}$ с соответствующим сокращением интервала выходной функции на интервал задания системного оператора. Для симметричных операторов со значениями $-n$ (вперед по времени) такой же момент наступает и в конце входного массива, и требует задания конечных условий или сокращения размера выходного сигнала.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
11. Зиновьев А.Л., Филиппов Л.И. Введение в теорию сигналов и цепей: Учебное пособие для вузов. - М.: Высшая школа, 1975. - 264 с.

Тема 4: СПЕКТРАЛЬНОЕ ПРЕДСТАВЛЕНИЕ СИГНАЛОВ

Я согласен, что все состоит из атомов. Но какое нам до этого дело?
Ведь мы занимаемся вопросом о природе богов!

Марк Туллий Цицерон. О природе богов.
Римский философ и политик, 1 в.д.н.э.

Природа экономна. Если и богов она стряпает из атомов, то каждым сигналом в отдельности тем более заниматься не будет. А значит, они тоже из чего-то состоят!

Владимир Петухов. Взгляд с горы.
Осетинский геофизик Уральской школы, XX в.

Введение.

4.1. Разложение сигналов по гармоническим функциям. Понятие собственных функций. Ряды Фурье. Тригонометрическая форма. Параметры эффекта Гиббса.

4.2. Непрерывные преобразования Фурье и Лапласа. Интеграл Фурье. Тригонометрическая форма. Преобразование Лапласа. Обобщенный ряд Фурье.

4.3. Основные свойства преобразований Фурье. Линейность. Свойства четности. Изменение аргумента функции. Теорема запаздывания. Преобразование производной. Преобразование интеграла. Преобразование свертки. Преобразование произведения. Умножение сигнала на гармоническую функцию. Спектры мощности. Равенство Парсевала.

4.4. Спектры некоторых сигналов. Единичные импульсы. Гребневая функция. Спектр прямоугольного импульса. Треугольные импульсы. Экспоненциальный импульс. Функции Лапласа и Гаусса. Спектр косинусоиды.

Литература.

ВВЕДЕНИЕ.

Спектральная (частотная) форма представления сигналов использует разложение сигнальных функций на периодические составляющие.

Периодичность гармонических колебаний исследовал еще в VI веке до нашей эры Пифагор и даже распространил его на описание гармонического движения небесных тел. Термин "spectrum" ("спектр") впервые применил И. Ньютон в 1571 году при описании разложения солнечного света, пропущенного через стеклянную призму, на многоцветную полосу. Он же дал и первую математическую трактовку периодичности волновых движений. В 18-м веке решениями волновых уравнений (в приложении к струнам) занимались Даниил Бернулли и Леонард Эйлер. По существу, уже Бернулли и Эйлер показали, что произвольные периодические функции представляют собой суммы простейших гармонических функций – синусов и косинусов кратных частот. Эти суммы получили название *рядов Фурье*, после того как в 1807 году французский инженер Жан Батист Фурье обосновал метод вычисления коэффициентов тригонометрического ряда, которым можно отображать с абсолютной точностью (при бесконечном числе членов ряда) или аппроксимировать с заданной точностью (при ограничении числа членов ряда) любую периодическую функцию, определенную на интервале одного периода $T = b-a$, и удовлетворяющую условиям Дирехле (ограниченная, кусочно-непрерывная, с конечным числом разрывов 1-го рода). Ряды Фурье в вещественной форме имеют следующий вид:

$$y(x) = (a_0/2) + \sum_{k=1}^{\infty} (a_k \cos(2\pi k f_1 x) + b_k \sin(2\pi k f_1 x)), \quad f_1 = 1/T.$$

$$a_k = (2/T) \int_a^b y(x) \cos(2\pi k f_1 x) dx, \quad b_k = (2/T) \int_a^b y(x) \sin(2\pi k f_1 x) dx.$$

Жан-Батист Жозеф ФУРЬЕ. Jean-Baptiste Joseph Fourier, 1768–1830.

Французский математик. Родился в Осере, в девять лет остался сиротой. Получил образование в церковной школе и военном училище, затем работал преподавателем математики. На протяжении всей жизни активно занимался политикой, арестован в 1794 году за защиту жертв террора, выпущен из тюрьмы после смерти Робеспьера. Принимал участие в создании знаменитой Политехнической школы в Париже. Сопровождал Наполеона в Египет, был назначен губернатором Нижнего Египта. По возвращении во Францию в 1801 году назначен губернатором одной из провинций. В 1822 году стал постоянным секретарем Французской академии наук.

На первых этапах своего развития данное направление разложения функций, полу-

чившее название *гармонического анализа*, имело больше теоретический характер и использовалось, в основном, в естественных науках для выявления и изучения периодичности и состава периодических составляющих (в том числе скрытых) в различных явлениях и процессах (активность солнца, девиация магнитного поля Земли, метеорологические наблюдения, и т.п.). Положение резко изменилось с появлением электротехнических и радиотехнических отраслей науки и техники и развитием радиосвязи, где гармонический состав сигналов приобрел конкретный физический смысл, а математический аппарат спектрального преобразования функций стал основным инструментом анализа и синтеза сигналов и систем. В настоящее время спектральный анализ является одним из основных методов обработки экспериментальных данных во многих отраслях науки и техники.

В качестве базисных функций разложения сигналов, в общем случае, принимаются комплексные экспоненциальные функции $\exp(j2\pi ft)$, от которых с использованием формул Эйлера можно перейти к вещественным синус - косинусным функциям. Термин "частотная" обязан происхождением обратной переменной $f = 1/|t|$ временного представления сигналов и функций. Понятие частотного преобразования не следует связывать только с временными аргументами, т.к. математический аппарат преобразования не зависит от физического смысла независимых переменных. Так, например, при переменной "x", как единице длины, значение f будет представлять собой пространственную частоту с размерностью $1/|x|$ - число периодических изменений сигнала на единице длины.

В математическом аппарате частотного анализа удобно использовать угловую частоту $\omega = 2\pi f$. Для процессов по другим независимым переменным в технической литературе вместо индекса частоты f часто используется индекс ν , а для угловой частоты индекс $k = 2\pi\nu$, который называют *волновым числом*.

4.1. РАЗЛОЖЕНИЕ СИГНАЛОВ ПО ГАРМОНИЧЕСКИМ ФУНКЦИЯМ [1, 21, 24, 25].

Понятие собственных функций. Удобство использования частотного представления сигналов заключается в том, что гармонические функции являются собственными функциями операций переноса, интегрирования, дифференцирования и других линейных операций, инвариантных по координатам.

Допустим, что исходная функция является линейной комбинацией функций синуса и косинуса:

$$s(x) = A \sin(x) + B \cos(x).$$

Осуществим произвольный сдвиг функции по аргументу на величину h . При этом получаем:

$$\begin{aligned} s(x+h) &= C \sin(x) + D \cos(x), \\ C &= A \cos(h) - B \sin(h), \\ D &= A \sin(h) + B \cos(h), \end{aligned}$$

где коэффициенты C и D , как и в исходном выражении коэффициенты A и B , не зависят от аргумента, при этом $C^2 + D^2 = A^2 + B^2$. Таким образом, при произвольном переносе функции по аргументу (а равно и при интегрировании, дифференцировании и других линейных преобразованиях) любую линейную комбинацию синуса и косинуса можно представить линейной комбинацией этих же функций.

Экспоненциальная комплексная запись гармонических функций делает это свойство еще нагляднее. Для произвольной гармонической функции имеем:

$$\cos(\omega t - \varphi) = A \cos(\omega t) + B \sin(\omega t),$$

где $A = \cos(\varphi)$, $B = \sin(\varphi)$, φ - начальный фазовый угол колебания при $t = 0$. Переходя к комплексной записи данной функции с использованием тождеств Эйлера

$$\cos(\omega t) = [\exp(j\omega t) + \exp(-j\omega t)]/2, \quad \sin(\omega t) = [\exp(j\omega t) - \exp(-j\omega t)]/2j,$$

получаем:

$$\cos(\omega t - \varphi) = C \exp(j\omega t) + C^* \exp(-j\omega t),$$

где: $C = 0,5 \exp(-j\varphi)$, $C^* = 0,5 \exp(j\varphi)$ - величина, комплексно сопряженная с C . Применяя в качестве гармонической составляющей разложения сигнала функцию $\exp(j\omega t)$, можно рассматривать вторую функцию $\exp(-j\omega t)$, комплексно сопряженную с первой, как такую же со-

ставляющую, но с отрицательной частотой. Естественно, что отрицательная частота является чисто математической абстракцией, но нужно помнить, что пара таких комплексно сопряженных составляющих в сумме всегда дает вещественную функцию.

Экспоненциальные функции также являются собственными функциями линейных операций. Для операции переноса с использованием экспоненциальных функций:

$$\exp[j\omega(t+h)] = \exp(j\omega h) \cdot \exp(j\omega t) = H(\omega) \exp(j\omega t),$$

где $H(\omega) = \exp(j\omega h)$ - собственное значение операции переноса, независимое от переменной.

Для операции дифференцирования:

$$d[\exp(j\omega t)]/dt = j\omega \exp(j\omega t), \quad H(\omega) = j\omega.$$

Для операции интегрирования:

$$\int \exp(j\omega t) dt = (1/j\omega) \exp(j\omega t), \quad H(\omega) = 1/j\omega.$$

В общей форме, для любых линейных операций преобразования:

$$T[\exp(j\omega t)] = H(\omega) \exp(j\omega t),$$

где $T[.]$ - произвольный линейный оператор, $H(\omega)$ - собственное значение операции, независимое от аргумента.

У специалистов - практиков существует предубеждение против использования комплексных функций с их мнимыми частотами. Поэтому в дальнейшем будем использовать и вещественные функции, и их комплексные аналоги, по крайней мере, до тех пор, пока простота и удобство использования последних не станет очевидным.

Ряды Фурье. Разложению в ряды Фурье подвергаются периодические сигналы. Периодическую функцию любой формы, заданную на интервале одного периода $T = b-a$ и удовлетворяющую на этом интервале условиям Дирихле (ограниченная, кусочно-непрерывная, с конечным числом разрывов 1-го рода), можно представить в виде ряда Фурье:

$$s(t) = \sum_{n=-\infty}^{\infty} S_n \exp(jn\Delta\omega t), \quad S_n = S(n\Delta\omega), \quad \Delta\omega = 2\pi/T, \quad (4.1.1)$$

где весовые коэффициенты S_n ряда определяются по формуле:

$$S_n = (1/T) \int_a^b s(t) \exp(-jn\Delta\omega t) dt. \quad (4.1.2)$$

Ряд Фурье представляет собой ансамбль комплексных экспонент $\exp(jn\Delta\omega t)$ с частотами, образующими арифметическую прогрессию. Функцию весовых коэффициентов $S(n\Delta\omega)$ принято называть комплексным спектром периодического сигнала или фурье-образом функции $s(t)$. Спектр периодического сигнала является дискретной функцией, т.к. он определен только для целых значений n с шагом по частоте, обратным периоду: $\Delta\omega = 2\pi/T$ (или $\Delta f = 1/T$). Первую частотную составляющую спектра при $n = 1$, равную $\omega_1 = 1 \cdot \Delta\omega = 2\pi/T$ (или $f_1 = 1/T$), называют *основной* частотой сигнала (первой гармоникой), остальные частоты дискретного спектра $n\omega_1$ при $n > 1$ называют гармониками сигнала. Значения $S(n\Delta\omega)$ по положительным и отрицательным значениям n являются комплексно сопряженными.

С чисто математических позиций множество функций $\exp(jn\Delta\omega t)$, $-\infty < n < \infty$ образует бесконечномерный базис линейного пространства $L^2[a,b]$ ортогональных синус-косинусных функций, а коэффициенты S_n по (4.1.2) представляют собой проекции сигнала $s(t)$ на эти базисные функции. Соответственно, сигнал $s(t)$ в форме ряда Фурье (4.1.1) – это бесконечномерный вектор в пространстве $L^2[a,b]$, точка с координатами S_n по базисным осям пространства $\exp(jn\Delta\omega t)$.

Подынтегральную функцию экспоненты в выражении (4.1.2) с использованием тождества Эйлера

$$\exp(\pm j\omega t) = \cos(\omega t) \pm j \cdot \sin(\omega t)$$

можно разложить на косинусную и синусную составляющие и выразить комплексный спектр в виде действительной и мнимой части:

$$S_n = (1/T) \int_a^b s(t) [\cos(n\Delta\omega t) - j \sin(n\Delta\omega t)] dt = A_n - jB_n. \quad (4.1.3)$$

$$A_n \equiv A(n\Delta\omega) = (1/T) \int_a^b s(t) \cos(n\Delta\omega t) dt, \quad (4.1.4)$$

$$B_n \equiv B(n\Delta\omega) = (1/T) \int_a^b s(t) \sin(n\Delta\omega t) dt. \quad (4.1.5)$$

На рис. 4.1.1 приведен пример периодического сигнала (прямоугольный импульс на интервале (1-3.3), повторяющийся с периодом $T=40$) и форма действительной и мнимой части его спектра. Обратим внимание, что действительная часть спектра является четной относительно нуля функцией $A(n\Delta\omega) = A(-n\Delta\omega)$, так как при вычислении значений $A(n\Delta\omega)$ по формуле (4.1.4) используется четная косинусная функция $\cos(n\Delta\omega t) = \cos(-n\Delta\omega t)$. Мнимая часть спектра является нечетной функцией $B(n\Delta\omega) = -B(-n\Delta\omega)$, так как для ее вычисления по (4.1.5) используется нечетная синусная функция $\sin(n\Delta\omega t) = -\sin(-n\Delta\omega t)$.

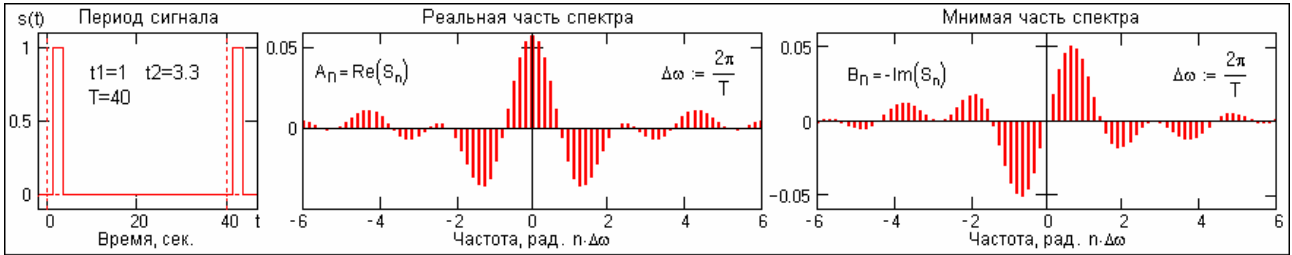


Рис. 4.1.1. Сигнал и его комплексный спектр.

Комплексные числа дискретной функции (4.1.3) могут быть представлены в виде модулей и аргументов комплексной экспоненты, что дает следующую форму записи комплексного спектра:

$$S_n = R_n \exp(j\varphi_n), \quad (4.1.3')$$

$$R_n^2 \equiv R^2(n\Delta\omega) = A^2(n\Delta\omega) + B^2(n\Delta\omega),$$

$$\varphi_n \equiv \varphi(n\Delta\omega) = \arctg(-B(n\Delta\omega)/A(n\Delta\omega)).$$

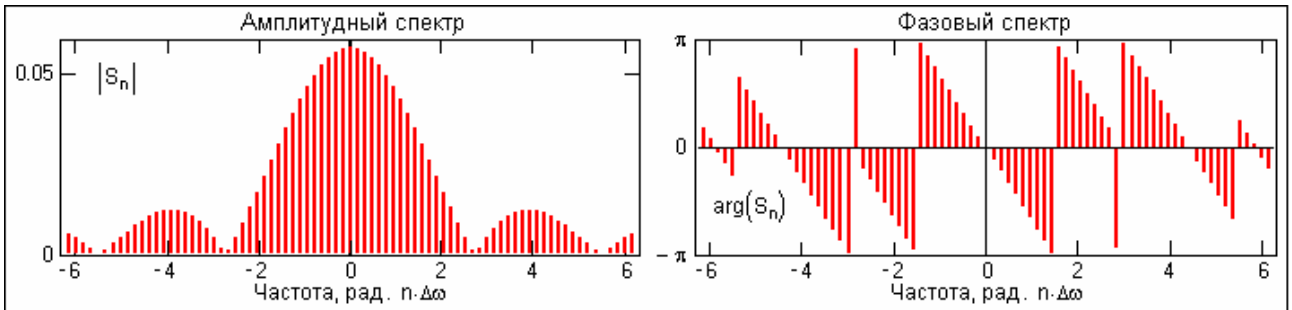


Рис. 4.1.2. Модуль и аргумент спектра.

Модуль спектра $R(n\Delta\omega)$ называют двусторонним спектром амплитуд или АЧХ - амплитудно-частотной характеристикой сигнала, а аргумент спектра (последовательность фазовых углов $\varphi(n\Delta\omega)$) - двусторонним спектром фаз или ФЧХ - фазово-частотной характеристикой. Спектр амплитуд всегда представляет собой четную функцию: $R(n\Delta\omega) = R(-n\Delta\omega)$, а спектр фаз нечетную: $\varphi(n\Delta\omega) = -\varphi(-n\Delta\omega)$. Пример спектра в амплитудном и фазовом представлении для сигнала, показанного на рис. 4.1.1, приведен на рис. 4.1.2. При рассмотрении спектра фаз следует учитывать периодичность 2π угловой частоты (при уменьшении фазового значения до величины менее $-\pi$ происходит сброс значения -2π).

Если функция $s(t)$ является четной, то все значения $B(n\Delta\omega)$ по (4.1.5) равны нулю, т.к. четные функции ортогональны синусным гармоникам и подынтегральное произведение $s(t) \cdot \sin(n\Delta\omega t)$ дает нулевой интеграл. Следовательно, спектр функции будет представлен только вещественными коэффициентами. Напротив, при нечетности функции $s(t)$ обнуляются все значения коэффициентов $A(n\Delta\omega)$ (нечетные функции ортогональным косинусным гармоникам) и спектр является чисто мнимым. Этот фактор не зависит от выбора границ задания периода функции на числовой оси. На рис. 4.1.3(А) можно наглядно видеть ортогональность первой гармоники синуса и четной функции, а на рис. 4.1.3(В) соответственно косинуса и

нечетной функции в пределах одного периода. Учитывая кратность частот последующих гармоник первой гармонике спектра, ортогональность сохраняется для всех гармоник ряда Фурье.

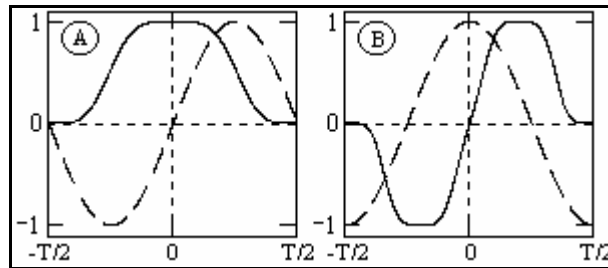


Рис. 4.1.3. Ортогональность функций.

При $n = 0$ имеем $B_0 = 0$, и получаем постоянную составляющую сигнала:

$$S_0 \equiv A_0 \equiv R_0 \equiv (1/T) \int_a^b s(t) dt.$$

Тригонометрическая форма рядов Фурье. Объединяя в (4.1.1) комплексно сопряженные составляющие (члены ряда, симметричные относительно центрального члена ряда S_0), можно перейти к ряду Фурье в тригонометрической форме:

$$s(t) = A_0 + 2 \sum_{n=1}^{\infty} (A_n \cos(n\Delta\omega t) + B_n \sin(n\Delta\omega t)), \quad (4.1.6)$$

$$s(t) = A_0 + 2 \sum_{n=1}^{\infty} R_n \cos(n\Delta\omega t + \varphi_n). \quad (4.1.6')$$

Значения A_n , B_n вычисляются по формулам (4.1.4-4.1.5), значения R_n и φ_n – по формулам (4.1.3').

Ряд (4.1.6) представляют собой разложение периодического сигнала $s(t)$ на сумму вещественных элементарных гармонических функций (косинусных и синусных) с весовыми коэффициентами, удвоенные значения которых (т.е. значения $2 \cdot A_n$, $2 \cdot B_n$) не что иное, как амплитуды соответствующих гармонических колебаний с частотами $n\Delta\omega$. Совокупность амплитудных значений этих гармоник образует односторонний физически реальный (только для положительных частот $n\Delta\omega$) спектр сигнала. Для сигнала на рис. 4.1.1, например, он полностью повторяет правую половину приведенных на рисунке спектров с удвоенными значениями амплитуд (за исключением значения A_0 на нулевой частоте, которое, как это следует из (4.1.6), не удваивается). Но такое графическое отображение спектров используется довольно редко (за исключением чисто технических приложений).

Более широкое применение для отображения физически реальных спектров находит формула (4.1.6'). Спектр амплитуд косинусных гармоник при таком отображении называется амплитудно-частотным составом сигнала, а спектр фазовых углов гармоник – фазовой характеристикой сигнала. Форма спектров повторяет правую половину соответствующих двусторонних спектров (см. рис. 4.1.2) также с удвоенными значениями амплитуд. Для четных сигналов отсчеты фазового спектра могут принимать только значения 0 или π , для нечетных соответственно $\pm\pi/2$.

Ряды Фурье произвольных аналоговых периодических сигналов могут содержать бесконечно большое количество членов. Однако одним из важных достоинств преобразования Фурье является то, что при ограничении (усечении) ряда Фурье до любого конечного числа его членов обеспечивается наилучшее по средней квадратической погрешности приближение к исходной функции (для данного количества членов).

На рис. 4.1.4 показано разложение в комплексный ряд Фурье модельного сигнала, выполненное в среде Mathcad. Модель сигнала задана с тремя разрывами первого рода (скачками). Любой скачок функции содержит все частоты диапазона до бесконечности, в связи с чем ряд Фурье также бесконечен и очень медленно затухает. На рисунке приведены значения только первых 100 членов ряда, при этом график спектра сигнала, как это обычно принято на

практике, построен в виде огибающей значений модулей коэффициентов ряда S_n и только по области положительных значений n .

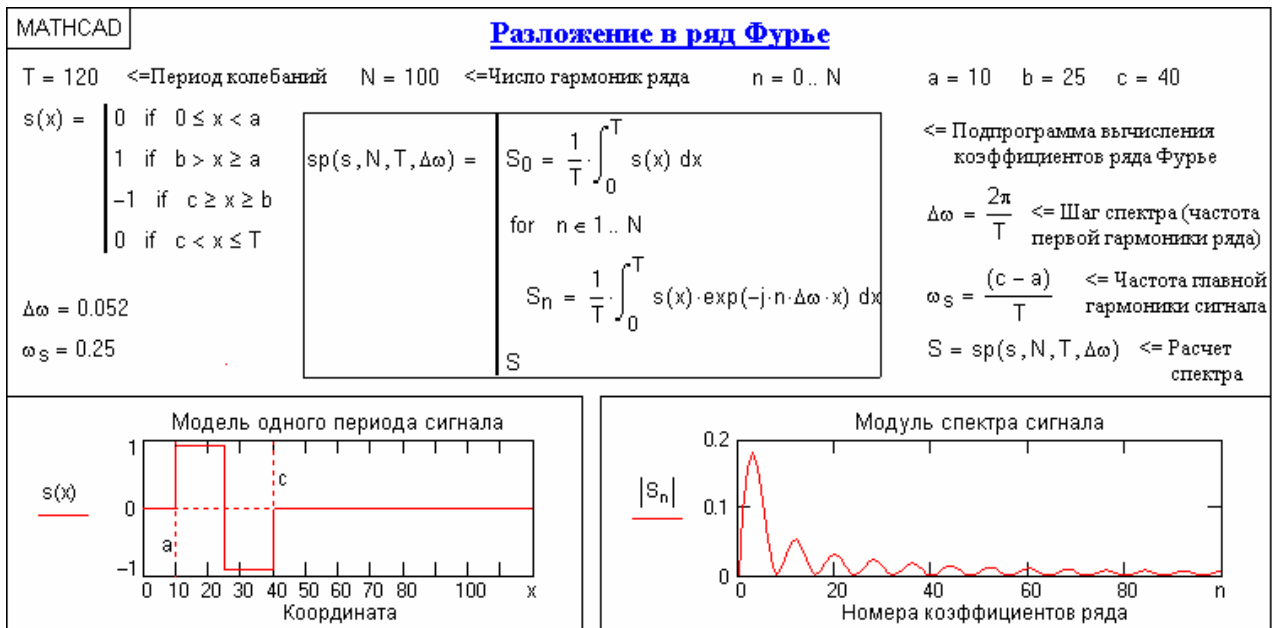


Рис. 4.1.4. Разложение сигнала в комплексный ряд Фурье.

Программа на рис. 4.1.5 продолжает программу рис. 4.1.4 и показывает реконструкцию сигнала по его спектру при ограничении числа членов ряда Фурье.

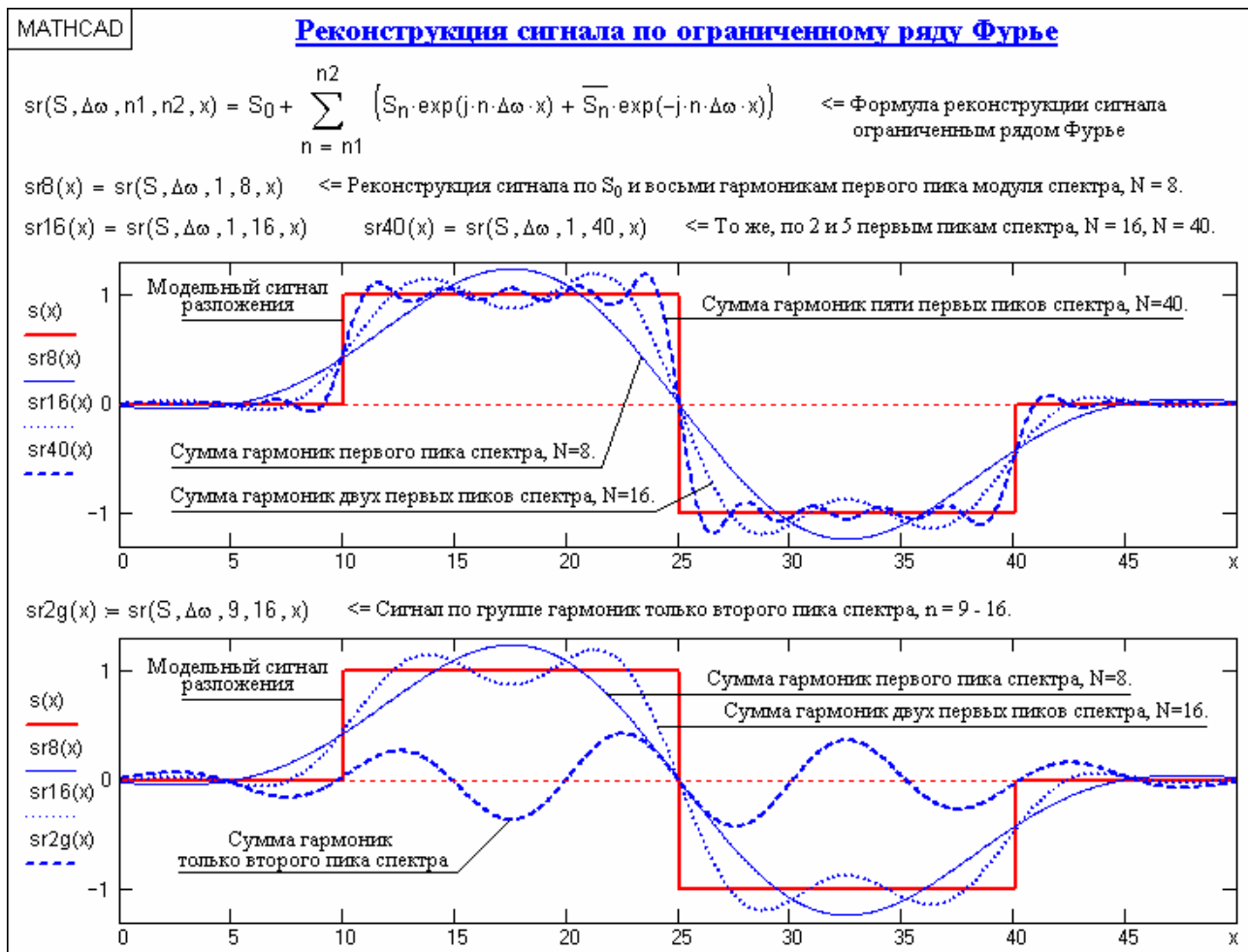


Рис. 4.1.5. Реконструкция сигнала (продолжение программы на рис. 4.1.4)

На верхнем графике рисунка приведен реконструированный сигнал при $N = 8$ (гармоники первого пика спектра, центр которого соответствует главной гармонике сигнала и члену

ряда $n = \omega_s/\Delta\omega$), $N = 16$ (гармоники двух первых пиков) и $N=40$ (пять первых пиков спектра). Естественно, что чем больше членов ряда включено в реконструкцию, тем ближе реконструированный сигнал к форме исходного сигнала.

Принцип последовательного приближения к исходной форме наглядно виден на нижнем графике рисунка. На нем же можно видеть и причины появления пульсаций на реконструкции скачков функций, которые носят название *эффекта Гиббса*. При изменении количества суммируемых членов ряда эффект Гиббса не исчезает. Не изменяется также относительная амплитуда пульсаций (по отношению к амплитуде скачка) и относительное затухание (по коэффициенту последовательного уменьшения амплитуды пульсаций по отношению к максимальному выбросу), изменяется только частота пульсаций, которая определяется частотой последних суммируемых гармоник.

Эффект Гиббса имеет место всегда при резких нарушениях монотонности функций. На скачках эффект максимален, во всех других случаях амплитуда пульсаций зависит от характера нарушения монотонности функции. Пример явления Гиббса для радиоимпульса приведен на рис. 4.1.6 (использована программа на рис. 4.1.4, точками показан реконструированный сигнал с увеличением масштаба в 10 раз).

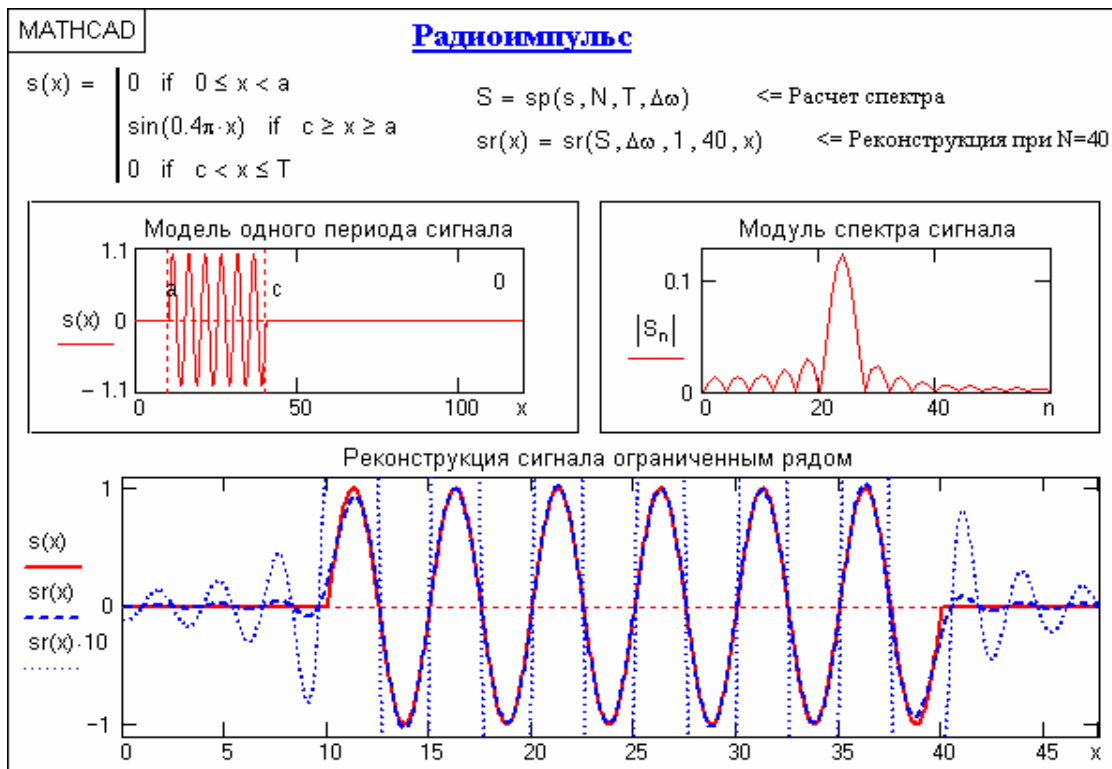


Рис. 4.1.6.

На рис. 4.1.7 приведен пример разложения в ряд Фурье одного периода $T=(a,c)$ модельного периодического сигнала $sq(x)$, представленного информационным сигналом $s(x)$ в сумме с шумовым сигналом. Спектр шумов близок к спектру белого шума (равномерное распределение энергии шумов по всем частотам спектра).

На спектре модельного сигнала достаточно четко выделяется диапазон частот информационного сигнала. Реконструкция сигнала с ограничением ряда Фурье гармониками только информационного сигнала (сигнал $sr5(x)$, $N=5$) дает сглаженную форму сигнала по минимуму среднеквадратического расхождения с модельным сигналом для данного количества членов ряда, но только по периоду разложения (a, c), и наиболее точное приближение к информационному сигналу. При увеличении в реконструкции количества членов ряда Фурье восстановленный сигнал начинает приближаться к модельному сигналу, но только по данному периоду $T=(a,c)$, при этом расхождение с информационным сигналом увеличивается. Заметим, что спектр сигнала может определяться и по нескольким периодам сигнала, что повышает точность реконструкции информационного сигнала.

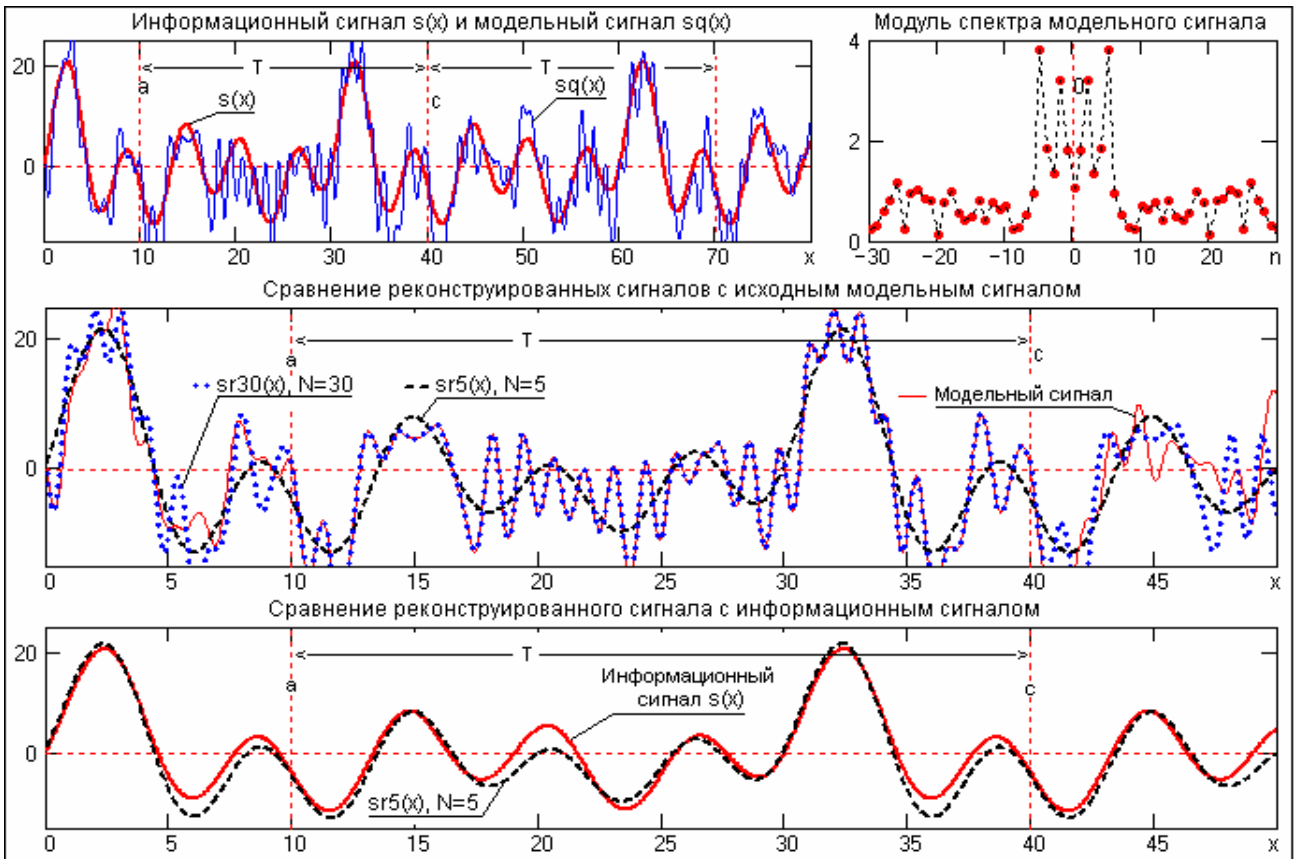


Рис. 4.1.7.

В ряд Фурье может разлагаться и произвольная непериодическая функция, заданная (ограниченная, вырезанная из другого сигнала, и т.п.) на интервале (a, b) , если нас не интересует ее поведение за пределами данного интервала. Однако следует помнить, что применение формул (4.1.1-4.1.6) автоматически означает периодическое продолжение данной функции за пределами заданного интервала (в обе стороны от него) с периодом $T = b - a$. Однако при этом на краях интервала может возникнуть явление Гиббса, если уровень сигнала на краях не совпадает и образуются скачки сигнала при его периодическом повторении, как это видно на рис. 4.1.8. При разложении исходной функции в ограниченный ряд Фурье и его обработке в частотной области на самом деле при этом обрабатывается не исходная функция,

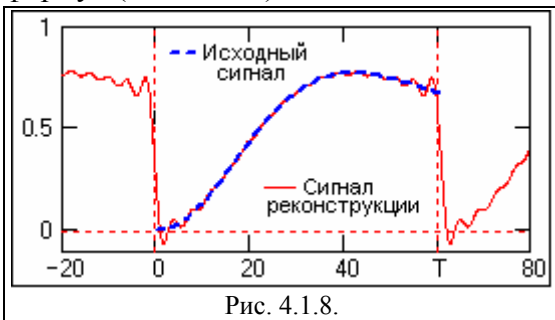


Рис. 4.1.8.

а реконструированная из ограниченного ряда Фурье.

При усечении рядов Фурье определенное искажение функций существует всегда. Но при малой доле энергии отсекаемой части сигнала (при быстром затухании спектров функций) этот эффект может быть и мало заметен. На скачках и разрывах функций он проявляется наиболее ярко.

Параметры эффекта Гиббса. Большинство методов анализа и обработки сигналов представляют собой или имеют в своем составе операцию свертки сигналов с функцией оператора свертки. Как сигнал, так и оператор свертки, выполняющий определенную задачу обработки данных и реализующий определенную частотную функцию системы обработки, могут быть бесконечно большими. Практика же обработки на ЭВМ может иметь дело только с ограниченными множествами и данных, и коэффициентов оператора. В общем случае, эти ограниченные множества "вырезаются" из бесконечных множеств, а разложение в ряды Фурье, также ограниченные по размерам, является одной из самых распространенных операций обработки цифровых множеств. С учетом этого рассмотрим явление Гиббса более подробно, т.к. при любых ограничениях рядов Фурье оно всегда может весьма существенно сказаться на

качестве и точности обработки сигналов.

Очевидно, что при усечении ряда Фурье (4.1.1) любой функции до конечного числа членов N мы будем иметь усеченный ряд Фурье:

$$s_N(x) = \sum_{n=-N}^N S(n) \exp(jxn\Delta\omega), \quad (4.1.7)$$

при этом происходит усечение спектральной характеристики функции до частоты $n\Delta\omega$ и сходимость суммы остающихся членов ряда $s_N(x)$ к исходной функции $s(x)$ ухудшается в тем большей степени, чем меньше значение N . Особенно ярко это проявляется на крутых перепадах (разрывах, скачках) функций:

- крутизна перепадов "размывается", т.к. она не может быть больше, чем крутизна (в нулевой точке) последней сохраненной гармоники ряда (4.1.7);

- по обе стороны "размытых" перепадов появляются выбросы и затухающие осцилляции с частотой, равной частоте последнего сохраненного или первого отброшенного члена ряда (4.1.7).

Рассмотрим явление Гиббса на примере разложения в ряд Фурье функции единичного скачка $s(x)$, которая имеет разрыв величиной 1 в точке $x = 0$. Уравнение функции:

$$s(x) = -0.5 \text{ при } -T/2 \leq x < 0; \quad s(x) = 0.5 \text{ при } 0 \leq x \leq T/2.$$

Поскольку функция является нечетной, ее ряд Фурье не содержит косинусных членов, и коэффициенты ряда в односторонней тригонометрической форме определяются выражением (с учетом соотношения $\Delta\omega = 2\pi/T$):

$$b_n = (2/T) \int_{-T/2}^{T/2} s(x) \sin(xn\Delta\omega) dx = (2/T) \int_0^{T/2} \sin(xn\Delta\omega) dx.$$

$$b_n = 2/(n\cdot\pi), \quad n\text{- нечетное,}$$

$$b_n = 0, \quad n\text{- четное.}$$

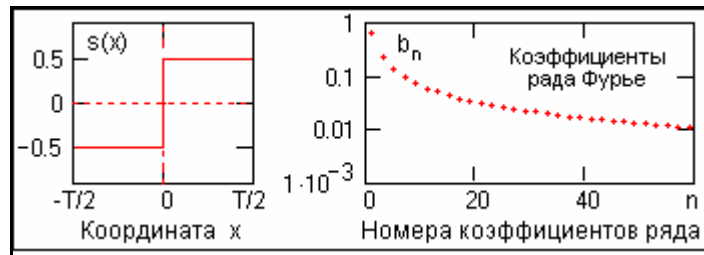


Рис. 4.1.9. Значения коэффициентов b_n .

Как видно на рис. 4.1.9, ряд коэффициентов b_n затухает очень медленно. Соответственно, медленно будет затухать и ряд Фурье функции $s(x)$:

$$s(x) = (2/\pi)[\sin x\Delta\omega + (1/3)\cdot\sin x3\Delta\omega + (1/5)\cdot\sin x5\Delta\omega + \dots].$$

$$s(x) = (2/\pi) \sum_{n=0}^{\infty} \sin[x(2n+1)\Delta\omega]/(2n+1). \quad (4.1.8)$$

Этот ряд при усечении до M нечетных членов можно записать в следующем виде:

$$s(x) = (2\Delta\omega/\pi) \sum_{n=0}^M \int_0^x \cos(x(2n+1)\Delta\omega) dx = (2\Delta\omega/\pi) \int_0^x \left[\sum_{n=0}^M \cos(x(2n+1)\Delta\omega) \right] dx.$$

Сумма косинусного ряда равна $\sin[2(M+1)x\Delta\omega]/(2\sin x\Delta\omega)$. Отсюда:

$$s_M(x) = \frac{\Delta\omega}{\pi} \int_0^x \frac{\sin 2(M+1)x\Delta\omega}{\sin x\Delta\omega} dx. \quad (4.1.9)$$

Для определения местоположения максимумов и минимумов возникающих осцилляций функции, приравняем к нулю ее первую производную (подинтегральную функцию) выражения (4.1.9), при этом:

$$x_k = \pm k\pi/(2\Delta\omega(M+1)) = \pm kT/(4(M+1)), \quad k = 1, 2, \dots$$

Соответственно, амплитудные значения первых (максимальных) осцилляций функции приходится на точки $x_{k=1} = \pm T/(4(M+1))$, вторых (противоположных по полярности) - на точки $x_{k=2} = \pm T/(2(M+1))$. Период пульсаций равен $x_{k=3} - x_{k=1} \equiv 2x_{k=1} = \pm T/(2(M+1))$, т.е. на одном

периоде задания сигнала появляется $2(M+1)$ пульсация с частотой, обратным периоду и равной $2(M+1)\Delta f$ – частоте последнего сохраненного в суммировании члена ряда Фурье. Функция пульсаций (при ее выделении) является нечетной относительно скачка. Соответственно, при скачке функции $s(x)$ на точке периода T значения x_k являются значениями Δx_k относительно точки скачка. Амплитудные значения функции в точках x_1 и x_2 (при подстановках x_1 и x_2 верхним пределом в (4.1.9)) практически не зависят от количества членов ряда M и равны:

$$s_M(x_1) \approx 0.5+0.09, \quad s_M(x_2) \approx 0.5-0.05.$$

Амплитуда последующих осцилляций постепенно затухает.

Реконструкция скачка при трех значениях ряда приведена на рис. 4.1.10. Как и положено, функция продолжается периодически за пределами заданного интервала $(-T/2, T/2)$, при этом на границах периодов также образуются скачки. Скачки являются центрами возникающих осцилляций. Наложение осцилляций друг на друга в зависимости от расстояния между их центрами может как уменьшать амплитуду пульсаций, так и увеличивать.

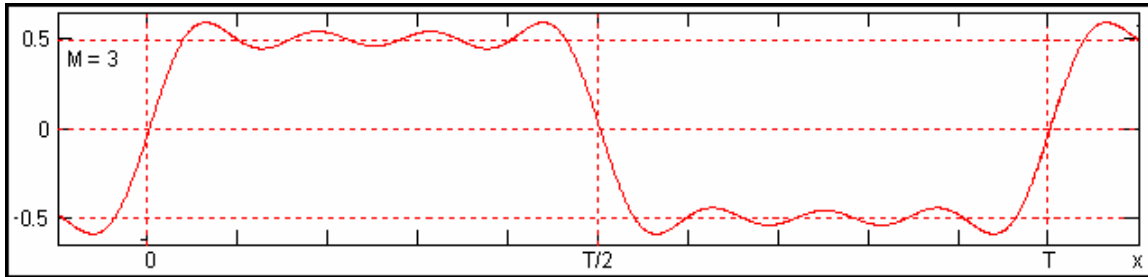


Рис. 4.1.10. Реконструкция скачка по ограниченному ряду Фурье при $M=3$.

Таким образом, для усеченных рядов Фурье предельные значения максимальных выбросов по обе стороны от скачка и следующих за ними обратных выбросов при единичной амплитуде разрыва функции достигают соответственно 9% и 5% значения амплитуды скачка. Кроме того, сам скачок функции из собственно скачка преобразуется в переходную зону, длина которой между точками максимальных выбросов по обе стороны скачка равна $T/(2(M+1))$, а по уровню исходных значений функции на скачке (в данном случае от -0.5 до 0.5) порядка $(2/3)T/(2(M+1))$. Это явление типично для всех функций с разрывами.

4.2. НЕПРЕРЫВНЫЕ ПРЕОБРАЗОВАНИЯ ФУРЬЕ И ЛАПЛАСА [1,24,25].

Интеграл Фурье. Спектры непериодических сигналов конечной длительности (финитных), зарегистрированных на интервале T , могут быть получены из уравнений для рядов Фурье как предельные значения функций суммирования при расширении периода T до бесконечности.

Зададим периодическую последовательность импульсов и разложим импульс на одном периоде T в ряд Фурье (формула 4.1.2). Не меняя положения импульса на интервале T , увеличим значение T в два раза (продлеваем интервал нулями), при этом выражение (4.1.2) для вычисления спектра остается без изменения, но по ней рассчитывается в 2 раза большее количество гармоник с уменьшением в 2 раза частоты первой гармоники и шага $\Delta\omega=2\pi/T$. Увеличение интервала T не влияет на результаты вычисления интеграла функции (4.1.2), т.к. интервал продления заполнен нулевыми значениями сигнала.

По существу, при увеличении периода T без изменения финитного сигнала форма спектра по оси частот остается без изменения, изменяется только шаг дискретизации спектра и, за счет множителя $1/T$, в 2 раза уменьшаются значения спектра. Новые гармоники располагаются в интервалах между гармониками первого ряда. Пример изменения спектра при увеличении периода T в 2 раза приведен на рис. 4.2.1.

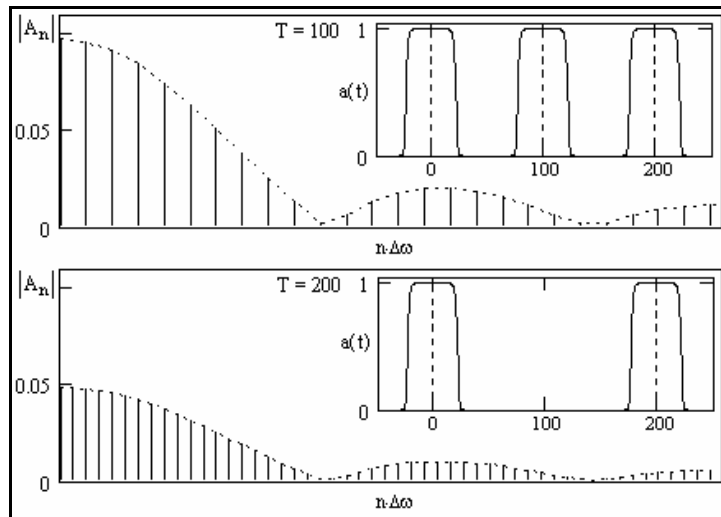


Рис. 4.2.1.

Процесс можно продолжить дальнейшим последовательным увеличением периода, при этом спектр будет приближаться к непрерывной функции. В пределе, при $T \rightarrow \infty$, периодическая последовательность импульсов заменяется одиночным финитным сигналом, дискретные частоты $n\Delta\omega$ обращаются в непрерывные текущие значения ω ($\Delta\omega = 2\pi/T \rightarrow 0$), суммирование амплитудных значений заменится интегрированием, а сами значения спектра становятся бесконечно малыми ($1/T \rightarrow 0$). Для исключения последнего множителя $1/T$ из расчетной формулы $S(\omega)$ исключается, и интегральное преобразование Фурье приобретает следующий вид:

$$s(t) = (1/2\pi) \int_{-\infty}^{\infty} S(\omega) \exp(j\omega t) d\omega, \quad (4.2.1)$$

$$S(\omega) = \int_{-\infty}^{\infty} s(t) \exp(-j\omega t) dt. \quad (4.2.2)$$

Формулу (4.2.2) обычно называют формулой прямого преобразования Фурье, а формулу (4.2.1) – обратного преобразования Фурье. Этими выражениями устанавливается взаимно однозначная связь сигнала и его спектра, а точнее – плотности спектра сигнала в последовательной полосе малых (стремящихся к нулю) полосах частот. Эту величину называют спектральной плотностью сигнала. Спектральные функции содержат ровно столько информации, сколько и исходный сигнал.

При преобразовании сигнала в пространство гармонических частот и обратно формулы прямого и обратного преобразований Фурье тождественны за исключением знака аргументов экспоненты:

$$s(t) = \int_{-\infty}^{\infty} S(f) \exp(j2\pi ft) df, \quad (4.2.1')$$

$$S(f) = \int_{-\infty}^{\infty} s(t) \exp(-j2\pi ft) dt. \quad (4.2.2')$$

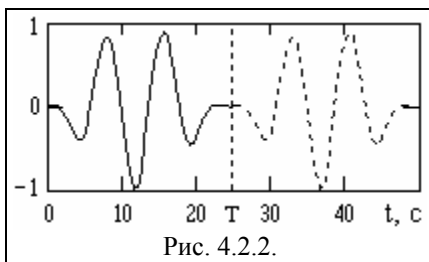


Рис. 4.2.2.

На рис. 4.2.2 сплошной кривой приведен пример непрерывного сигнала $s(t)$, энергия которого сосредоточена на конечном интервале $T = (0,25)$. Если нас не интересует форма данного сигнала за пределами интервала T , то спектр сигнала в виде ряда Фурье можно определить по формуле (4.1.2). При обратном преобразовании Фурье по формуле (4.1.1), т.е. при восстановлении сигнала по его спектру, в интервале T будет восстановлен исходный сигнал $s(t)$. Но если интервал для восстановления будет задан больше интервала T , например равным $0-2T$, то за пределами этого интервала начнется периодическое повторение исходного сигнала, как это показано пунктиром на рис. 4.2.2. Если такой процесс нежелателен и за пределами интервала T должны быть сохранены нулевые значения сигнала, то необходимо использовать интегральное преобразование Фурье (4.2.1, 4.2.2). При этом следует учитывать особенности интегрального преобразования.

Спектральная функция $S(\omega)$ представляет собой комплексную спектральную плотность сигнала, непрерывную на частотном интервале от $-\infty$ до ∞ . Если $s(t)$ – вещественная функция, то спектр этой функции является сопряжено симметричным относительно нулевой частоты

$$S(-\omega) = S^*(\omega)$$

и содержит четную действительную и нечетную мнимую части:

$$S(\omega) = A(\omega) - jB(\omega), \quad (4.2.3)$$

$$A(\omega) = \int_{-\infty}^{\infty} s(t)\cos(\omega t) dt, \quad (4.2.4)$$

$$B(\omega) = \int_{-\infty}^{\infty} s(t)\sin(\omega t) dt. \quad (4.2.5)$$

Как и в случае рядов Фурье, вещественные четные функции имеют вещественный четный спектр, представленный только спектральной функцией $A(\omega)$, а вещественные нечетные – нечетный и только мнимый спектр, представленный спектральной функцией $B(\omega)$.

Пример спектральной функции $S(f)$ для сигнала $s(t)$ на рис. 4.2.2 приведен на рис. 4.2.3. Как правило, графическое отображение спектральных функций выполняется в виде модуля и аргумента спектральной функции (амплитудного и фазового спектра), приведенных на рис. 4.2.4.

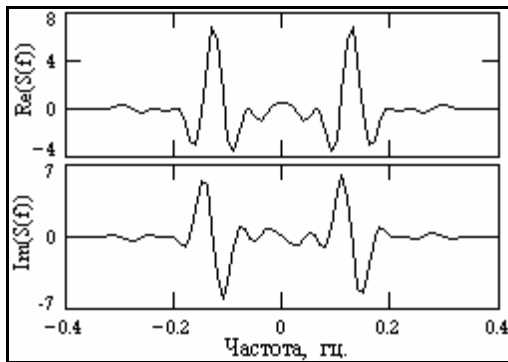


Рис. 4.2.3.

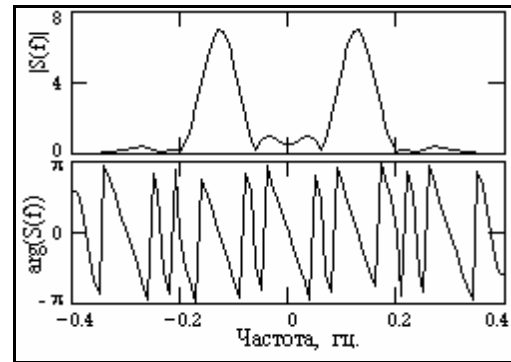


Рис. 4.2.4.

Такое представление аналогично (4.1.3'):

$$R(\omega) = \sqrt{A^2(\omega) + B^2(\omega)}, \quad (4.2.6)$$

$$\varphi(\omega) = \arctg(-B(\omega)/A(\omega)), \quad (4.2.7)$$

но в отношении функции модуля также имеет смысл спектральной плотности модуля.

Заметим также, что сопряженная симметричность спектральной функции позволяет в формулах (4.2.1)-(4.2.2) менять местами знаки аргументов в экспонентах, при этом изменяется только знак мнимой части и аргумента спектра.

Еще раз подчеркнем различие между спектрами и спектральными функциями сигналов. При практическом использовании формулы (4.2.2) для вычисления спектральных функций конечных сигналов, заданных на определенном интервале T , пределы интегрирования обычно устанавливаются по границам интервала T , так как нет необходимости выполнять интегрирование в бесконечных пределах, если за пределами интервала T мы имеем нулевые (или незначимые) значения сигнала. Однако при сравнении формулы (4.2.2) с выражением (4.1.2) можно наглядно видеть, что значения интеграла (4.2.2) не нормируются на величину интервала T . Отсюда следует, что числовые отсчеты значений модуля функции $S(\omega)$ для определенных значений ω_i не являются амплитудными значениями соответствующих гармонических колебаний с частотой ω_i . Значения $S(\omega)$ по сравнению со значениями функции $S(n\Delta\omega)$ по (4.1.2) при $n\Delta\omega = \omega_i$ завышены на множитель T . Это можно объяснить тем, что обратное преобразование Фурье по (4.1.1) представляет собой прямое суммирование гармоник с соответствующими амплитудами колебаний, в то время как интегрирование по (4.2.1) представляет собой предельное суммирование значений $S(\omega_i) \cdot d\omega_i$, где $d\omega = 2\pi/T$ (или, в обычном частотном представлении, $df = 1/T$) при $T \Rightarrow \infty$.

Что касается спектра фазовых углов, то значения по (4.2.7) и по (4.1.3') при $n\Delta\omega = \omega_i$

полностью совпадают, так как их вычисление производится по отношению мнимой и действительной части спектра, наличие (или отсутствие) постоянного множителя в которых не меняет значение отношения.

Тригонометрическая форма интеграла Фурье (при объединении комплексно сопряженных частей спектральных функций):

$$s(t) = (1/2\pi) \int_{-\infty}^{\infty} [A(\omega)\cos(\omega t) + B(\omega)\sin(\omega t)] d\omega. \quad (4.2.8)$$

$$s(t) = (1/2\pi) \int_{-\infty}^{\infty} R(\omega)\cos(\omega t - \varphi(\omega)) d\omega. \quad (4.2.8')$$

Прямое и обратное преобразование Фурье подобны. Любая теорема, доказанная для прямого преобразования Фурье, справедлива и для обратного преобразования, и наоборот. Это непосредственно следует из выражений прямого и обратного преобразования Фурье, которые различаются только знаком в экспоненте. Особенно наглядно (см. рис. 4.2.5) это видно для четных сигналов (заданных функциями, симметричными относительно $t = 0$), для которых $B(\omega) = 0$ и, соответственно, фазовый спектр равен нулю:

$$s(t) = 2 \int_0^{\infty} S(f)\cos(2\pi ft)df, \quad S(f) = 2 \int_0^{\infty} s(t)\cos(2\pi ft)dt.$$

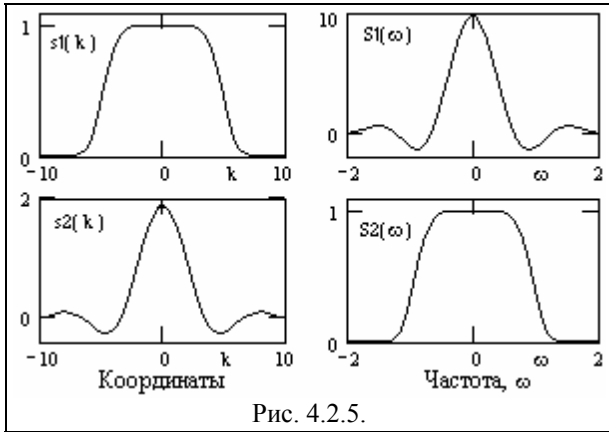


Рис. 4.2.5.

В математическом анализе для упрощения записей используют символическую форму обозначения преобразования Фурье:

$$s(t) \leftrightarrow S(f), \quad s(t) \leftrightarrow S(\omega),$$

где, в общем случае, как фурье-образ функции, так и она сама могут быть комплексными.

Для физических сигналов и их достаточно корректных математических моделей преобразование Фурье, как правило, всегда существует. С чисто математических позиций сигналу $s(t)$ можно сопоставить спектральную плотность $S(\omega)$, если существует интеграл:

$$\int_{-\infty}^{\infty} |s(t)| dt < \infty. \quad (4.2.9)$$

Преобразование Лапласа. Если условие (4.2.9) не выполняется, то определенные приближения спектральных плотностей вычисляются с использованием специальных методов, одним из которых является преобразование Лапласа.

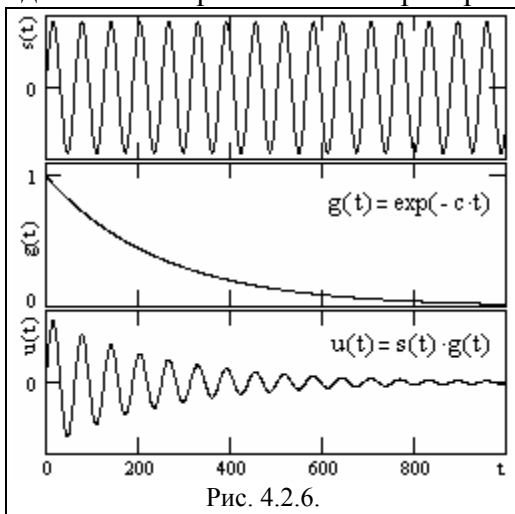


Рис. 4.2.6.

Допустим, что функция $s(t)$ задана на интервале $(0, \infty)$, равна нулю при $t < 0$, а интеграл спектральной функции (4.2.2) расходится. Умножим $s(t)$ на экспоненциальную функцию $\exp(-ct)$, где c - положительная константа, и выберем значение 'c' таким, чтобы произведение $u(t) = s(t) \cdot \exp(-ct)$ удовлетворяло условию абсолютной интегрируемости. Сущность данной операции хорошо видна на рис. 4.2.6. Интегрируемость функции $u(t)$ может быть установлена для любой функции $s(t)$ соответствующим выбором коэффициента 'c'. При этом спектральная плотность функции $u(t)$ может быть вычислена по формуле (4.2.2):

$$U(\omega, c) = \int_0^{\infty} [s(t) \exp(-ct)] \exp(-j\omega t) dt.$$

После объединения экспоненциальных функций это выражение можно переписать следующим образом:

$$U(c+j\omega) = \int_0^{\infty} s(t) \exp[-(c+j\omega)t] dt. \quad (4.2.10)$$

Соответствующее обратное преобразование Фурье функции $U(c+j\omega)$:

$$(1/2\pi) \int_{-\infty}^{\infty} S(c+j\omega) \exp(j\omega t) d\omega = s(t) \exp(-ct).$$

Для восстановления функции $s(t)$ достаточно умножить обе части данного выражения на $\exp(ct)$, объединить экспоненциальные множители под интегралом и заменить переменную интегрирования ω на $c+j\omega$:

$$s(t) = (1/2\pi) \int_{c-j\infty}^{c+j\infty} S(c+j\omega) \exp[(c+j\omega)t] d(c+j\omega). \quad (4.2.11)$$

Обозначим комплексную переменную $c+j\omega$ в выражениях (4.2.10,4.2.11) через p и получим общепринятую форму прямого и обратного преобразования Лапласа:

$$S(p) = \int_0^{\infty} s(t) \exp[-pt] dt. \quad (4.2.10')$$

$$s(t) = (1/2\pi) \int_{c-j\infty}^{c+j\infty} S(p) \exp(pt) d(p). \quad (4.2.11')$$

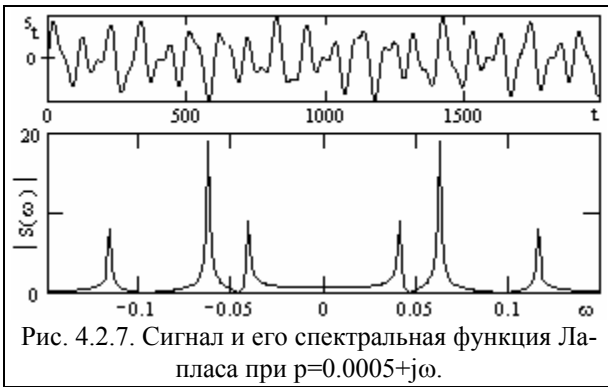


Рис. 4.2.7. Сигнал и его спектральная функция Лапласа при $p=0.0005+j\omega$.

Сигнальную функцию $s(t)$ в преобразованиях Лапласа обычно называют *оригиналом*, а ее спектральную функцию $S(p)$ - *изображением оригинала*. Пример спектральной функции Лапласа для оригинала - сложного и неограниченного во времени сигнала, состоящего из каузальной суммы трех гармоник, приведен на рис. 4.2.7. По спектральной функции Лапласа можно выделить эти три основных частоты сигнала и оценить соотношение их амплитуд. Ширина пиков спектральной функции при выделении "чистых" гармоник зависит от значения коэффициента 'с' и уменьшается при его уменьшении.

Если вместо p в изображениях оригинала подставить переменную $j\omega$, то будут получены спектральные функции, полностью идентичные преобразованию Фурье каузальных функций (имеющих нулевые значения при $t < 0$).

Обобщенный ряд Фурье. Тригонометрические функции не являются единственно возможными функциями разложения сигналов. В общем случае разложение сигнала $s(t)$ на интервале (a, b) в ряд вида $\sum_{k=0}^N c_k \varphi_k(t)$ может быть выполнено по произвольным функциям $\varphi_k(t)$.

При задании минимальной погрешности приближения

$$\Delta s = \int_a^b [s(t) - \sum_{k=0}^N c_k \varphi_k(t)]^2 dt$$

коэффициенты c_k могут быть найдены из системы линейных уравнений:

$$\frac{\partial \Delta s}{\partial c_k} = \int_a^b \frac{\partial}{\partial c_k} [s(t) - \sum_{k=0}^N c_k \varphi_k(t)]^2 dt = 0, \quad k = 0, 1, 2, \dots, N.$$

При линейной независимости функций $\varphi_k(t)$ данная система уравнений имеет единственное решение. Если все функции $\varphi_k(t)$ взаимно ортогональны и соответствующей нормировкой обеспечена их ортонормированность

$$\int_a^b \varphi_m(t) \varphi_n(t) dt = \begin{cases} 1, & n = m \\ 0, & n \neq m \end{cases},$$

то процесс нахождения коэффициентов c_k оказывается наиболее простым:

$$c_k = \int_a^b s(t) \varphi_k(t) dt,$$

и для принятого значения N погрешность приближения Δs является минимальной. Если при

$N \rightarrow \infty$ имеет место $\Delta s \rightarrow 0$, система функций $\varphi_k(t)$ называется базисной системой координат пространства сигналов $L^2[a, b]$. При этом имеет место равенство:

$$s(t) = \sum_{k=0}^{\infty} c_k \varphi_k(t).$$

Разложение по ортонормированной системе базисных функций называется *обобщенным рядом Фурье*, а набор коэффициентов c_k представляет собой спектр функции $s(t)$ в соответствующем базисе. В зависимости от специфики решаемых задач применяются различные системы базисных функций. В частности, используются разложения по полиномам Лежандра, Чебышева, Лагерра, Эрмита, функциям Хаара и Уолша и т.п.

4.3. СВОЙСТВА ПРЕОБРАЗОВАНИЙ ФУРЬЕ [1,17,21].

Свойствами преобразований Фурье определяется взаимное соответствие трансформации сигналов и их спектров.

1. Линейность. Преобразование Фурье относится к числу линейных интегральных операций, т.е. спектр суммы сигналов равен сумме спектров этих сигналов.

$$\sum_n a_n s_n(t) \Leftrightarrow \sum_n a_n S_n(\omega). \quad (4.3.1)$$

Пример суммирования сигналов и его отображения в суммирования спектров приведен на рис. 4.3.1:

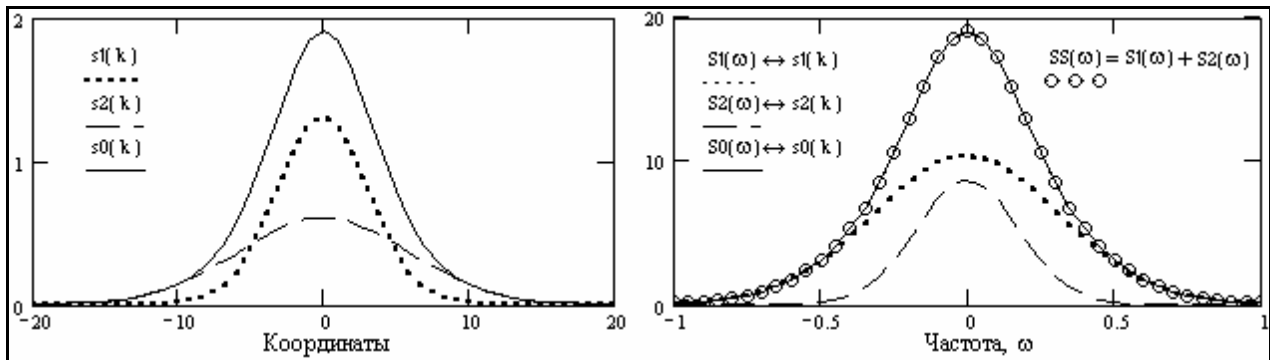


Рис. 4.3.1. Сигналы и их спектры. $s_0(k) = s_1(k) + s_2(k) \Leftrightarrow S_1(\omega) + S_2(\omega) = S_0(\omega)$.

Сигнал $s(t)$	Спектр $S(\omega)$
Четный	Вещественный, четный
Нечетный	Мнимый, нечетный
Произвольный	Действительная часть – четная. Мнимая часть - нечетная

2. Свойства четности преобразования определяются косинусными (четными, действительными) и синусными (нечетными, мнимыми) частями разложения и подобием прямого и обратного преобразований.

На рис. 4.3.2. приведены примеры, поясняющие свойства четности преобразования. Сигнал $s_1(k)$ является четным, $s_1(k) = s_1(-k)$, и имеет только вещественный четный спектр (мнимая часть спектральной функции представлена нулевыми значениями). Сигнал $s_2(k) = -s_2(-k)$ нечетный и имеет мнимый нечетный спектр, а нулевыми значениями представлена его действительная часть. Сигнал $s_3(k)$ образован суммой сигналов $s_1(k)$ и $s_2(k)$. Соответственно, спектральная функция сигнала представлена и действительной четной частью (принадлежащей $s_1(k)$), и мнимой нечетной частью (принадлежащей $s_2(k)$). При обратном преобразовании Фурье отдельно действительной и мнимой части спектра $S_3(\omega)$, равно как и любых других комплексных спектров, будут отдельно восстановлены четная и нечетная части исходного сигнала.

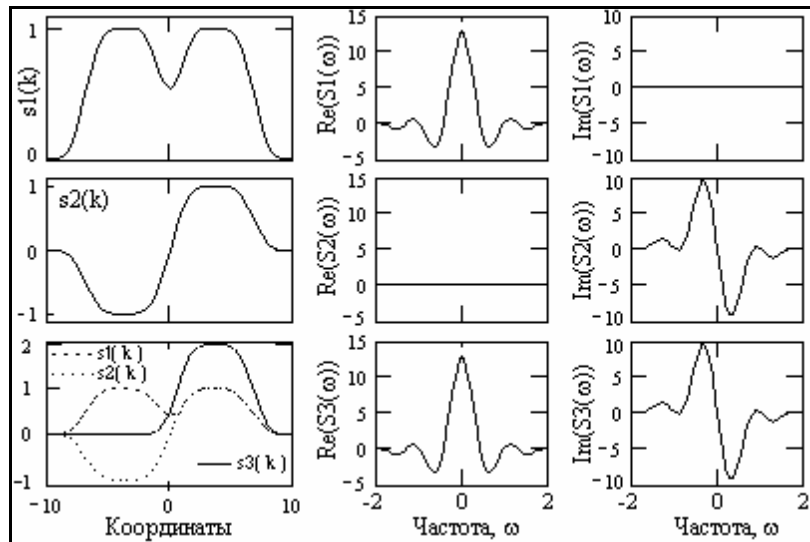


Рис. 4.3.2. Свойства четности преобразования.

Заметим, что произвольный исходный сигнал может быть задан в одностороннем варианте (в интервале $0-T$), но четная и нечетная части этого сигнала занимают интервал от $-T$ до T , при этом на левой половине числовой оси (от $-T$ до 0) эти два сигнала компенсируют друг друга, давая нулевые значения.

3. Изменение аргумента функции (сжатие или расширение сигнала) приводит к обратному изменению аргумента ее фурье-образа и обратно пропорциональному изменению его модуля. Действительно, если $s(t) \Leftrightarrow S(\omega)$, то при изменении длительности сигнала с сохранением его формы (растяжении сигнала по временной оси), т.е. для сигнала с новым аргументом $s(x) = s(at)$ при $x=at$, получаем:

$$s(at) \Leftrightarrow \int_{-\infty}^{\infty} s(at) \exp(-j\omega t) dt = (1/a) \int_{-\infty}^{\infty} s(x) \exp(-jx\omega/a) dx$$

$$s(at) \Leftrightarrow (1/a) S(\omega/a). \quad (4.3.2')$$

Выражение (4.3.2') действительно при $a > 0$. При $a < 0$ происходит зеркальный поворот сигнала относительно вертикальной оси, а замена переменной $t=x/a$ вызывает перестановку пределов интегрирования и, соответственно, изменение знака спектра:

$$s(at) \Leftrightarrow -(1/a) S(\omega/a). \quad (4.3.2'')$$

Обобщенная формула изменения аргумента:

$$s(at) \Leftrightarrow (1/|a|) S(\omega/a), \quad a \neq 0 \quad (4.3.2)$$

Если под аргументом функции и ее спектра понимать определенные физические единицы, например, время - частота, то отсюда следует: чем короче по своей длительности сигнал, тем шире по частоте его спектр, и наоборот. Это можно наглядно видеть на рис. 4.3.1. для сигналов $s1(k)$ и $s2(k)$ и их спектров $S1(\omega)$ и $S2(\omega)$.

От изменения аргумента функций следует отличать изменение масштаба представления функций. Изменение масштаба аргументов изменяет только оцифровку числовых осей отображения сигналов и их спектров, но не изменяет самих сигналов и спектров. Так, при масштабе оси времен $t=1$ секунда, масштаб оси частот $f=1/t=1$ герц, а при $t=1$ мксек $f=1/t=1$ МГц ($t=at$, $f=1/at$, $a=10^{-6}$).

4. Теорема запаздывания. Запаздывание (сдвиг, смещение) сигнала по аргументу функции на интервал t_0 приводит к изменению фазочастотной функции спектра (фазового угла всех гармоник) на величину $-\omega t_0$ без изменения модуля (амплитудной функции) спектра. Применяя замену переменной $t-t_0 = x$, получаем:

$$s(t-t_0) \Leftrightarrow \int_{-\infty}^{\infty} s(t-t_0) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} s(x) \exp(-j\omega x) \exp(-j\omega t_0) dx = S(\omega) \exp(-j\omega t_0) \quad (4.3.3)$$

Совершенно очевидно, что амплитуды гармоник сигнала при его сдвиге изменяться не должны. С учетом того, что $|\exp(-j\omega t_0)|=1$, это следует и из (4.3.3):

$$|S(\omega) \exp(-j\omega t_0)| = |S(\omega)|.$$

Фазовый спектр сдвигается на $-\omega t_0$ с линейной зависимостью от частоты:

$$S(\omega) \exp(-j\omega t_0) = R(\omega) \exp[j(\varphi(\omega))] \exp(-j\omega t_0) = R(\omega) \exp[j(\varphi(\omega) - \omega t_0)]. \quad (4.3.4)$$



Рис. 4.3.3. Изменение спектра сигнала при его сдвиге.

Пример двух одинаковых сигналов, сдвинутых относительно друг друга на $t_0=1$, и соответствующих данным сигналам спектров приведен на рис. 4.3.3.

5. Преобразование производной (дифференцирование сигнала):

$$\begin{aligned} s(t) &= d[y(t)]/dt = d\left[\int_{-\infty}^{\infty} Y(\omega) \exp(j\omega t) d\omega\right]/dt = \int_{-\infty}^{\infty} Y(\omega) [d(\exp(j\omega t))/dt] d\omega = \\ &= \int_{-\infty}^{\infty} j\omega Y(\omega) \exp(j\omega t) d\omega \Leftrightarrow j\omega Y(\omega). \end{aligned} \quad (4.3.5)$$

Таким образом, дифференцирование сигнала отображается в спектральной области простым умножением спектра сигнала на *оператор дифференцирования сигнала в частотной области* $j\omega$, что эквивалентно дифференцированию каждой гармоники спектра. Умножение на $j\omega$ приводит к обогащению спектра производной сигнала высокочастотными составляющими (по сравнению с исходным сигналом) и уничтожает составляющие с нулевой частотой.

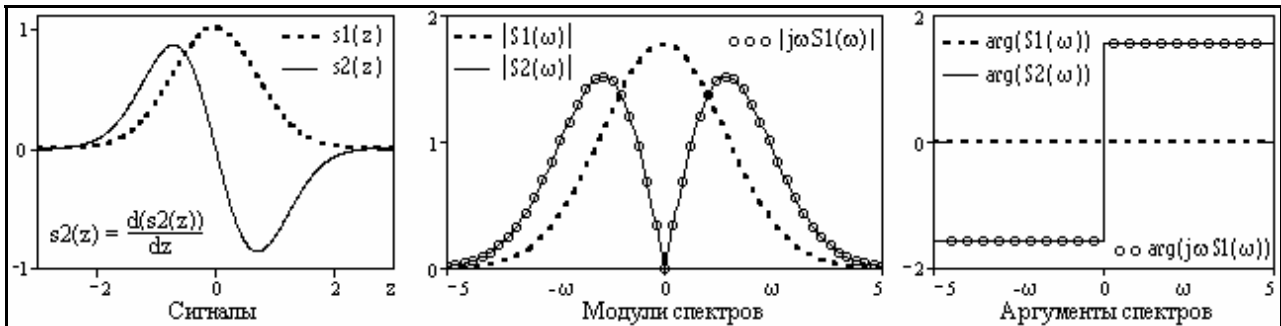


Рис. 4.3.4. Спектры сигнала и его производной.

Пример сигнала, его производной и соответствующих им спектров приведен на рис. 4.3.4. По изменению аргумента спектра (для четного исходного сигнала он был нулевым) можно видеть, что для всех гармоник спектра появляется сдвиг фаз на $\pi/2$ (90°) для положительных частот, и на $-\pi/2$ (-90°) для отрицательных частот.

В общем случае, для кратных производных:

$$d^n[y(t)]/dt^n = (j\omega)^n Y(\omega). \quad (4.3.6)$$

6. Преобразование интеграла сигнала в частотной области при известном спектре сигнала может быть получено из следующих простых соображений. Если имеет место $s(t) = d[y(t)]/dt \Leftrightarrow j\omega Y(\omega) = S(\omega)$, то должна выполняться и обратная операция: $y(t) = \int s(t) dt \Leftrightarrow Y(\omega) = S(\omega)/j\omega$. Отсюда следует:

$$\int s(t) dt \Leftrightarrow (1/j\omega)S(\omega). \quad (4.3.7)$$

Оператор интегрирования в частотной области $(1/j\omega)$ при $\omega > 1$ ослабляет в амплитудном спектре высокие частоты и при $\omega < 1$ усиливает низкие. Фазовый спектр сигнала смещается на -90° для положительных частот и на 90° для отрицательных. Пример модуля спектра сигнала и его интегральной функции приведены на рис. 4.3.5.

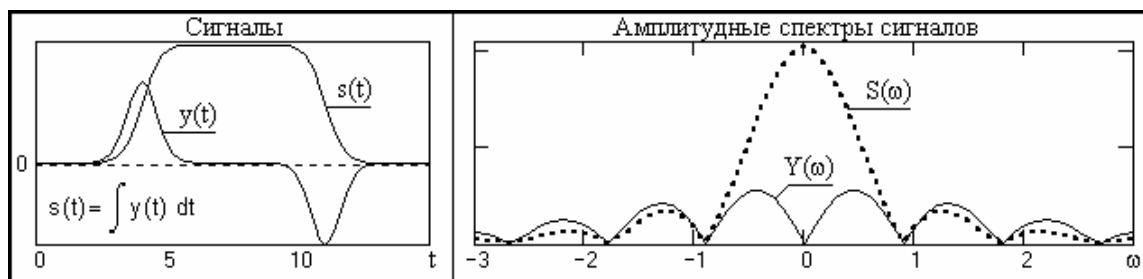


Рис. 4.3.5. Сигналы и амплитудные спектры сигналов.

Формула (4.3.7) справедлива для сигналов с нулевой постоянной составляющей. При интегрировании сигналов с определенным значением постоянной составляющей $C = \text{const}$ в правой части выражения (4.3.7) появляется дополнительное слагаемое преобразования Фурье постоянной составляющей C , которое представляет собой, как будет показано ниже, дельта-функцию на нулевой частоте с весовым коэффициентом, равным значению C :

$$y(t) = (1/j\omega)S(\omega) + C\delta(\omega_0).$$

7. Преобразование свертки сигналов $y(t) = s(t) * h(t)$:

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\tau) h(t-\tau) \exp(-j\omega t) d\tau dt.$$

$$Y(\omega) = \int_{-\infty}^{\infty} s(\tau) d\tau \int_{-\infty}^{\infty} h(t-\tau) \exp(-j\omega t) dt.$$

По теореме запаздывания (4.3.3):

$$\int_{-\infty}^{\infty} h(t-\tau) \exp(-j\omega t) dt = H(\omega) \exp(-j\omega\tau).$$

Отсюда: $Y(\omega) = \int_{-\infty}^{\infty} H(\omega) s(\tau) \exp(-j\omega\tau) d\tau = H(\omega) \cdot S(\omega).$

$$s(t) * h(t) \Leftrightarrow S(\omega) H(\omega). \quad (4.3.8)$$

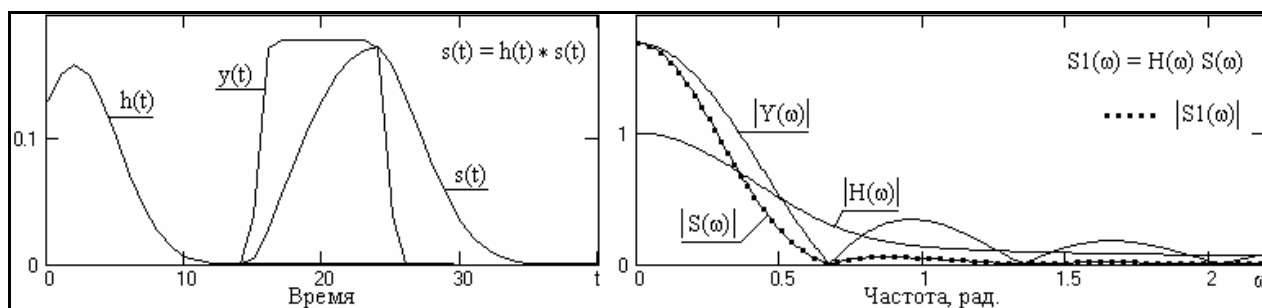


Рис. 4.3.6. Сигналы и амплитудные спектры сигналов.

Пример выполнения свертки в частотной области приведен на рис. 4.3.6. Отметим, что частотное представление $H(\omega)$ импульсного отклика $h(t)$ линейной системы (или соответствующей линейной операции) имеет смысл частотной передаточной функции системы и позволяет определить сигнал на выходе системы (в частотной форме представления) при задании произвольного сигнала (в частотной форме) на ее входе. По существу, функция $H(\omega)$ представляет собой распределение по частоте коэффициента пропускания частотных составляющих сигнала с входа на выход системы (операции).

Таким образом, свертка функций в координатной форме отображается в частотном представлении произведением фурье-образов этих функций.

8. Преобразование произведения сигналов $y(t) = s(t) \cdot h(t)$:

$$\begin{aligned} Y(\omega) &= \int_{-\infty}^{\infty} s(t) h(t) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} s(t) \left[(1/2\pi) \int_{-\infty}^{\infty} H(\omega') \exp(j\omega' t) d\omega' \right] dt = \\ &= (1/2\pi) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(t) H(\omega') \exp(-j(\omega - \omega') t) d\omega' dt = (1/2\pi) \int_{-\infty}^{\infty} H(\omega') d\omega' \int_{-\infty}^{\infty} s(t) \exp(-j(\omega - \omega') t) dt = \\ &= (1/2\pi) \int_{-\infty}^{\infty} H(\omega') S(\omega - \omega') d\omega' = (1/2\pi) H(\omega) * S(\omega). \end{aligned} \quad (4.3.9)$$

Таким образом, *произведение функций в координатной форме отображается в частотном представлении сверткой фурье-образов этих функций* (с нормировочным множителем $(1/2\pi)$, учитывающем несимметричность прямого и обратного преобразования Фурье функций $s(t)$ и $h(t)$ при использовании угловых частот).

9. Умножение сигнала на гармоническую функцию заполняет сигнал гармонической частотой и формирует радиоимпульс. Без учета начальной фазы гармоник:

$$s(t) = u(t) \cos(\omega_0 t).$$

Спектр радиоимпульса:

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{\infty} u(t) \cos(\omega_0 t) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} u(t) \frac{1}{2} [\exp(j\omega_0 t) + \exp(-j\omega_0 t)] \exp(-j\omega t) dt = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} u(t) \exp(j\omega_0 t) \exp(-j\omega t) dt + \frac{1}{2} \int_{-\infty}^{\infty} u(t) \exp(-j\omega_0 t) \exp(-j\omega t) dt = \\ &= \frac{1}{2} U(\omega) \exp(j\omega_0 t) + \frac{1}{2} U(\omega) \exp(-j\omega_0 t). \end{aligned} \quad (4.3.10)$$

Спектры сигналов обычно низкочастотные и сосредоточены в центре частотной оси. Частота гармоник заполнения, как правило, много больше максимальной частоты гармоник сигнала. Из (4.3.10) следует, что спектр сигнала раздвигается (с коэффициентом $1/2$) и смешается влево и вправо по оси частот на частоты $\pm\omega_0$. Особенно наглядно это видно для четных сигналов и приведено на рис. 4.3.7.

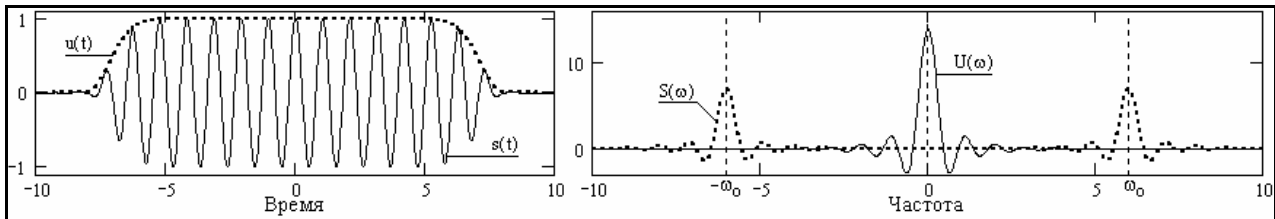


Рис. 4.3.7. Радиоимпульс и его амплитудный спектр.

10. Спектры мощности. Если функция $s(t)$ имеет фурье-образ $S(\omega)$, то спектральная плотность мощности данной функции определяется выражениями:

$$w(t) = s(t) s^*(t) = |s(t)|^2 \Leftrightarrow |S(\omega)|^2 = S(\omega) S^*(\omega) = W(\omega). \quad (4.3.11)$$

Спектр мощности - вещественная неотрицательная четная функция, которую очень часто называют энергетическим спектром. Спектр мощности, как квадрат модуля спектра сигнала, не содержит фазовой информации о частотных составляющих, а, следовательно, восстановление сигнала по спектру мощности невозможно. Это означает также, что сигналы с различными фазовыми характеристиками могут иметь одинаковые спектры мощности. В частности, сдвиг сигнала не отражается на его спектре мощности.

Для функций мощности взаимодействия сигналов в частотной области соответственно имеем частотные спектры мощности взаимодействия сигналов:

$$\begin{aligned} W_{xy}(\omega) &= X(\omega) Y^*(\omega), \\ W_{yx}(\omega) &= Y(\omega) X^*(\omega), \\ W_{xy}(\omega) &= W_{yx}^*(\omega). \end{aligned}$$

Функции мощности взаимодействия сигналов комплексные, даже если обе функции $x(t)$ и $y(t)$ вещественны, при этом $\text{Re}[W_{xy}(\omega)]$ - четная функция, а $\text{Im}[W_{xy}(\omega)]$ - нечетная. Отсюда полная энергия взаимодействия сигналов при интегрировании функций мощности взаимодействия определяется только реальной частью спектра:

$$E_{xy} = (1/2\pi) \int_{-\infty}^{\infty} W_{xy}(\omega) d\omega = (1/\pi) \int_{-\infty}^{\infty} \text{Re}[W_{xy}] d\omega,$$

и всегда является вещественным числом.

11. Равенство Парсеваля. Полная энергия спектра сигнала:

$$E_s = \int_{-\infty}^{\infty} W(f) df = \int_{-\infty}^{\infty} |S(f)|^2 df. \quad (4.3.12)$$

Так как координатное и частотное представление по существу только разные математические отображения одного и того же сигнала, то равной должна быть и энергия сигнала в

двух представлениях, откуда следует равенство Парсеваля:

$$\int_{-\infty}^{\infty} |s(t)|^2 dt = \int_{-\infty}^{\infty} |S(f)|^2 df,$$

т.е. энергия сигнала равна интегралу модуля его частотного спектра - сумме энергий всех частотных составляющих сигнала. Аналогично для энергии взаимодействия сигналов:

$$\int_{-\infty}^{\infty} x(t) y^*(t) dt = \int_{-\infty}^{\infty} X(f) Y^*(f) df.$$

Из равенства Парсеваля следует инвариантность скалярного произведения сигналов и нормы относительно преобразования Фурье:

$$(x(t), y(t)) = (X(f), Y(f)), \quad \|x(t)\|^2 = \|X(f)\|^2.$$

Не следует забывать, что при представлении спектров в круговых частотах (по ω) в правой части приведенных равенств должен стоять множитель $1/2\pi$.

4.4. СПЕКТРЫ НЕКОТОРЫХ СИГНАЛОВ [1,16].

1. Единичные импульсы. Функция $\delta(t)$, центрированная относительно $t = 0$, значения которой по определению равны нулю при $t \neq 0$, а интеграл от $-\infty$ до ∞ равен 1, имеет равномерное спектральное распределение в бесконечной полосе частот от 0 до ∞ :

$$TF[\delta(t)] = \int_{-\infty}^{\infty} \delta(t) \exp(-j\omega t) dt = 1. \quad (4.4.1)$$

Это следует и из свойства свертки функций, поскольку свертка функции с бесконечно коротким импульсом с единичной площадью не должна приводить к изменению функции:

$$s(t) * \delta(t) = s(t).$$

Выполняя преобразование Фурье правой и левой части данного выражения, имеем:

$$S(\omega) H(\omega) = S(\omega),$$

что может быть реализовано только при $H(\omega) = 1$.

Отсюда следует также, что дельта-функцию можно записать в виде обратного преобразования Фурье:

$$\delta(t) = (1/2\pi) \int_{-\infty}^{\infty} \exp(j\omega t) d\omega. \quad (4.4.2)$$



Рис. 4.4.1. Спектр функции $\delta(t-2)$

С учетом теоремы запаздывания (4.3.3), для обобщенной функции Дирака соответственно имеем:

$$\delta(t-\tau) \Leftrightarrow \exp(-j\omega\tau). \quad (4.4.1')$$

$$\delta(t-\tau) = (1/2\pi) \int_{-\infty}^{\infty} \exp(j\omega(t-\tau)) d\omega. \quad (4.4.1'')$$

Пример спектра функции приведен на рис.

4.4.1.

Для единичного импульса с площадью, равной P:

$$P \cdot \delta(t) \Leftrightarrow P.$$

Если спектром весовой дельта-функции является константа, то на основе дуальности преобразования Фурье спектром константы должна быть весовая дельта-функция в нуле частотной оси.

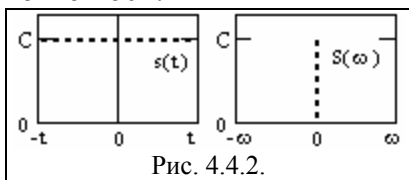


Рис. 4.4.2.

$$C \Leftrightarrow C \cdot \delta(\omega).$$

Представить графически эту операцию для непрерывных функций невозможно. Но для дискретных спектральных функций с использованием весового импульса Кронекера она имеет вполне реальный смысл (рис. 4.4.2).

2. Гребневая функция $\text{Ш}_T(t)$ представляет собой последовательность импульсов Дирака с периодом $T = 1/F$, где F- частота следования импульсов:

$$\text{Ш}_T(t) = \sum_{k=-\infty}^{\infty} \delta(t-kT).$$

Спектр гребневой функции (с учетом теоремы запаздывания при $\Delta f = 1/T = F$) также представляет собой последовательность импульсов Дирака:

$$\text{Ш}_T(t) = (1/T) \sum_{k=-\infty}^{\infty} \exp(-2\pi j n \Delta f t) \Leftrightarrow (1/T) \sum_{k=-\infty}^{\infty} \delta(f-kF) = F \cdot \text{Ш}_F(f). \quad (4.4.3)$$



Рис. 4.4.2. П - импульсы

3. Спектр прямоугольного импульса $\Pi_r(t)$ амплитудой U и длительностью r (рис. 4.4.2). При расположении начала координат по центру импульса:

$$\Pi_r(\omega) = \int_{-\infty}^{\infty} \Pi_r(t) \exp(-j\omega t) dt = U \int_{-r/2}^{r/2} \exp(-j\omega t) dt,$$

$$\Pi_r(\omega) = rU \sin(\omega r/2) / (\omega r/2) = rU \text{sinc}(\omega r/2). \quad (4.4.4)$$

Вид функций $\Pi_r(\omega)$ приведен на рис. 4.4.3. Представлена только часть частотного диапазона, в остальной части диапазона постепенно затухающие флюктуации спектров, которые, чисто теоретически, простираются до бесконечности.

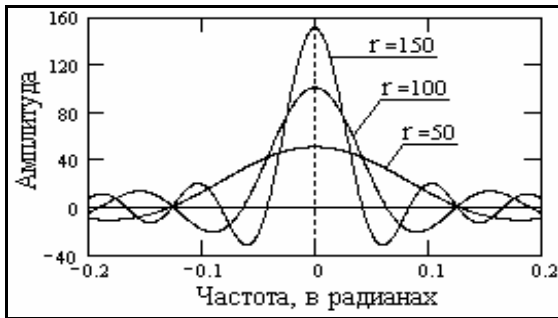


Рис. 4.4.3. Спектры П - импульсов.

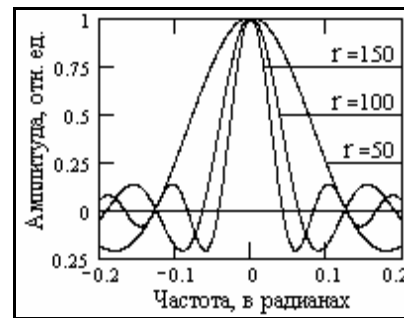


Рис. 4.4.4. Спектры П-импульсов

Как и следовало ожидать, для вещественной и четной динамической П-функции спектр сигнала также является вещественной и четной функцией частоты. Спектр имеет лепестковый характер, и ширина лепестков (по пересечениям нулевой линии) обратно пропорциональна длительности импульсов и равна $2\pi/r$. Значение спектральной плотности на нулевой частоте равно площади импульсов.

Функция вида $\sin(x)/x$ в анализе сигналов встречается довольно часто и имеет специальное обозначение: $\text{sinc}(x) = \sin(x)/x$. Она называется интегральным синусом или функцией отсчетов.

На рис. 4.4.4 приведены нормированные по площади спектры этих же импульсов. При сравнении спектров с рис. 4.4.2 можно наглядно видеть характер зависимости ширины спектров (по ширине главного максимума) от длительности импульсов. Нетрудно также заметить, что форма спектра П - импульсов остается практически постоянной и только "растягивается" по шкале частоты при уменьшении длительности импульсов. Чем шире сигнал, тем короче его спектр.

Если прямоугольный импульс начинается в момент времени t_0 , то имеем:

$$\Pi(\omega) = U \int_{t_0}^{t_0+r} \exp(-j\omega t) dt = rU \text{sinc}(\omega r/2) \exp[-j\omega(t_0+r/2)]. \quad (4.4.5)$$

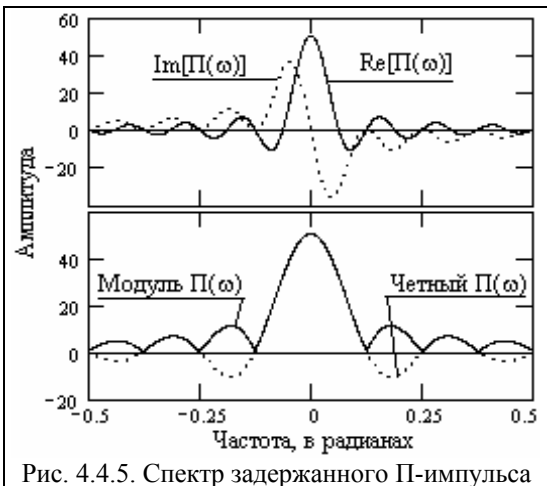


Рис. 4.4.5. Спектр задержанного П-импульса

Это выражение может быть получено непосредственно из (4.4.4) с использованием теоремы смещения. Вид функций $\Pi(\omega)$ при $r = 50$ и $t_0 = 50$ приведен на рис. 4.4.5.

Как видно на рисунке, спектр сигнала, несимметричного относительно $t = 0$, имеет две части: четную действительную $A(\omega) = \text{Re}(\Pi(\omega))$, и нечетную мнимую $B(\omega) = \text{Im}(\Pi(\omega))$. Модуль спектра $R(\omega) = |\Pi(\omega)|$ всегда четный, имеет только положительные значения и полностью повторяет $|\Pi_r(\omega)|$ четного импульса (при смещении начала координат в центр импульса).

При изменении величины сдвига импульса модуль спектра остается без изменений, т.к. амплитуда частотных составляющих сигнала зависит только от его формы и не меняется от места расположения сигнала на координатной оси. Сдвиг сигнала определяет его фазовый спектр, пример которого для задержанного П-импульса приведен на рис. 4.4.6. Заметим, что фактический фазовый спектр имеет непрерывный характер. Пилообразная форма кривых на рис. 4.4.6 объясняется периодическим сбросом действительных значений фазы сигнала на величину π .



Учитывая, что значения функций на отрицательных частотах спектра всегда в одном и том же порядке однозначно соотносятся со значениями на положительных частотах (четные функции $A(\omega)$ и $R(\omega)$, нечетные функции $B(\omega)$ и $\varphi(\omega)$), в дальнейшем двусторонние спектры сигналов будем приводить только для области положительных частот.

Для количественной характеристики соотношения формы сигналов и их спектров применяют понятие *базы сигнала*, под которой понимают произведение эффективных значений длительности сигнала и ширины его спектра. Конкретное значение базы сигнала зависит от способа определения этих параметров. Для сигналов простой формы значение базы обычно составляет несколько единиц. Если для прямоугольного импульса эффективную ширину спектра принять по длительности центрального пика ($2\pi/r$), то значение базы сигнала будет равно 2π .

Длительности сигналов и ширина их спектров связаны *принципом неопределенности*, которым устанавливается, что значение их произведения (база сигналов) не может быть меньше 1. Этим устанавливается, что не может существовать коротких сигналов с узким спектром. Напротив, максимальное значение базы сигналов не ограничивается, и могут существовать *сигналы с большой базой*, имеющие как большую длительность, так и широкий спектр.

Если прямоугольные импульсы повторяются с периодом T , то соответственно при $\Delta\omega = 1/T$ имеем:

$$\Pi_r(k\Delta\omega) = (rU/T) \operatorname{sinc}(k\Delta\omega r/2) \exp(-jk\Delta\omega(t_0 - r/2)). \quad (4.4.6)$$

Как и положено, спектр периодического сигнала дискретен по ω , а при снятии нормировки спектра на длительность периода (умножением на T) огибающая спектра повторяет выражение (4.4.5).

4. Треугольные импульсы длительностью r по основанию с площадью, равной P , могут быть получены сверткой двух прямоугольных импульсов длительностью $r/2$ с амплитудой $2P/r$, откуда:

$$s(t) = \Pi_{r/2}(t) * \Pi_{r/2}(t) \Leftrightarrow S(\omega) = \Pi_{r/2}(\omega) \Pi_{r/2}(\omega), \\ S(\omega) = P \operatorname{sinc}^2(\omega r/4).$$

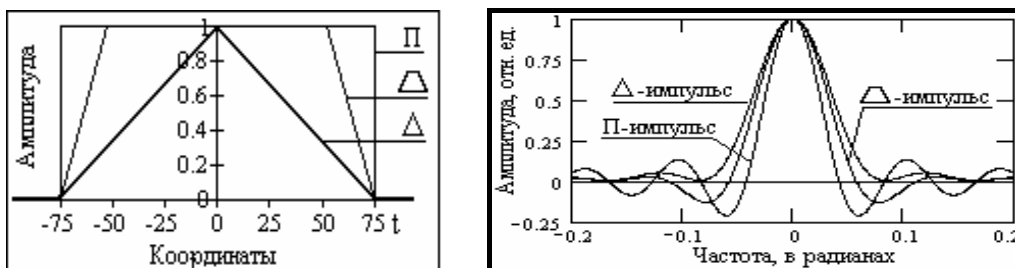


Рис. 4.4.7. Форма и спектры импульсов.

Спектр треугольного импульса также имеет лепестковую структуру с шириной лепестков $4\pi/r$. Соответственно, база треугольного импульса равна 4π . Спектральная функция за счет квадратирования интегрального синуса имеет только положительные значения.

Аналогично можно получить и спектры трапеций (при разной длительности П-импульсов).

Примеры импульсов и сопоставление формы их нормированных спектров (делением значений $S(\omega)$ на площадь импульсов - значение $S(0)$) приведены на рис. 4.4.6.

Заметим, что обратная операция – аппроксимация спектра сигнала произведением спектров простых сигналов с последующим переводом спектров в координатную область, позволяет представить сложный исходный сигнал в виде свертки более простых сигналов.

5. Экспоненциальный импульс $s(t) = U \exp(-at)$, $t \geq 0$, $a > 0$. Функция $\exp(-at)$ только условно может быть названа импульсом, т.к. определена и при $t \Rightarrow \infty$, но при $a > 0$ она достаточно быстро затухает. Преобразование Фурье экспоненциального импульса, как произвольного одностороннего сигнала, имеет действительную и мнимую части:

$$S(\omega) = U \int_0^{\infty} \exp(-(a+j\omega)t) dt = U/(a+j\omega).$$

Функция $S(\omega)$ бесконечна по частоте. Форма импульса, модуль и аргумент его спектра (фазовая характеристика в градусах) приведены на рис. 4.4.7.

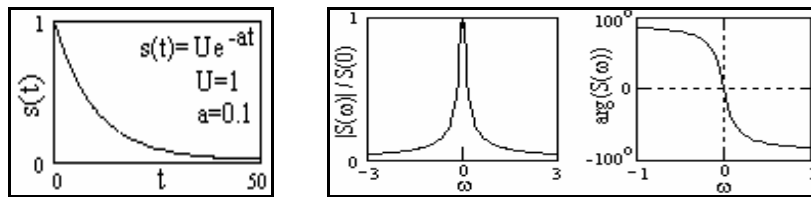


Рис. 4.4.7. Форма и спектр экспоненциального импульса.

6. Функции Лапласа и Гаусса. Для функции Лапласа (экспонента по модулю t) имеем следующее преобразование:

$$U \exp(-a|t|) \Leftrightarrow 2aU/(a^2+\omega^2), a>0. \quad (4.4.8)$$

Форма функции (при $U=1$, $a=0.1$) и ее вещественный спектр (функция четная и представлена только действительной частью) приведены на рис. 4.4.8.

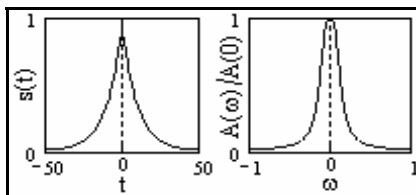


Рис. 4.4.8. Функция Лапласа

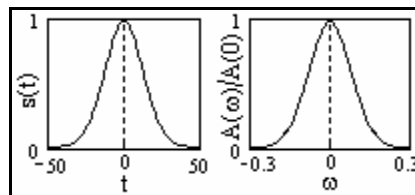


Рис. 4.4.9. Функция Гаусса.

Преобразование для центрированной функции Гаусса:

$$U \exp(-pt^2) \Leftrightarrow U \sqrt{\pi/p} \exp(-\omega^2/4\pi). \quad (4.4.9)$$

Спектр центрированной функции Гаусса - также функция Гаусса. Форма функции (при $U=1$, $a=0.003$) и ее вещественный спектр приведены на рис. 4.4.9. Если эффективную длительность и ширину спектра гауссовых функций определять по уровню $1/e$ от максимума ($\tau=2/a$, $\Delta\omega=a$), то база сигнала равна 4.

Сравнивая на рисунках 4.4.8 и 4.4.9 функции Лапласа и Гаусса и их спектры (с учетом масштаба последних), нетрудно заметить, что чем более плавно изменяются значения сигнала (меньше его дифференциал), тем более низкочастотным является спектр сигнала.

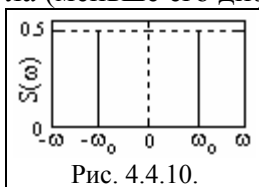


Рис. 4.4.10.

7. Спектр косинусоиды. Одним из условий применения интегрального преобразования Фурье функций является их абсолютная интегрируемость. Применение \square -функций позволяет получить спектральную плотность и для неинтегрируемых функций. Так, для косинусоиды с частотой ω_0 спектр может быть определен с помощью свойства смещения:

$$\begin{aligned} (1/2) \delta(\omega-\omega_0) &\Leftrightarrow (1/2) \exp(-j\omega_0 t), \\ (1/2) \delta(\omega+\omega_0) &\Leftrightarrow (1/2) \exp(j\omega_0 t), \\ (1/2) [\delta(\omega-\omega_0)+\delta(\omega+\omega_0)] &\Leftrightarrow \cos(\omega_0 t). \end{aligned}$$

Спектральная плотность косинусоиды вещественна и представляет собой два импульса Дирака, расположенных симметрично относительно $\omega = 0$ на частотах $-\omega_0$ и ω_0 (рис. 4.4.10, с условной нормировкой по амплитуде).

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
11. Зиновьев А.Л., Филиппов Л.И. Введение в теорию сигналов и цепей: Учебное пособие для вузов. - М.: Высшая школа, 1975. - 264 с.
16. Макс Ж. Методы и техника обработки сигналов при физических измерениях: В 2-х томах. - М.: Мир, 1983.
17. Никитин А.А. Теоретические основы обработки геофизической информации: Учебник для вузов. - М.: Недра, 1986.- 342 с.
21. Рапопорт М.Б. Вычислительная техника в полевой геофизике: Учебник для вузов. - М.: Недра, 1993.- 350 с.
24. Хемминг Р.В. Цифровые фильтры. – М.: Недра, 1987. – 221 с.
25. Сергиенко А.Б. Цифровая обработка сигналов. – СПб.: Питер, 2003. – 608 с.

Главный сайт автора ♦ Лекции по сигналам ♦ Практикум

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 5: ДИСКРЕТИЗАЦИЯ СИГНАЛОВ

Все вещи таковы, каков дух того, кто ими владеет. Если он умеет ими пользоваться, они хороши. Если не умеет - плохи.

Публий Теренций. Римский драматург, II в.д.н.э.

Студентом не раз пытался равномерно дискретизировать палку ливерной колбасы. Никогда не получалось. Свой кусок всегда был на рубль длиннее и на полтинник толще.

Александр Кудрявцев. Уральский геофизик и конструктор, XX в.

Содержание:**Введение.**

5.1. Задачи дискретизации функций. Принципы дискретизации. Воспроизведение сигнала.

5.2. Равномерная дискретизация. Спектр дискретного сигнала. Интерполяционный ряд Котельникова-Шеннона. Дискретизация с усреднением. Дискретизация спектров. Информационная тождественность динамической и спектральной формы сигнала. Дискретизация усеченных сигналов. Соотношение спектров одиночного и периодического сигналов.

5.3. Дискретизация по критерию наибольшего отклонения.

5.4. Адаптивная дискретизация.

5.5. Квантование сигналов.

5.6. Децимация и интерполяция данных.

Литература.

ВВЕДЕНИЕ.

В первой половине XX века при регистрации и обработке информации использовались, в основном, измерительные приборы и устройства аналогового типа, работающие в реальном масштабе времени, при этом даже для величин, дискретных в силу своей природы, применялось преобразование дискретных сигналов в аналоговую форму. Положение изменилось с широким распространением микропроцессорной техники и ЭВМ. Цифровая регистрация и обработка информации оказалась более совершенной и точной, более универсальной, многофункциональной и гибкой. Мощь и простота цифровой обработки сигналов настолько преобладают над аналоговой, что преобразование аналоговых по природе сигналов в цифровую форму давно стало производственным стандартом.

Под дискретизацией сигналов понимают преобразование функций непрерывных переменных в функции дискретных переменных, по которым исходные непрерывные функции могут быть восстановлены с заданной точностью. Роль отсчетов выполняют квантованные значения функций в дискретной шкале координат. Под квантованием понимают преобразование непрерывной по значениям величины в величину с дискретной шкалой значений из конечного множества разрешенных, которые называют уровнями квантования. Если уровни квантования нумерованы, то результатом преобразования является число, которое может быть выражено в любой числовой системе.

Для производственных задач обработки данных обычно требуется значительно меньше информации, чем ее поступает от измерительных датчиков в виде непрерывного аналогового сигнала. При статистических флуктуациях измеряемых величин и конечной погрешности средств измерений информация о величине сигнала всегда ограничена. Рациональное выполнение дискретизации и квантования исходных данных дает возможность снизить затраты на хранение и обработку информации. Использование дискретных сигналов позволяет применять методы кодирования информации с возможностью последующего обнаружения и исправления ошибок при обращении информации. Цифровая форма сигналов облегчает также унификацию операций преобразования информации на всех этапах ее обращения.

Сущность дискретизации аналоговых сигналов заключается в том, что непрерывность во времени аналоговой функции $s(t)$ заменяется последовательностью коротких импульсов, амплитудные значения которых s_n в общем случае определяются с помощью дискретных ве-

совых функций, либо непосредственно выборками (отсчетами) мгновенных значений сигнала $s(t)$ в моменты времени t_n .

5.1. ЗАДАЧИ ДИСКРЕТИЗАЦИИ ФУНКЦИЙ [10, 21].

Принципы дискретизации. Представление сигнала $s(t)$ на интервале T совокупностью дискретных значений c_n записывается в виде:

$$(c_1, c_2, \dots, c_N) = A[s(t)],$$

где A - оператор дискретизации. Запись операции восстановления сигнала $s(t)$:

$$s'(t) = B[(c_1, c_2, \dots, c_N)].$$

Выбор операторов A и B определяется требуемой точностью восстановления сигнала. Наиболее простыми являются линейные операторы. В общем случае:

$$c_n = \int_t q_n(t) s(t) dt, \quad (5.1.1)$$

где $q_n(t)$ - система весовых функций. Воспроизводящая функция $s'(t)$ соответственно представляется аппроксимирующим полиномом:

$$s'(t) = \sum_n c_n v_n(t), \quad (5.1.2)$$

где $v_n(t)$ - система базисных функций. Базисные функции ортогональны и обеспечивают сходимость ряда к $s(t)$ при $n \Rightarrow \infty$. Оптимальными являются методы дискретизации, обеспечивающие минимальный числовой ряд при заданной погрешности воспроизведения сигнала.

Отсчеты в выражении (5.1.1) связаны с операцией интегрирования, что обеспечивает высокую помехоустойчивость дискретизации. Однако в силу сложности технической реализации "взвешенного" интегрирования, последнее используется достаточно редко, при высоких уровнях помех. Более широкое распространение получили методы, при которых сигнал $s(t)$ заменяется совокупностью его мгновенных значений $s(t_n)$ в моменты времени t_n . Роль весовых функций в этом случае выполняют гребневые (решетчатые) функции. Отрезок времени Δt между соседними отсчетами называют шагом дискретизации. Дискретизация называется равномерной с частотой $F=1/\Delta t$, если значение Δt постоянно по всему диапазону преобразования сигнала. При неравномерной дискретизации значение Δt между выборками может изменяться по определенной программе или в зависимости от изменения каких-либо параметров сигнала.

Воспроизведение непрерывного сигнала по выборкам может проводиться как на основе ортогональных, так и неортогональных базисных функций. При неортогональных базисных функциях используются, в основном, степенные алгебраические полиномы вида:

$$s'(t) = \sum_{n=0}^N c_n t^n. \quad (5.1.3)$$

Если значения аппроксимирующего полинома совпадают со значениями выборок в моменты их отсчета, то такой полином называют интерполирующим. В качестве интерполирующих полиномов обычно используются многочлены Лагранжа. Для реализации интерполирующих полиномов необходима задержка сигнала на интервал дискретизации, что в системах реального времени требует определенных технических решений. В качестве экстраполирующих полиномов используют, как правило, многочлены Тейлора.

Естественным требованием к выбору частоты дискретизации является внесение минимальных искажений в динамику изменения сигнальных функций. Логично полагать, что искажения информации будут тем меньше, чем выше частота дискретизации F . С другой стороны также очевидно, что чем больше значение F , тем большим количеством цифровых данных будут отображаться сигналы, и тем большее время будет затрачиваться на их обработку. В оптимальном варианте значение частоты дискретизации сигнала F должно быть необходимым и достаточным для обработки информационного сигнала с заданной точностью, т.е. обеспечивающим допустимую погрешность восстановления аналоговой формы сигнала (среднеквадратическую в целом по интервалу сигнала, либо по максимальным отклонениям от истинной формы в характерных информационных точках сигналов).

5.2. РАВНОМЕРНАЯ ДИСКРЕТИЗАЦИЯ [16,21].

Спектр дискретного сигнала. Допустим, что для обработки задается произвольный аналоговый сигнал $s(t)$, имеющий фурье-образ $S(f)$. Равномерная дискретизация непрерывного сигнала $s(t)$ с частотой F (шаг $\Delta t = 1/F$) с математических позиций означает умножение функции $s(t)$ на гребневую функцию $\text{Ш}_{\Delta t}(t) = \sum_k \delta(t-k\Delta t)$:

$$s_{\Delta t}(t) = s(t) \cdot \text{Ш}_{\Delta t}(t) = s(t) \sum_{k=-\infty}^{\infty} \delta(t-k\Delta t) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \delta(t-k\Delta t). \quad (5.2.1)$$

С учетом известного преобразования Фурье гребневой функции

$$\text{Ш}_{\Delta t}(t) \Leftrightarrow (1/T) \sum_{n=-\infty}^{\infty} \delta(f-nF) = F \cdot \text{Ш}_F(f), \quad (5.2.2)$$

фурье-образ дискретной функции $s_{\Delta t}(t)$:

$$S_F(f) = S(f) * F \cdot \text{Ш}_F(f). \quad (5.2.3)$$

Отсюда, для спектра дискретного сигнала имеем:

$$S_F(f) = F \cdot S(f) * \sum_{n=-\infty}^{\infty} \delta(f-nF) = F \sum_{n=-\infty}^{\infty} S(f-nF). \quad (5.2.4)$$

Из выражения следует, что спектр дискретного сигнала представляет собой *непрерывную периодическую функцию* с периодом F , совпадающую с функцией $F \cdot S(f)$ непрерывного сигнала $s(t)$ в пределах центрального периода от $-f_N$ до f_N , где $f_N = 1/2\Delta t = F/2$. Частоту f_N (или для круговой частоты $\omega_N = \pi/\Delta t$) называют частотой Найквиста. Центральный период функции $S_F(f)$ называют главным частотным диапазоном.

Как правило, шаг дискретизации сигнала (шаг числовых массивов) условно принимают равным $\Delta t = 1$, при этом главный частотный диапазон занимает интервал $-0.5 \leq f \leq 0.5$, или, в шкале угловых частот, соответственно $-\pi \leq \omega \leq \pi$. Примеры равномерной дискретизации аналоговых сигналов $s_1(t) = \exp(-a|t|)$ и $s_2(t) = \exp(-bt^2)$ (дискретные отсчеты нанесены кружками) и спектры этих дискретных сигналов приведены на рис. 5.2.1 и 5.2.2.

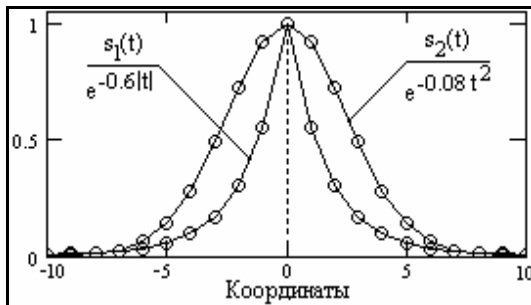


Рис. 5.2.1. Дискретные сигналы.

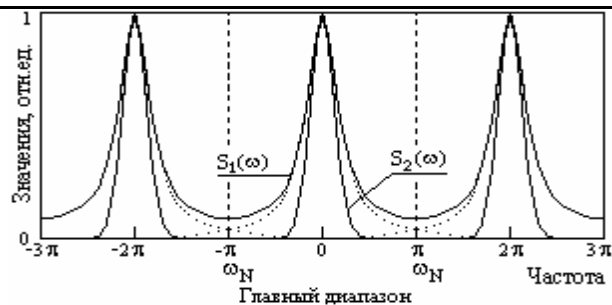


Рис. 5.2.2. Спектры дискретных сигналов.

Для того чтобы периодическое повторение спектра, вызванное дискретизацией аналогового сигнала, не изменяло спектр в главном частотном диапазоне (по отношению к спектру исходного аналогового сигнала), необходимо и достаточно, чтобы максимальные частотные составляющие f_{\max} в спектре аналогового сигнала не превышали частоты Найквиста ($f_{\max} \leq f_N = F/2$). Это означает, что частота дискретизации сигнала должна быть минимум в два раза выше максимальной частотной составляющей в спектре сигнала:

$$F = 1/\Delta t \geq 2f_{\max}, \quad (5.2.5)$$

что обеспечивает выход спектра на нулевые значения на концах главного диапазона, как это имеет место для спектра $S_2(\omega)$ на рис. 5.2.2.

Другими словами, на одном периоде колебаний с частотой f_{\max} должно быть минимум две точки отсчета. Это и понятно – по одной точке отсчета на периоде гармонического сигнала определение частоты и фазы данной гармоники невозможно. Если условие (5.2.5) нарушается, искажения частотного спектра исходного аналогового сигнала неизбежны. На рис. 5.2.2 наглядно видно, что частота дискретизации для сигнала $s_1(t)$ данному условию не удовлетворяет, спектры периодов перекрылись, и результирующий спектр дискретных отсчетов

сигнала $s_1(t)$ отличается от фактического спектра сигнала $s_2(t)$ (фактический спектр и его периодические повторения в области перекрытия спектра главного частотного диапазона со спектрами боковых диапазонов показаны пунктиром).

Характер возникающих искажений во временной области при нарушении условия (5.2.5) можно наглядно видеть на рис. 5.2.3. На рисунке показаны три возможных варианта соотношения частот гармонических сигналов с постоянной частотой их дискретизации.

1. График А – частота гармонического сигнала меньше частоты Найквиста. Дискретным отсчетам может соответствовать только исходная гармоника, амплитуда, частота и фаза которой могут быть однозначно определены по любым трем последовательным точкам (три уравнения, три неизвестных).

2. График В – частота гармонического сигнала равна частоте Найквиста. Это означает периодическое повторение каждой пары последовательных отсчетов, а, следовательно, для решения имеется только два уравнения с тремя неизвестными с возможностью определения только частоты, и то при условии, что начальная фаза сигнала не совпадает с начальной фазой частоты дискретизации (в этом случае все отсчеты нулевые). Амплитуда и фаза сигнала определяются однозначно только при условии совпадения отсчетов с экстремумами гармоники.

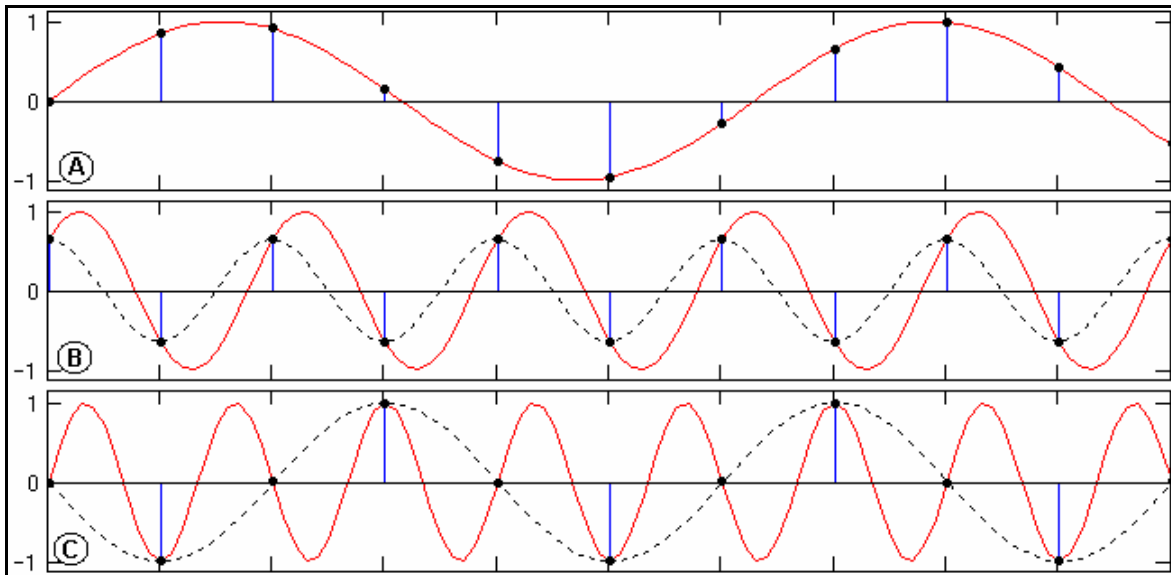


Рис. 5.2.3. Дискретизация гармоник с разной частотой.

3. График С – частота гармонического сигнала больше частоты Найквиста. Решение трех уравнений по трем последовательным точкам позволяет определить амплитуду гармоники, но дает искаженные значения частоты и фазы колебания (показано пунктиром). Это так называемый эффект *появления ложных (кажущихся) частот (aliasing)*. Частоты гармонических колебаний выше частоты Найквиста как бы зеркально "отражаются" в главный частотный диапазон от его границ (на частоте Найквиста), что можно видеть на рис. 5.2.2 для действительного спектра сигнала $S_1(\omega)$, показанного точками. Этот эффект аналогичен всем известному эффекту обратного вращения колес автомобиля (и любых других быстро вращающихся объектов) на экранах кино и телевизоров, когда скорость их вращения начинает превышать частоту смены кадров.

Интерполяционный ряд Котельникова-Шеннона. Спектр дискретизированного сигнала (5.2.4) представляет собой сумму сдвинутых копий исходного аналогового сигнала с шагом сдвига, равным частоте дискретизации. Очевидно, что если спектры копий не перекрываются, то по центральной копии дискретного спектра можно восстановить исходный аналоговый сигнал с абсолютной точностью. Умножая функцию (5.2.3) на прямоугольную весовую функцию $\Pi_F(f)$, равную 1 в пределах главного частотного диапазона $[-F/2, F/2]$ и нулю за его пределами, получаем непрерывный спектр в бесконечных по частоте границах, равный спектру $F \cdot S(f)$ в пределах главного частотного диапазона:

$$F \cdot S(f) = F \cdot [S(f) * \text{Ш}_F(f)] \cdot \Pi_F(f). \quad (5.2.6)$$

Обратное преобразование Фурье такого спектра должно давать конечный и непрерывный сигнал. Произведем обратное преобразование обеих частей равенства (5.2.6):

$$F \cdot [S(f) * \text{Ш}_F(f)] \Leftrightarrow s_{\Delta t}(t), \quad \Pi_F(f) \Leftrightarrow F \cdot \text{sinc}(\pi Ft).$$

$$F \cdot s(t) = s_{\Delta t}(t) * F \cdot \text{sinc}(\pi Ft).$$

$$s(t) = \text{sinc}(\pi Ft) * \sum_{k=-\infty}^{\infty} s(k\Delta t) \delta(t-k\Delta t),$$

Дискретизированный сигнал $s_{\Delta t}(t) = \sum_k s(k\Delta t) \delta(t-k\Delta t)$ представляет собой сумму последовательных весовых импульсов Кронекера, сдвинутых на интервал Δt , со значениями веса, равными значениям отсчетов функции $s(t)$ в моменты $k\Delta t$. При прохождении такого сигнала через систему с импульсным откликом $h(t) = \text{sinc}(\pi Ft) = \sin(\pi Ft)/\pi Ft$ каждый весовой импульс Кронекера возбudit на выходе соответствующую последовательную серию сдвинутых и масштабированных копий оператора фильтра. Отсюда, с учетом очевидного равенства

$$\delta(t-k\Delta t) * \text{sinc}(\pi Ft) = \text{sinc}[\pi F(t-k\Delta t)],$$

выходной сигнал будет представлять собой сумму сдвинутых весовых импульсных откликов системы, где значение веса определяется отсчетами дискретного сигнала:

$$s(t) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \text{sinc}[\pi F(t-k\Delta t)] = \sum_{k=-\infty}^{\infty} s(k\Delta t) \text{sinc}[\pi(t/\Delta t - k)]. \quad (5.2.7)$$

Эта конечная формула носит название интерполяционного ряда Котельникова-Шеннона. Из нее следует, что если наибольшая частота в спектре произвольной непрерывной функции $s(t)$ не превышает частоты ее дискретизации, то она без потери точности может быть представлена в виде числовой последовательности дискретных значений $s(k\Delta t)$, $k = 0, 1, 2, \dots$, и однозначно восстановлена по этой последовательности. В этом и состоит сущность теоремы отсчетов Котельникова. В зарубежной литературе она называется также теоремой Шеннона или теоремой дискретизации (sampling theorem).

Академик В.А.Котельников, 1908-2005. Крупнейший ученый в области радиотехники, радиофизики и информатики. Окончил Московский энергетический институт в 1931 году. С 1931 г. по 1941 г. преподает в МЭИ и ведет научную работу в ЦНИИ связи. В 1933 г. формулирует знаменитую теорему отсчетов, которая носит его имя. В период Великой Отечественной войны (1941-1945 гг.) работал над созданием специальной аппаратуры связи. С 1948 г. по 1953 г. директор и главный конструктор ОКБ МЭИ. В 1953 году избран академиком АН СССР. С 1954 года - директор Института радиотехники и электроники АН СССР. Занимался теорией помехоустойчивой радиосвязи и радиолокации, радиолокационным исследованием планет. Лауреат Ленинской премии, дважды лауреат Государственной премии СССР. Дважды удостоен звания Героя Социалистического труда, награжден шестью орденами Ленина, орденом "За заслуги перед Отечеством" I степени.

По существу, ряд (5.2.7) представляет собой частный случай разложения сигнала в соответствии с формулой (5.1.2) по системе ортогональных функций интегрального синуса $v(t, k\Delta t) = \text{sinc}(\pi F(t-k\Delta t)) = \text{sinc}(\pi(t/\Delta t - k))$, образующих базис пространства сигналов $s(t)$. Для проверки ортогональности достаточно вычислить скалярное произведение базисных функций:

$$\int_{-\infty}^{\infty} v(t, n\Delta t) v(t, m\Delta t) dt = \begin{cases} \Delta t, & n = m \\ 0, & n \neq m \end{cases}.$$

Разложение (5.2.7) проще и понятнее, чем разложение в ряды Фурье, что можно видеть на рис. 5.2.4. Вес каждой функции отсчетов $\text{sinc}[\pi F(t-k\Delta t)]$ формирует пиковое значение интегрального синуса в каждой текущей точке $t = k\Delta t$, равное значению сигнала $s(k\Delta t)$, при этом во всех остальных точках дискретных отсчетов $\text{sinc}[\pi F(t-(k \pm j)\Delta t)]$, $j = 1, 2, \dots$ значения интегрального синуса равны нулю. Ряд числовых значений интегрального синуса для дискретных значений $t = n\Delta t$ при суммировании по k полностью эквивалентен гребневой функции:

$$\sum_{k=-\infty}^{\infty} \text{sinc}[\pi F(n\Delta t - k\Delta t)] \equiv \text{Ш}_{\Delta t}(t).$$

Однако, в отличие от гребневой функции, в интервале между дискретными отсчетами интегральный синус имеет не нулевые, а определенные осциллирующие значения. Суперпозицией этих значений по текущим значениям t от всех интегральных синусов, осцилляции которых доходят до данного значения t , и образуются значения аналогового сигнала в интервалах между отсчетами.

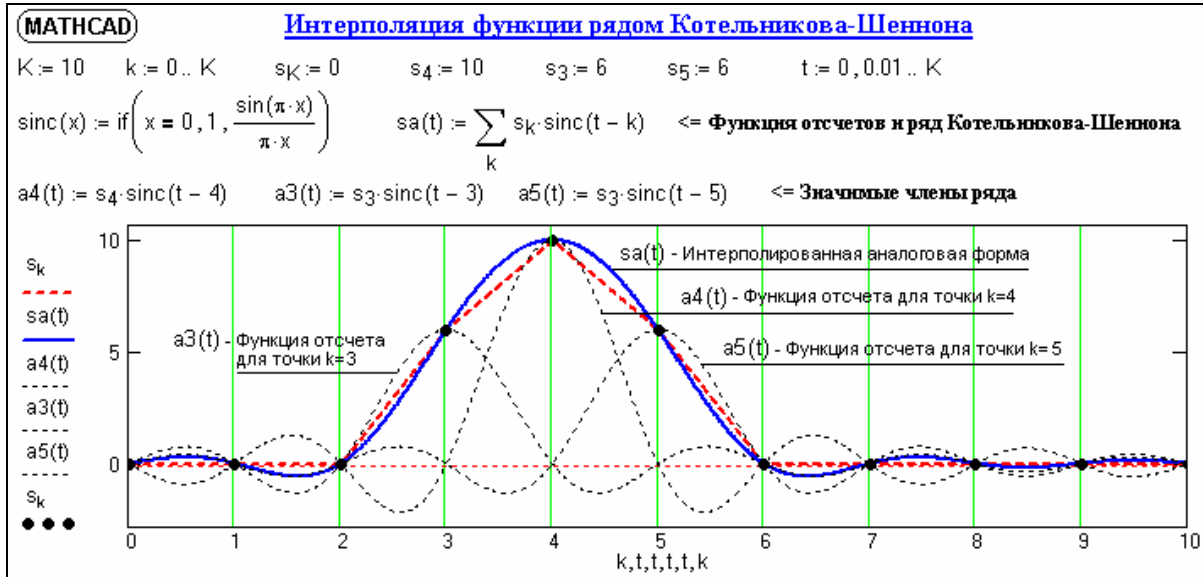


Рис. 5.2.4. Восстановление непрерывного сигнала по дискретным отсчетам.

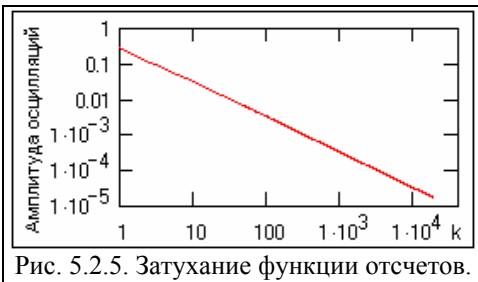


Рис. 5.2.5. Затухание функции отсчетов.

В принципе, функции отсчетов имеют бесконечные осцилляции, и восстанавливают аналоговый сигнал, бесконечный по аргументу. Амплитуда осцилляций функций отсчетов затухает достаточно медленно (см. рис. 5.2.5). Однако на рис. 5.2.4 нетрудно заметить, что, в силу знакопеременности функций отсчетов по интервалам дискретизации, осцилляции восстанавливаемых кривых с финитным спектром затухают достаточно быстро,

и для данных без существенных выбросов и больших перепадов значений определяются, в основном, отсчетами, ближайшими к интерполируемому интервалу. Это позволяет ограничивать интервал суммирования в формуле (2.5.7) определенными окрестностями текущих точек интерполяции.

Курсовая работа 1 – Исследование и разработка основных правил ограничения интервала суммирования при интерполяции данных рядом Котельникова-Шеннона.

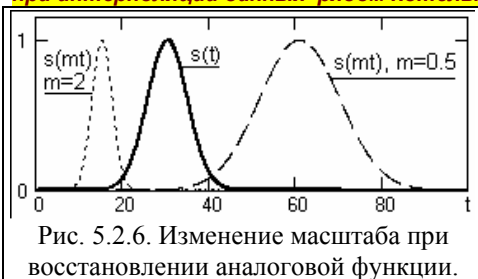


Рис. 5.2.6. Изменение масштаба при восстановлении аналоговой функции.

Ряд (5.2.7) позволяет простым введением масштабного множителя в аргумент интегрального синуса изменять представление сигнала на временной оси, растягивать или сжимать сигнал:

$$s(t) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \text{sinc}[\pi F(m \cdot t - k\Delta t)].$$

По аналогичной формуле может выполняться пересчет дискретных данных на другой интервал дискретизации:

$$s(n \cdot \Delta t_{\text{new}}) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \text{sinc}[\pi F(n \cdot \Delta t_{\text{new}} - k\Delta t)].$$

Примеры восстановления аналоговой форму произвольного финитного сигнала и изменения шага дискретизации данных приведены на рис. 5.2.7.

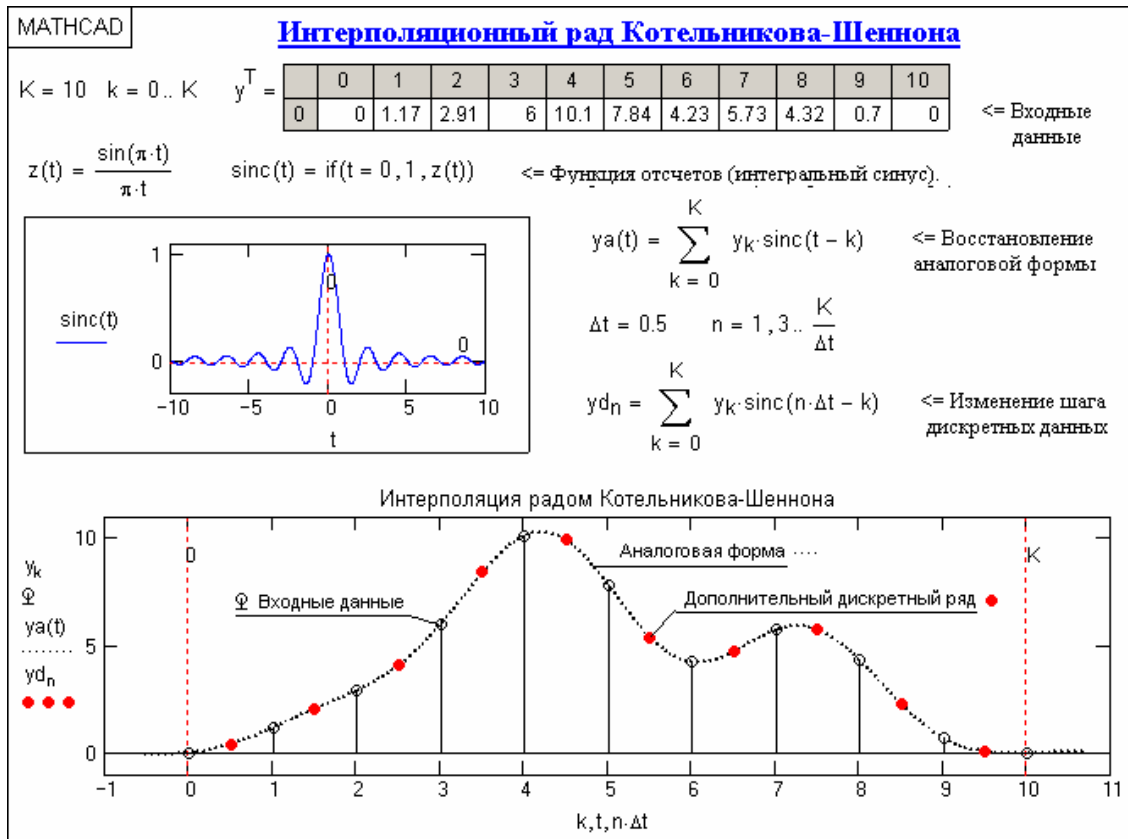


Рис. 5.2.7. Интерполяция по Котельникову-Шеннону.

На рис. 5.2.8 приведено моделирование дискретизации аналогового сигнала, влияние наложение спектров боковых периодов на спектр главного диапазона дискретного сигнала и восстановление из этого спектра аналоговой формы сигнала.

Графики А и Б рисунка – модельный аналоговый сигнал, точки его дискретизации и модуль спектра дискретного сигнала. Вычисление спектра выполнено быстрым преобразованием Фурье (БПФ) и отображает, соответственно, частотный диапазон $0-2f_N$. Дискретизация выполнена корректно, с выполнением условия (5.2.5), о чем можно судить и по спектру дискретного сигнала (график Б, выход на незначимые значения к частоте Найквиста f_N).

Кривая S1 на графике В – спектр модельного дискретного сигнала при нарушении условия (5.2.5). В данном случае это произойдет при увеличении шага дискретизации в 2 раза, что вызовет уменьшение в 2 раза новой частоты Найквиста и перемещение границы главного диапазона на отметку $0.5f_N$ на графике Б, при этом произойдет перекрытие спектров поддиапазонов. На графике приведены кривые S1a и S1b, которые являются отдельными спектрами правой половины главного диапазона без сложения со спектром правого бокового диапазона (интервал $0-2f_N$, где f_N – частота Найквиста новой дискретизации), и левой половины правого бокового диапазона на том же интервале $0-2f_N$ без сложения со спектром главного диапазона. Хорошо видны «хвосты» спектров, выходящие за границы интервала Найквиста от центров диапазонов и заходящие в соседние диапазоны. Сложением этих спектров в интервале $0-2f_N$ нетрудно убедиться, что полученный результат будет полностью соответствовать спектру S1 новой дискретизации исходного сигнала. Обратим внимание, что сложение спектров рядом расположенных диапазонов может вызывать не только увеличение высокочастотных составляющих (как это можно было видеть на рис. 5.2.2 – спектр S1), ни и их взаимную компенсацию, как имеет место для спектра S1 в данном случае (кривая точками на графике В).

Перекрытие спектров диапазонов вызовет искажение аналоговой формы сигнала, восстановленного из его дискретных отсчетов, что можно видеть на графике Г – кривая s2. В данном случае, при частичной взаимной компенсации перекрывающихся частей спектров, наиболее сильное искажение произошло во второй, высокочастотной части сигнала.

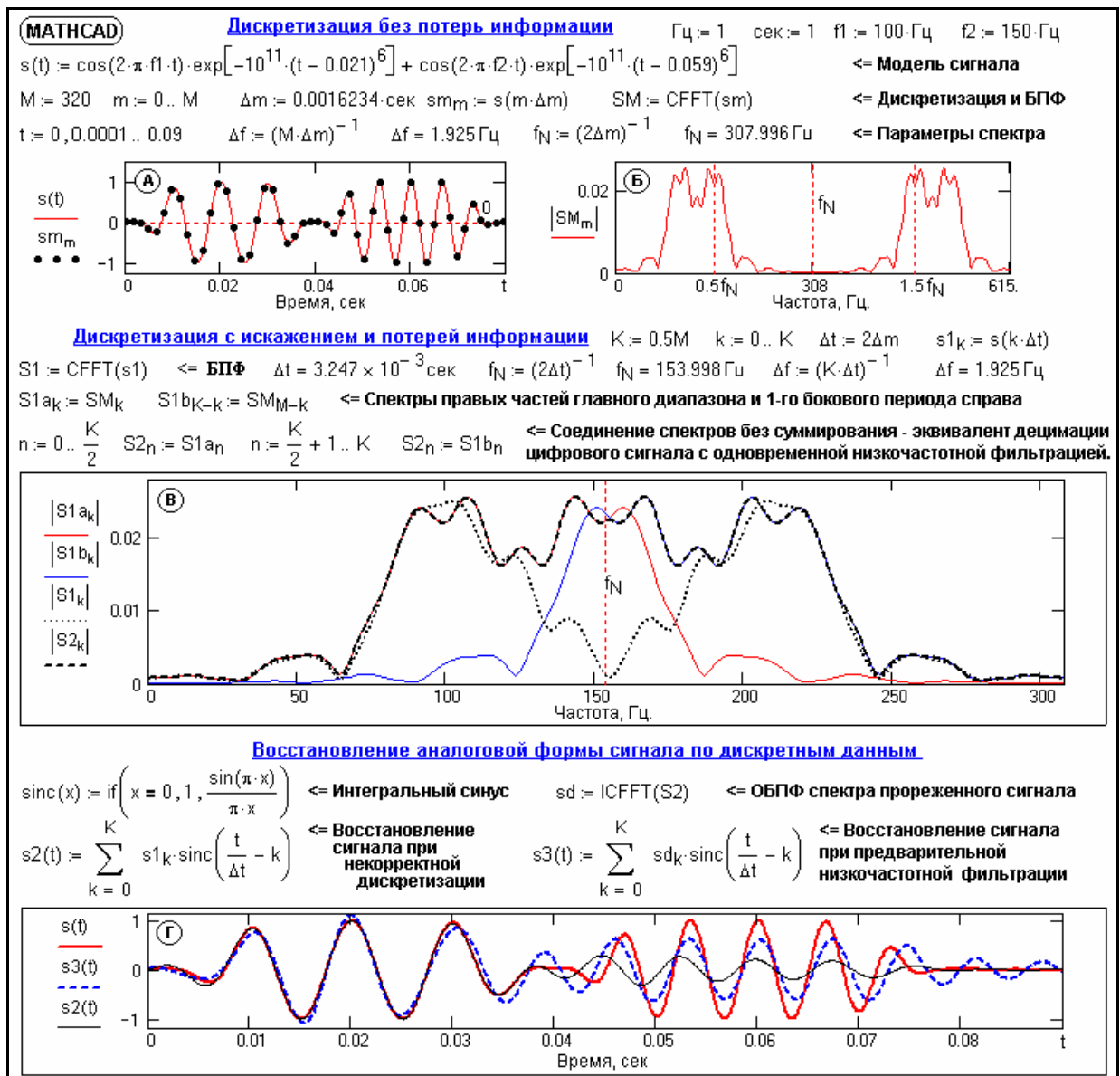


Рис. 5.2.8. Моделирование дискретизации аналогового сигнала.

Дискретизируемые сигналы, как правило, содержат широкополосные шумы, высокочастотные составляющие которых неизбежно перекрываются при периодизации спектра, и увеличивают погрешность восстановления сигналов. Для исключения этого фактора перед проведением дискретизации должно быть обеспечено подавление всех частот выше частоты Найквиста, т.е. выполнена низкочастотная фильтрация сигнала. Если последнее не проведено, то при дискретизации целесообразно в 2-4 раза уменьшить интервал дискретизации относительно оптимального и первой операцией обработки данных выполнить низкочастотную цифровую фильтрацию, после чего можно провести децимацию данных.

Увеличение интервала дискретизации сигналов является довольно распространенной операцией при цифровой обработке данных, и не только при подготовке данных для хранения с целью сокращения их количества. При комплексной обработке данных различной природы интервалы дискретизации этих данных могут оказаться различными, и производится их приведение к одному значению. Аналогичная операция выполняется, как правило, и при создании многослойных информационных пакетов. В таких случаях снижение частоты дискретизации каких-либо данных является вынужденной необходимостью даже с потерей части высокочастотных составляющих информации. Предварительное отфильтровывание отбрасываемых данных перед децимацией (для исключения их попадания в главный частотный диапазон и искажения основной информации) в этом случае является обязательным, особенно при достаточно высокой энергии этих составляющих сигнала. Пример такой децимации

приведен на рис. 5.2.8 на графиках В и Г - спектр $S_2(f)$ децимированных данных и аналоговый сигнал $s_2(t)$, восстановленный по дискретным отсчетам $sd(k\Delta t) \leftrightarrow S_2(f)$. Децимация выполнена непосредственно в частотной области путем смыкания на частотной части $0-f_N$ спектра $SM(f)$ исходного сигнала $sm(m\Delta m)$ с сопряженной частью на интервале $1.5f_N - f_N$, что сокращает новый интервал Найквиста в 2 раза и формирует спектр $S_2(f)$, соответствующий дискретному сигналу с увеличенным в два раза интервалом дискретизации данных с полностью подавленной частью спектральных составляющих от $0.5f_N$ до $1.5f_N$. Такой метод может применяться для децимации (передискретизации) данных с любой кратностью.

Дискретизация с усреднением. Если дискретизация сигнала производится импульсами конечной ширины, то таким импульсам соответствуют средние значения сигнала на интервале длительности импульсов. При длительности импульсов τ имеем:

$$s(k\Delta t) = (1/\tau) \int_{k\Delta t - \tau/2}^{k\Delta t + \tau/2} s(t) dt. \quad (5.2.8)$$

С использованием селектирующей и гребневой функций эта операция отображается следующим образом:

$$s_{\Delta t}(t) = (1/\tau)[s(t) * \Pi_{\tau}(t)]\text{Ш}_{\Delta t}(t). \quad (5.2.9)$$

Соответственно спектр дискретной функции:

$$S_F(f) = [S(f) \cdot \text{sinc}(\pi f \tau)] * F \cdot \text{Ш}_F(f). \quad (5.2.10)$$

Отсюда следует, что при дискретизации с усреднением спектр $S(f)$ заменяется спектром $S(f) \cdot \text{sinc}(\pi f \tau)$, периодическое продолжение которого и образует спектр дискретной функции. При обратном преобразовании Фурье и при использовании интерполяционной формулы Котельникова-Шеннона, вместо исходной функции $s(t)$ получаем функцию $s'(t) = s(t) * \Pi_{\tau}(t)/\tau$, что эквивалентно пропусканию сигнала через фильтр с откликом $h(t) = \Pi_{\tau}(t)/\tau$, т.е. через низкочастотный сглаживающий фильтр "скользящего" среднего с окном τ .

Допустим $\tau = \lambda \Delta t$, $\lambda \leq 1$, $F = 2\alpha f_{\max}$, $\alpha \geq 1$. Для этих условий частотная передаточная функция фильтра записывается в следующем виде: $H(f) = \text{sinc}[(\pi \lambda / 2\alpha)(f/f_{\max})]$. Если потеря составляющих сигнала на всех частотах не должна превышать 3%, необходимо выполнить условие: $\text{sinc}(\pi \lambda / 2\alpha) \geq 0,97$. При $\alpha = 1$ отсюда следует, что значение λ должно быть равно $\lambda \leq 0,27$, т.е. ширина импульса дискретизации может составлять до 27 % интервала дискретизации.

Отметим, что в выражении (5.2.8) значения отсчетов относятся к центру интервалов τ импульсов дискретизации. Если отсчет будет относиться к концу интервалов τ , что имеет место при обработке информации в режиме реального времени, то в выходной функции (5.2.9) появится сдвиг на интервал $\tau/2$, а в ее спектре соответственно сдвиг фаз на $\omega \tau/2$ (в правой части выражения (5.2.10) добавится множитель $\exp(-j\pi f \tau)$).

Дискретизация спектров. Теоремы, доказанные для прямого преобразования Фурье, в такой же мере действительны и для обратного. При дискретизации спектра сигнала с шагом Δf динамическое представление сигнала также становится периодическим с периодом $T = 1/\Delta f$. Для сохранения возможности точного восстановления сигнала в пределах главного периода (без наложения сигналов соседних периодов) частотный шаг дискретизации должен удовлетворять условию:

$$\Delta f \leq 1/T. \quad (5.2.11)$$

Попутно отметим, что для временной формы каузального сигнала главным периодом принимают интервал от 0 до T , хотя при обработке данных на ЭВМ это не имеет значения и главный период может устанавливаться от $-T/2$ до $T/2$.

Информационная тождественность динамической и частотной форм дискретного представления сигнала непосредственно следует из теоремы Котельникова-Шеннона.

Основой любых преобразований при обработке данных обычно является финитный (конечный по длительности) сигнал, зарегистрированный на интервале $0-T$ и состоящий из определенных частотных составляющих от 0 до f_{\max} . Оптимальная дискретизация аналогового сигнала без потери точности его восстановления, как рассмотрено выше, соответствует двум отсчетам на периоде максимальной частотной составляющей:

$$\Delta t = 1/2f_{\max}, \quad N_t = T/\Delta t. \quad (5.2.12)$$

где N_t – общее количество отсчетов на интервале T задания сигнала. Если сигнал зарегистрирован непосредственно в дискретной форме, то он автоматически ограничен по максимальной частоте, т.е. максимальные частоты в таком сигнале равны $f_{\max} \leq 1/2\Delta t$.

При переводе дискретного сигнала в частотную форму спектр сигнала непрерывен и периодичен с периодом $1/\Delta t = 2f_N$. Для оптимальной дискретизации по частоте без потери точности восстановления непрерывного спектра должны выполняться условия:

$$\Delta f = 1/T = 1/(\Delta t N_t), \quad f_N = 1/2\Delta t, \quad (5.2.13)$$

$$N_f = 2f_N/\Delta f = N_t. \quad (5.2.14)$$

Спектр сигнала подвергается каким-либо преобразованиям (обработке), как правило, только в главном частотном диапазоне и тем самым превращается в непериодический сигнал, существующий только в интервале $2f_N$ (от $-f_N$ до f_N). Значения спектра за пределами главного диапазона по умолчанию полагаются равными нулю. При обратном переводе такого сигнала из частотной формы в динамическую сигнал также является непрерывным и периодическим с периодом $1/\Delta f = T$, при этом оптимальная дискретизация по координатам без потери точности восстановления непрерывной формы соответствует условиям:

$$\Delta t = 1/2f_N, \quad T = 1/\Delta f, \quad (5.2.15)$$

$$N_t = T/\Delta t = N_f. \quad (5.2.16)$$

При осуществлении преобразований $s(k\Delta t) \Leftrightarrow S(n\Delta f)$, равно как и $S(n\Delta f) \Leftrightarrow s(k\Delta t)$, условие $N_f = N_t$ является необходимым и достаточным для полного сохранения информации при преобразованиях сигнала из одной формы представления в другую. Условия (5.2.12-5.2.16) задают оптимальность преобразований без потерь информации. Если исходный сигнал дискретизирован оптимально и представлен N отсчетами, то уменьшение количества отсчетов при преобразовании неизбежно приводит к определенным потерям информации.

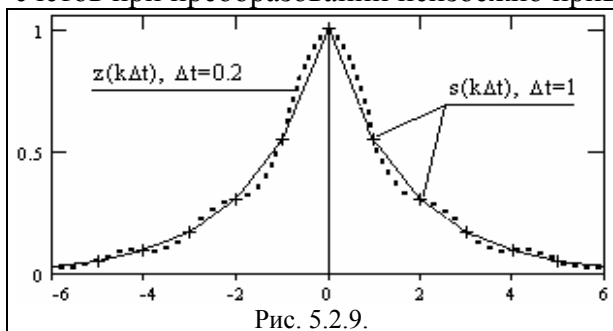


Рис. 5.2.9.

Что касается увеличения числа отсчетов при преобразовании функций (уменьшение интервалов дискретизации), то оно всегда возможно, т.к. выходной сигнал преобразования финитных сигналов является непрерывной функцией и, в принципе, интервал дискретизации может быть установлен бесконечно малым. Однако увеличение числа отсчетов не увеличивает ни количества информации, заключенной в

исходном сигнале, ни точности ее представления. По существу, такая операция полностью эквивалентна интерполяции исходного сигнала рядом Котельникова-Шеннона. Пример такой операции приведен на рис. 5.2.9.

Отсчеты $s(k\Delta t)$ и огибающая их кривая на рисунке 5.2.9 повторяют (в более детальном масштабе) сигнал $s_1(t)$ на рис. 5.2.1, дискретизированный с шагом $\Delta t = 1$. Как уже отмечалось, интервал дискретизации данного сигнала оказался завышенным, и спектр сигнала искажен (рис. 5.2.2). При выполнении операции $s(k\Delta t) \Rightarrow S(n\Delta f)$ количество точек дискретизации спектра $S(n\Delta f)$ было увеличено в 5 раз по отношению к количеству точек сигнала $s(k\Delta t)$, т.е. $N_f = 5N_t$. При обратном преобразовании $S(n\Delta f) \Rightarrow z(k\Delta t)$, были выполнены условия (5.2.15-5.2.16), при этом шаг дискретизации сигнала при его восстановлении оказался также в 5 раз меньше исходного ($\Delta t = 0.2$). Результат можно видеть на рис. 5.2.9 (кривая $z(k\Delta t)$). Абсолютно такой же результат дает и интерполяция сигнала $s(k\Delta t)$ рядом Котельникова-Шеннона с переводом на шаг $\Delta t = 0.2$. Искажение аналогового сигнала закладывается при его дискретизации, если шаг дискретизации не удовлетворяет условию (5.2.5), и при любых дальнейших преобразованиях уже не может быть исправлено, т.к. информация о первоначальной форме аналогового сигнала при некорректной дискретизации утрачивается безвозвратно.

Дискретизация усеченных сигналов. При выполнении условия (5.2.5) для сигналов с

ограниченным спектром аналоговая форма сигнала может быть восстановлена по дискретным отсчетам, если сигнал на интервале T его задания является финитным или, по крайней мере, настолько быстро затухающим, что отсчеты сигнала за пределами интервала T практически равны нулю. Задача дискретизации усложняется для медленно затухающих сигналов, сигналов бесконечной длительности и сигналов со спектром, неограниченным по частоте. Последнее имеет место, если в сигнале присутствуют разрывы и резкие скачки.

В общем случае, длительность сигнала и ширина его спектра не могут быть одновременно ограничены конечными интервалами. Если длительность сигнала ограничена и сигнал урезан в области не нулевых значений, то спектр сигнала неограничен и наоборот. Однако обработка реальных сигналов возможна только с их ограничением, как по координатам, так и по ширине спектра. При этом в качестве оценки корректности ограничения сигналов используется энергетический критерий, согласно которому длительность сигнала T и практическую ширину спектра Ω устанавливают такими, чтобы в них была сосредоточена подавляющая часть энергии сигнала. Это достигается при выполнении условий:

$$\int_0^T |s(t)|^2 dt = k \int_{-\infty}^{\infty} |s(t)|^2 dt, \quad (5.2.17)$$

$$\int_0^{\Omega} |S(\omega)|^2 d\omega = k \int_0^{\infty} |S(\omega)|^2 d\omega, \quad (5.2.17')$$

где k - коэффициент представительности (качества) задания сигнала, значение которого, в зависимости от целевых задач обработки сигналов, может устанавливаться от 0,9 до 0,99.

Допустим, что произвольный сигнал $s(t)$ рассматривается в пределах конечного интервала $[0, T]$ и принимается равным нулю за его пределами. Такой сигнал может быть получен умножением сигнала $s(t)$ на прямоугольную весовую функцию $\Pi_T(t)$:

$$s_T(t) = s(t) \Pi_T(t).$$

Для спектра $S_T(f)$ функции $s_T(t)$ соответственно имеем:

$$S_T(f) = S(f) * T \cdot \text{sinc}(\pi f T). \quad (5.2.18)$$

Спектр $S_T(f)$ неограничен, поскольку неограничен носитель функции $\text{sinc}(\pi f T)$. Отсюда следует, что частота дискретизации функции $s_T(t)$ в принципе должна быть бесконечно большой, т.е. корректная дискретизация невозможна. На практике полагают, что спектр $S_T(f)$ также определен в конечной области $[-\Omega, \Omega]$:

$$S'_T(f) = S_T(f) \cdot \Pi_{2\Omega}(f),$$

при этом вне этой области, по оценке Шеннона, для спектра $S_T(f)$ справедлива формула:

$$|S_T(f)| \approx 1/\Omega T, \quad f \notin (-\Omega, \Omega). \quad (5.2.19)$$

Но усеченная часть спектра определяет разность значений между исходной функцией $s_T(t)$ и функцией $s'_T(t)$, восстановленной по усеченному спектру $S'_T(f)$, т.к. отсеченных гармоник спектра будет недоставать для полного восстановления функции $s_T(t)$:

$$\varepsilon_T(t) = s_T(t) - s'_T(t).$$

Соответственно, оценка дисперсии погрешности аппроксимации определяется выражением:

$$\sigma^2 = \int_{-\infty}^{\infty} \varepsilon_T^2(t) dt \approx 1/\Omega T, \quad \sigma \approx 1/\sqrt{\Omega T}. \quad (5.2.20)$$



Рис. 5.2.10. Вид функции погрешности аппроксимации

Эти выражения определяют порядок среднеквадратической погрешности аппроксимации, которая является интегральной по интервалу T , а не локальной разностью значений $s_T(t) - s'_T(t)$. Типичный вид погрешности аппроксимации усеченных сигналов приведен на рис. 5.2.10. В точках дискретизации погрешность равна нулю, максимальна на центрах интервалов дискретизации и нарастает при приближении к границам интервала T .

Физические данные обычно регистрируются по определенным интервалам T и, как правило, не выходят на нулевые значения на границах интервалов. В этом случае ограничение ширины спектра можно проводить по (5.2.20) с учетом допустимой среднеквадратиче-

ской погрешности аппроксимации данных. Частота Ω при усечении спектра может рассматриваться в качестве частоты Найквиста для сигнала $s_T(t)$ при его дискретизации, что определяет частоту дискретизации не менее $F = 2\Omega$ и количество точек дискретизации не менее $N = TF = 2\Omega T$.

В силу тождественности свойств прямого и обратного преобразования Фурье аналогичная методика может применяться и для оценки условий дискретизации спектров.

Таким образом, дискретизация усеченных сигналов возможна, однако при обработке усеченных сигналов необходимо проявлять осторожность и контролировать как значение среднеквадратической ошибки искажений, так и характер возникающих искажений сигнала и его спектра. Так, например, при усечении функции автокорреляции в спектре мощности сигнала могут появиться отрицательные значения, т.к. функция отсчетов $\text{sinc}(\pi f T)$ в (5.2.18) является знакопеременной. Другой пример - проектирование частотных полосовых фильтров. При задании передаточной функции фильтра $H(f)$ в частотной области в виде П-образной функции $H(f) = \Pi_r(f)$ обратное преобразование Фурье дает импульсный отклик фильтра $h(t) \Leftrightarrow H(f)$ бесконечно большой длины. Усечение отклика $h_T(t) = h(t)\Pi_T(t)$ вызывает изменение передаточной функции фильтра (явление Гиббса): $H_T(f) = \Pi_r(f) * \Pi_T(f) \Rightarrow \Pi_r(f) \cdot T \cdot \text{sinc}(\pi f T)$, при этом по краям скачков П-функции появляются затухающие флюктуации с амплитудой первого выброса до 9% от значений коэффициента передачи фильтра в полосе пропускания.

Так как частотный характер искажений, возникающих при усечении сигнала, определяется весовой функцией $\Pi_T(t) \Leftrightarrow T \cdot \text{sinc}(\pi f T)$, то допустимый уровень и форму искажения сигнала можно устанавливать не только подбором интервала T , но и применением других весовых функций. Так, для исключения появления отрицательных значений в спектрах мощности усечение функций автокорреляции целесообразно выполнять весовыми функциями, которые не имеют отрицательных значений в своих спектрах. Одной из таких функций является, например, треугольная весовая функция (окно Бартлетта).

Соотношение спектров одиночного и периодического сигналов. Спектр $S_T(f) = S(k\Delta f)$ периодического сигнала $s_T(t)$ с периодом T дискретен ($\Delta f = 1/T$). Спектр $S(f)$ одиночного сигнала $s(t)$, заданного на интервале T , непрерывен и представляет собой спектральную плотность сигнала при $T \Rightarrow \infty$. Но периодический сигнал можно представить и в виде свертки одного периода с гребневой функцией Дирака:

$$s_T(t) = s(t) * \text{Ш}_T(t).$$

При переходе в частотную область получаем:

$$S_T(f) = (1/T) \cdot S(f) \cdot \text{Ш}_{1/T}(f) = S(k\Delta f),$$

$$S_T(f) = (1/T) \sum_{k=-\infty}^{\infty} S(f) \delta(f - k/T). \quad (5.2.21)$$

Отсюда следует, что спектр периодического сигнала представляет собой дискретизированный спектр одиночного сигнала, нормированный на длительность периода.

С другой стороны, одиночный сигнал $s(t)$ может быть получен из периодического сигнала $s_T(t)$ умножением на селектирующий прямоугольный импульс $\Pi_T(t)$:

$$s(t) = s_T(t) \cdot \Pi_T(t).$$

Спектр одиночного сигнала:

$$S(f) = T \cdot S_T(f) * \Pi_T(f) = T \sum_{k=-\infty}^{\infty} S(k\Delta f) \cdot \text{sinc}[\pi T(f - k/T)], \quad (5.2.22)$$

т.е. непрерывный спектр одиночного сигнала однозначно устанавливается по спектру периодического сигнала (интерполяция рядом Котельникова-Шеннона в частотной области).

5.3. ДИСКРЕТИЗАЦИЯ ПО КРИТЕРИЮ НАИБОЛЬШЕГО ОТКЛОНЕНИЯ [10].

Задача абсолютно точного восстановления сигнала на практике обычно не ставится, в отличие от задачи минимального физического объема информации, при котором сохраняется возможность ее восстановления в непрерывной форме с определенным допустимым значе-

нием погрешности. Такая задача актуальна всегда, и особенно при дистанционных методах регистрации и обработки информации, передаче сигналов по каналам связи и при подготовке информации к длительному хранению. Одним из методов решения этой задачи является дискретизация сигналов по критерию наибольшего отклонения.

В процессе дискретизации по критерию наибольшего отклонения задается допустимое значение погрешности восстановления сигнала σ . При восстановлении сигнала непрерывная функция $s(t)$ аппроксимируется, как правило, степенными полиномами n -го порядка. Погрешность восстановления функции $s(t)$ полиномом $s_a(t)$ определяется остаточным членом $L(t)$:

$$L(t) = s(t) - s_a(t) = \sigma(t).$$

Шаг дискретизации выбирается из условия обеспечения $L(t) < \sigma$ по всему интервалу определения функции $s(t)$. Как правило, динамика функции $s(t)$ может существенно изменяться в различные моменты времени по интервалу регистрации, при этом шаг дискретизации также может изменяться, при условии не превышения заданной погрешности на каждом шаге. При установленном значении σ уменьшение числа отсчетов обеспечивается повышением степени аппроксимирующего многочлена. На практике обычно ограничиваются ступенчатой, линейной и параболической аппроксимацией полиномами соответственно нулевой, первой и второй степеней.

В качестве интерполирующих многочленов используют многочлены Лагранжа. Для многочленов Лагранжа нулевой степени значение $s_a(t)$ в момент времени t на интервале $t_i < t < t_{i+1}$ между двумя последовательными отсчетами функции принимается равным отсчету $s(t_{i+1})$. Если восстановление сигнала $s(t)$ проводить по двум отсчетам: $s_a(t_i) = [s(t_{i+1}) - s(t_i)]/2$, то при том же шаге дискретизации погрешность восстановления сигнала уменьшается вдвое. Но при использовании двух последовательных отсчетов лучше использовать многочлены Лагранжа первой степени, т.е. соединение двух последовательных отсчетов прямой линией, что дает еще большее уменьшение погрешности восстановления аналоговой формы сигнала.

В качестве экстраполирующих многочленов используется многочлены Тейлора. Для многочлена Тейлора нулевой степени условия восстановления сигнала практически не отличаются от многочлена Лагранжа, за исключением направления (от текущего зарегистрированного отсчета и вперед по t). Для многочленов Тейлора более высоких степеней при восстановлении сигнала помимо отсчета $s(t_i)$ используются также соответствующие значения производных в точке отсчета. Восстановление сигнала многочленами Тейлора происходит без задержки во времени. Однако при использовании многочленов выше нулевой степени для точного восстановления сигнала по сравнению с интерполяционными методами требуется в два раза более высокая частота дискретизации.

5.4. АДАПТИВНАЯ ДИСКРЕТИЗАЦИЯ [10].

Частота равномерной дискретизации информации рассчитывается по предельным значениям частотных характеристик сигналов. Адаптивная дискретизация ориентирована на динамические характеристики сигнала, что позволяет обеспечивать его восстановление при минимальном числе выборок. В основе принципов адаптивной дискретизации лежит слежение за текущей погрешностью восстановления сигнала. Наиболее широкое применение получили алгоритмы дискретизации с адаптацией по длине интервала аппроксимации. Сущность дискретизации заключается в последовательном наращивании интервала аппроксимации с непрерывным сравнением сигнала $s(t)$ с воспроизводящей функцией $s_a(t)$. При достижении заданного значения σ наращивание интервала прекращается, и производится отсчет значения $s(t_i)$, т.е. дискретизация является неравномерной. Для воспроизведения сигналов нерегулярной дискретизации обычно используются степенные алгебраические полиномы нулевой и первой степени в интерполяционном или в экстраполяционном вариантах.

Наиболее простой является техника адаптивной дискретизации с использованием многочлена нулевой степени. На момент t_i начала каждого интервала аппроксимирующий

полином $s_a(t)$ принимается равным $s(t_i)$, вычисляется текущая разность $L(t) = s(t) - s_a(t)$ и производится сравнение ее значения с заданным значением σ . При фиксировании равенства $L(t) = \sigma$ производится очередной отсчет и начинается следующий интервал.

При использовании аппроксимирующего многочлена первой степени вычисляется значение $s_a(t) = s(t_i) + s'(t_i)(t - t_i)$, где $s'(t)$ - производная сигнала. Момент очередного отсчета определяется выполнением равенства $s(t) - s(t_i) - s'(t_i)(t - t_i) = \sigma$. Следует иметь в виду, что данный алгоритм неэффективен при наличии высокочастотных помех, к которым весьма чувствительна операция дифференцирования.

Курсовая работа 2 – Исследовать и разработать программу оценки спектра дискретного сигнала при неравномерном шаге дискретизации.

Самыми простыми способами восстановления сигналов при адаптивной дискретизации являются линейная и квадратичная интерполяции, которые выполняются по уравнениям:

$$f(x)_{\text{лин}} = a_0 + a_1x. \quad f(x)_{\text{кв}} = a_0 + a_1x + a_2x^2.$$

Эти уравнения являются частным случаем полиномиальной интерполяции с помощью аппроксимирующего полинома:

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n = \sum_{i=0}^n a_i x^i. \quad (5.4.1)$$

Для выполнения полиномиальной интерполяции достаточно по выражению (5.4.1) составить систему линейных уравнений для n последовательных отсчетов и определить n значений коэффициентов a_i . При *глобальной* интерполяции, по всем N точкам задания функции, степень полинома равна $N-1$. Глобальная интерполяция обычно выполняется для достаточно коротких (не более 8-10 отсчетов) массивов данных. Пример выполнения глобальной интерполяции приведен на рис. 5.4.1.

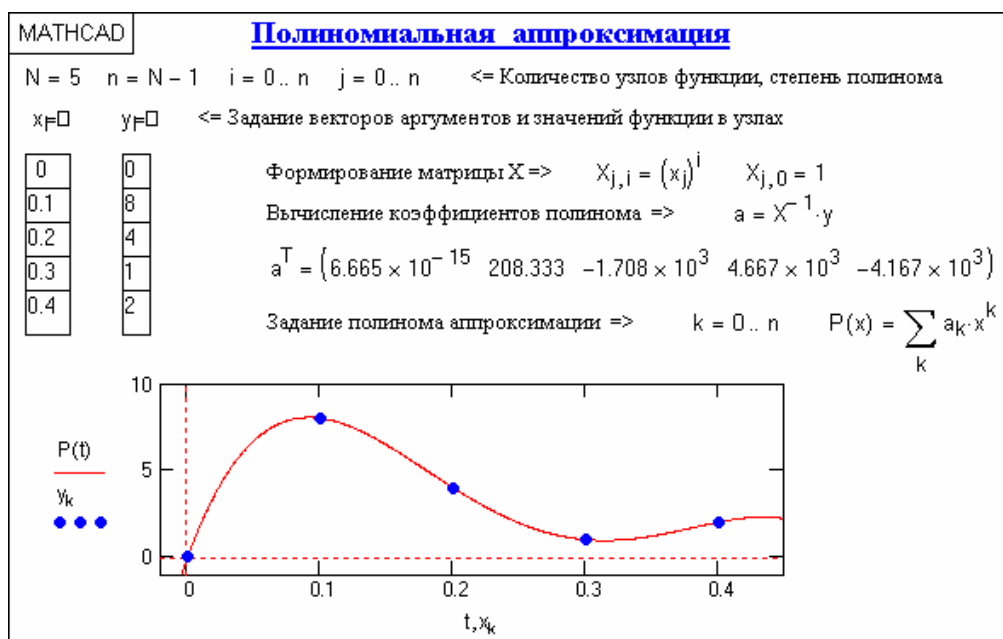


Рис. 5.4.1. Интерполяция данных.

Большие массивы данных интерполируются последовательными *локальными* частями или в скользящем по массиву данных окне интерполяции, как правило, с нечетным значением N и вычислением требуемых значений сигнала в определенном интервале центральной части окна.

Курсовая работа 3 – Исследовать и обосновать оптимальный метод полиномиальной интерполяции произвольных данных с равномерным шагом дискретизации. Разработать программу интерполяции.

Курсовая работа 4 – Исследовать и обосновать оптимальный метод полиномиальной интерполяции произвольных данных с неравномерным шагом дискретизации. Разработать программу интерполяции.

Для практического использования более удобны формулы аппроксимации, не требующие предварительного определения коэффициентов аппроксимирующих полиномов. К числу таких формул относится интерполяционный многочлен по Лагранжу /30/. При аппрок-

симации функции $y(x)$ многочленом n -ой степени $f(x)$:

$$f(x) = \frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_0-x_n)} y_0 + \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)} y_1 + \dots$$

$$\dots + \frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\dots(x_n-x_{n-1})} y_n. \quad (5.4.2)$$

Пример интерполяции по Лагранжу приведен на рис. 5.4.2.

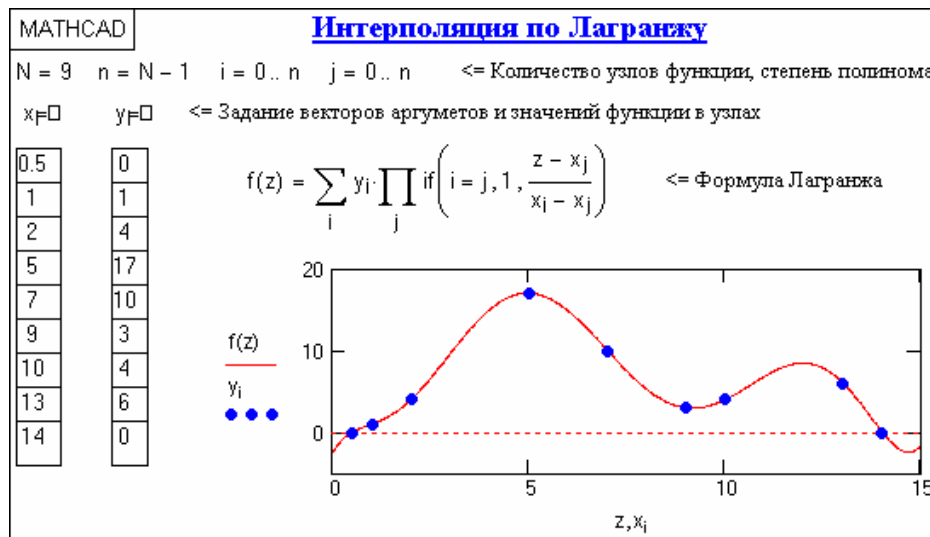


Рис. 1.2.2. Интерполяция по Лагранжу.

Курсовая работа 5 – Исследовать и обосновать оптимальный метод интерполяции по Лагранжу произвольных данных с неравномерным шагом дискретизации. Разработать программу интерполяции.

5.5. КВАНТОВАНИЕ СИГНАЛОВ [5,21].

Дискретизация аналоговых сигналов с преобразованием в цифровую форму связана с квантованием сигналов. Сущность квантования состоит в замене несчетного множества возможных значений функции, в общем случае случайных, конечным множеством цифровых отсчетов, и выполняется округлением мгновенных значений входной функции $s(t_i)$ в моменты времени t_i до ближайших значений $s_i(t_i) = n_i \Delta\sigma$, где $\Delta\sigma$ - шаг квантования шкалы цифровых отсчетов. Квантование с постоянным шагом $\Delta\sigma$ называется *равномерным*. Математически операция квантования может быть выражена формулой:

$$s_i(t_i) = \left[\frac{s(t_i)}{\Delta\sigma} + \frac{1}{2} \right] \cdot \Delta\sigma,$$

где скобки означают целую часть значения в скобках.

При квантовании сигналов в большом динамическом диапазоне значений шаг квантования может быть и неравномерным, например, логарифмическим, т.е. пропорциональным логарифму значений входного сигнала. Установленный диапазон шкалы квантования от s_{\min} до s_{\max} и шаг квантования $\Delta\sigma$ определяют число делений шкалы $N_\sigma = (s_{\max} - s_{\min}) / \Delta\sigma$ и соответственно цифровую разрядность квантования. В результате дискретизации и квантования непрерывная функция $s(t)$ заменяется числовой последовательностью $\{s(k\Delta t)\}$. Погрешность округления $\varepsilon_i = s(k\Delta t) - s_i(k\Delta t)$ заключена в пределах $-\Delta\sigma/2 < \varepsilon < \Delta\sigma/2$ и называется шумом квантования. Требуемая точность квантования оценивается по влиянию возникающего шума квантования на последующую обработку сигналов.

При достаточно малом шаге квантования любое значение в его пределах можно считать равновероятным, при этом значения ε распределены по равномерному закону:

$$p(\varepsilon) = 1/\Delta\sigma, \quad -\Delta\sigma/2 \leq \varepsilon \leq \Delta\sigma/2.$$

Соответственно, дисперсия и среднее квадратическое значение шума квантования:

$$\varepsilon^2 = \Delta\sigma^2/12, \quad \bar{\varepsilon} \approx 0.3 \Delta\sigma. \quad (5.5.1)$$

При задании уровня шума квантования с использованием выражения (5.5.1) нетрудно определить допустимое значение шага квантования.

Входной сигнал содержит, как правило, аддитивную смесь собственно сигнала $s(t)$ и помехи $q(t)$ с дисперсией соответственно σ_q^2 . Если помехи не коррелированы с сигналом, то после квантования суммарная дисперсия шумов:

$$\sigma^2 = \sigma_q^2 + \varepsilon^2.$$

На практике шаг квантования выбирают обычно таким, чтобы не происходило заметного изменения отношения сигнал/шум, т.е. $\varepsilon^2 \ll \sigma_q^2$.

5.6. ДЕЦИМАЦИЯ И ИНТЕРПОЛЯЦИЯ ДАННЫХ [4,5,17].

Децимацией (прореживанием, сокращением) цифровых данных принято называть уплотнение данных с удалением избыточной информации. Последнее имеет место, если шаг дискретизации данных был установлен излишне подробным и $f_N = 1/2\Delta t \gg f_{\max}$ сигнала. Информация высокочастотной части сигнала может быть ненужной, если основная энергия полезной части сигнала заключена в низкочастотной области. Децимация может потребоваться и в том случае, если массивы данных представлены с разным шагом дискретизации.

Децимации должна предшествовать низкочастотная фильтрация данных. Это связано с тем, что в процессе децимации шаг дискретизации Δt заменяется на новый шаг $\Delta t' = p\Delta t$, где $p > 1$, с соответствующим сжатием главного частотного диапазона, при этом появляется опасность отражения отбрасываемых частотных составляющих и высокочастотных шумов в главный диапазон (как и при неправильном выборе шага дискретизации). Точка отсечки низкочастотного фильтра устанавливается по новой частоте Найквиста: $f_N' = 1/(2p\Delta t)$.

Значение коэффициента p при децимации может быть произвольным, но, как правило, используются целочисленные значения, и децимация выливается в простое прореживание данных. При нецелочисленном значении p децимация может проводиться с использованием интерполяционного ряда Котельникова-Шеннона (равно как и любого другого интерполяционного многочлена) или преобразования Фурье. Последнее выполняется путем перевода сигнала в частотную форму и возвращением в координатную форму с новым шагом $\Delta t' = p\Delta t$, при этом низкочастотная фильтрация может производиться непосредственно в частотном диапазоне. Возможно также и прямое усечение главного частотного диапазона с N точек до $N' = N/p$ с возвратом из нового частотного диапазона в координатную форму с количеством точек N' , но при этом следует учитывать последствия усечения спектральной функции (умножения на прямоугольное селектирующее окно) на форму восстанавливаемого по ней сигнала (свертка исходного сигнала с фурье-образом прямоугольного селектирующего окна).

Интерполяция данных отличается от децимации только значением коэффициента $p < 1$, с соответствующим увеличением частоты Найквиста, и не требует низкочастотной фильтрации.

Для децимации и интерполяции данных разработаны также специальные высокоскоростные методы и алгоритмы (цифровые фильтры) - экспандеры и компрессоры [4,5/].

ЛИТЕРАТУРА

4. Гольденберг Л.М. и др. Цифровая обработка сигналов: Справочник. - М.: Радио и связь, 1985.
5. Гольденберг Л.М. и др. Цифровая обработка сигналов: Учебное пособие для вузов.- М.: Радио и связь, 1990.- 256 с.
10. Дмитриев В.И. Прикладная теория информации: Учебник для вузов. - М.: Высшая школа, 1989.
16. Макс Ж. Методы и техника обработки сигналов при физических измерениях. - М.: Мир, 1983.
17. Никитин А.А. Теоретические основы обработки геофизической информации: Учебник для вузов.- М.: Недра, 1986.- 342 с.
21. Рапопорт М.Б. Вычислительная техника в полевой геофизике: Учебник для вузов.- М.: Недра, 1993.- 350 с.
30. Корн Г., Корн Е. Справочник по математике для научных работников и инженеров. - М.: Наука, 1984.

Главный сайт автора ~ Лекции по сигналам ~ Практикум

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 6: ДИСКРЕТНЫЕ ПРЕОБРАЗОВАНИЯ СИГНАЛОВ

Нет ничего, сколь бы великим и изумительным оно не показалось с первого взгляда, на что мало-помалу не начинаешь смотреть с меньшим изумлением.

Лукреций. О природе вещей.

Римский поэт и философ, I в.д.н.э.

Преобразование должно быть осознанным и целенаправленным действием. Был баран - стал шашлык. Что может быть прекраснее такого преобразования!

Виль Ибрагимов. Зри в корень.

Ташкентский геофизик Уральской школы, XX в.

Содержание: Введение. 6.1. Преобразование Фурье. Дискретное преобразование Фурье. Быстрое преобразование Фурье. 6.2. Преобразование Лапласа. 6.3. Z - преобразование сигналов. Определение преобразования. Примеры z-преобразования. Связь с преобразованиями Фурье и Лапласа. Свойства z-преобразования. Отображение z-преобразования. Аналитическая форма z-образов. Обратное z-преобразование. 6.4. Дискретная свертка (конволюция). Уравнение дискретной свертки. Техника свертки. Литература.

ВВЕДЕНИЕ

Цифровая обработка сигналов оперирует с дискретными преобразованиями сигналов и обрабатывающих данные сигналы систем. Математика дискретных преобразований зародилась в недрах аналоговой математики еще в 18 веке в рамках теории рядов и их применения для интерполяции и аппроксимации функций, однако ускоренное развитие она получила в 20 веке после появления первых вычислительных машин. В принципе, в своих основных положениях математический аппарат дискретных преобразований подобен преобразованиям аналоговых сигналов и систем. Однако дискретность данных требует учета этого фактора, и его игнорирование может приводить к существенным ошибкам. Кроме того, ряд методов дискретной математики не имеет аналогов в аналитической математике.

6.1. ПРЕОБРАЗОВАНИЕ ФУРЬЕ [5,17,21].

Дискретное преобразование Фурье может быть получено непосредственно из интегрального преобразования дискретизаций аргументов ($t_k = k\Delta t$, $f_n = n\Delta f$):

$$S(f) = \int_{-\infty}^{\infty} s(t) \exp(-j2\pi ft) dt, \quad S(f_n) = \Delta t \sum_{k=-\infty}^{\infty} s(t_k) \exp(-j2\pi f_n k\Delta t), \quad (6.1.1)$$

$$s(t) = \int_{-\infty}^{\infty} S(f) \exp(j2\pi ft) df, \quad s(t_k) = \Delta f \sum_{n=-\infty}^{\infty} S(f_n) \exp(j2\pi n\Delta f t_k). \quad (6.1.2)$$

Напомним также, что дискретизация функции по времени приводит к периодизации ее спектра, а дискретизация спектра по частоте - к периодизации функции. Для дискретных преобразований $s(k\Delta t) \Leftrightarrow S(n\Delta f)$, и функция, и ее спектр дискретны и периодичны, а числовые массивы их представления соответствуют заданию на главных периодах $T = N\Delta t$ (от 0 до T или от $-T/2$ до $T/2$), и $2f_N = N\Delta f$ (от $-f_N$ до f_N), где N – количество отсчетов, при этом:

$$\Delta f = 1/T = 1/(N\Delta t), \quad \Delta t = 1/2f_N = 1/(N\Delta f), \quad \Delta t\Delta f = 1/N, \quad N = 2Tf_N. \quad (6.1.3)$$

Соотношения (6.1.3) являются условиями информационной равноценности динамической и частотной форм представления сигналов. Другими словами: число отсчетов функции и ее спектра должны быть одинаковыми. Но каждый отсчет комплексного спектра представляется двумя вещественными числами и, соответственно, число отсчетов комплексного спектра в 2 раза больше отсчетов функции? Это так. Однако представление спектра в комплексной форме - не более чем удобное математическое представление спектральной функции, реальные отсчеты которой образуются сложением двух сопряженных комплексных отсчетов, а полная информация о спектре функции в комплексной форме заключена только в одной его половине - отсчетах действительной и мнимой части комплексных чисел в частотном интервале от 0 до f_N , т.к. информация второй половины диапазона от 0 до $-f_N$ является сопряженной с первой половиной и никакой дополнительной информации не несет.

При дискретном представлении сигналов аргумент t_k обычно проставляется номерами

отсчетов k (по умолчанию $\Delta t = 1$, $k = 0, 1, \dots, N-1$), а преобразования Фурье выполняются по аргументу n (номер шага по частоте) на главных периодах. При значениях N , кратных 2:

$$S(f_n) \equiv S_n = \sum_{k=0}^{N-1} s_k \exp(-j2\pi kn/N), \quad n = -N/2, \dots, 0, \dots, N/2. \quad (6.1.4)$$

$$s(t_k) \equiv s_k = (1/N) \sum_{n=-N/2}^{N/2-1} S_n \exp(j2\pi kn/N), \quad k = 0, 1, \dots, N-1. \quad (6.1.5)$$

Главный период спектра в (6.1.4) для циклических частот от -0.5 до 0.5 , для угловых частот от $-\pi$ до π . При нечетном значении N границы главного периода по частоте (значения $\pm f_N$) находятся на половину шага по частоте за отсчетами $\pm(N/2)$ и, соответственно, верхний предел суммирования в (6.1.5) устанавливается равным $N/2$.

В вычислительных операциях на ЭВМ для исключения отрицательных частотных аргументов (отрицательных значений номеров n) и использования идентичных алгоритмов прямого и обратного преобразования Фурье главный период спектра обычно принимается в интервале от 0 до $2f_N$ ($0 \leq n \leq N$), а суммирование в (6.1.5) производится соответственно от 0 до $N-1$. При этом следует учитывать, что комплексно сопряженным отсчетам S_n^* интервала $(-N, 0)$ двустороннего спектра в интервале $0-2f_N$ соответствуют отсчеты S_{N+1-n} (т.е. сопряженными отсчетами в интервале $0-2f_N$ являются отсчеты S_n и S_{N+1-n}).

Пример: На интервале $T = [0, 99]$, $N=100$, задан дискретный сигнал $s(k) = \sum_{i=3}^8 \delta(k-i)$ - прямоугольный импульс с единичными значениями на точках k от 3 до 8. Форма сигнала и модуль его спектра в главном частотном диапазоне, вычисленного по формуле $S(n) = \sum_{k=0}^{99} s(k) \cdot \exp(-j2\pi kn/100)$ с нумерацией по n от -50 до $+50$ с шагом по частоте, соответственно, $\Delta\omega = 2\pi/100$, приведены на рис. 6.1.1.

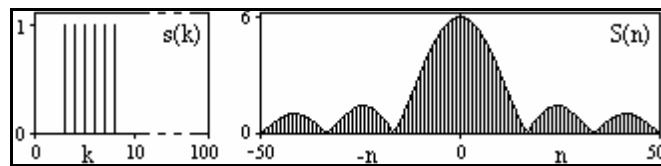


Рис. 6.1.1. Дискретный сигнал и модуль его спектра.

На рис. 6.1.2 приведена огибающая значений другой формы представления главного диапазона спектра. Независимо от формы представления спектр периодичен, в чем нетрудно убедиться, если вычислить значения спектра для большего интервала аргумента n с сохранением того же шага по частоте, как это показано на рис. 6.1.3 для огибающей значений спектра.

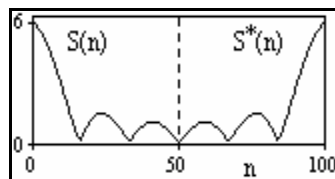


Рис. 6.1.2. Модуль спектра.

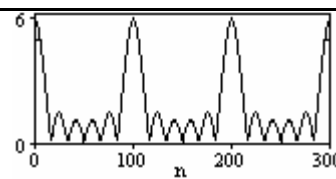


Рис. 6.1.3. Модуль спектра.

На рис. 6.1.4. показано обратное преобразование Фурье для дискретного спектра, выполненное по формуле $s'(k) = (1/100) \sum_{n=0}^{99} S(n) \cdot \exp(j2\pi kn/100)$, которое показывает периодизацию исходной функции $s(k)$, но главный период $k = \{0, 99\}$ этой функции полностью совпадает с исходным сигналом $s(k)$.

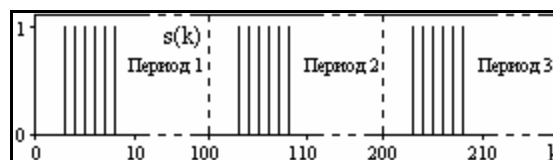


Рис. 6.1.4. Обратное преобразование Фурье.

Преобразования (6.1.4-6.1.5) называют дискретными преобразованиями Фурье (ДПФ). Для ДПФ, в принципе, справедливы все свойства интегральных преобразований Фурье, однако при этом следует учитывать периодичность дискретных функций и спектров. Произведению спектров двух дискретных функций (при выполнении каких-либо операций при обработке сигналов в частотном представлении, как, например, фильтрации сигналов непосредственно в частотной форме) будет соответствовать свертка периодизированных функций во временном представлении (и наоборот). Такая свертка называется циклической (см. раздел 6.4) и ее результаты на концевых участках информационных интервалов могут существенно отличаться от свертки финитных дискретных функций (линейной свертки).

Из выражений ДПФ можно видеть, что для вычисления каждой гармоники нужно N операций комплексного умножения и сложения и соответственно N^2 операций на полное выполнение ДПФ. При больших объемах массивов данных это может приводить к существенным временным затратам. Ускорение вычислений достигается при использовании быстрого преобразования Фурье.

Быстрое преобразование Фурье (БПФ). В основе БПФ лежит прореживание по частоте и пирамидальный алгоритм, которым исключаются повторные вычисления периодически повторяющихся членов ряда Фурье.

Допустим, что массив чисел s_k содержит $N = 2^r$ отсчетов (r - целое). Разделим исходный массив на два первых промежуточных массива с четными и нечетными отсчетами:

$$s_k' = s_{2k}, \quad s_k'' = s_{2k+1}, \quad 0 \leq k \leq N/2-1.$$

Выполним ДПФ каждого массива с учетом того, что шаг функций равен 2 (при $\Delta t=1$), а период промежуточных спектров будет соответственно равен $N/2$:

$$s_k' \Rightarrow S_n', \quad s_k'' \Rightarrow S_n'', \quad 0 \leq n \leq N/2-1.$$

Для получения одной половины искомого спектра S_n сложим полученные спектры с учетом теоремы запаздывания, т.к. отсчеты функции s_k'' сдвинуты относительно s_k' на один шаг дискретизации:

$$S_n = S_n' + S_n'' \cdot \exp(-j2\pi n/N). \quad (6.1.6)$$

Вторая половина спектра, комплексно сопряженная с первой, с учетом периода повторения $N/2$ промежуточных спектров определяется выражением:

$$S_{n+N/2} = S_n' + S_n'' \cdot \exp(-j2\pi(n+N/2)/N) = S_n' - S_n'' \cdot \exp(-j2\pi n/N). \quad (6.1.7)$$

Нетрудно видеть, что для вычисления полного спектра в данном случае потребуется $N^2/4$ операций для вычисления промежуточных спектров плюс еще N операций комплексного умножения и сложения, что создает ощутимый эффект по сравнению с ДПФ.

Но деление массивов на две части может быть применено и к первым промежуточным массивам, и ко вторым, и т.д. до тех пор, пока в массивах не останется по одному отсчету, фурье - преобразование которых равно самому отсчету. Тем самым, алгоритм преобразования превращается в пирамидальный алгоритм перестановок со сложением/вычитанием и с единичным умножением на значение $\exp(-j2\pi n/N)$ соответствующего уровня пирамиды. Первый алгоритм БПФ на данном принципе (из множества модификаций, существующих в настоящее время) был разработан Кули-Тьюки в 1965 г. и позволил повысить скорость вычислений в N/r раз по сравнению с ДПФ. Чем больше N , тем больше эффект БПФ. Так, при $N = 1024$ имеем $r = 10$ и соответственно $N/r \approx 100$. Что касается условия по количеству точек $N = 2^r$, то оно рассматривается в варианте $N_k \leq 2^r$, где r - минимальное целое. Массивы с $N_k < 2^r$ дополняются до 2^r нулями, что не изменяет форму спектра. Изменяется только шаг $\Delta\omega$ по представлению спектра ($\Delta\omega = 2\pi/2^r < 2\pi/N$), который несколько избыточен по адекватному представлению сигнала в частотной области. В настоящее время существуют и алгоритмы БПФ с другими основаниями и их комбинациями, при которых не требуется дополнения сигналов нулями до 2^r .

Заметим, что в соответствии с (6.1.7) отсчеты, сопряженные с правой половиной главного частотного диапазона $(0, \pi)$, относятся не к диапазону $(-\pi, 0)$, а к диапазону $(\pi, 2\pi)$, что, учитывая периодичность спектра дискретных данных, значения не имеет. Т.е. выходной час-

тотный диапазон БПФ равен $(0, 2\pi)$. Общее количество отсчетов комплексного спектра в этом условно главном диапазоне равно количеству точек исходного сигнала (с учетом нулевых точек при дополнении сигнала до $N=2^l$). Алгоритм быстрого обратного преобразования (ОБПФ) тождественен алгоритму прямого БПФ.

Алгоритмы прямого и обратного БПФ широко используются в современном программном обеспечении для анализа и обработки цифровых данных. Пример выполнения БПФ приведен на рис. 6.1.5.

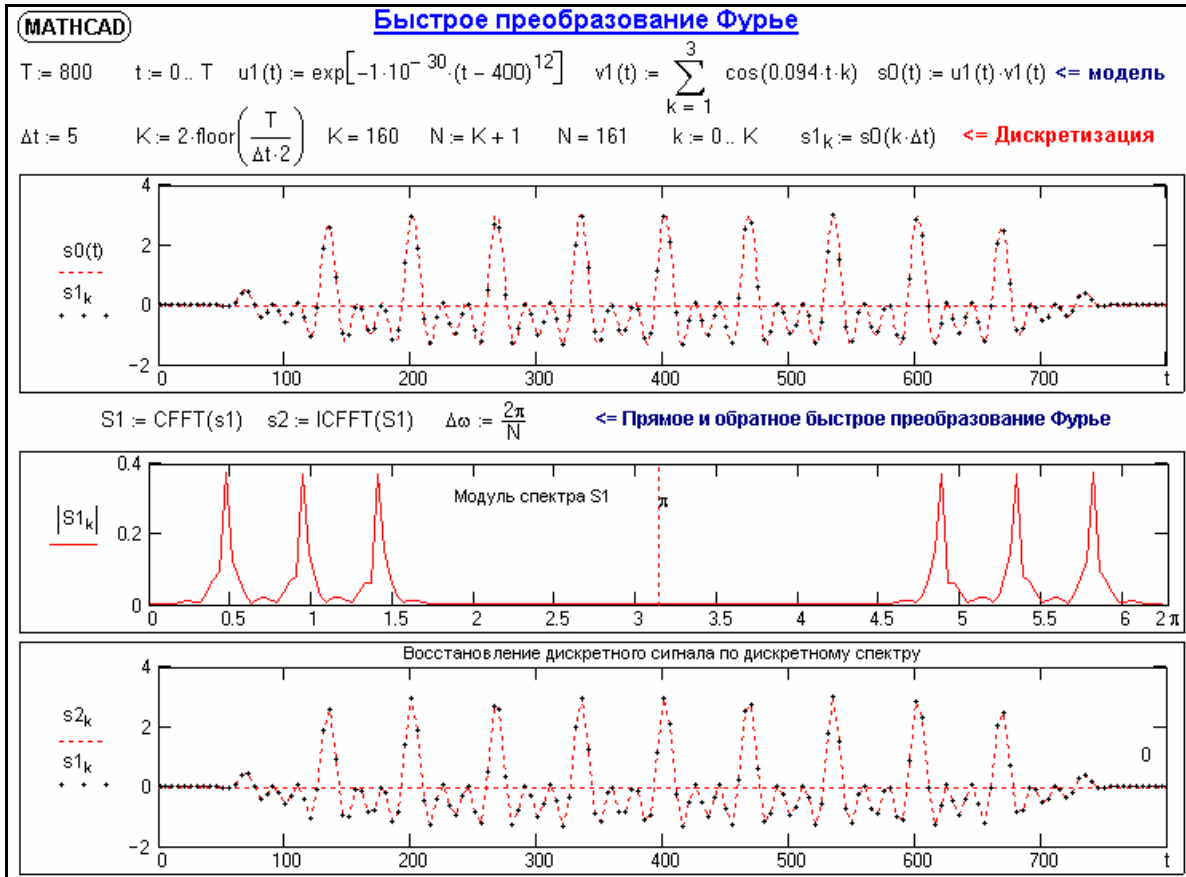


Рис. 6.1.5. Пример БПФ.

6.2. ПРЕОБРАЗОВАНИЕ ЛАПЛАСА.

Дискретное преобразование Лапласа (ДПЛ), как и ДПФ, может быть получено из интегрального преобразования дискретизаций аргументов ($t_k = k\Delta t$, $\omega_n = n\Delta\omega$):

$$Y(p) = \int_0^{\infty} y(t) \exp(-pt) dt, \quad Y(p_n) = \Delta t \sum_{k=0}^{\infty} y(t_k) \exp(-p_n t_k), \quad (6.2.1)$$

где $p = \sigma + j\omega$ - комплексная частота, $\sigma \geq 0$.

$$y(t) = (1/2\pi j) \int_{\sigma - j\omega}^{\sigma + j\omega} Y(p) \exp(pt) dp. \quad y(t_k) = \Delta t \sum_{n=-\infty}^{\infty} Y(p_n) \exp(p_n t_k). \quad (6.2.2)$$

Функцию $Y(p)$ называют изображением Лапласа функции $y(t)$ - оригинала изображения. Нетрудно заметить, что при $\sigma = 0$ преобразование Лапласа превращается в одностороннее преобразование Фурье, а для каузальных сигналов - в полную аналогию ПФ. Наиболее существенной особенностью преобразования Лапласа является возможность его применения для спектрального анализа функций, не имеющих фурье-образов из-за расходимости интегралов Фурье. Для понимания последнего запишем интеграл Лапласа в развернутой форме:

$$Y(p) = \int_0^{\infty} y(t) \exp(-\sigma t - j\omega t) dt = \int_0^{\infty} y(t) \exp(-\sigma t) \exp(-j\omega t) dt = \int_0^{\infty} y'(t) \exp(-j\omega t) dt.$$

Правый интеграл для каузальных сигналов представляет собой преобразование Фурье,

при этом сам сигнал $y'(t)$ за счет экспоненциального множителя $\exp(-\sigma t)$ соответствующим выбором значения $\sigma > 0$ превращается в затухающий, конечный по энергии, что и требуется для существования его фурье-образа. Все свойства и теоремы преобразований Фурье имеют соответствующие аналоги и для преобразований Лапласа.

Пример сопоставления преобразований Фурье и Лапласа приведен на рис. 6.2.1.

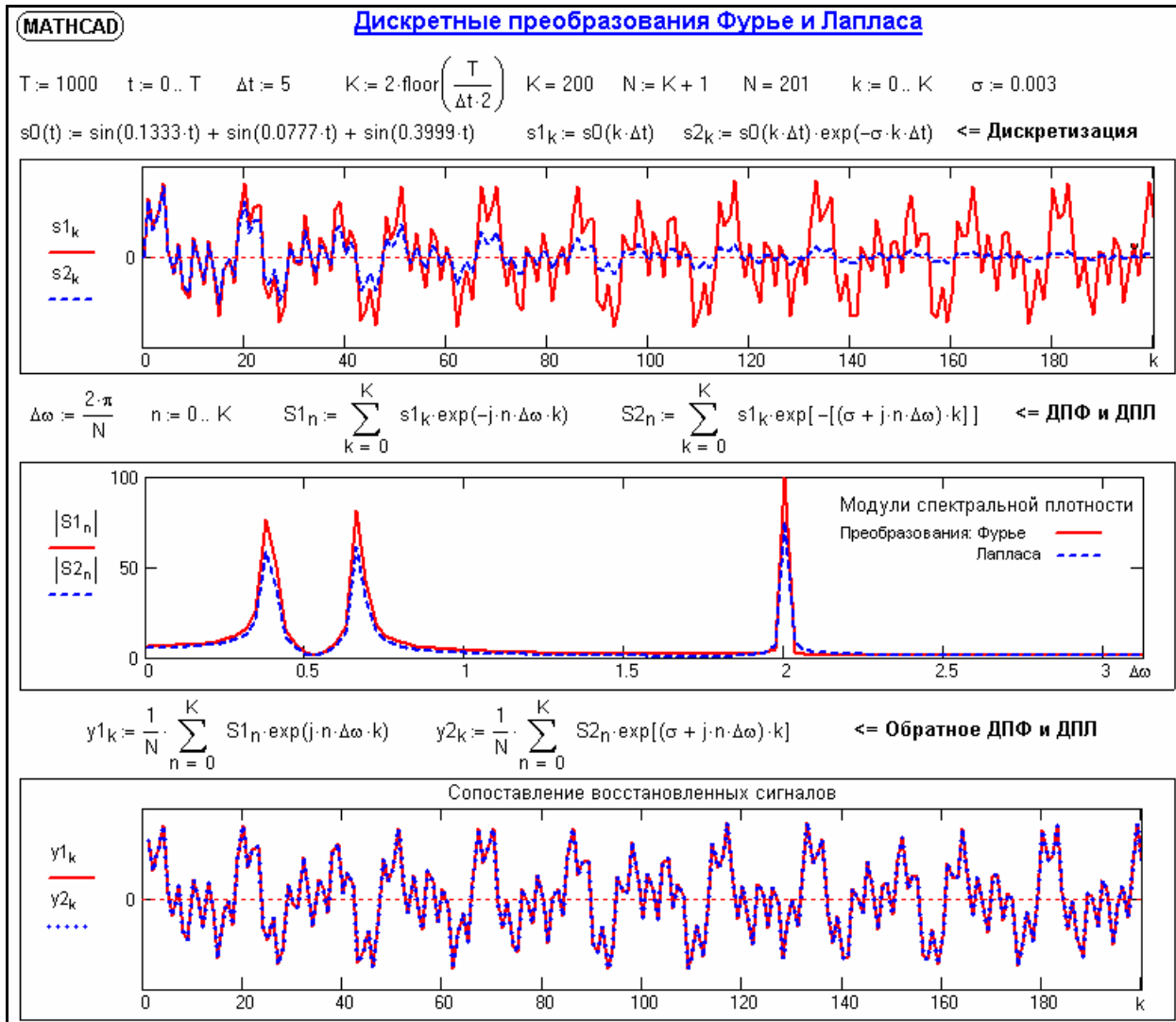


Рис. 6.2.1. Сопоставление преобразований Фурье и Лапласа.

6.3. Z - ПРЕОБРАЗОВАНИЕ СИГНАЛОВ [2,13,21].

Определение преобразования. Распространенным способом анализа дискретных цифровых последовательностей является z-преобразование (z-transform). Оно играет для дискретных сигналов и систем такую же роль, как для аналоговых – преобразование Лапласа.

Произвольной непрерывной функции $s(t)$, равномерно дискретизированной и отображенной отсчетами $s_k = s(k\Delta t)$, равно как и непосредственно дискретной функции, можно поставить в соответствие степенной полином по z , последовательными коэффициентами которого являются значения s_k :

$$s_k = s(k\Delta t) \Leftrightarrow \text{TZ}[s(k\Delta t)] = \sum_{k=-\infty}^{\infty} s_k z^k = S(z). \quad (6.3.1)$$

где $z = \sigma + j\nu = r \cdot \exp(-j\varphi)$ - произвольная комплексная переменная. Полином $S(z)$ называют z-образом или z-изображением функции $s(k\Delta t)$. Преобразование имеет смысл для области тех значений z , в которой ряд $S(z)$ сходится, т.е. сумма ряда представляет собой аналитическую функцию переменной z , не имеющую полюсов и особых точек.

Пример: $s_k = \{1, 2, 0, -1, -2, -1, 0, 0\}$.

$$S(z) = 1z^0 + 2z^1 + 0z^2 - 1z^3 - 2z^4 - 1z^5 + 0z^6 + 0z^7 = 1 + 2z - z^3 - 2z^4 - z^5.$$

Впервые z -преобразование введено в употребление П.Лапласом в 1779 и повторно "открыто" В.Гуревичем в 1947 году с изменением символики на z^{-1} . В настоящее время в технической литературе имеют место оба вида символики. На практическое использование преобразования это не влияет, т.к. смена знака только зеркально изменяет нумерацию членов полинома (относительно z^0), числовое пространство которых в общем случае от $-\infty$ до $+\infty$. В дальнейшем в качестве основной будем использовать символику положительных степеней z , давая пояснения по особенностям отрицательной символики, если таковая имеется.

По заданному или полученному в результате анализа какой-либо системы z -полиному однозначно восстанавливается соответствующая этому полиному функция путем идентификации коэффициентов степеней при z^k с k -отсчетами функции.

$$\text{Пример: } S(z) = 1 + 3z^2 + 8z^3 - 4z^6 - 2z^7 = 1z^0 + 0z^1 + 3z^2 + 8z^3 + 0z^4 + 0z^5 - 0z^6 - 2z^7.$$

$$s_k = \{1, 0, 3, 8, 0, 0, -4, -2\}.$$

Смысл величины z в z -полиноме заключается в том, что она является оператором единичной задержки по координатам функции. Умножение z -образа сигнала $s(k)$ на величину z^n означает задержку сигнала на n интервалов: $z^n S(z) \Leftrightarrow s(k-n)$.

Z -образы с положительными степенями z соответствуют каузальным (физически реализуемым) процессам и системам, которые работают в реальном масштабе времени с текущими и "прошлыми" значениями сигналов. При обработке информации на ЭВМ каузальность сигналов не относится к числу ограничений и возможно использование отрицательных степеней z , соответствующих отсчетам сигналов "вперед". Последнее применяется, например, при синтезе симметричных операторов фильтров, что позволяет производить обработку информации без внесения в сигнал фазовых искажений. При использовании символики z^{-1} "прошлым" значениям соответствуют значения с отрицательными степенями z , "будущим" – с положительными.

Основное достоинство z -преобразований заключается в простоте математических операций со степенными полиномами, что имеет немаловажное значение при расчетах цифровых фильтров и спектральном анализе.

Примеры z -преобразования часто встречающихся на практике дискретных сигналов.

Импульсы Кронекера. В общем случае, в произвольной точке числовой оси:

$$\delta(k-n) = 1 \text{ при } k=n, \quad \delta(k-n) = 0 \text{ при } k \neq n.$$

$$X_\delta(z) = \sum_{k=-\infty}^{\infty} \delta(k-n) z^k = z^n.$$

Для импульса Кронекера в нулевой точке соответственно $X_\delta(z) = z^0 = 1$.

Функция Хевисайда (единичный скачок).

$$x(k) = 0 \text{ при } k < 0, \quad x(k) = 1 \text{ при } k \geq 0.$$

$$X(z) = \sum_{k=0}^{\infty} z^k = z^k.$$

Ряд сходится при $|z| < 1$, при этом его сумма равна:

$$X(z) = 1/(1-z), \quad |z| < 1.$$

При использовании символики z^{-1} :

$$X(z) = 1/(1-z^{-1}) = z/(z-1), \quad |z| > 1.$$

Экспоненциальная функция:

$$x(k) = 0 \text{ при } k < 0, \quad x(k) = a^k \text{ при } k \geq 0.$$

$$X(z) = \sum_{k=-\infty}^{\infty} x(k) z^k = \sum_{k=0}^{\infty} a^k z^k = \sum_{k=0}^{\infty} (az)^k.$$

Как и в предыдущем случае, ряд сходится при $|az| < 1$, т.е. при $|z| < |a|$, при этом:

$$X(z) = 1/(1-az), \quad |z| > |a|.$$

Связь с преобразованиями Фурье и Лапласа. Запишем дискретный сигнал s_k в виде суммы весовых импульсов Кронекера:

$$s_k = s(k\Delta t) = \sum_{n=-\infty}^{\infty} s(n\Delta t) \delta(k\Delta t - n\Delta t).$$

Определим спектр сигнала по теореме запаздывания:

$$S(\omega) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \exp(-j\omega k\Delta t).$$

Выполним замену переменных, $z = \exp(-j\omega\Delta t)$, и получим:

$$S(\omega) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \cdot z^k = S(z).$$

Отсюда следует, что дискретное преобразование Фурье является частным случаем z -преобразования при $z = \exp(-j\omega\Delta t)$. Аналогичной подстановкой $z = \exp(-p)$ может осуществляться переход к дискретному преобразованию Лапласа. В общем виде:

$$S(\omega) = S(z), \quad z = \exp(-j\omega\Delta t); \quad S(p) = S(z), \quad z = \exp(-p\Delta t). \quad (6.3.2)$$

Обратное преобразование:

$$S(z) = S(\omega), \quad \omega = \ln z / j\Delta t; \quad S(z) = S(p), \quad p = \ln z / \Delta t. \quad (6.3.3)$$

При отрицательной символике z связь между представлениями осуществляется соответственно подстановками $z^{-1} = \exp(j\omega\Delta t)$ и $z^{-1} = \exp(p)$.

Свойства z -преобразования. Без углубления в теорию, можно констатировать, что все свойства ДПФ действительны и для z -преобразования. Отметим некоторые из них.

Линейность: Если $S(k) = a \cdot x(k) + b \cdot y(k)$, то $S(z) = aX(z) + bY(z)$. Соответственно, z -преобразование допустимо только для анализа линейных систем и сигналов, удовлетворяющих принципу суперпозиции.

Задержка на n тактов: $y(k) = x(k-n)$.

$$Y(z) = \sum_{k=-\infty}^{\infty} y(k) \cdot z^k = \sum_{k=-\infty}^{\infty} x(k-n) \cdot z^k = z^n \sum_{k=-\infty}^{\infty} x(k-n) \cdot z^{k-n} = z^n \sum_{m=-\infty}^{\infty} x(m) \cdot z^m = z^n X(z).$$

Соответственно, умножение z -образа сигнала на множитель z^n вызывает сдвиг сигнала на n тактов дискретизации.

Для z -преобразования действительны все известные теоремы о спектрах. В частности, свертка двух сигналов отображается в z -области произведением их z -образов, и наоборот:

$$s(k) * h(k) \Leftrightarrow S(z)H(z), \quad s(k) \cdot h(k) \Leftrightarrow S(z) \cdot H(z).$$

При $z = \exp(-j\omega\Delta t)$ z -преобразование представляет собой особую форму представления дискретных сигналов, при которой на полином $S(z)$ можно сослаться как на временную функцию (по значениям коэффициентов $k\Delta t$), так и на функцию частотного спектра сигнала (по значениям аргумента ω).

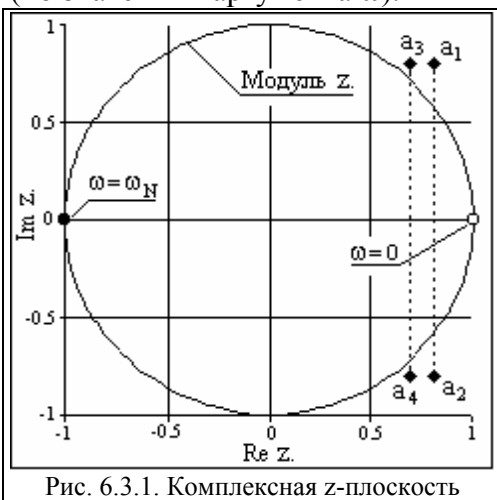


Рис. 6.3.1. Комплексная z -плоскость

Отображение z -преобразования выполняются на комплексной z -плоскости с $\text{Re } z$ и $\text{Im } z$ по осям координат (рис. 6.3.1). Спектральной оси частот ω на z -плоскости соответствует окружность радиуса:

$$|z| = |\exp(-j\omega\Delta t)| = \sqrt{\cos^2(\omega\Delta t) + \sin^2(\omega\Delta t)} = 1.$$

Подстановка значения какой-либо частоты ω в $z = \exp(-j\omega\Delta t)$ отображается точкой на окружности. Частоте $\omega = 0$ соответствует точка $\text{Re } z = 1$ и $\text{Im } z = 0$ на правой стороне оси абсцисс. При повышении частоты точка смещается по окружности против часовой стрелки, и занимает крайнее левое положение на частоте Найквиста $\omega_N = \pi/\Delta t$ ($\text{Re } z = -1$, $\text{Im } z = 0$). Отрицательные частоты спектра отображаются аналогично по ча-

совой стрелке на нижней полуокружности. Точки $\pm \omega_N$ совпадают, а при дальнейшем повышении или понижении частоты значения начинают повторяться в полном соответствии с периодичностью спектра дискретной функции. Проход по полной окружности соответствует одному периоду спектра, а любая гармоника спектра сигнала задается на плоскости двумя точками, симметричными относительно оси абсцисс.

Z-преобразование позволяет производить разложение сигналов и функций, например передаточных функций фильтров, на короткие составляющие операции свертки, для чего достаточно приравнять z-полином к нулю, найти его корни a_i , и переписать полином в виде произведения двучленов:

$$S(z) = a_0(z-a_1)(z-a_2)\dots,$$

где a_0 - последний отсчет сигнала (коэффициент при старшей степени z).

Но произведению в z-области соответствует свертка в координатной области, и при обратном преобразовании двучлены $(z-a_i)$ превращаются в двухточечные диполи $\{-a_i, 1\}$, а сигнал длиной N представляется сверткой (N-1) диполей:

$$s_k = a_0 \{-a_1, 1\} * \{-a_2, 1\} * \{-a_3, 1\} * \dots$$

Пример. $s_k = \{1.4464, -2.32, 3.37, -3, 1\}$. $S(z) = z^4 - 3z^3 + 3.37z^2 - 2.32z + 1.4464$. $a_0 = 1$.

Корни полинома S(z): $a_1 = 0.8 + 0.8j$, $a_2 = 0.8 - 0.8j$, $a_3 = 0.7 + 0.8j$, $a_4 = 0.7 - 0.8j$,

$$S(z) = (z - 0.8 - 0.8j)(z - 0.8 + 0.8j)(z - 0.7 - 0.8j)(z - 0.7 + 0.8j).$$

Корни полинома представлены на z-плоскости на рис. 6.3.1. Корни полинома комплексные и четыре двучлена в координатной области также будут комплексными. Но они являются сопряженными, и для получения вещественных функций следует перемножить сопряженные двучлены и получить биквадратные блоки: $S(z) = (z^2 - 1.4z + 1.13)(z^2 - 1.6z + 1.28)$.

При переходе в координатную область: $s_k = \{1.13, -1.4, 1\} * \{1.28, -1.6, 1\}$.

Таким образом, исходный сигнал разложен на свертку двух трехчленных сигналов (функций).

Аналитическая форма z-образов существует для z-преобразований, если возможно свертывание степенного ряда в аналитическое выражение. Выше, в примерах z-преобразования, уже приводилось приведение к аналитической форме z-образов функции Хевисайда и экспоненциальной функции.

Обратное z-преобразование в общем случае производится интегрированием по произвольному замкнутому контуру C, расположенному в области сходимости и окружающему все особые точки (нули и полюсы) z-образа:

$$s_k = (1/2\pi j) \oint_C S(z) z^{k+1} dz$$

Способом, удобным для практического применения, является разложение рациональных S(z) на простые дроби. С учетом линейности преобразования:

$$S(z) = \sum_{n=1}^N a_n / (1 - b_n z) \Leftrightarrow \sum_{n=1}^N a_n (b_n)^k = s_k.$$

Пример. $S(z) = 1/(1-5z+6z^2) = 3/(1-3z) - 3/(1-2z) \Leftrightarrow 3 \cdot 3^k - 3 \cdot 2^k = s(k)$.

При разложении функции S(z) по степеням z обратное z-преобразование не вызывает затруднений.

6.4. ДИСКРЕТНАЯ СВЕРТКА (КОНВОЛЮЦИЯ) [5,17,21].

Уравнение дискретной свертки двух функций (сигналов) может быть получено непосредственно из интегрального уравнения свертки при замене интегрирования суммированием мгновенных значений функций с шагом Δt :

$$y(k\Delta t) = \Delta t \sum_n s(n\Delta t) h(k\Delta t - n\Delta t) = \Delta t \sum_n h(n\Delta t) s(k\Delta t - n\Delta t). \quad (6.4.1)$$

При выполнении дискретной свертки мы имеем дело с цифровыми массивами, при этом шаг дискретизации для массивов по физическому аргументу свертки должен быть равным и принимается за 1, а в качестве аргумента используется нумерация отсчетов в массивах:

$$y(k) = \sum_n h(n)s(k-n) \equiv \sum_n h_n s_{k-n} \equiv y_k. \quad (6.4.1')$$

$$y(k) = h(n) * s(k-n) \equiv s(k) * h(n) \equiv s_k * h_n.$$

Техника свертки приведена на рис. 6.4.1. Для вычисления свертки массив одной из функций (s_k - входного сигнала) располагается по ходу возрастания номеров. Массив второй функции (h_n - более короткой, оператор свертки), строится параллельно первому массиву в обратном порядке (по ходу уменьшения номеров, в режиме обратного времени). Для вычисления y_k значение h_0 располагается против s_k , все значения s_{k-n} перемножаются с расположенными против них значениями h_n и суммируются. Результаты суммирования являются выходным значением функции y_k , после чего оператор h_n сдвигается на один номер k вперед (или функция s_k сдвигается ему навстречу) и вычисление повторяется для номера $k+1$ и т.д.

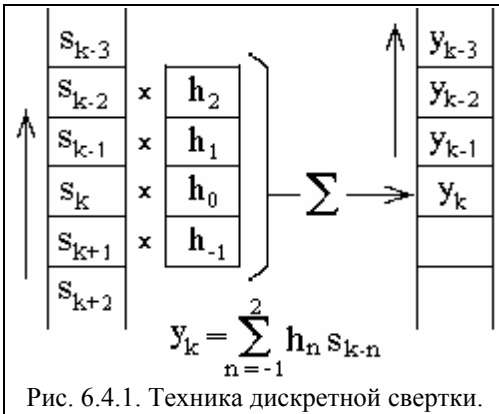


Рис. 6.4.1. Техника дискретной свертки.

В начальный момент свертки при вычислении значений y_k оператор h_n , построенный в режиме обратного времени, "зависает" для значений $k-n$ при $n > k$ против отсутствующих отсчетов входной функции. "Зависание" исключают либо заданием начальных условий - дополнительных отсчетов, чаще всего нулевых или равных первому отсчету входной функции, либо началом свертки с отсчета входной функции $k = n$ с соответствующим сокращением интервала выходной функции. Для операторов со значениями $-n$ (вперед по времени) такой же момент может наступать и в конце входного массива.

Пример. Уравнение свертки: $y_k = \sum_{n=0}^2 b_n x_{k-n} = b_0 x_k + b_1 x_{k-1} + b_2 x_{k-2}$. Значения оператора b_n :

$b_0 = 5, b_1 = 3, b_2 = 2$. Входной сигнал: $x_k = \{0, 1, 0, 0, 0\}$, начальные условия: $x_{-n} = 0$.

Расчет выходного сигнала:

$$y_0 = 5x_0 + 3x_{-1} + 2x_{-2} = 5 \cdot 0 + 3 \cdot 0 + 2 \cdot 0 = 0, \quad y_1 = 5x_1 + 3x_0 + 2x_{-1} = 5 \cdot 1 + 3 \cdot 0 + 2 \cdot 0 = 5,$$

$$y_2 = 5x_2 + 3x_1 + 2x_0 = 5 \cdot 0 + 3 \cdot 1 + 2 \cdot 0 = 3, \quad y_3 = 5x_3 + 3x_2 + 2x_1 = 5 \cdot 0 + 3 \cdot 0 + 2 \cdot 1 = 2,$$

$$y_4 = 5x_4 + 3x_3 + 2x_2 = 5 \cdot 0 + 3 \cdot 0 + 2 \cdot 0 = 0, \quad y_5 = 5x_5 + 3x_4 + 2x_3 = 5 \cdot 0 + 3 \cdot 0 + 2 \cdot 0 = 0$$

Выходной сигнал: $y_k = \{0, 5, 3, 2, 0\}$

Заметим: свертка функции оператора с единичным входным сигналом представляет собой повторение функции оператора свертки на выходе.

На рис. 6.4.2 приведен пример выполнения дискретной свертки каузальным (односторонним) и четным (симметричным, двусторонним) оператором одного и того же сигнала.

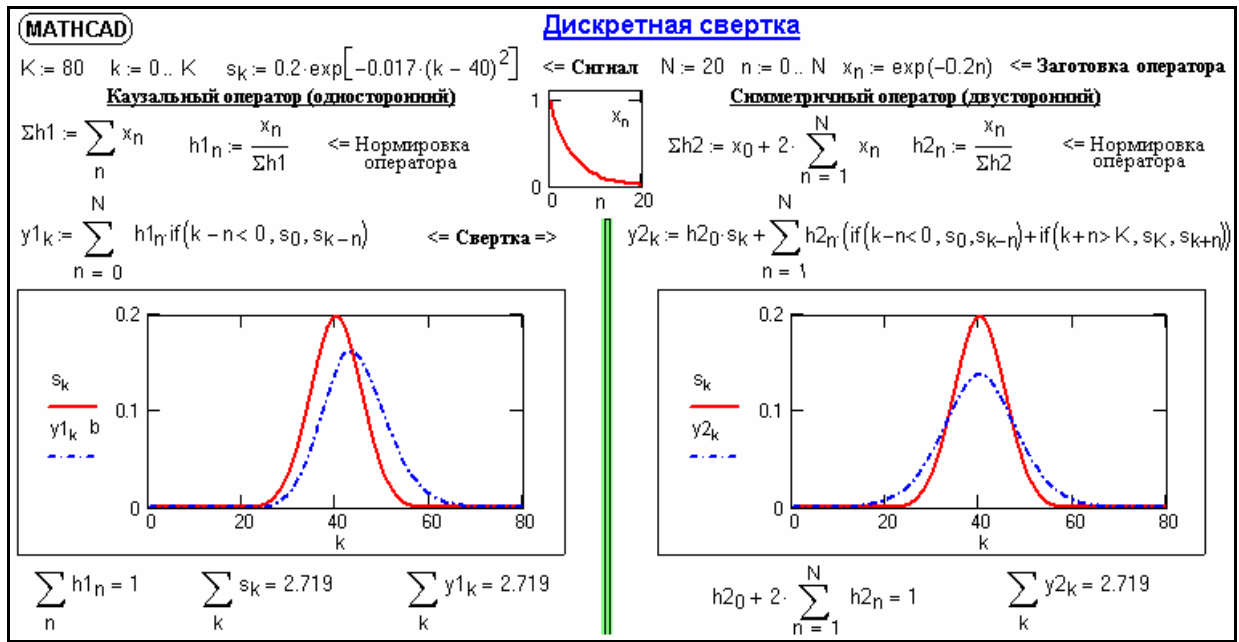


Рис. 6.4.2. Примеры выполнения дискретной свертки.

Для дискретной свертки действительны все свойства и теоремы интегральной свертки. В частности, свертка функций в координатной области отображается произведением их спектров в частотной области, что позволяет использовать ДПФ для вычисления свертки при большой длине функций по следующей схеме:

$$s(k) \Leftrightarrow S(\omega), \quad h(n) \Leftrightarrow H(\omega), \quad Y(\omega) = S(\omega) \cdot H(\omega), \quad Y(\omega) \Leftrightarrow y(k).$$

Выполнение произведения спектров может производиться только при одинаковой их длине, и оператор h перед ДПФ обычно дополняется нулями до размера функции $s(k)$.

Второй фактор, который следует принимать во внимание, это цикличность свертки при ее выполнении в спектральной области, обусловленная периодизацией дискретных функций. Перемножаемые спектры являются спектрами периодических функций, и результат на конечных интервалах может не совпадать с дискретной линейной сверткой, где условия продления интервалов (начальные условия) задаются, а не повторяют главный период.

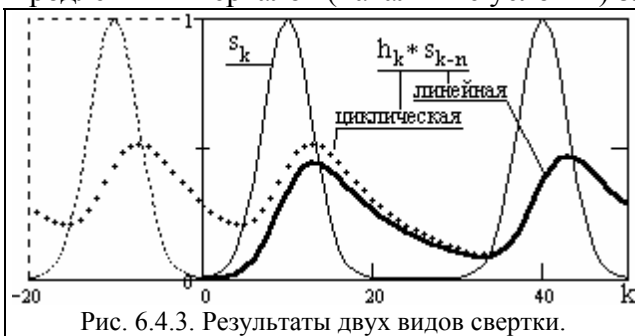


Рис. 6.4.3. Результаты двух видов свертки.

На рис. 6.4.3 приведены результаты свертки сигнала s_k , заданного на интервале $k=(0-50)$, с функцией $h_n = a \cdot \exp(-a \cdot n)$, $a = 0.1$. Свертка, выполненная через ДПФ, в левой части интервала резко отличается от линейной свертки. Характер искажения становится понятным, если дополнить главный интервал с левой стороны его периодическим продолжением (на рисунке показана часть левого бокового периода, свертка с которым заходит в главный период).

Для операторов h_n со значениями n , вперед по положению, аналогичные искажения появятся и в правой части главного периода. Для устранения таких искажений функции должны продлеваться нулями на размер оператора h , что исключит наложение боковых периодов главной трассы функции.

При выполнении свертки через БПФ ощутимое повышение скорости вычислений появляется только при большой длине функций и операторов (например, $M > 1000, N > 100$). Следует также обращать внимание на разрядность результатов, т.к. перемножение чисел дает увеличение разрядности в 2 раза. При ограниченной разрядности числового представления с соответствующим округлением это может приводить к погрешностям суммирования.

ЛИТЕРАТУРА

2. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. – М.: Мир, 1989. – 540 с.

5. Гольденберг Л.М. и др. Цифровая обработка сигналов: Учебное пособие для вузов. - М.: Радио и связь, 1990.- 256 с.
13. Канасевич Э.Р. Анализ временных последовательностей в геофизике. - М.: Недра, 1985.- 300 с.
17. Никитин А.А. Теоретические основы обработки геофизической информации: Учебник для вузов. - М.: Недра, 1986.- 342 с.
18. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. – М.: Связь, 1979. – 416 с.
21. Рапопорт М.Б. Вычислительная техника в полевой геофизике: Учебник для вузов. - М.: Недра, 1993.- 350 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 7: ЭНЕРГЕТИЧЕСКИЕ СПЕКТРЫ СИГНАЛОВ

Математик может говорить все, что взбретет ему в голову, но физик обязан сохранять хотя бы крупицу здравого смысла.

Джосая Гиббс. Американский физик, XIX в.

Хотела бы я знать, где эта самая крупица здравого смысла в энергии информационного сигнала, если сама информация считается понятием нематериальным.

Маргарита Пястолова. Иркутский геофизик Уральской школы, XX в.

Содержание: Введение. 7.1. Мощность и энергия сигналов. 7.2. Энергетические спектры сигналов. Скалярное произведение сигналов. Взаимный энергетический спектр. Энергетический спектр сигнала. Литература.

ВВЕДЕНИЕ

Понятия мощности и энергии в теории сигналов не относятся к характеристикам каких-либо физических величин сигналов, а являются их количественными характеристиками, отражающими определенные свойства сигналов и динамику изменения их значений во времени, в пространстве или по любым другим аргументам.

Для произвольного, в общем случае комплексного, сигнала мгновенная мощность *по определению* равна квадрату функции его модуля, для вещественных сигналов - квадрату функции амплитуд. Энергия сигнала, также по определению, равна интегралу от мощности по всему интервалу существования или задания сигнала.

Энергия сигналов может быть конечной или бесконечной. Конечную энергию имеют финитные сигналы и сигналы, затухающие по своим значениям в пределах конечной длительности, которые не содержат дельта-функций и особых точек (разрывов второго рода и ветвей, уходящих в бесконечность). В противном случае их энергия равна бесконечности. Бесконечна также энергия периодических сигналов.

7.1. МОЩНОСТЬ И ЭНЕРГИЯ СИГНАЛОВ [1,3,16].

Частотное представление применяется не только для спектрального анализа сигналов, но и для упрощения вычислений энергии сигналов и их корреляционных характеристик.

Как уже рассматривалось ранее, для произвольного сигнала $s(t) = a(t) + jb(t)$, где $a(t)$ и $b(t)$ - вещественные функции, мгновенная мощность сигнала (плотность распределения энергии) определяется выражением:

$$w(t) = s(t)s^*(t) = a^2(t) + b^2(t) = |s(t)|^2.$$

Энергия сигнала равна интегралу от мощности по всему интервалу существования сигнала. В пределе:

$$E_s = \int_{-\infty}^{\infty} w(t) dt = \int_{-\infty}^{\infty} |s(t)|^2 dt.$$

По существу, мгновенная мощность является плотностью мощности сигнала, так как измерения мощности возможны только через энергию, выделяемую на определенных интервалах ненулевой длины:

$$w(\tau) = (1/\Delta t) \int_{\tau - \Delta t/2}^{\tau + \Delta t/2} |s(t)|^2 dt.$$

Сигнал $s(t)$ изучается, как правило, на определенном интервале T (для периодических сигналов - в пределах одного периода T), при этом средняя мощность сигнала:

$$W_T(\tau) = (1/T) \int_{\tau}^{\tau+T} w(t) dt = (1/T) \int_{\tau}^{\tau+T} |s(t)|^2 dt.$$

Понятие средней мощности может быть распространено и на незатухающие сигналы, энергия которых бесконечно велика. В случае неограниченного интервала T строго корректное определение средней мощности сигнала производится по формуле:

$$W_s = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T w(t) dt.$$

Энергия и норма сигналов связаны соотношениями:

$$E_s = \|s(t)\|^2, \quad \|s\| = \sqrt{E_s}.$$

7.2. ЭНЕРГЕТИЧЕСКИЕ СПЕКТРЫ СИГНАЛОВ [1].

Скалярное произведение сигналов. Энергия суммы двух произвольных сигналов $u(t)$ и $v(t)$ определяется выражением:

$$E = \int_{-\infty}^{\infty} [u(t)+v(t)]^2 dt = E_u + E_v + 2 \int_{-\infty}^{\infty} u(t)v(t) dt. \quad (7.2.1)$$

Как следует из этого выражения, энергии сигналов, в отличие от самих сигналов, в общем случае не обладают свойством аддитивности. Энергия суммарного сигнала $u(t)+v(t)$, кроме суммы энергий составляющих сигналов, содержит в себе и так называемую энергию взаимодействия сигналов или взаимную энергию:

$$E_{uv} = 2 \int_{-\infty}^{\infty} u(t)v(t) dt. \quad (7.2.2)$$

Интеграл выражения (7.2.2) для двух вещественных сигналов является фундаментальной характеристикой, пропорциональной взаимной энергии сигналов. Его называют *скалярным произведением* сигналов:

$$\Pi_{uv} = (u(t), v(t)) = \int_{-\infty}^{\infty} u(t)v(t) dt = \|u\| \cdot \|v\| \cos \varphi, \quad (7.2.3)$$

Скалярное произведение обладает следующими свойствами:

1. $(u, v) \geq 0$;
2. $(u, v) = (v, u)$;
3. $(au, v) = a(u, v)$, где a – вещественное число;
4. $(u+v, a) = (u, a) + (v, a)$.

Линейное пространство сигналов с таким скалярным произведением называется *гильбертовым пространством* H . С учетом того, что $\cos \varphi \leq 1$, в гильбертовом пространстве справедливо неравенство Коши-Буняковского:

$$|\Pi_{uv}| \leq \|u\| \cdot \|v\|. \quad (7.2.4)$$

Для комплексного гильбертова пространства скалярное произведение также представляет собой вещественное число и вычисляется по формуле:

$$\Pi_{uv} = \int_{-\infty}^{\infty} u(t)v^*(t) dt \equiv \int_{-\infty}^{\infty} u^*(t)v(t) dt. \quad (7.2.3')$$

Из выражения (7.2.3) следует, что косинус угла между сигналами:

$$\cos \varphi = \Pi_{uv} / (\|u\| \cdot \|v\|). \quad (7.2.5)$$

При полной тождественности сигналов (равенстве амплитуд и временных координат) имеем $\varphi = 0$, $\cos \varphi = 1$, и скалярное произведение становится равным энергии сигналов:

$$\Pi_{uv} = \int_{-\infty}^{\infty} u(t)^2 dt \equiv \int_{-\infty}^{\infty} v(t)^2 dt \equiv \|u\|^2 \equiv \|v\|^2.$$

Дискретные сигналы обычно рассматриваются в *пространстве Евклида* (обозначение пространства - R_2). Скалярное произведение двух сигналов в пространстве Евклида:

$$\Pi_{uv} = (u_k, v_k) = \sum_{k=1}^n u_k v_k,$$

где n - размерность пространства.

Взаимный энергетический спектр. Из очевидной однозначности энергии взаимодействия сигналов независимо от формы их математического представления (в динамической и частотной модели) следует выражение для скалярного произведения произвольных вещественных сигналов $u(t)$ и $v(t)$ через спектральные плотности сигналов $U(\omega)$ и $V(\omega)$ в комплексном гильбертовом пространстве:

$$P_{uv} = (1/2\pi) \int_{-\infty}^{\infty} U(\omega)V^*(\omega) d\omega \equiv (1/2\pi) \int_{-\infty}^{\infty} U^*(\omega)V(\omega) d\omega. \quad (7.2.6)$$

Функции

$$W_{uv}(\omega) = U(\omega)V^*(\omega), \quad W_{vu}(\omega) = U^*(\omega)V(\omega), \quad W_{uv}(\omega) = W_{vu}^*(\omega), \quad (7.2.7)$$

для которых справедливо выражение (7.2.6), называется *взаимными энергетическими спектрами* вещественных сигналов, и являются функциями распределения плотности энергии взаимодействия сигналов (мощности взаимодействия) по частоте.

В общем случае, за исключением спектров четных функций, взаимные энергетические спектры также являются комплексными функциями:

$$U(\omega) = A_u(\omega) + j B_u(\omega), \quad V(\omega) = A_v(\omega) + j B_v(\omega).$$

$$W_{uv} = A_u A_v + B_u B_v + j (B_u A_v - A_u B_v) = \text{Re } W_{uv}(\omega) + j \text{Im } W_{uv}(\omega). \quad (7.2.7')$$

С учетом четности реальной части и нечетности мнимой части энергетических спектров, интеграл мнимой части выражения (7.2.7') равен нулю, а, следовательно, скалярное произведение сигналов всегда является вещественным и неотрицательным, как и энергия сигналов:

$$P_{uv} = (1/2\pi) \int_{-\infty}^{\infty} W_{uv}(\omega) d\omega \equiv (1/\pi) \int_0^{\infty} \text{Re } W_{uv}(\omega) d\omega. \quad (7.2.8)$$



Рис. 7.2.1. Форма и энергетические спектры сигналов.

На рис. 7.2.1 приведена форма двух одинаковых сдвинутых во времени и частично перекрывающихся лапласовских импульсов $u(t)$ и $v(t)$, а также суммарный импульс $z(t)=u(t)+v(t)$. Плотности энергии сигналов $W(f)$ приведены в относительных единицах плотности энергии суммарного сигнала $W_z(f)$ на нулевой частоте.

Как видно из графиков, плотности энергии сигналов являются вещественными неотрицательными функциями и содержат только реальные части. В отличие от них, плотность взаимной энергии сигналов является комплексной функцией, при этом модуль плотности по своим значениям на шкале частот соизмерим со средними значениями плотности энергии сигналов на этих частотах и не зависит от их взаимного расположения на временной оси. Для сигналов, одинаковых по форме, модуль взаимной плотности равен значениям плотности энергии сигналов.

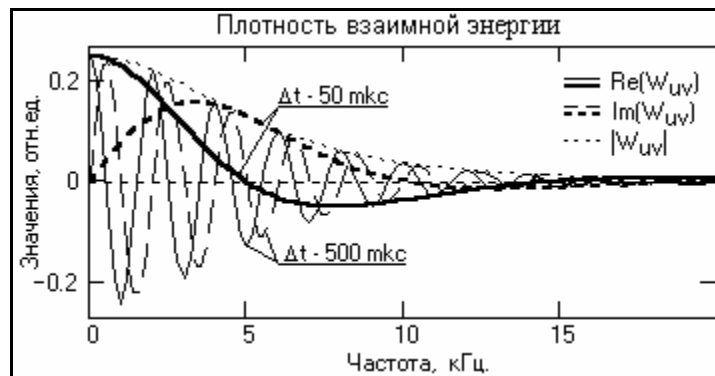


Рис. 7.2.2. Взаимные энергетические спектры сигналов.

На рис. 7.2.2 приведены плотности взаимной энергии тех же сигналов при разной величине временного сдвига Δt между сигналами. Однако при постоянном значении модуля взаимной энергии сигналов действительная и мнимая функции спектра мощности существенно изменяются при изменении сдвига между сигналами. При незначительной величине временного перекрытия сигналов частота осцилляций реальной и мнимой части плотности взаимной энергии достаточно велика, а относительный коэффициент затухания колебаний (уменьшение амплитудных значений от периода к периоду) достаточно мал. Соответственно, при вычислении скалярного произведения по формуле (7.2.8) положительные амплитудные значения осцилляций $\text{Re}(W_{uv})$ практически полностью компенсируются отрицательными значениями и результирующий интеграл, а равно и энергия взаимодействия сигналов (удвоенное значение скалярного произведения), близка к нулевой (стремится к нулю по мере увеличения сдвига между сигналами).

При увеличении степени взаимного перекрытия сигналов частота осцилляций плотности взаимной энергии уменьшается ($\Delta t = 50$ мкс на рис. 7.2.2) и основным по энергии реальной части спектра становится центральный низкочастотный пик, площадь которого не компенсируется площадью последующей отрицательной полуволны осцилляции. Соответственно, возрастает и энергия взаимодействия сигналов. При полном перекрытии сигналов (при нулевом фазовом угле между сигналами) осцилляции исчезают, и энергия взаимодействия сигналов максимальна.

Энергетический спектр сигнала. Если функция $s(t)$ имеет фурье-образ $S(\omega)$, то плотность мощности сигнала (*спектральная плотность энергии сигнала*) определяется выражением:

$$w(t) = s(t)s^*(t) = |s(t)|^2 \Leftrightarrow |S(\omega)|^2 = S(\omega)S^*(\omega) = W(\omega). \quad (7.2.9)$$

Спектр мощности $W(\omega)$ - вещественная неотрицательная четная функция, которую обычно называют *энергетическим спектром*. Спектр мощности, как квадрат модуля спектральной плотности сигнала, не содержит фазовой информации о его частотных составляющих, а, следовательно, восстановление сигнала по спектру мощности невозможно. Это означает также, что сигналы с различными фазовыми характеристиками могут иметь одинаковые спектры мощности. В частности, сдвиг сигнала не отражается на его спектре мощности. Последнее позволяет получить выражение для энергетического спектра непосредственно из выражений (7.2.7). В пределе, для одинаковых сигналов $u(t)$ и $v(t)$ при сдвиге $\Delta t \Rightarrow 0$, мнимая часть спектра $W_{uv}(\omega)$ стремится к нулевым значениям, а реальная часть - к значениям модуля спектра. При полном временном совмещении сигналов имеем:

$$W_{uv}(\omega) = U(\omega)V^*(\omega) = U(\omega)U^*(\omega) = |U(\omega)|^2 = W_u(\omega). \quad (7.2.10)$$

Соответственно, полная энергия сигнала:

$$E_u = \int_{-\infty}^{\infty} u(t)^2 dt = (1/2\pi) \int_{-\infty}^{\infty} W_u(t) dt = (1/2\pi) \int_{-\infty}^{\infty} |U(\omega)|^2 d\omega, \quad (7.2.11)$$

т.е. энергия сигнала равна интегралу квадрата модуля его частотного спектра - сумме энергии его частотных составляющих, и всегда является вещественной величиной.

Для произвольного сигнала $s(t)$ равенство

$$\int_{-\infty}^{\infty} |s(t)|^2 dt = \int_{-\infty}^{\infty} |S(f)|^2 df$$

обычно называют равенством Парсеваля (в математике - теоремой Планшереля, в физике - формулой Релея). Равенство очевидно, так как координатное и частотное представления по существу только разные математические отображения одного и того же сигнала. Аналогично для энергии взаимодействия двух сигналов:

$$\int_{-\infty}^{\infty} u(t) v^*(t) dt = \int_{-\infty}^{\infty} U(f) V^*(f) df.$$

Из равенства Парсеваля следует инвариантность скалярного произведения сигналов и нормы относительно преобразования Фурье:

$$(u(t), v(t)) = (U(f), V(f)), \quad \|s(t)\|^2 = \|S(f)\|^2.$$

В целом ряде чисто практических задач регистрации и передачи сигналов энергетический спектр сигнала имеет весьма существенное значение.

Периодические сигналы переводятся в спектральную область в виде рядов Фурье. Запишем периодический сигнал с периодом T в виде ряда Фурье в комплексной форме:

$$s(t) = \sum_{k=-\infty}^{\infty} S_k \exp(j2\pi kt/T),$$

и вычислим среднюю мощность сигнала за один период:

$$W_T = (1/T) \int_0^T s^2(t) dt = (1/T) \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} S_k S_m \int_0^T \exp(j2\pi(k+m)t/T) dt.$$

Интервал $0-T$ содержит целое число периодов всех подынтегральных экспонент, и равен нулю, за исключением экспоненты при $k = -m$, для которой интеграл равен T . Соответственно, средняя мощность периодического сигнала равна сумме квадратов модулей коэффициентов его ряда Фурье:

$$W_T = \sum_{k=-\infty}^{\infty} |S_k|^2.$$

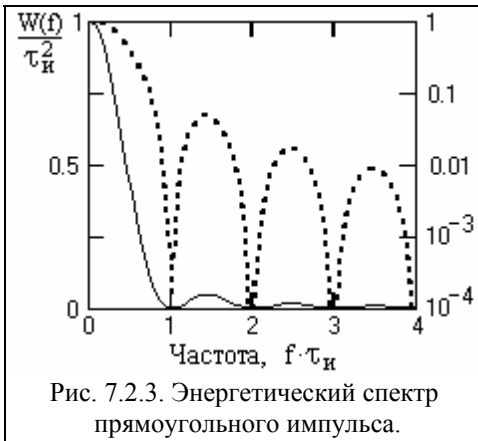


Рис. 7.2.3. Энергетический спектр прямоугольного импульса.

Как правило, спектры сигналов с крутыми фронтами (например, кодовых сигналов при передаче цифровых данных) являются многолепестковыми с постепенным затуханием энергии в последовательных лепестках. Пример нормированного энергетического спектра прямоугольного импульса длительностью τ_n приведен на рис. 7.2.3. Спектры выполнены в линейном (сплошная линия) и логарифмическом (пунктир) масштабе по оси значений. Для четкого разделения лепестков функции спектров приведены по безразмерной частотной переменной $f \cdot \tau_n$.

Интегрированием энергетического спектра по интервалам лепестков спектра нетрудно вычислить, что в пределах первого лепестка сосредоточено 90.2% энергии всего сигнала, в пределах второго – 4.8%, в пределах третьего – 1.7%, и т.д. Если форма сигналов в пункте их приема (детектирования) существенного значения не имеет, а регистрация сигналов идет на уровне статистических шумов, равномерно распределенных по всему частотному диапазону, то такие сигналы целесообразно пропускать через фильтр нижних частот с выделением только первого энергетического лепестка сигнала. Естественно, что при этом фронты регистрируемого сигнала будут сглажены. Но при расширении полосы пропускания фильтра на два или три лепестка энергия принимаемого сигнала будет увеличена соответственно на 4.8 или 6.5%, в то время как энергия шумов в 2 или 3 раза.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
3. Васильев Д.В. Радиотехнические цепи и сигналы: Учебное пособие для вузов. - М.: Радио и связь, 1982. - 528 с.
16. Макс Ж. Методы и техника обработки сигналов при физических измерениях. - М.: Мир, 1983.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 8: КОРРЕЛЯЦИЯ СИГНАЛОВ

Предельный страх и предельный пыл храбрости одинаково расстраивают желудок и вызывают понос.

Мишель Монтень. Опыты.
Французский юрист-мыслитель, XVI в.

Вот это номер! Две функции имеют стопроцентную корреляцию с третьей и ортогональны друг другу. Ну и шуточки были у Всевышнего при сотворении Мира.

Анатолий Пышминцев. Мысли.
Русский геофизик-практик, XX в.

Содержание: Введение. 8.1. Автокорреляционные функции сигналов (АКФ). Фinitные сигналы. Сигналы, неограниченные во времени. Периодические сигналы. Дискретные сигналы. Кодовые сигналы. Функции автоковариации. 8.2. Взаимнокорреляционные функции сигналов (ВКФ). 8.3. Спектральные плотности корреляционных функций. Спектральная плотность АКФ. Интервал корреляции сигнала. Спектральная плотность ВКФ. Литература.

ВВЕДЕНИЕ

Корреляция (correlation), и ее частный случай для центрированных сигналов – ковариация, является методом анализа сигналов. Приведем один из вариантов использования метода. Допустим, что имеется сигнал $s(t)$, в котором может быть (а может и не быть) некоторая последовательность $x(t)$ конечной длины T , временное положение которой нас интересует. Для поиска этой последовательности в скользящем по сигналу $s(t)$ временном окне длиной T вычисляются скалярные произведения сигналов $s(t)$ и $x(t)$. Тем самым мы "прикладываем" искомый сигнал $x(t)$ к сигналу $s(t)$, скользя по его аргументу, и по величине скалярного произведения оцениваем степень сходства сигналов в точках сравнения. В функциональном пространстве сигналов эта степень сходства может выражаться в нормированных единицах коэффициента корреляции, т.е. в косинусе угла между векторами сигналов, и, соответственно, будет принимать значения от 1 (полное совпадение сигналов) до -1 (полная противоположность).

В варианте автокорреляции (autocorrelation) по аналогичной методике производится определение скалярного произведения сигнала $s(t)$ с собственной копией, скользящей по аргументу. Автокорреляция позволяет оценить среднестатистическую зависимость текущих отсчетов сигнала от своих предыдущих и последующих значений (так называемый радиус корреляции значений сигнала), а также выявить в сигнале наличие периодически повторяющихся элементов.

Особое значение методы корреляции имеют при анализе случайных процессов для выявления неслучайных составляющих и неслучайных параметров этих процессов.

Заметим, что в терминах "корреляция" и "ковариация" в настоящее время существует изрядная путаница. В иностранной литературе термин "ковариация" применяется к центрированным функциям, а "корреляция" – к произвольным. В отечественной литературе, и особенно в литературе по сигналам и их обработке, довольно часто применяется прямо противоположная терминология. Однако при переводах иностранной литературы терминология, как правило, не изменяется, и начинает все шире проникать и в отечественную литературу. Принципиального значения это не имеет, но при знакомстве с литературными источниками стоит обращать внимание на принятое назначение данных терминов.

При разработке настоящих лекций было принято решение использовать общепринятую международную терминологию, как согласованную по понятиям с основными положениями теории вероятностей и математической статистики.

8.1. АВТОКОРРЕЛЯЦИОННЫЕ ФУНКЦИИ СИГНАЛОВ [1,25].

Автокорреляционные функции (АКФ) фinitных сигналов. АКФ (correlation function, CF) сигнала $s(t)$, локализованного во времени и конечного по энергии, является количест-

венной интегральной характеристикой формы сигнала, и определяется интегралом от произведения двух копий сигнала $s(t)$, сдвинутых относительно друг друга на время τ :

$$B_s(\tau) = \int_{-\infty}^{\infty} s(t) s(t+\tau) dt = \langle s(t), s(t+\tau) \rangle = \|s(t)\| \|s(t+\tau)\| \cos \varphi(\tau). \quad (8.1.1)$$

Как следует из этого выражения, АКФ является скалярным произведением сигнала и его копии в функциональной зависимости от переменной величины значения сдвига τ . Соответственно, АКФ имеет физическую размерность энергии, а при $\tau = 0$ значение АКФ непосредственно равно энергии сигнала:

$$B_s(0) = \int_{-\infty}^{\infty} s(t)^2 dt = E_s.$$

АКФ относится к четным функциям, в чем нетрудно убедиться заменой переменной $t = t - \tau$ в выражении (8.1.1):

$$B_s(\tau) = \int_{-\infty}^{\infty} s(t) s(t+\tau) dt = \int_{-\infty}^{\infty} s(t-\tau) s(t) dt = B_s(-\tau).$$

Максимум АКФ, равный энергии сигнала при $\tau=0$, всегда положителен, а модуль АКФ при любом значении временного сдвига не превосходит энергии сигнала. Последнее прямо вытекает из свойств скалярного произведения (как и неравенство Коши-Буняковского):

$$\langle s(t), s(t+\tau) \rangle = \|s(t)\| \cdot \|s(t+\tau)\| \cdot \cos \varphi(\tau),$$

$$\cos \varphi(\tau) = 1 \text{ при } \tau = 0, \quad \langle s(t), s(t+\tau) \rangle = \|s(t)\| \cdot \|s(t)\| = E_s,$$

$$\cos \varphi(\tau) < 1 \text{ при } \tau \neq 0, \quad \langle s(t), s(t+\tau) \rangle = \|s(t)\| \cdot \|s(t+\tau)\| \cdot \cos \varphi(\tau) < E_s.$$

В качестве примера на рис. 8.1.1 приведены два сигнала – прямоугольный импульс и радиоимпульс одинаковой длительности, и соответствующие данным сигналам формы их АКФ. Амплитуда колебаний радиоимпульса установлена равной \sqrt{T} амплитуды прямоугольного импульса, при этом энергии сигналов также будут одинаковыми, что подтверждается равными значениями центральных максимумов АКФ. При конечной длительности значимых значений импульсов длительности АКФ также конечны, и равны удвоенным значениям длительности импульсов (при сдвиге копии конечного импульса на интервал его длительности как влево, так и вправо, произведение импульса со своей копией становится равным нулю). Частота колебаний АКФ радиоимпульса равна частоте колебаний заполнения радиоимпульса (боковые минимумы и максимумы АКФ возникают каждый раз при последовательных сдвигах копии радиоимпульса на половину периода колебаний его заполнения).

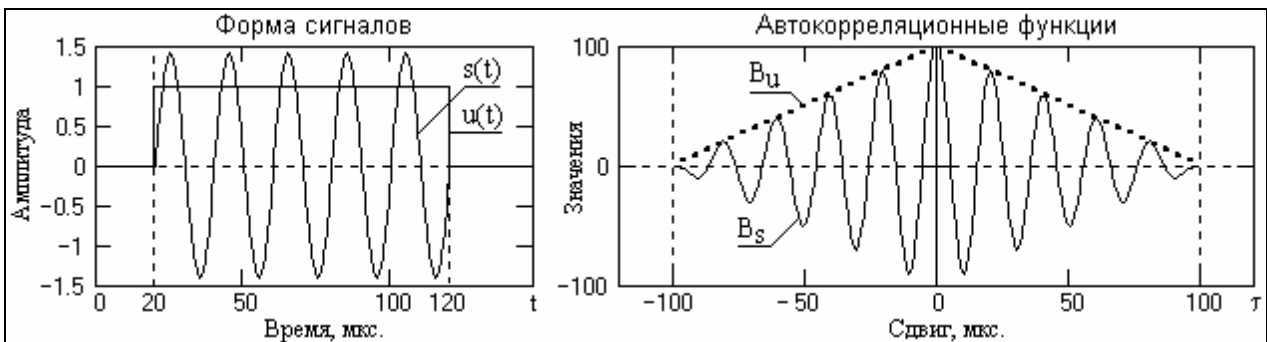


Рис. 8.1.1.

С учетом четности, графическое представление АКФ обычно производится только для положительных значений τ . Знак $+\tau$ в выражении (8.1.1) означает, что при увеличении значений τ от нуля копия сигнала $s(t+\tau)$ сдвигается влево по оси t . На практике сигналы обычно задаются на интервале положительных значений аргументов от $0-T$, что дает возможность продления интервала нулевыми значениями, если это необходимо для математических операций. В этих границах вычислений более удобным для построения вычислительных алгоритмов является сдвиг копии сигнала вправо по оси аргументов, т.е. применение в выражении (8.1.1) функции копии $s(t-\tau)$:

$$B_s(\tau) = \int_{-\infty}^{\infty} s(t) s(t-\tau) dt. \quad (8.1.1')$$

По мере увеличения значения величины сдвига τ для финитных сигналов временное перекрытие сигнала с его копией уменьшается, а, соответственно, косинус угла взаимодействия и скалярное произведение в целом стремятся к нулю:

$$\lim_{|\tau| \rightarrow \infty} B_s(\tau) = 0.$$

АКФ, вычисленная по центрированному значению сигнала $s(t)$, и нормированная на центральное значение, изменяется в интервале от -1 до 1 и представляет собой функцию автокорреляционных коэффициентов:

$$\rho_s(\tau) = B_s(\tau)/B_s(0) = B_s(\tau)/\|s(t)\|^2.$$

Иногда эту функцию называют "истинной" автокорреляционной функцией.

АКФ неограниченных во времени сигналов. Для сигналов, заданных на определенном интервале $[a, b]$, вычисление АКФ производится с нормировкой на длину интервала $[a, b]$:

$$B_s(\tau) = \frac{1}{b-a} \int_a^b s(t) s(t+\tau) dt. \quad (8.1.1'')$$

В пределе, для сигналов с бесконечной энергией, АКФ может быть получена как среднее значение скалярного произведения сигнала и его копии при устремлении интервала длительности сигнала T к бесконечности:

$$B_s(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T s(t) s(t + \tau) dt.$$

АКФ по данному выражению имеет физическую размерность мощности, и равна средней взаимной мощности сигнала и его копии в функциональной зависимости от сдвига копии.

АКФ периодических сигналов. Энергия периодических сигналов бесконечна, поэтому АКФ периодических сигналов вычисляется по одному периоду T , с усреднением скалярного произведения сигнала и его сдвинутой копии в пределах этого периода:

$$B_s(\tau) = (1/T) \int_0^T s(t) s(t-\tau) dt. \quad (8.1.2')$$

При $\tau=0$ значение нормированной на период АКФ в этом случае равно не энергии, а средней мощности сигналов в пределах периода. При этом АКФ периодических сигналов является периодической функцией с тем же периодом T . Так, для сигнала $s(t) = A \cos(\omega_0 t + \varphi_0)$ при $T=2\pi/\omega_0$ имеем:

$$B_s(\tau) = \frac{\omega_0}{2\pi} \int_{-\pi/\omega_0}^{\pi/\omega_0} A \cos(\omega_0 t + \varphi_0) A \cos(\omega_0(t-\tau) + \varphi_0) = (A^2/2) \cos(\omega_0 \tau).$$

Отметим, что полученный результат не зависит от начальной фазы гармонического сигнала, что характерно для любых периодических сигналов и является одним из свойств КФ.

АКФ дискретных сигналов. При интервале дискретизации данных $\Delta t = \text{const}$ вычисление АКФ выполняется по интервалам $\Delta \tau = \Delta t$ и обычно записывается, как дискретная функция номеров n сдвига отсчетов $n\Delta t$:

$$B_s(n) = \Delta t \sum_{k=-\infty}^{\infty} s_k \cdot s_{k-n}. \quad (8.1.3)$$

Дискретные сигналы обычно задаются в виде числовых массивов определенной длины с нумерацией отсчетов $k = 0, 1, \dots, N$, а вычисление дискретной АКФ выполняется в одностороннем варианте с учетом длины массивов по формуле:

$$B_s(n) = \frac{N \cdot \Delta t}{N+1-n} \sum_{k=0}^{N-n} s_k \cdot s_{k-n}. \quad (8.1.3')$$

Множитель $N/(N+1-n)$ в данной функции является поправочным коэффициентом на постепенное уменьшение числа перемножаемых и суммируемых значений (от N до $N-n$) по мере увеличения сдвига n . Без этой поправки для нецентрированных сигналов в значениях

АКФ появляется тренд суммирования средних значений.

Практически, дискретная АКФ имеет такие же свойства, как и непрерывная АКФ. Она также является четной, а ее значение при $n = 0$ равно энергии дискретного сигнала.

Кодовые сигналы являются разновидностью дискретных сигналов. На определенном интервале кодового слова $M \cdot \Delta t$ они могут иметь только два амплитудных значения: 0 и 1 или 1 и -1 . При выделении кодов на существенном уровне шумов форма АКФ кодового слова имеет особое значение. С этой позиции наилучшими считаются такие коды, значения боковых лепестков АКФ которых минимальны по всей длине интервала кодового слова при максимальном значении центрального пика. К числу таких кодов относится код Баркера, приведенный в таблице 8.1. Как видно из таблицы, амплитуда центрального пика кода численно равна значению M , при этом амплитуда боковых осцилляций при $n \neq 0$ не превышает 1.

Таблица 8.1.

M	Сигнал Баркера	АКФ сигнала
2	1, -1	2, -1
3	1, 1, -1	3, 0, -1
4	1, 1, 1, -1	4, 1, 0, -1
	1, 1, -1, 1	4, -1, 0, 1
5	1, 1, 1, -1, 1	5, 0, 1, 0, 1
7	1, 1, 1, -1, -1, 1, -1	7, 0, -1, 0, -1, 0, -1
11	1, 1, 1, -1, -1, -1, 1, -1, 1, -1	11, 0, -1, 0, -1, 0, -1, 0, -1, 0, -1
13	1, 1, 1, 1, 1, -1, -1, 1, 1, -1, 1, -1, 1	13, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1

Функции автоковариации (ФАК). В практике обработки и статистического анализа дискретных данных и функций вместо корреляционных функций обычно используются ковариационные функции, хотя эти два термина очень часто используются как синонимы. Строго корректно, под функциями автоковариации (ФАК) следует понимать вторые центральные моменты распределения числовых величин в цифровых массивах:

$$C_s(n) = \frac{N \cdot \Delta t}{N+1-n} \sum_{k=0}^{N-n} (s_k - \bar{s}) \cdot (s_{k-n} - \bar{s}), \quad (8.1.4)$$

где \bar{s} - среднее значение функций. Максимум ФАК, как и АКФ, соответствует $n = 0$ и равен дисперсии сигнала s_k . Из сравнения (8.1.4) с выражением (8.1.3) отсюда следует, что дисперсия сигналов (квадрат стандартного отклонения от среднего значения) равна средней энергии сигнала по интервалу его задания. Соответственно, связь ФАК с энергетическим спектром (плотностью распределения мощности сигнала) через преобразование Фурье сохраняется аналогичной АКФ.

8.2. ВЗАИМНЫЕ КОРРЕЛЯЦИОННЫЕ ФУНКЦИИ СИГНАЛОВ [1,19].

Взаимная корреляционная функция (ВКФ) разных сигналов (cross-correlation function, CCF) описывает как степень сходства формы двух сигналов, так и их взаимное расположение друг относительно друга по координате (независимой переменной). Обобщая формулу (8.1.1) автокорреляционной функции на два различных сигнала $s(t)$ и $u(t)$, получаем следующее скалярное произведение сигналов:

$$B_{su}(\tau) = \int_{-\infty}^{\infty} s(t) u(t+\tau) dt. \quad (8.2.1)$$

Взаимная корреляция сигналов характеризует определенную корреляцию явлений и физических процессов, отображаемых данными сигналами, и может служить мерой "устойчивости" данной взаимосвязи при отдельной обработке сигналов в различных устройствах. Для конечных по энергии сигналов ВКФ также конечна, при этом:

$$|B_{su}(\tau)| \leq \|s(t)\| \cdot \|u(t)\|,$$

что следует из неравенства Коши-Буняковского и независимости норм сигналов от сдвига по координатам.

При замене переменной $t = t - \tau$ в формуле (8.2.1), получаем:

$$B_{su}(\tau) = \int_{-\infty}^{\infty} s(t-\tau) u(t) dt = \int_{-\infty}^{\infty} u(t) s(t-\tau) dt = B_{us}(-\tau). \quad (8.2.1')$$

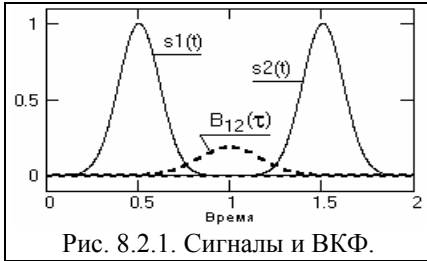


Рис. 8.2.1. Сигналы и ВКФ.

Отсюда следует, что для ВКФ не выполняется условие четности, $B_{su}(\tau) \neq B_{su}(-\tau)$, и значения ВКФ не обязаны иметь максимум при $\tau = 0$. Это можно наглядно видеть на рис. 8.2.1, где заданы два одинаковых сигнала с центрами на точках 0.5 и 1.5. Вычисление по формуле (2.4.3) с постепенным увеличением значений τ означает последовательные сдвиги сигнала $s_2(t)$ влево по оси времени (для каждого значения $s_1(t)$ для подынтегрального умножения берутся значения $s_2(t+\tau)$). При $\tau=0$ сигналы ортогональны и значение $B_{12}(\tau)=0$. Максимум $B_{12}(\tau)$ будет наблюдаться при сдвиге сигнала $s_2(t)$ влево на значение $\tau=1$, при котором происходит полное совмещение сигналов $s_1(t)$ и $s_2(t+\tau)$.

Одни и те же значения ВКФ по формулам (8.2.1) и (8.2.1') наблюдаются при одном и том же взаимном положении сигналов: при сдвиге на интервал τ сигнала $u(t)$ относительно $s(t)$ право по оси ординат и сигнала $s(t)$ относительно сигнала $u(t)$ влево, т.е. $B_{su}(\tau) = B_{us}(-\tau)$.

На рис. 8.2.2 приведены примеры ВКФ для прямоугольного сигнала $s(t)$ и двух одинаковых треугольных сигналов $u(t)$ и $v(t)$. Все сигналы имеют одинаковую длительность T , при этом сигнал $v(t)$ сдвинут вперед на интервал $T/2$.

Сигналы $s(t)$ и $u(t)$ одинаковы по временному расположению и площадь "перекрытия" сигналов максимальна при $\tau=0$, что и фиксируется функцией B_{su} . Вместе с тем функция B_{su} резко асимметрична, так как при асимметричной форме сигнала $u(t)$ при симметричной форме $s(t)$ (относительно центра сигналов) площадь "перекрытия" сигналов изменяется по-разному в зависимости от направления сдвига (знака τ при увеличении значения τ от нуля). При смещении исходного положения сигнала $u(t)$ влево по оси ординат (на опережение сигнала $s(t)$ - сигнал $v(t)$) форма ВКФ остается без изменения и сдвигается вправо на такое же значение величины сдвига – функция B_{sv} на рис. 8.2.2. Если поменять местами выражения функций в (8.2.1), то новая функция B_{vs} будет зеркально повернутой относительно $\tau=0$ функцией B_{sv} .

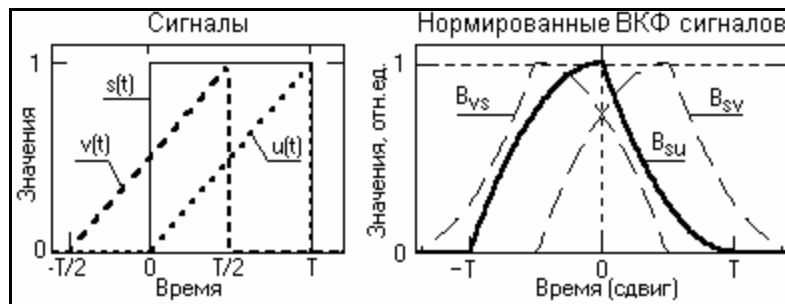


Рис. 8.2.2. Взаимнокорреляционные функции сигналов.

Для периодических сигналов понятие ВКФ обычно не применяется, за исключением сигналов с одинаковым периодом, например, сигналов входа и выхода систем при изучении характеристик систем.

Количественный показатель степени сходства сигналов $s(t)$ и $u(t)$ - функция взаимных корреляционных коэффициентов:

$$\rho_{su}(\tau) = B_{su}(\tau) / \sqrt{B_s(0)B_u(0)} = B_{su}(\tau) / (\|s(t)\| \|u(t)\|).$$

Интервал изменения значений функций при сдвигах τ_i может изменяться от -1 (полная обратная корреляция) до 1 (полное сходство или стопроцентная корреляция). При сдвигах τ_i , на которых наблюдаются нулевые значения $\rho_{su}(\tau_i)$, сигналы независимы друг от друга (некоррелированы). Величина

$$\sigma_{su} = \sqrt{1 - \rho_{su}^2(i)_{\max}}$$

представляет собой среднеквадратическую величину разности между значениями сигналов, которую обычно называют "стандартной ошибкой оценки".

8.3. СПЕКТРАЛЬНЫЕ ПЛОТНОСТИ КОРРЕЛЯЦИОННЫХ ФУНКЦИЙ [1,25].

Спектральная плотность АКФ может быть определена из следующих простых соображений.

В соответствии с выражением (8.1.1) АКФ представляет собой функцию скалярного произведения сигнала и его копии, сдвинутой на интервал τ , при $-\infty < \tau < \infty$:

$$B_s(\tau) = \langle s(t), s(t-\tau) \rangle.$$

С другой стороны, скалярное произведение может быть определено через спектральные плотности сигнала и его копии, произведение которых представляет собой спектральную плотность взаимной мощности:

$$\langle s(t), s(t-\tau) \rangle = (1/2\pi) \int_{-\infty}^{\infty} S(\omega) S_{\tau}^*(\omega) d\omega.$$

Смещение сигнала по оси абсцисс на интервал τ отображается в спектральном представлении умножением спектра сигнала на $\exp(-j\omega\tau)$, а для сопряженного спектра на множитель $\exp(j\omega\tau)$:

$$S_{\tau}^*(\omega) = S^*(\omega) \cdot \exp(j\omega\tau).$$

С учетом этого получаем:

$$\begin{aligned} B_s(\tau) &= (1/2\pi) \int_{-\infty}^{\infty} S(\omega) S^*(\omega) \cdot \exp(j\omega\tau) d\omega = \\ &= (1/2\pi) \int_{-\infty}^{\infty} |S(\omega)|^2 \exp(j\omega\tau) d\omega. \end{aligned} \quad (8.3.1)$$

Но последнее выражение представляет собой обратное преобразование Фурье энергетического спектра сигнала (спектральной плотности энергии). Следовательно, энергетический спектр сигнала и его автокорреляционная функция связаны преобразованием Фурье:

$$B_s(\tau) \Leftrightarrow |S(\omega)|^2 = W_s(\omega). \quad (8.3.2)$$

Аналогичный результат может быть получен и прямым преобразованием Фурье автокорреляционной функции:

$$\begin{aligned} \int_{-\infty}^{\infty} B_s(\tau) \exp(-j\omega\tau) d\tau &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(t) s(t-\tau) \exp(-j\omega\tau) dt d\tau = \\ &= \int_{-\infty}^{\infty} s(t) \exp(-j\omega t) \int_{-\infty}^{\infty} s(t-\tau) \exp(j\omega(t+\tau)) d(t+\tau) dt = S(\omega) S^*(\omega) = W_s(\omega). \end{aligned}$$

Таким образом, спектральная плотность АКФ есть не что иное, как спектральная плотность мощности сигнала, которая, в свою очередь, может определяться прямым преобразованием Фурье через АКФ:

$$|S(\omega)|^2 = \int_{-\infty}^{\infty} B_s(\tau) \cdot \exp(-j\omega\tau) d\tau. \quad (8.3.3)$$

Последнее выражение накладывает определенные ограничения на форму АКФ и методику их ограничения по длительности.

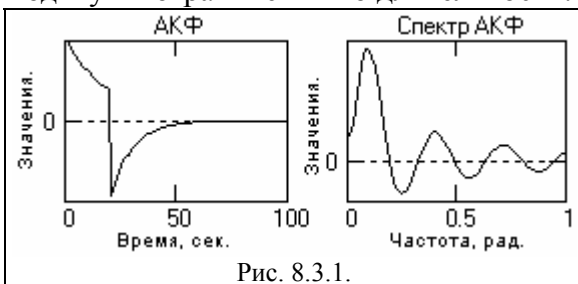


Рис. 8.3.1.

Энергетический спектр сигналов всегда положителен, мощность сигналов не может быть отрицательной. Следовательно, АКФ не может иметь формы прямоугольного импульса, т.к. преобразование Фурье прямоугольного импульса – знакопеременный интегральный синус. На АКФ не должно быть и разрывов первого рода (скачков), т.к. с учетом четности АКФ любой симметричный скачок по координате $\pm\tau$ порождает “разделение” АКФ на сумму определенной не-

прерывной функции и прямоугольного импульса длительностью 2τ с соответствующим появлением отрицательных значений в энергетическом спектре. Пример последнего приведен на рис. 8.3.1 (графики функций приведены, как принято для четных функций, только своей правой частью).

АКФ достаточно протяженных сигналов обычно ограничиваются по размерам (исследуются ограниченные интервалы корреляции данных от $-T/2$ до $T/2$). Однако усечение АКФ, это умножение АКФ на прямоугольный селектирующий импульс длительностью T , что в частотной области отображается сверткой фактического спектра мощности со знакопеременной функцией интегрального синуса $\text{sinc}(\omega T/2)$. С одной стороны, это вызывает определенное сглаживание спектра мощности, что зачастую бывает полезным, например, при исследовании сигналов на значительном уровне шумов. Но, с другой стороны, может происходить и существенное занижение величины энергетических пиков, если в сигнале имеются какие-либо гармонические составляющие, и появление отрицательных значений мощности на краевых частях пиков и скачков. Пример проявления данных факторов приведен на рис. 8.3.2.

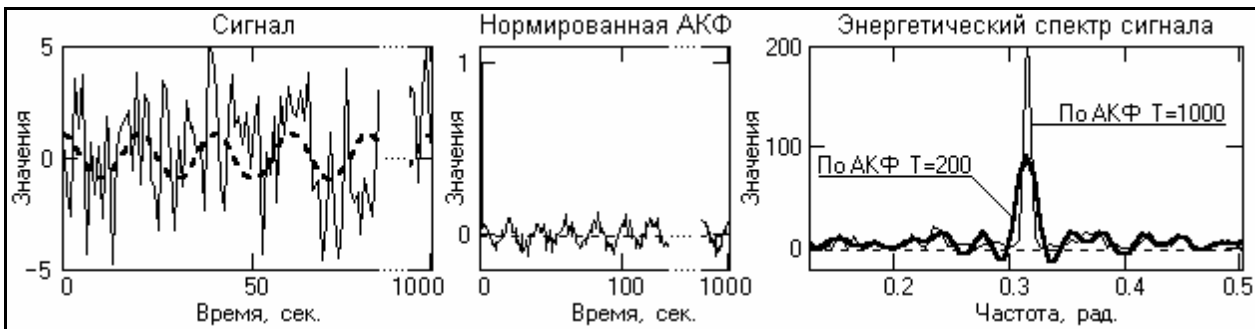


Рис. 8.3.2. Вычисление энергетического спектра сигнала по АКФ разной длины.

Как известно, спектры мощности сигналов не имеют фазовой характеристики и по ним невозможно восстановление сигналов. Следовательно, АКФ сигналов, как временное представление спектров мощности, также не имеет информации о фазовых характеристиках сигналов и восстановление сигналов по АКФ невозможно. Сигналы одной формы и сдвинутые во времени имеют одинаковые АКФ. Больше того, сигналы разной формы могут иметь сходные АКФ, если имеют близкие спектры мощности.

Перепишем уравнение (8.3.1) в следующей форме

$$\int_{-\infty}^{\infty} s(t) s(t-\tau) dt = (1/2\pi) \int_{-\infty}^{\infty} S(\omega) S^*(\omega) \cdot \exp(j\omega\tau) d\omega,$$

и подставим в это выражение значение $\tau=0$. Полученное равенство хорошо известно и называется *равенством Парсеваля*

$$\int_{-\infty}^{\infty} s^2(t) dt = (1/2\pi) \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega.$$

Оно позволяет вычислять энергию сигнала, как по временной, так и по частотной области.

Интервал корреляции сигнала является числовым параметром оценки ширины АКФ и степени значимой корреляции значений сигнала по аргументу.

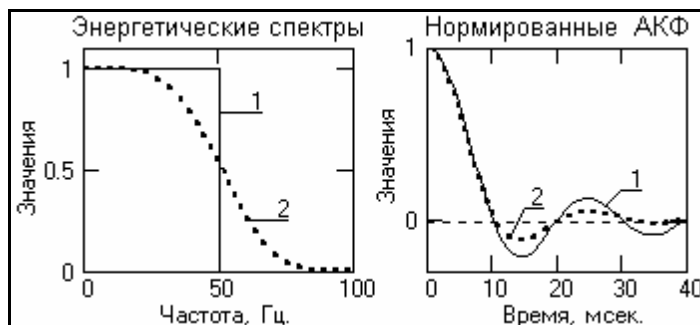


Рис. 8.3.3.

Если допустить, что сигнал $s(t)$ имеет примерно равномерный энергетический спектр с верхней граничной частотой до ω_B (форма центрированного прямоугольного импульса, как, например, сигнал 1 на рис. 8.3.3 с $f_B=50$ Гц в одностороннем представлении), то АКФ сигнала определится выражением:

$$B_s(\tau) = (W_0/\pi) \int_0^{\omega_B} \cos(\omega\tau) d\omega = (W_0\omega_B/\pi) \sin(\omega_B\tau)/(\omega_B\tau).$$

Интервалом корреляции сигнала τ_K считается величина ширины центрального пика АКФ от максимума до первого пересечения нулевой линии. В данном случае для прямоугольного спектра первое пересечение нуля соответствует $\sin(\omega_B\tau) = 0$ при $\omega_B\tau = \pi$, откуда:

$$\tau_K = \pi/\omega_B = 1/2f_B. \quad (8.3.4)$$

Интервал корреляции тем меньше, чем выше верхняя граничная частота спектра сигнала. Для сигналов с плавным срезом по верхней граничной частоте роль параметра ω_B играет средняя ширина спектра (сигнал 2 на рис. 8.3.3).

Спектральная плотность мощности статистических шумов при единичном измерении представляет собой случайную функцию $W_q(\omega)$ со средним значением $W_q(\omega) \Rightarrow \sigma_q^2$, где σ_q^2 – дисперсия шумов. В пределе, при равномерном спектральном распределении шумов от 0 до ∞ , АКФ шумов стремится к значению $B_q(\tau) \Rightarrow \sigma_q^2$ при $\tau \Rightarrow 0$, $B_q(\tau) \Rightarrow 0$ при $\tau \neq 0$, т.е. статистические шумы не коррелированы ($\tau_K \Rightarrow 0$).

Спектральная плотность ВКФ может быть получена на основании тех же соображений, что и для АФК, или непосредственно из формулы (8.3.1) заменой спектральной плотности сигнала $S^*(\omega)$ на спектральную плотность второго сигнала $U^*(\omega)$:

$$B_{su}(\tau) = (1/2\pi) \int_{-\infty}^{\infty} S(\omega)U^*(\omega) \cdot \exp(j\omega\tau) d\omega. \quad (8.3.5)$$

Или, при смене порядка сигналов:

$$B_{us}(\tau) = (1/2\pi) \int_{-\infty}^{\infty} U(\omega)S^*(\omega) \cdot \exp(j\omega\tau) d\omega. \quad (8.3.5')$$

Произведение $S(\omega)U^*(\omega)$ представляет собой взаимный энергетический спектр $W_{su}(\omega)$ сигналов $s(t)$ и $u(t)$. Соответственно, $U(\omega)S^*(\omega) = W_{us}(\omega)$. Следовательно, как и АКФ, взаимнокорреляционная функция и спектральная плотность взаимной мощности сигналов связаны между собой преобразованиями Фурье:

$$B_{su}(\tau) \Leftrightarrow W_{su}(\omega) \equiv W_{us}^*(\omega). \quad (8.3.6)$$

$$B_{us}(\tau) \Leftrightarrow W_{us}(\omega) \equiv W_{su}^*(\omega). \quad (8.3.6')$$

В общем случае, за исключением спектров четных функций, из условия несоблюдения четности для функций ВКФ следует, что взаимные энергетические спектры являются комплексными функциями:

$$U(\omega) = A_u(\omega) + j B_u(\omega), \quad V(\omega) = A_v(\omega) + j B_v(\omega).$$

$$W_{uv} = A_u A_v + B_u B_v + j (B_u A_v - A_u B_v) = \text{Re } W_{uv}(\omega) + j \text{Im } W_{uv}(\omega),$$

и содержат определенную фазовую характеристику гармонических составляющих ВКФ, которой и формируется сдвиг максимума ВКФ.

На рис. 8.3.4 можно наглядно видеть особенности формирования ВКФ на примере двух одинаковых по форме сигналов, сдвинутых относительно друг друга.

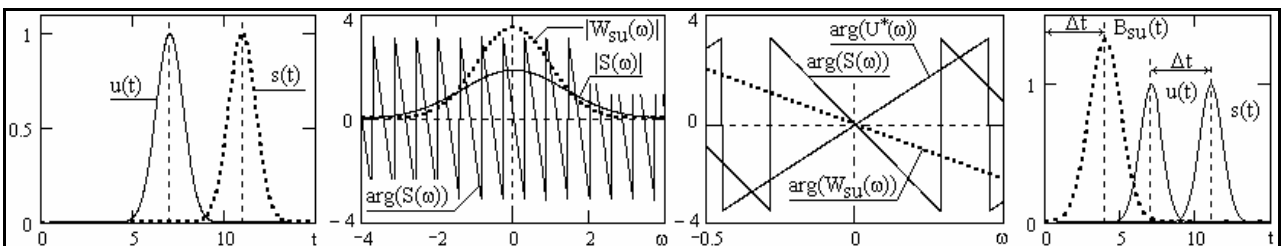


Рис. 8.3.4. Формирование ВКФ.

Форма сигналов и их взаимное расположение приведены на виде А. Модуль и аргумент спектра сигнала $s(t)$ приведены на виде В. Модуль спектра $u(t)$ тождественен модулю $S(\omega)$. На этом же виде приведен модуль спектра взаимной мощности сигналов $W_{su}(\omega) = S(\omega) \cdot U^*(\omega)$. Как известно, при перемножении комплексных спектров модули спектров перемножаются, а фазовые углы складываются, при этом для сопряженного спектра $U^*(\omega)$ фазовый угол меняет знак. Если первым в формуле вычисления ВКФ (8.2.1) стоит сигнал $s(t)$, а сигнал $u(t-\tau)$ на оси ординат стоит впереди $s(t)$, то фазовые углы $S(\omega)$ по мере увеличения частоты нарастают в сторону отрицательных значений углов (без учета периодического сброса значений на 2π), а фазовые углы $U^*(\omega)$ по абсолютным значениям меньше фазовых углов $s(t)$ и нарастают (за счет сопряжения) в сторону положительных значений. Результатом умножения спектров (как это видно на рис. 8.3.4, вид С) является вычитание из фазовых углов $S(\omega)$ значений углов $U^*(\omega)$, при этом фазовые углы спектра $W_{su}(\omega)$ остаются в области отрицательных значений, что обеспечивает сдвиг всей функции ВКФ (и ее пиковых значений) вправо от нуля по оси τ на определенную величину (для одинаковых сигналов – на величину разности между сигналами по оси ординат). При смещении начального положения сигнала $u(t)$ в сторону сигнала $s(t)$ фазовые углы $W_{su}(\omega)$ уменьшаются, в пределе до нулевых значений при полном совмещении сигналов, при функции $B_{su}(t)$ смещается к нулевым значениям τ , в пределе до обращения в АКФ (при одинаковых сигналах $s(t)$ и $u(t)$).

При анализе дискретных данных и числовых рядов соответственно используется функция взаимной ковариации (ФВК):

$$C_{su}(n) = \frac{N \cdot \Delta t}{N+1-n} \sum_{k=0}^{N-n} (s_k - \bar{s}) \cdot (u_{k+n} - \bar{u}), \quad (8.3.7)$$

Как известно для детерминированных сигналов, если спектры двух сигналов не перекрываются и, соответственно, взаимная энергия сигналов равна нулю, такие сигналы ортогональны друг другу. Связь энергетических спектров и корреляционных функций сигналов показывает еще одну сторону взаимодействия сигналов. Если спектры сигналов не перекрываются и их взаимный энергетический спектр равен нулю на всех частотах, то при любых временных сдвигах τ друг относительно друга их ВКФ также равна нулю. А это означает, что такие сигналы являются некоррелированными. Это действительно как для детерминированных, так и для случайных сигналов и процессов.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
19. Отнес Р., Эноксон Л. Прикладной анализ временных рядов. – М.: Мир, 1982. – 428 с.
25. Сергиенко А.Б. Цифровая обработка сигналов. / Учебник для вузов. – СПб.: Питер, 203. – 608 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 9: МОДУЛИРОВАННЫЕ СИГНАЛЫ

Мир создан ради богов и людей.

Хрисипп. Греческий философ, стоик, III в.д.н.э.

Все создается с какой-то целью. Придется разобраться, для чего же созданы модулированные сигналы. И почему на них не обратили внимание ни древние греки, ни Михайло Васильевич Ломоносов.

Евгений Прокопчук. Иркутский геофизик Уральской школы, казак, XX в.

Содержание: Введение. 9.1. Амплитудная модуляция. Однотональная модуляция. Энергия однотонального АМ-сигнала. Многотональный модулирующий сигнал. Демодуляция АМ-сигналов. Балансная амплитудная модуляция. Однополосная амплитудная модуляция. Полярная модуляция. 9.2. Сигналы с угловой модуляцией. Фазовая модуляция (ФМ). Частотная модуляция (ЧМ). Однотональная угловая модуляция. Спектры сигналов с угловой модуляцией. Сигналы с многотональной угловой модуляцией. Демодуляция УМ – сигналов. Квадратурная модуляция. Пример моделирования квадратурной модуляции в системе Mathcad. Демодуляция квадратурного сигнала. 9.3. Внутриимпульсная частотная модуляция. ЛЧМ-сигналы. Спектр прямоугольного ЛЧМ-сигнала. 9.4. Импульсно-модулированные сигналы. Амплитудно-импульсная модуляция. Широко-импульсная модуляция. Временная импульсная модуляция. Кодово-импульсная модуляция. 9.5. Модуляция символьных и кодовых данных. Амплитудно-манипулированные сигналы. Угловая манипуляция. Литература.

ВВЕДЕНИЕ

Сигналы от измерительных датчиков и любых других источников информации передаются по линиям связи к приемникам - измерительным приборам, в измерительно-вычислительные системы регистрации и обработки данных, в любые другие центры накопления и хранения данных. Как правило, информационные сигналы являются низкочастотными и ограниченными по ширине спектра, в отличие от широкополосных высокочастотных каналов связи, рассчитанных на передачу сигналов от множества источников одновременно с частотным разделением каналов. Перенос спектра сигналов из низкочастотной области в выделенную для их передачи область высоких частот выполняется операцией *модуляции*.

Допустим, что низкочастотный сигнал, подлежащий передаче по какому-либо каналу связи, задается функцией $s(t)$. В канале связи для передачи данного сигнала выделяется определенный диапазон высоких частот. На входе канала связи в специальном передающем устройстве формируется вспомогательный, как правило, непрерывный во времени периодический высокочастотный сигнал $u(t) = f(t; a_1, a_2, \dots, a_m)$. Совокупность параметров a_i определяет форму вспомогательного сигнала. Значения параметров a_i в отсутствие модуляции являются величинами постоянными. Если на один из этих параметров перенести сигнал $s(t)$, т.е. сделать его значение пропорционально зависимым от значения $s(t)$ во времени (или по любой другой независимой переменной), то форма сигнала $u(t)$ приобретает новое свойство. Она несет информацию, тождественную информации в сигнале $s(t)$. Именно поэтому сигнал $u(t)$ называют несущим сигналом, *несущим колебанием* или просто *несущей* (carrier), а физический процесс переноса информации на параметры несущего сигнала – его *модуляцией* (modulation). Исходный информационный сигнал $s(t)$ называют *модулирующим* (modulating signal), результат модуляции – *модулированным сигналом* (modulated signal). Обратную операцию выделения модулирующего сигнала из модулированного колебания называют *демодуляцией* (demodulation).

Основным видом несущих сигналов являются гармонические колебания:

$$u(t) = U \cdot \cos(\omega t + \varphi),$$

которые имеют три свободных параметра: U , ω и φ . В зависимости от того, на какой из данных параметров переносится информация, различают *амплитудную* (АМ), *частотную* (ЧМ) или *фазовую* (ФМ) *модуляцию* несущего сигнала. Частотная и фазовая модуляция тесно взаимосвязаны, поскольку изменяют аргумент функции косинуса, и их обычно объединяют под общим названием - *угловая модуляция* (angle modulation). В каналах передачи цифровой ин-

формации получила также распространение *квадратурная* модуляция, при которой одновременно изменяются амплитуда и фаза несущих колебаний.

При использовании в качестве несущих сигналов периодических последовательностей импульсов (например, прямоугольных) свободными параметрами модуляции могут быть амплитуда, длительность, частота следования и фаза (положение импульса относительно тактовой точки) импульсов. Это дает четыре основных вида импульсной модуляции: АИМ, ДИМ, ЧИМ и ФИМ.

В качестве несущих сигналов можно использовать не только периодические колебания, но и стационарные случайные процессы. В качестве модулируемых параметров случайных сигналов используются моменты случайных процессов. Так, например, модуляция второго момента случайных последовательностей (модуляция по мощности) представляет собой аналогию амплитудной модуляции.

9.1. АМПЛИТУДНАЯ МОДУЛЯЦИЯ [1,25].

Амплитудная модуляция (amplitude modulation, АМ) исторически была первым видом модуляции, освоенным на практике. В настоящее время АМ применяется в основном только для радиовещания на сравнительно низких частотах (не выше коротких волн) и для передачи изображения в телевизионном вещании. Это обусловлено низким КПД использования энергии модулированных сигналов.

АМ соответствует переносу информации $s(t) \Rightarrow U(t)$ при постоянных значениях параметров несущей частоты ω и φ . АМ – сигнал представляет собой произведение информационной огибающей $U(t)$ и гармонического колебания ее заполнения с более высокими частотами. Форма записи амплитудно-модулированного сигнала:

$$u(t) = U(t) \cdot \cos(\omega_0 t + \varphi_0), \quad (9.1.1)$$

$$U(t) = U_m \cdot [1 + M \cdot s(t)], \quad (9.1.2)$$

где U_m – постоянная амплитуда несущего колебания при отсутствии входного (модулирующего) сигнала $s(t)$, M – коэффициент амплитудной модуляции

Значение M характеризует *глубину* амплитудной модуляции. В простейшем случае, если модулирующий сигнал представлен одночастотным гармоническим колебанием с амплитудой S_0 , то коэффициент модуляции равен отношению амплитуд модулирующего и несущего колебания $M = S_0 / U_m$. Значение M должно находиться в пределах от 0 до 1 для всех гармоник модулирующего сигнала. При значении $M < 1$ форма огибающей несущего колебания полностью повторяет форму модулирующего сигнала $s(t)$, что можно видеть на рис. 9.1.1 (сигнал $s(t) = \sin(\omega_s t)$). Малую глубину модуляции для основных гармоник модулирующего сигнала ($M \ll 1$) применять нецелесообразно, т.к. при этом мощность передаваемого информационного сигнала будет много меньше мощности несущего колебания, и мощность передатчика используется неэкономично.

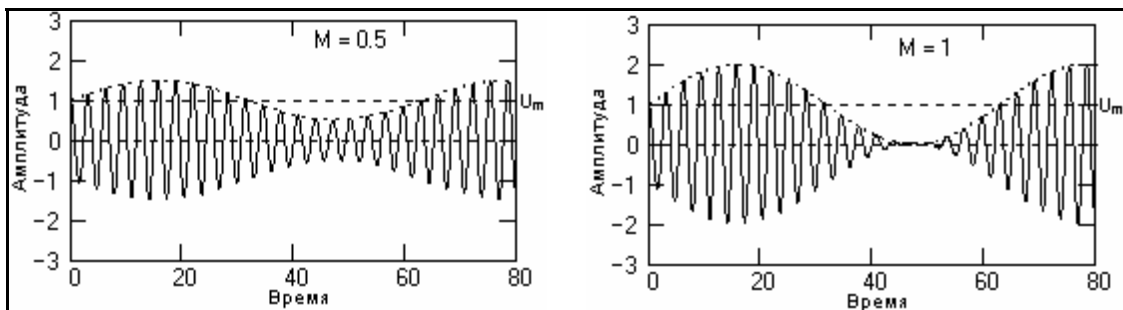


Рис. 9.1.1. Модулированный сигнал.

Рис. 9.1.2. Глубокая модуляция

На рис. 9.1.2 приведен пример так называемой *глубокой модуляции*, при которой значение M стремится к 1 в экстремальных точках функции $s(t)$. При глубокой модуляции используются также понятия относительного коэффициента модуляции вверх: $M_v = (U_{\max} - U_m) / U_m$, и модуляции вниз: $M_n = (U_m - U_{\min}) / U_m$, которые обычно выражаются в %.

Стопроцентная модуляция ($M=1$) может приводить к искажениям сигналов при пере-

грузках передатчика, если последний имеет ограниченный динамический диапазон по амплитуде несущих частот или ограниченную мощность передатчика (увеличение амплитуды несущих колебаний в пиковых интервалах сигнала $U(t)$ в два раза требует увеличения мощности передатчика в четыре раза).

При $M > 1$ возникает так называемая *перемодуляция*, пример которой приведен на рис. 9.1.3. Форма огибающей при перемодуляции искажается относительно формы модулирующего сигнала и после демодуляции, если применяются ее простейшие методы, информация может исказиться.

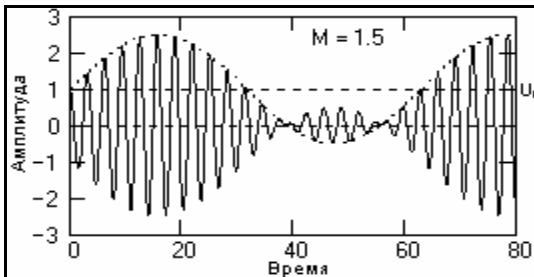


Рис. 9.1.3. Перемодуляция сигнала.

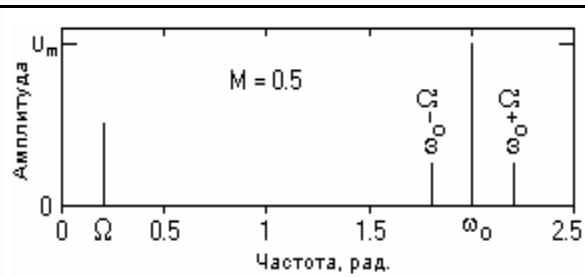


Рис. 9.1.4. Физические спектры сигналов.

Однотональная модуляция. Простейшая форма модулированного сигнала создается при *однотональной* амплитудной модуляции – модуляции несущего сигнала гармоническим колебанием с одной частотой Ω :

$$u(t) = U_m [1 + M \cdot \cos(\Omega t)] \cdot \cos(\omega_0 t). \quad (9.1.3)$$

Значения начальных фазовых углов несущего и модулирующего колебания здесь и в дальнейшем, если это не имеет принципиального значения, для упрощения получаемых выражений будем принимать равными нулю. С учетом формулы $\cos(x) \cdot \cos(y) = (1/2)[\cos(x+y) + \cos(x-y)]$ из выражения (9.1.3) получаем:

$$u(t) = U_m \cos(\omega_0 t) + (U_m M / 2) \cos[(\omega_0 + \Omega)t] + (U_m M / 2) \cos[(\omega_0 - \Omega)t]. \quad (9.1.4)$$

Отсюда следует, что модулирующее колебание с частотой Ω перемещается в область частоты ω_0 и расщепляется на два колебания, симметричные относительно частоты ω_0 , с частотами соответственно $(\omega_0 + \Omega)$ – верхняя боковая частота, и $(\omega_0 - \Omega)$ – нижняя боковая частота (рис. 9.1.4 для сигнала, приведенного на рис. 9.1.1). Амплитуды колебаний на боковых частотах равны друг другу, и при 100%-ной модуляции равны половине амплитуды колебаний несущей частоты. Если получить уравнение (9.1.4) с учетом начальных фаз несущей и модулирующей частоты, то правило изменения фаз аналогично изменению частоты: начальная фаза модулирующего колебания для верхней боковой частоты складывается с начальной фазой несущей, для нижней – вычитаются из фазы несущей. Физическая ширина спектра модулированного сигнала в два раза больше ширины спектра сигнала модуляции.

Энергия однотонального АМ-сигнала. Обозначим отдельными индексами (нес- несущая, вб- верхняя боковая, нб- нижняя боковая) составляющие колебания однотонального АМ-сигнала (9.1.4) и определим функцию его мгновенной мощности:

$$p(t) = u^2(t) = u_{\text{нес}}^2(t) + u_{\text{вб}}^2(t) + u_{\text{нб}}^2(t) + 2u_{\text{нес}}(t)u_{\text{вб}}(t) + 2u_{\text{нес}}(t)u_{\text{нб}}(t) + 2u_{\text{вб}}(t)u_{\text{нб}}(t). \quad (9.1.5)$$

Для определения средней мощности сигнала выполним усреднение функции $p(t)$:

$$P_u = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T p(t) dt$$

Все взаимные мощности модулированного сигнала при усреднении становятся равными нулю (спектры не перекрываются), при этом:

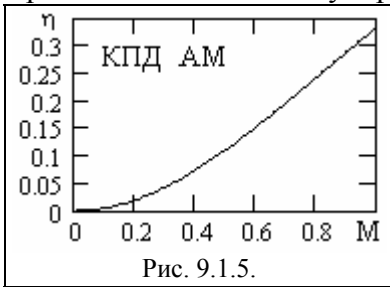
$$P_u = P_{\text{нес}} + P_{\text{вб}} + P_{\text{нб}} = U_m^2 / 2 + (U_m M)^2 / 4. \quad (9.1.6)$$

Доля мощности боковых частот в единицах мощности несущей частоты:

$$(P_{\text{вб}} + P_{\text{нб}}) / P_{\text{нес}} = M^2 / 2, \quad (9.1.7)$$

т.е. не превышает 50% даже при 100%-ной модуляции.

Под полезной мощностью модулированных сигналов понимают мощность, несущую информацию, т.е. в данном случае мощность боковых частот. Коэффициент полезного действия данного типа модуляции определяется отношением мощности боковых частот к общей средней мощности модулированного сигнала:



$$\eta_{AM} = (U_m^2 M^2/4) / P_u = M^2/(M^2+2). \quad (9.1.8)$$

Как можно видеть на рис. 9.1.5, даже при $M=1$ КПД амплитудной модуляции составляет только 33%, а при практическом использовании обычно меньше 20%.

Для модулированных сигналов применяют также понятие пиковой мощности P_{max} . Значение пиковой мощности для однотонового АМ-сигнала:

$$P_{max} = U_m^2 (1+M)^2.$$

Многотональный модулирующий сигнал имеет произвольный спектральный состав. Математическая модель такого сигнала, в том числе непрерывного по частоте, может быть аппроксимирована тригонометрической суммой, в пределе бесконечной:

$$s(t) = \sum_{n=1}^N a_n \cos(\Omega_n t + \Phi_n), \quad (9.1.9)$$

где значения амплитуд a_n и начальных фаз Φ_n упорядоченной возрастающей последовательности гармоник Ω_n произвольны. Подставляя (9.1.9) в (9.1.2) и заменяя произведения $M \cdot a_n$ парциальными (частичными) коэффициентами модуляции $M_n = M \cdot a_n$, получим обобщенное уравнение амплитудно-модулированного сигнала и его физического спектра:

$$u(t) = U_m \left[1 + \sum_{n=1}^N M_n \cos(\Omega_n t + \Phi_n) \right] \cdot \cos(\omega_0 t + \varphi_0). \quad (9.1.10)$$

$$u(t) = U_m \cos(\omega_0 t + \varphi_0) + (U_m/2) \sum_{n=1}^N M_n \cos[(\omega_0 + \Omega_n)t + \varphi_0 + \Phi_n] + (U_m/2) \sum_{n=1}^N M_n \cos[(\omega_0 - \Omega_n)t + \varphi_0 - \Phi_n]. \quad (9.1.11)$$



Рис. 9.1.6. Многотональная модуляция.

На рис. 9.1.6 приведен схематический пример амплитудных спектров модулирующего и АМ-сигналов при многотональной модуляции. Он также содержит полосы верхних и нижних боковых частот относительно несущей частоты ω_0 , являющихся прямой и зеркальной масштабными копиями модулирующего сигнала. Соответственно, полная ширина спектра АМ-сигнала равна удвоенной ширине спектра модулирующего сигнала.

Пример.

Частотный диапазон каротажного кабеля длиной 1 км 0-200 кГц. Частотный диапазон измерительных датчиков скважинного прибора 5-5000 Гц. От какого количества датчиков одновременно может передаваться информация по данному каротажному кабелю?

Минимальная несущая частота должна быть на порядок выше максимальной частоты модулирующего сигнала, т.е. порядка 50 кГц. Для передачи сигнала от одного датчика требуется полоса частот $2 \cdot 5 = 10$ кГц плюс пустой защитный интервал для исключения перекрестных помех порядка 1 кГц, т.е. 11 кГц. Общее количество каналов передачи информации: $(200-50+5)/11 = 14$ каналов.

Автокорреляционная функция АМ-сигналов:

$$B_u(\tau) = \int_{-\infty}^{\infty} u(t) u(t-\tau) dt. \quad (9.1.12)$$

С учетом того, что постоянная фаза не влияет на форму АКФ, при $u(t)=U(t)\cdot\cos\omega_0t$ получаем:

$$\cos\omega_0t\cdot\cos\omega_0(t-\tau)=0.5\cos\omega_0\tau+0.5\cos\omega_0(2t-\tau).$$

$$B_u(\tau)=\frac{\cos\omega_0\tau}{2}\int_{-\infty}^{\infty}U(t)U(t-\tau)dt+0.5\int_{-\infty}^{\infty}U(t)U(t-\tau)\cos\omega_0(2t-\tau)dt. \quad (9.1.13)$$

Второй интеграл в формуле АКФ существенно меньше первого - произведение медленно меняющейся функции $U(t)U(t-\tau)$ и сильно осциллирующего члена с частотой $2\omega_0$, и им можно пренебречь. Первый интеграл представляет собой АКФ сигнала $U(t)$. Отсюда:

$$B_u(\tau)\cong\frac{\cos\omega_0\tau}{2}B_U(\tau). \quad (9.1.14)$$

Полная энергия сигнала за счет усреднения по высокочастотным колебаниям:

$$B_u(0)\cong(1/2)B_U(0).$$

При бесконечной энергии сигнала:

$$B_u(\tau)=\lim_{T\rightarrow\infty}\frac{1}{T}\int_0^T u(t)u(t-\tau)dt. \quad (9.1.15)$$

Демодуляция АМ-сигналов может выполняться несколькими способами.

Самый простейший способ – двухполупериодное детектирование (вычисление модуля сигнала) с последующим сглаживанием полученных однополярных полупериодов несущей фильтром низких частот.

На рис. 9.1.7 приведен пример изменения однотонального амплитудно-модулированного сигнала и его физического спектра при детектировании (в реальной одно-сторонней шкале частот и в реальной шкале амплитудных значений гармоник колебаний). Параметры представленного сигнала: несущая частота 30 Гц, частота модуляции 3 Гц, коэффициент модуляции $M=1$.

Как видно на рисунке, при детектировании спектр модулированного сигнала становится однополярным, переходит на основную несущую частоту $2\omega_0$ и уменьшается по энергии почти в 5 раз. Основная часть энергии (более 4/5) трансформируется в область низких частот и распределяется между постоянной составляющей и выделенной гармоникой сигнала модуляции. Между постоянной составляющей и выделенной гармоникой энергия распределяется в зависимости от значения коэффициента модуляции M . При $M=1$ энергии равны, при $M=0$ (в отсутствие сигнала модуляции) вся энергия переходит на постоянную составляющую.

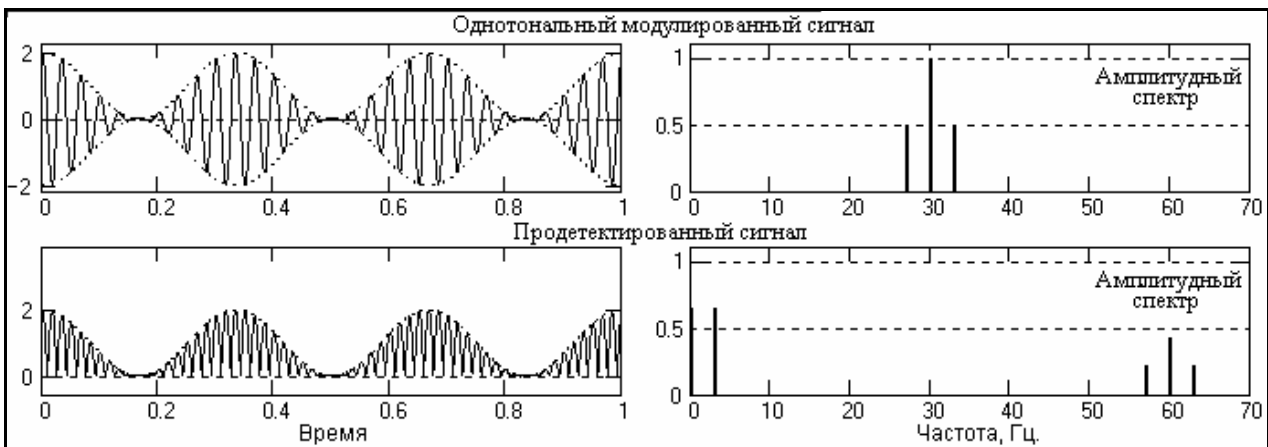


Рис. 9.1.7. Изменение однотонального модулированного сигнала при детектировании

Кроме этих составляющих в спектре появляются также 2-я, 3-я и более высокие гармоники детектированного модулированного сигнала (т.е. в данном случае на частотах {117, 120, 123} Гц, {177, 180, 183} Гц и т.д.), которые не показаны на рисунке. Энергия второй гармоники не превышает 2%, а остальных и вовсе незначительна. Демодуляторы сигнала выделяют после детектирования только низкочастотный информационный сигнал и подавляют все остальные частоты, включая постоянную составляющую.

Очевидно также, что в случае перемодуляции сигнала исходный информационный сигнал будет восстанавливаться с ошибкой.

Другой распространенный метод – *синхронное детектирование*. При синхронном детектировании модулированный сигнал умножается на опорное колебание с частотой несущего колебания. Без учета фазовых углов колебаний:

$$y(t) = u(t) \cos(\omega_0 t) = U(t) \cos(\omega_0 t) \cos(\omega_0 t) = \frac{1}{2} U(t) + \frac{1}{2} U(t) \cos(2\omega_0 t). \quad (9.1.16)$$

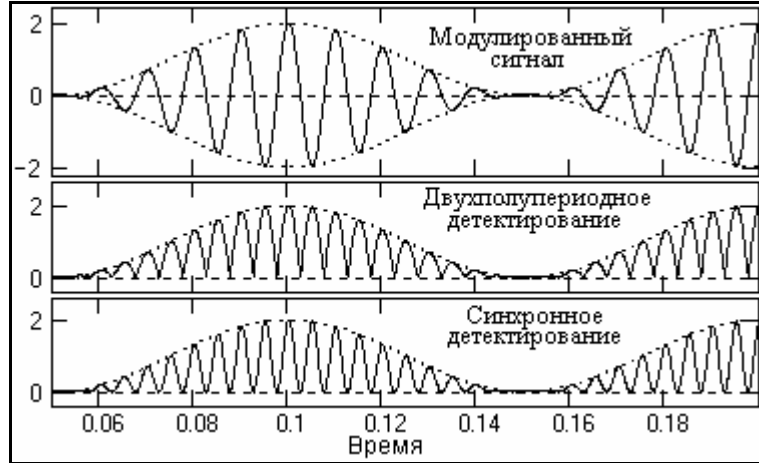


Рис. 9.1.8.

Как следует из этого выражения, сигнал разделяется на два слагаемых, первое из которых повторяет исходный модулирующий сигнал, а второе повторяет модулированный сигнал на удвоенной несущей частоте $2\omega_0$. На рис. 9.1.8 приведено визуальное сопоставление двухполупериодного и синхронного детектирования, которое наглядно показывает практически полное подобие процессов. Но форма новой несущей при синхронном детектировании является чистой гармоникой, в отличие от двухполупериодного детектирования, где новая несущая явно содержит дополнительные гармоники более высоких частот. Физический амплитудный спектр сигналов после демодуляции подобен спектру двухполупериодного детектирования, но однозначно соотносится со спектром входного модулированного сигнала: амплитуды гармоник модулированного сигнала на частоте $2\omega_0$ в два раза меньше амплитуд входного сигнала, постоянная составляющая равна амплитуде несущей частоты ω_0 и не зависит от глубины модуляции, амплитуда информационного демодулированного сигнала в 2 раза меньше амплитуды исходного модулирующего сигнала. Замечательной особенностью синхронного детектирования является полная независимость от глубины модуляции, т.е. коэффициент модуляции сигнала может быть больше 1.

Однако при синхронном детектировании требуется точное совпадение фаз и частот опорного колебания демодулятора и несущей гармоники АМ-сигнала.

При сдвиге фазы опорного колебания на $\Delta\omega$ относительно несущей частоты выходной сигнал демодулятора оказывается умноженным на косинус фазовой ошибки:

$$y(t) = U(t) \cos(\omega_0 t) \cos(\omega_0 t - \Delta\varphi) = \frac{1}{2} U(t) \cos(-\Delta\varphi) + \frac{1}{2} U(t) \cos(2\omega_0 t - \Delta\varphi).$$

и амплитуда сигнала занижается, а при $\Delta\omega = \pi/2$ становится равной нулю.

При сдвиге частоты между несущим и опорным колебаниями сигнал демодулятора оказывается умноженным на гармоническое колебание с разностной частотой:

$$y(t) = U(t) \cos(\omega_0 t) \cos(\omega_0 t - \Delta\omega) = \frac{1}{2} U(t) \cos(-\Delta\omega t) + \frac{1}{2} U(t) \cos((2\omega_0 - \Delta\omega)t),$$

при этом выходной сигнал демодулятора начинает пульсировать (beat – биения) с частотой биений $\Delta\omega$.

Для частотной и фазовой синхронизации между несущим и опорным колебаниями в составе демодуляторов обычно используются следящие системы фазовой автоподстройки опорной частоты.



Рис. 9.1.9. Балансная модуляция.

Балансная амплитудная модуляция или АМ с подавлением несущей частоты (АМ-ПН). Как следует из вышеприведенных данных, основная доля мощности АМ – сигнала приходится на несущую частоту. При балансной модуляции производится перемножение двух сигналов – модулирующего и несущего, при котором происходит подавление несущего колебания, соответственно, КПД модуляции становится равным 100%. Так, для однотонального

сигнала (без учета начальных фаз колебаний) при $U(t) = M \cdot \cos(\Omega t)$ имеем:

$$u(t) = U_m M \cdot \cos(\Omega t) \cdot \cos(\omega_0 t) = (U_m M / 2) \{ \cos[(\omega_0 + \Omega)t] + \cos[(\omega_0 - \Omega)t] \}, \quad (9.1.17)$$

т.е. биения двух одинаковых по амплитуде гармонических сигналов с верхней и нижней боковыми частотами. По существу, однотональный модулирующий сигнал переносится на биения двух высоких частот. Пример сигнала с балансной модуляцией приведен на рис. 9.1.9. Амплитудный спектр сигнала подобен приведенному на рис. 9.1.4 с отсутствующей несущей частотой ω_0 . Аналогично, многотональный балансно - модулированный сигнал имеет две симметричные относительно частоты ω_0 группы верхних и нижних боковых колебаний:

$$u(t) = (U_m / 2) \left\{ \sum_{n=1}^N M_n \cos[(\omega_0 + \Omega_n)t + \Phi_n] + \sum_{n=1}^N M_n \cos[(\omega_0 - \Omega_n)t - \Phi_n] \right\}. \quad (9.1.18)$$

Физическая сущность подавления несущей частоты заключается в следующем. При переходе огибающей биений $U(t)$ через нуль фаза несущей частоты высокочастотного заполнения скачком изменяется на 180° , поскольку функция косинуса огибающей имеет разные знаки слева и справа от нуля. При этом в достаточно высокочастотной системе (с малыми потерями энергии), настроенной на частоту ω_0 , колебания, возбужденные одним периодом биений, будут гаситься последующим периодом.

Однако балансная модуляция не получила широкого распространения в связи с трудностями, возникающими при демодуляции сигналов. В принципе, синхронное детектирование, как и для АМ, позволяет выполнять демодуляцию без каких-либо проблем, но при условии известной частоты несущей сигнала и точной фазовой синхронизации опорной частоты с несущей. Но во входном сигнале демодулятора при АМ-ПН несущая частота отсутствует, и автоматическая синхронизация становится невозможной. Для снятия этой трудности обычно применяют неполное подавление несущей и оставляют в модулированном сигнале определенный "остаток" несущей (пилот-сигнал), который и используется для фазочастотной авто-синхронизации при демодуляции.

Однополосная амплитудная модуляция. При идентичности информации в группах верхних и нижних боковых частот нет никакой необходимости в их одновременной передаче. Одна из них перед подачей сигнала в канал связи может быть удалена, чем достигается двукратное сокращение полосы занимаемых сигналом частот. Уравнение сигнала с одной боковой полосой (ОБП – сигнал, single side band - SSB) может быть получено непосредственно из 9.1.11. Для верхней (знаки '+' во втором слагаемом) или нижней (знаки '-') боковой полосы:

$$u(t) = U_m \cos(\omega_0 t + \varphi_0) + (U_m / 2) \sum_{n=1}^N M_n \cos[(\omega_0 \pm \Omega_n)t + \varphi_0 \pm \Phi_n]. \quad (9.1.19)$$

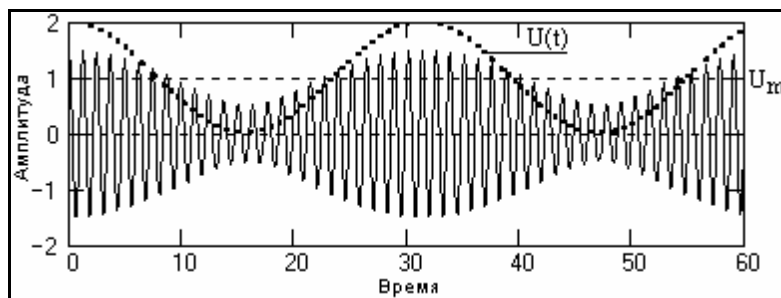


Рис. 9.1.10. Однополосная амплитудная модуляция.

Внешняя форма ОБП – сигнала после удаления одной боковой полосы, пример которой приведен на рис. 9.1.10 для однотонального сигнала, сходна с обычным АМ – сигналом, но ее огибающая, как это нетрудно заметить, отличается от огибающей $U(t)$, заданной при модуляции по при $M = 1$ (показана пунктиром).

Для демодуляции ОБП – сигнала может использоваться как двухполупериодное, так и синхронное детектирование, со всеми особенностями, присущими этим методам. Результаты демодуляции отличаются от демодуляции АМ – сигналов только в 2 раза меньшей амплитудой выходных сигналов.

При однополосной модуляции возможно также подавление несущей частоты (полное или частичное, с оставлением пилот-сигнала), что позволяет полнее использовать мощность передатчика.

Полярная модуляция решает конкретную техническую задачу – передачу двух сигналов одновременно, что требуется, например, в стереовещании или при передаче стереоснимков. Рассмотрим это на примере стереосигналов.

В системе стереовещания необходимо передавать два сигнала $s_1(t)$ и $s_2(t)$ одновременно (левый и правый каналы) при условии совмещения с монофоническими приемниками. Для выполнения этого условия создается специальный модулирующий сигнал. Процесс создания сигнала поясняется на рис. 9.1.11, где в качестве канальных сигналов приняты моногармонические сигналы s_1 и s_2 .

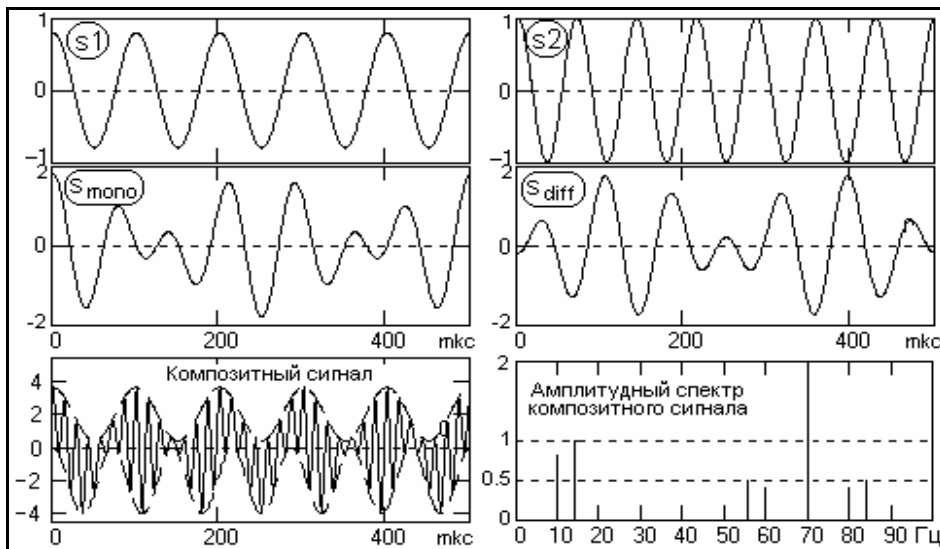


Рис. 9.1.11. Полярная модуляция.

Специальный модулирующий сигнал формируется из двух сигналов - монофонического и разностного. Монофонический сигнал образуется суммой сигналов в каналах, разностный – разностью сигналов:

$$s_{\text{mono}}(t) = s_1(t) + s_2(t), \quad s_{\text{diff}}(t) = s_1(t) - s_2(t),$$

что позволяет восстанавливать исходные сигналы каналов:

$$s_1(t) = (s_{\text{mono}}(t) + s_{\text{diff}}(t))/2, \quad s_2(t) = (s_{\text{mono}}(t) - s_{\text{diff}}(t))/2.$$

Монофонический сигнал является основным и не изменяется по частоте, что позволяет принимать его монофоническим приемникам. Для одновременной передачи разностного сигнала монофонический сигнал суммируется с *поднесущей* частотой ω_0 (subcarrier), которая располагается за звуковым диапазоном частот монофонических приемников (в области ультразвука), и модулируется разностным сигналом (с установкой коэффициента модуляции значением смещения A_0):

$$s(t) = s_{\text{mono}}(t) + (A_0 + s_{\text{diff}}(t)) \cdot \cos(\omega_0 t).$$

Полученный сигнал называют *композиционным стереосигналом*. Именно он используется в качестве модулирующего сигнала для любого метода модуляции, в том числе и для угло-

вой модуляции, которая будет рассматриваться ниже. Как видно на рис. 9.1.11, верхняя и нижняя огибающие композитного сигнала с точностью до постоянной составляющей соответствуют первому и второму сигналу стереоканалов, что позволяет достаточно просто выделять эти сигналы на приемной стороне. На практике поднесущую частоту композитного сигнала обычно частично или целиком подавляют. Подавление поднесущей выполняется изменением значения смещения $A_o \rightarrow 0$, при этом разностный сигнал переходит в режим перемодуляции, а динамический диапазон амплитуд композитного сигнала уменьшается в два раза.

9.2. СИГНАЛЫ С УГЛОВОЙ МОДУЛЯЦИЕЙ [1,25].

При угловой модуляции (angle modulation) в несущем гармоническом колебании $u(t) = U_m \cos(\omega t + \varphi)$ значение амплитуды колебаний U_m остается постоянным, а информация $s(t)$ переносится либо на частоту ω , либо на фазовый угол φ . И в том, и в другом случае текущее значение фазового угла гармонического колебания $u(t)$ определяет аргумент $\psi(t) = \omega t + \varphi$, который называют *полной фазой* колебания.

Фазовая модуляция (ФМ, phase modulation - PM). При фазовой модуляции значение фазового угла постоянной несущей частоты колебаний ω_0 пропорционально амплитуде модулирующего сигнала $s(t)$. Соответственно, уравнение ФМ – сигнала определяется выражением:

$$u(t) = U_m \cos[\omega_0 t + k \cdot s(t)], \quad (9.2.1)$$

где k – коэффициент пропорциональности. Пример однотонального ФМ – сигнала приведен на рис. 9.2.1.

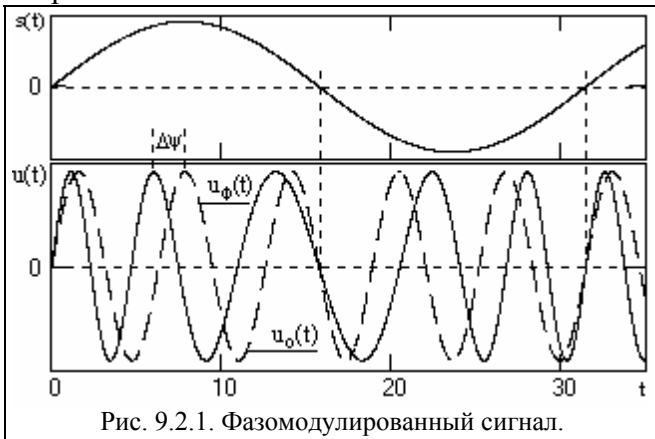


Рис. 9.2.1. Фазомодулированный сигнал.

При $s(t) = 0$, ФМ – сигнал является простым гармоническим колебанием и показан на рисунке функцией $u_0(t)$. С увеличением значений $s(t)$ полная фаза колебаний $\psi(t) = \omega_0 t + k \cdot s(t)$ нарастает во времени быстрее и опережает линейное нарастание $\omega_0 t$. Соответственно, при уменьшении значений $s(t)$ скорость роста полной фазы во времени спадает. В моменты экстремальных значений $s(t)$ абсолютное значение фазового сдвига $\Delta\psi$ между ФМ – сигналом и значением $\omega_0 t$ немодулированного колебания

также является максимальным и носит название *девиации фазы* (вверх $\Delta\varphi_v = k \cdot s_{\max}(t)$, или вниз $\Delta\varphi_n = k \cdot s_{\min}(t)$ с учетом знака экстремальных значений модулирующего сигнала).

Для колебаний с угловой модуляцией применяется также понятие мгновенной частоты (instantaneous frequency), под которой понимают производную от полной фазы по времени:

$$\omega(t) = \psi(t)/dt = \omega_0 + k ds(t)/dt.$$

Полная фаза колебаний в произвольный момент времени может быть определена интегрированием мгновенной частоты:

$$\psi(t) = \int_{-\infty}^t \omega(t) dt, \quad \text{или} \quad \psi(t) = \int_0^t \omega(t) dt + \varphi_0.$$

Частотная модуляция (ЧМ, frequency modulation - FM) характеризуется линейной связью модулирующего сигнала с мгновенной частотой колебаний, при которой мгновенная частота колебаний образуется сложением частоты высокочастотного несущего колебания ω_0 со значением амплитуды модулирующего сигнала с определенным коэффициентом пропорциональности:

$$\omega(t) = \omega_0 + k \cdot s(t). \quad (9.2.2)$$

Соответственно, полная фаза колебаний:

$$\psi(t) = \omega_0(t) + k \int_{-\infty}^t s(t) dt, \quad \text{или} \quad \psi(t) = \omega_0(t) + k \int_0^t s(t) dt + \varphi_0.$$

Уравнение ЧМ – сигнала:

$$u(t) = U_m \cos(\omega_0 t + k \int_0^t s(t) dt + \varphi_0). \quad (9.2.3)$$

Аналогично ФМ, для характеристики глубины частотной модуляции используются понятия *девиации частоты вверх* $\Delta\omega_B = k \cdot s_{\max}(t)$, и *вниз* $\Delta\omega_H = k \cdot s_{\min}(t)$.

Частотная и фазовая модуляция взаимосвязаны. Если изменяется начальная фаза колебания, изменяется и мгновенная частота, и наоборот. По этой причине их и объединяют под общим названием угловой модуляции (УМ). По форме колебаний с угловой модуляцией невозможно определить, к какому виду модуляции относится данное колебание, к ФМ или ЧМ, а при достаточно гладких функциях $s(t)$ формы сигналов ФМ и ЧМ вообще практически не отличаются.

Однотональная угловая модуляция. Рассмотрим гармонический модулирующий сигнал с постоянной частотой колебаний Ω . Начальная фаза колебаний:

$$\varphi(t) = \beta \sin(\Omega t),$$

где β - индекс угловой модуляции (modulation index), которым задается интенсивность колебаний начальной фазы. Полная фаза модулированного сигнала с учетом несущей частоты ω_0 :

$$\psi(\tau) = \omega_0 t + \beta \sin(\Omega t).$$

Уравнение модулированного сигнала:

$$u(t) = U_m \cos(\omega_0 t + \beta \sin(\Omega t)). \quad (9.2.4)$$

Мгновенная частота колебаний:

$$\omega(t) = d\psi(t)/dt = \omega_0 + \beta\Omega \cos(\Omega t).$$

Как следует из этих формул, и начальная фаза, и мгновенная частота изменяется по гармоническому закону. Максимальное отклонение от среднего значения ω_0 равно $\omega_d = \beta\Omega$, и получило название *девиации частоты* (frequency deviation). Отсюда, индекс угловой модуляции равен отношению девиации частоты к частоте модулирующего сигнала:

$$\beta = \omega_d / \Omega. \quad (9.2.5)$$

Различия между частотной и фазовой модуляцией проявляются при изменении частоты Ω модулирующего сигнала.

При фазовой модуляции девиация частоты прямо пропорциональна Ω , а индекс угловой модуляции от частоты модулирующего сигнала не зависит:

$$\beta = \text{const}, \quad \omega_d = \beta \Omega.$$

Напротив, при ЧМ постоянным параметром модуляции является девиация частоты, при этом индекс модуляции обратно пропорционален частоте модулирующего сигнала:

$$\omega_d = \text{const}, \quad \beta = \omega_d / \Omega.$$

Спектры сигналов с угловой модуляцией.

Формулу (9.2.4) однотональной модуляции можно преобразовать к виду:

$$u(t) = U_m \cos(\beta \cdot \sin(\Omega t)) \cos(\omega_0 t) - U_m \sin(\beta \cdot \sin(\Omega t)) \sin(\omega_0 t). \quad (9.2.6)$$

При малых значениях индекса угловой модуляции ($\beta \ll 1$, узкополосная модуляция) имеют место приближенные равенства:

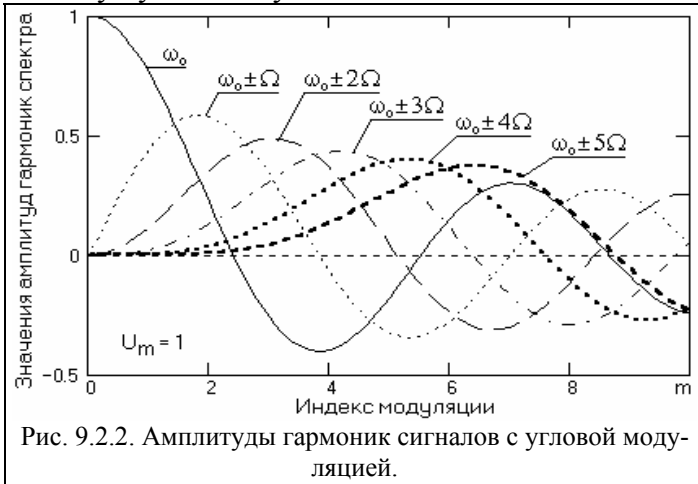
$$\cos(\beta \cdot \sin(\Omega t)) \approx 1, \quad \sin(\beta \cdot \sin(\Omega t)) \approx \beta \cdot \sin(\omega_0 t).$$

При их использовании в (9.2.6), получаем:

$$u(t) \approx U_m \cos(\omega_0 t) + (\beta U_m / 2) \cos[(\omega_0 + \Omega)t] + (-\beta U_m / 2) \cos[(\omega_0 - \Omega)t]. \quad (9.2.7)$$

Сравнение данного выражения с формулой АМ – сигнала (9.1.4) позволяет сделать вывод, что амплитудные спектры однотональных ФМ и ЧМ сигналов при $\beta \ll 1$ практически аналогичны АМ сигналам и также содержат верхнюю и нижнюю боковые частоты $\omega_0 + \Omega$ и $\omega_0 - \Omega$. Различие заключается только в смене знака амплитуды нижней боковой частоты на минус, т.е. в дополнительном фазовом сдвиге нижней боковой частоты на 180° относительно верхней боковой частоты. Соответственно, гармонические АМ сигналы могут быть трансформированы в ЧМ сигналы изменением на 180° начальной фазы одной из боковых полос. Заметим также, что при малых значениях индекса β основная мощность сигнала приходится

на несущую частоту.



Математическая модель однонаправленных ЧМ и ФМ сигналов с любым значением индекса модуляции β в общем случае получается разложением функции (9.2.4) в следующий ряд:

$$u(t) = U_m \sum_{k=-\infty}^{\infty} J_k(m) \cos[(\omega_0 + k\Omega)t],$$

где $J_k(m)$ – функция Бесселя k -го индекса от аргумента $m = \beta$. Из этого уравнения следует, что спектр сигнала содержит бесконечное число составляющих – нижних и верхних боковых колебаний, с частотами $\omega_0 \pm k\Omega$, которые соответствуют

гармоникам частоты модуляции, и с амплитудами, пропорциональными значениям $J_k(m)$. Амплитуды пяти первых гармоник и несущей частоты при $U_m = 1$ в зависимости от индекса модуляции приведены на рис. 9.2.2.

При малой величине индекса β значимые амплитудные значения имеют только первые гармоники. С ростом величины β количество значимых боковых составляющих увеличивается, а энергия сигнала перераспределяется на боковые составляющие. Функции Бесселя имеют колебательный характер, поэтому спектр при удалении от несущей частоты ω_0 спадает немонотонно. На рис. 9.2.2 можно также видеть, что при определенных значениях индекса модуляции (2.405, 5.52, 8.654 и т.д.) несущая частота ω_0 в спектре сигнала полностью отсутствует. Форма физических амплитудных спектров модулированных сигналов относительно несущей частоты при разных индексах модуляции приведена на рис. 9.2.3.

С ростом индекса модуляции полоса частот, занимаемая сигналом, расширяется. Практическая ширина спектра сигнала с угловой модуляцией определяется по формуле:

$$\Pi_{\text{практ}} = 2(\beta + 1)\Omega, \quad (9.2.8)$$

т.е. спектральными составляющими с номерами $k > (\beta + 1)$ пренебрегают. Формирование реальных сигналов, как правило, выполняется при $\beta \gg 1$, при этом эффективная ширина спектра равна удвоенной девиации частоты:

$$\Pi_{\text{практ}} \approx 2\beta\Omega = 2\omega_d. \quad (9.2.9)$$

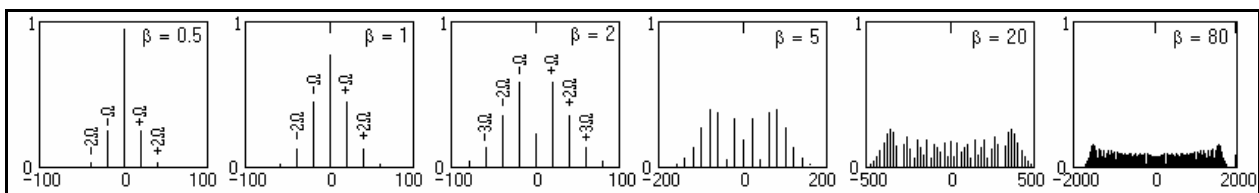


Рис. 9.2.3. Модули спектров ЧМ сигнала при разных индексах модуляции. (несущая частота 2500 Гц, гармоника модуляции 25 Гц, шкала частот в Гц относительно несущей)

Отсюда следует, что по сравнению с АМ – сигналами, полоса частот которых равна 2Ω , для передачи сигналов с угловой модуляцией требуется полоса частот, в β раз большая. С другой стороны, именно широкополосность ЧМ и ФМ сигналов обеспечивает их большую помехоустойчивость по сравнению с АМ сигналами.

Для функций Бесселя имеет также место: $J_{-k}(m) = (-1)^k J_k(m)$. Это означает, что начальные фазы боковых колебаний с частотами $\omega_0 + k\Omega$ и $\omega_0 - k\Omega$ совпадают при четных k , и отличаются на 180° при нечетных k .

Сигналы с многонаправленной угловой модуляцией отличаются еще большей сложностью спектрального состава. В их спектре присутствуют не только боковые частоты с гармониками частот модулирующего сигнала, но и боковые комбинационные частоты типа

$\omega_0 \pm \Omega_1 \pm \Omega_2 \pm \dots \pm \Omega_i$, со всеми возможными комбинациями частот модулирующего сигнала Ω_i . При непрерывном спектре модулирующего сигнала спектры ЧМ и ФМ сигналов также становятся непрерывными.

Демодуляция УМ – сигналов много сложнее демодуляции сигналов АМ.

При демодуляции полностью зарегистрированных цифровых сигналов обычно используется метод формирования комплексного аналитического сигнала с помощью преобразования Гильберта:

$$u_a(t) = u(t) + j u_h(t),$$

где $u_h(t)$ – аналитически сопряженный сигнал или квадратурное дополнение сигнала $u(t)$, которое вычисляется сверткой сигнала $u(t)$ с оператором Гильберта ($1/\pi t$):

$$u_h(t) = (1/\pi) \int_{-\infty}^{\infty} u(t') dt'/(t-t').$$

Полная фаза колебаний представляет собой аргумент аналитического сигнала:

$$\psi(t) = \arg(u_a(t)).$$

Дальнейшие операции определяются видом угловой модуляции. При демодуляции ФМ сигналов из фазовой функции вычитается значение немодулированной несущей $\omega_0 t$:

$$\varphi(t) = \psi(t) - \omega_0 t.$$

При частотной модуляции фазовая функция дифференцируется с вычитанием из результата значения частоты ω_0 :

$$\varphi(t) = \psi(t)/dt - \omega_0.$$

В принципе, данный метод может применяться и в реальном масштабе времени, но с определенной степенью приближения, поскольку оператор Гильберта слабо затухает.

Обычно в реальном масштабе времени используется квадратурная обработка, при которой входной сигнал умножается на два опорных колебания со сдвигом фазы между колебаниями в 90° :

$$u_1(t) = u(t) \cos(\omega_0 t) = U_m \cos(\omega_0 t + \varphi(t)) \cos(\omega_0 t) = \frac{1}{2} U_m \cos \varphi(t) + \frac{1}{2} \cos(2\omega_0 t + \varphi(t)),$$

$$u_2(t) = u(t) \sin(\omega_0 t) = U_m \cos(\omega_0 t + \varphi(t)) \sin(\omega_0 t) = -\frac{1}{2} U_m \sin \varphi(t) + \frac{1}{2} \sin(2\omega_0 t + \varphi(t)).$$

Из этих двух сигналов фильтрами низких частот выделяются низкочастотные колебания, и формируется аналитический сигнал:

$$u_a(t) = \frac{1}{2} U_m \cos \varphi(t) - \frac{1}{2} j U_m \sin \varphi(t).$$

Аргумент этого аналитического сигнала, как и в первом случае, представляет полную фазу колебаний, обработка которой выполняется аналогично.

Квадратурная модуляция позволяет модулировать несущую частоту одновременно двумя сигналами путем модуляции амплитуды несущей одним сигналом, и фазы несущей другим сигналом. Уравнение результирующих колебаний амплитудно-фазовой модуляции:

$$s(t) = u(t) \cos(\omega_0 t + \varphi(t)).$$

Сигнал $s(t)$ обычно формируют в несколько другой последовательности, с учетом последующей демодуляции. Раскроем косинус суммы и представим сигнал в виде суммы двух АМ-колебаний.

$$s(t) = u(t) \cos(\omega_0 t) \cos \varphi(t) - u(t) \sin(\omega_0 t) \sin \varphi(t).$$

При $a(t) = u(t) \cos \varphi(t)$ и $b(t) = -u(t) \sin \varphi(t)$, сигналы $a(t)$ и $b(t)$ могут быть использованы в качестве модулирующих сигналов несущих колебаний $\cos(\omega_0 t)$ и $\sin(\omega_0 t)$, сдвинутых по фазе на 90° относительно друг друга:

$$s(t) = a(t) \cos(\omega_0 t) + b(t) \sin(\omega_0 t).$$

Полученный сигнал называют квадратурным (quadrature), а способ модуляции - *квадратурной* модуляцией (КАМ).

Спектр квадратурного сигнала может быть получен непосредственно по уравнению балансной модуляции (9.1.17) для суммы двух сигналов:

$$S(\omega) = \frac{1}{2} A(\omega + \omega_0) + \frac{1}{2} A(\omega - \omega_0) - \frac{1}{2} j B(\omega + \omega_0) + \frac{1}{2} j B(\omega - \omega_0).$$

Демодуляция квадратурного сигнала соответственно выполняется умножением на два

опорных колебания, сдвинутых относительно друг друга на 90° :

$$s_1(t) = s(t) \cos \omega_0 t = \frac{1}{2} a(t) + \frac{1}{2} a(t) \cos 2\omega_0 t + \frac{1}{2} b(t) \sin 2\omega_0 t,$$

$$s_2(t) = s(t) \sin \omega_0 t = \frac{1}{2} b(t) + \frac{1}{2} a(t) \sin 2\omega_0 t - \frac{1}{2} b(t) \cos 2\omega_0 t.$$

Низкочастотные составляющие $a(t)$ и $b(t)$ выделяются фильтром низких частот. Как и при балансной амплитудной модуляции, для точной демодуляции сигналов требуется точное соблюдение частоты и начальной фазы опорного колебания.

Пример моделирования квадратурной модуляции в системе Mathcad.

Моделирование выполняется в дискретной форме.

$N := 2999$ $n := 0 .. N$ $\Delta t := 0.001$ 'Интервал и шаг дискретизации (в сек).

$f_0 := 50$ $f_1 := 2$ $f_2 := 3$

'Частоты в Гц несущей, первого и второго сигналов.

$s1_n := \sin(2 \cdot \pi \cdot f_1 \cdot n \cdot \Delta t)$

'Первый модулирующий сигнал (моногармоника с амплитудой 1).

$s2_n := \sin(2 \cdot \pi \cdot f_2 \cdot n \cdot \Delta t)$

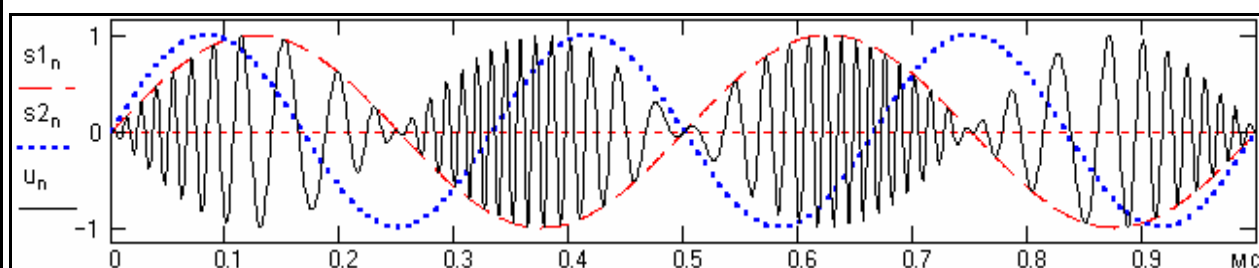
'Первый модулирующий сигнал (моногармоника с амплитудой 1).

$\beta := 10$ $\varphi_n := \beta \cdot s2_n$

'Перенос информации $s2_n$ на фазу

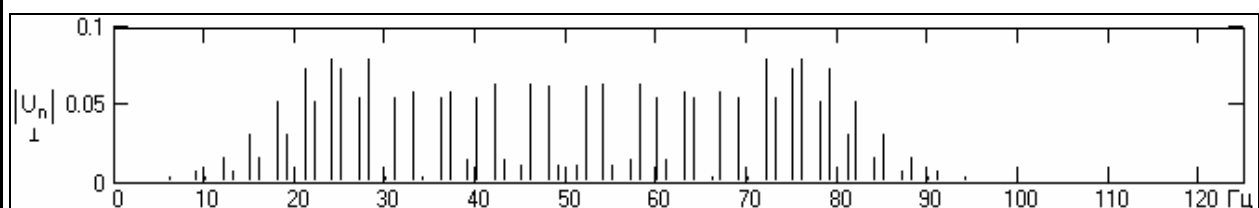
$u_n := s1_n \cdot \cos(2 \cdot \pi \cdot f_0 \cdot n \cdot \Delta t + \varphi_n)$

'Амплитудно-фазовая модуляция



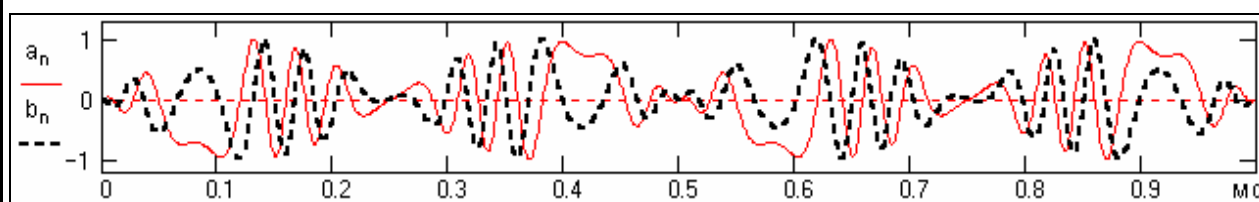
$U := \text{CFFT}(u)$ $\Delta f := 1/[(N+1) \cdot \Delta t]$

'БПФ и шаг по частоте



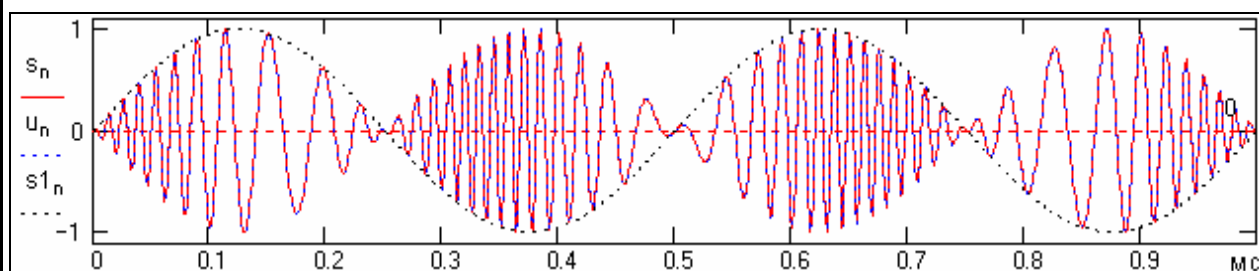
$a_n := s1_n \cdot \cos(\varphi_n)$ $b_n := s1_n \cdot \sin(\varphi_n)$

'Формирование модулирующих сигналов



$s_n := a_n \cdot \cos(2 \cdot \pi \cdot f_0 \cdot n \cdot \Delta t) + b_n \cdot \sin(2 \cdot \pi \cdot f_0 \cdot n \cdot \Delta t)$

'Квадратурный сигнал. Сравнением с сигналом u_n нетрудно убедиться в их идентичности, a , следовательно, идентичны и их спектры.



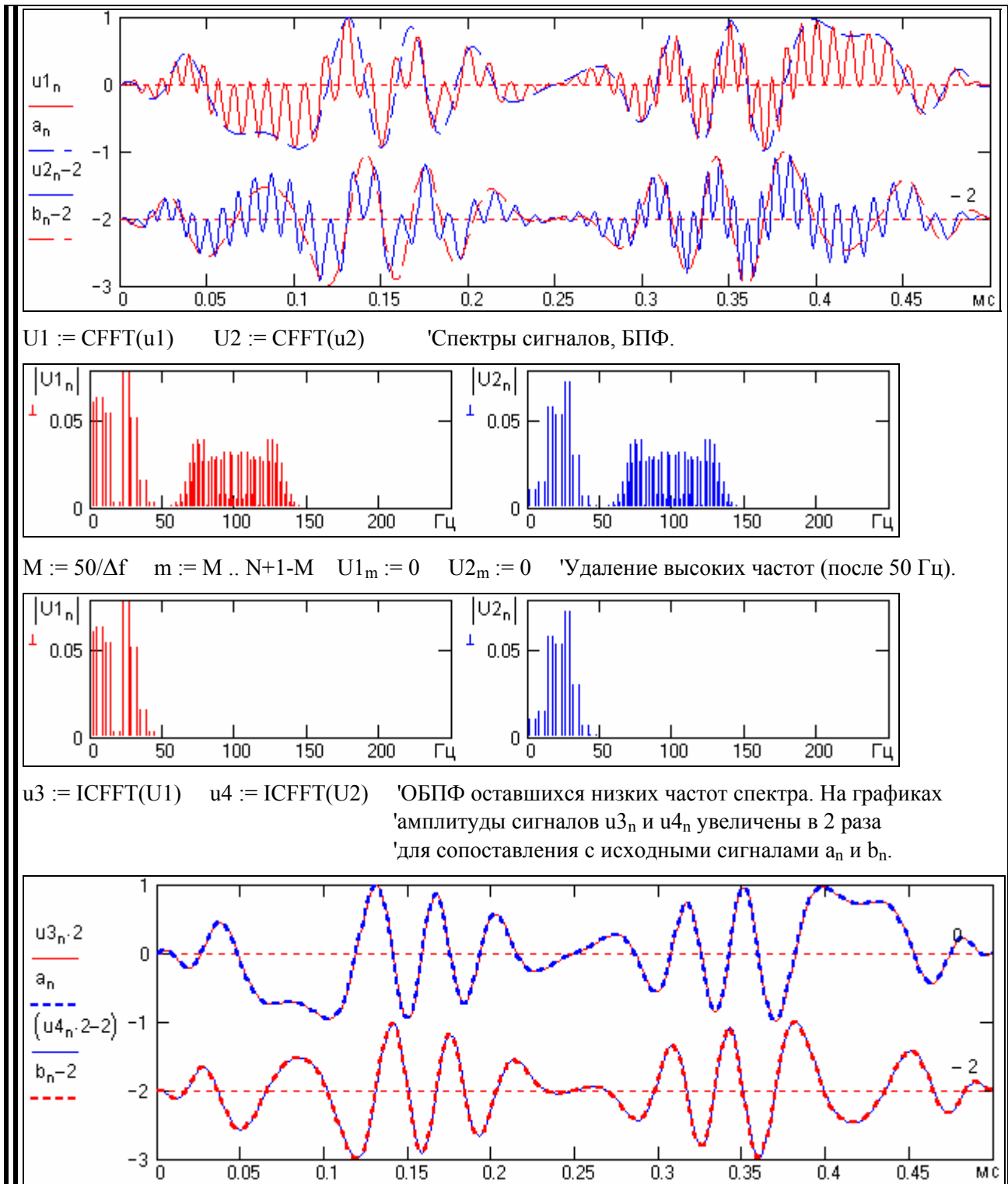
Демодуляция квадратурного сигнала.

$u1_n := s_n \cdot \cos(2 \cdot \pi \cdot f_0 \cdot n \cdot \Delta t)$

'Раздельная синхронная демодуляция сигналов a_n и b_n . Графики

$u2_n := s_n \cdot \sin(2 \cdot \pi \cdot f_0 \cdot n \cdot \Delta t)$

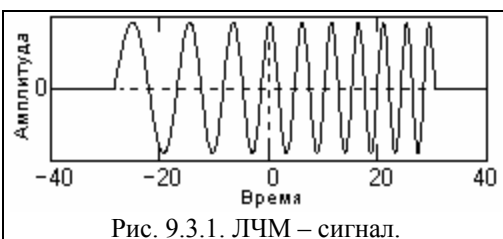
'сигналов $u2_n$ и b_n смешены на -2 для представления в одном поле.



9.3. ВНУТРИИМПУЛЬСНАЯ ЧАСТОТНАЯ МОДУЛЯЦИЯ [1].

Сигнал с внутриимпульсной частотной модуляцией – это радиоимпульс, высокочастотное заполнение которого имеет переменную частоту.

ЛЧМ – сигналы. Если закон изменения мгновенной частоты заполнения имеет линейный характер, то такие сигналы носят название ЛЧМ – сигналов (линейная частотная модуляция). Наиболее широкое применение они получили в радиолокации. Пример ЛЧМ – сигнала с огибающей прямоугольной формы приведен на рис. 9.3.1.



ЛЧМ – сигналы имеют одно замечательное

свойство. Если сигнал подать на частотно-зависимую линию задержки, время задержки сигнала которой велико на малых частотах (в начальной части ЛЧМ – сигнала) и уменьшается по мере нарастания частоты в ЛЧМ – сигнале, то на выходе такой линии происходит "сжатие" сигнала в один период высокочастотного колебания путем суммирования амплитудных значений всех периодов сигнала. При этом происходит увеличение амплитуды выходного сигнала и уменьшение статистических шумов, так как суммируемые одновременно по этим же периодам шумы не коррелированы.

Для модели радиоимпульса с прямоугольной огибающей примем его длительность равной τ_u , и точку $t = 0$ поместим в центр радиоимпульса. Допустим также, что частота заполнения линейно нарастает от начала импульса к его концу со скоростью μ (c^2), при этом:

$$\omega(\tau) = \omega_0 + \mu\tau. \quad (9.3.1)$$

Девияция частоты за время длительности импульса и полная фаза сигнала:

$$\Delta\omega = \mu \cdot \tau_u. \quad (9.3.2)$$

$$\psi(t) = \omega_0 t + \mu t^2 / 2. \quad (9.3.3)$$

Уравнение ЛЧМ – сигнала:

$$u(t) = \begin{cases} 0, & t < -\tau_u / 2 \\ U_m \cos(\omega_0 t + \mu t^2 / 2), & -\tau_u / 2 \leq t \leq \tau_u / 2 \\ 0, & t > \tau_u / 2 \end{cases} \quad (9.3.4)$$

Спектр прямоугольного ЛЧМ – сигнала вычисляется через преобразование Фурье. Девияция частоты за время длительности импульса по сравнению с несущей частотой обычно мала ($\Delta\omega \ll \omega_0$) и форма спектра зависит от так называемой базы импульса:

$$\Delta\omega \cdot \tau_u = \mu \cdot \tau_u^2. \quad (9.3.5)$$

На рис. 9.3.2 приведен пример формы спектральной плотности ЛЧМ – сигнала при малом значении базы в области несущей частоты сигнала.

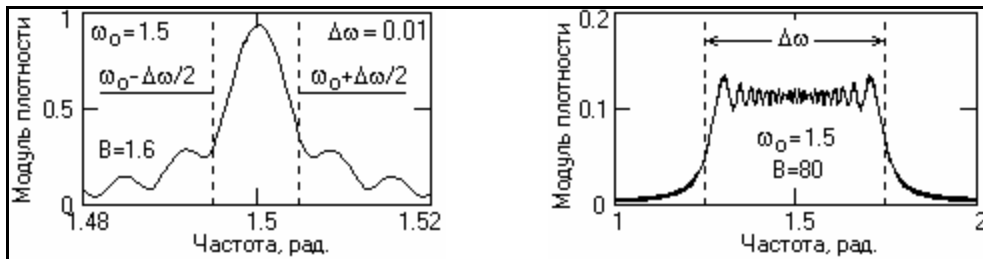


Рис.9.3.2. Спектр ЛЧМ- сигнала.

Рис. 9.3.3. Спектр при $B \gg 1$.

На практике значение базы сигналов обычно много больше 1. Увеличение базы сопровождается расширением полосы спектра $\Delta\omega$, при этом в пределах этой полосы модуль спектральной плотности практически постоянен и равен $U_m \cdot \sqrt{\pi/2\mu}$. Пример спектра приведен на рис. 9.3.3.

9.4. ИМПУЛЬСНО – МОДУЛИРОВАННЫЕ СИГНАЛЫ.

В импульсной модуляции в качестве носителя модулированных сигналов используются последовательности импульсов, как правило – прямоугольных. В беспроводных системах передачи данных (в радиосвязи) эти последовательности заполняются высокочастотными колебаниями, создавая тем самым двойную модуляцию. Как правило, эти виды модуляции применяются при передаче дискретизированных данных. Для прямоугольных импульсов наиболее широко используются амплитудно-импульсная (АИМ) и широтно-импульсная (ШИМ) модуляция.

Амплитудно-импульсная модуляция (АИМ) заключается в изменении приращения амплитуды импульсов пропорционально функции управляющего сигнала при постоянной длительности импульсов и периоде их следования:

$$U(t) = U_0 + k \cdot s(t), \quad \tau_u = \text{const}, \quad T = \text{const}. \quad (9.4.1)$$

Спектр АИМ рассмотрим на примере модулирования однотонального сигнала $s(t)$, приведенного на рис. 9.4.1. Напишем уравнение модулированного сигнала в следующей форме:

$$u(t) = (1 + M \cos \Omega t) \cdot f(t), \quad (9.4.2)$$

где $f(t)$ – периодическая последовательность прямоугольных импульсов с частотой ω_0 , которую можно аппроксимировать рядом Фурье (без учета фазы):

$$f(t) = U_0 + \sum_{n=1}^{\infty} U_n \cos n\omega_0 t. \quad (9.4.3)$$

Подставляя (9.4.3) в (9.4.2), получаем:

$$\begin{aligned} u(t) &= (1 + M \cos \Omega t) U_0 + \sum_{n=1}^{\infty} U_n \cos n\omega_0 t \cdot (1 + M \cos \Omega t) = \\ u(t) &= U_0 + U_0 M \cos \Omega t + \sum_{n=1}^{\infty} U_n \cos n\omega_0 t + \\ &+ 0.5M \sum_{n=1}^{\infty} U_n \cos (n\omega_0 + \Omega)t + 0.5M \sum_{n=1}^{\infty} U_n \cos (n\omega_0 - \Omega)t. \end{aligned} \quad (9.4.2)$$

Форма спектра, в начальной части спектрального диапазона, приведена на рис. 9.4.1. В целом, спектр бесконечен, что определяется бесконечностью спектра прямоугольных импульсов. Около каждой гармоники $n\omega_0$ спектра прямоугольных импульсов появляются боковые составляющие $n\omega_0 \pm \Omega$, соответствующие спектру моделирующей функции (при многотональном сигнале – боковые полосы спектров). При дополнительном высокочастотном заполнении импульсов весь спектр смещается в область высоких частот на частоту заполнения.

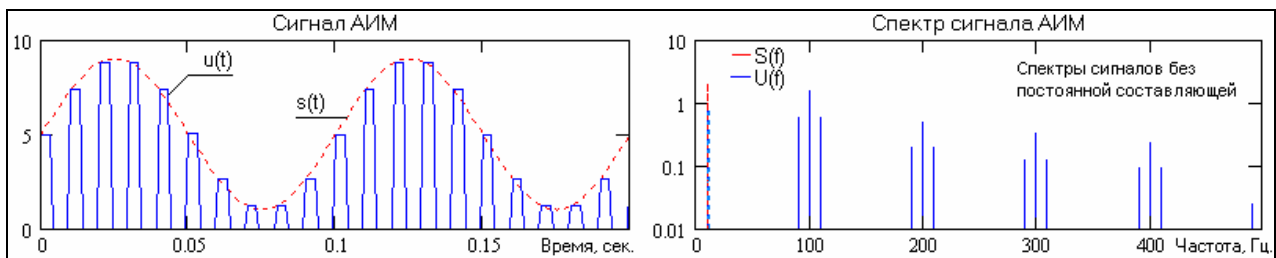


Рис. 9.4.1.

Широтно-импульсная модуляция (ШИМ, в английской терминологии pulse width modulation, PWM), которую иногда называют *модуляцией по длительности импульсов* (ДИМ), заключается в управлении длительностью импульсов пропорционально функции управляющего сигнала при постоянной амплитуде импульсов и периоде следования по фронту импульсов:

$$\tau(t) = t_0 + k \cdot s(t), \quad U = \text{const}, \quad T = \text{const}. \quad (9.4.3)$$

Рассмотрим выполнение ШИМ в простейшем варианте на примере гармонического колебания, приведенного на рис. 9.4.2.

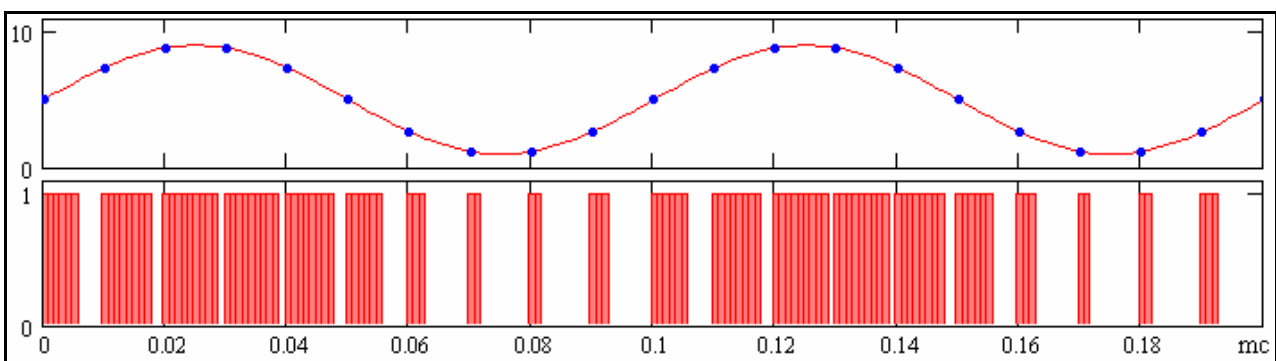


Рис. 9.4.2. Широтно-импульсная модуляция.

Передаваемая кривая дискретизируется, при этом имеет значение, как интервал дис-

кретизации, так и количество уровней квантования. При передаче данных прямоугольные импульсы начинаются в моменты дискретных отсчетов данных, а длительность импульсов устанавливается пропорциональной значению отсчетов, при этом максимальная длительность импульсов не должна превышать интервала дискретизации данных. Пример сформированных импульсов приведен на рис. 9.4.2 непосредственно под дискретизированной гармоникой, при этом число уровней квантования гармоникой принято равным 8.

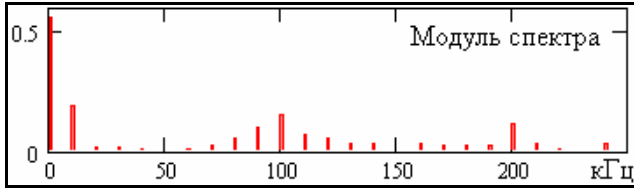


Рис. 9.4.3. Спектр ШИМ – сигнала.

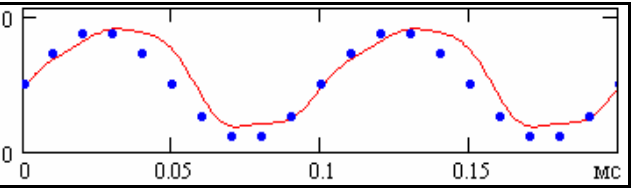


Рис. 9.4.4. Восстановленный сигнал.

На рис. 9.4.3 приведен спектр сформированного сигнала ШИМ. В начальной части спектра он содержит постоянную составляющую среднего уровня сигнала и пик частоты гармоник, закодированной в ШИМ – сигнале. Если выделить из спектра эти две составляющие, то восстанавливается исходный сигнал с погрешностью квантования, приведенный на рис. 9.4.4. Естественно, что при малом числе уровней квантования погрешность восстановления исходного гармонического сигнала очень велика.

Попутно заметим, что широтно-импульсная модуляция с последующим выделением постоянной составляющей может весьма эффективно использоваться (и используется) для слежения за средним уровнем сигнала и автоматического регулирования его динамического диапазона, как, например, в системах установки громкости звука и яркости цветов и изображения в целом в современных телевизионных установках.

Временная импульсная модуляция (ВИМ) представляет собой девиацию импульсов по временной оси по закону модулирующего сигнала, и по существу аналогична угловой модуляции гармонической несущей. Она также может быть фазовой (ФИМ) или частотной (ЧИМ).

Кодово-импульсная модуляция заключается в том, что в точках дискретизации модулирующего сигнала производится квантование его значений и кодирование квантованных значений, как правило, в двоичной системе исчисления. Кодированные значения затем передаются при помощи соответствующей кодовой последовательности стандартных символов.

9.5. МОДУЛЯЦИЯ СИМВОЛЬНЫХ И КODOVЫХ ДАННЫХ [25].

В настоящее время информация передается по каналам связи в основном в цифровой форме. Числа при передаче с периодом T поступают от источника информации и называются *символами* (symbol), а частота передачи символов – *символьной скоростью* (symbol rate) $f_T=1/T$. В практике передачи данных распространена двоичная (binary) последовательность символов, где числа передаются значениями 0 и 1.

Символьные последовательности являются дискретными квантованными сигналами, которые обычно передаются следующим образом. Каждому из возможных символов устанавливается определенный набор параметров несущего колебания, которые поддерживаются постоянными на интервале T до прихода следующего символа. Это означает преобразование последовательности чисел в ступенчатый сигнал (кусочно-постоянная интерполяция) который используется в качестве модулирующего сигнала. Соответственно, параметры несущего колебания, на которые переносится ступенчатый сигнал, также меняются скачкообразно. Такой способ модуляции несущей называется *манипуляцией* (keying), и может выполняться с использованием всех рассмотренных методов модулирования.

Амплитудно-манипулированные сигналы простейшего типа представляют собой последовательности радиоимпульсов, разделенные паузами. Такие сигналы используются в радиотелеграфии и в системах передачи дискретных данных. Форма огибающей радиоимпульсов в общем случае может быть произвольной, паузы могут отличаться по длительности от радиоимпульсов.

На рис. 9.5.1. приведен пример амплитудно-манипулированного сигнала:

$$u(t) = U_m \cdot \cos(2\pi f_0 t),$$

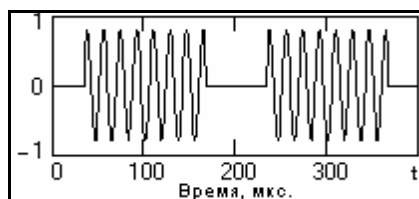


Рис. 9.5.1. АМП-сигнал.

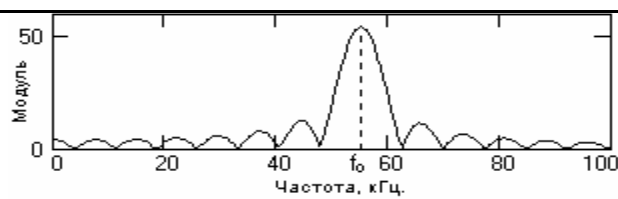


Рис. 9.5.2. Модуль спектра АМП-сигнала.

с прямоугольной П-формой огибающей. Соответственно, в частотной области спектр АМП – сигнала образуется сверткой спектра огибающей функции (в данном случае – спектра прямоугольного импульса) со спектром косинусного колебания (дельта - функции на частоте f_0). Модуль спектральной плотности сигнала приведен на рис. 9.5.2. Спектр прямоугольного импульса довольно слабо затухает и простирается неограниченно далеко, а поэтому его использование в качестве огибающей АМП - сигнала не рекомендуется, хотя и является наиболее простым по техническому исполнению.

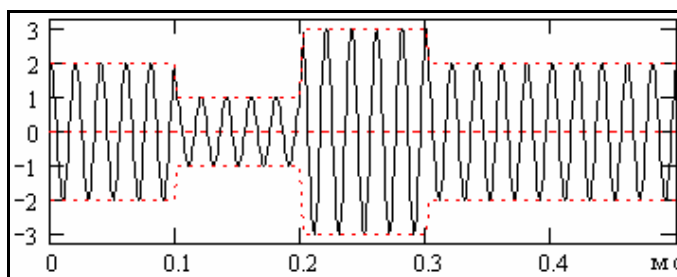


Рис. 9.5.3.

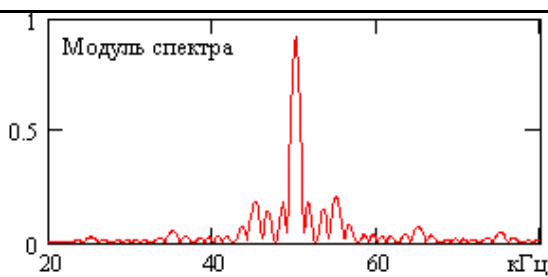


Рис. 9.5.4.

На рис. 9.5.3. приведен пример формы классического АМП сигнала при передаче нескольких символов, каждому из которых соответствует индивидуальная амплитуда несущей частоты при постоянной длительности интервалов посылки. Модуль спектра сигнала приведен на рис. 9.5.4 и тоже имеет достаточно большую ширину значимой части спектра вокруг несущей частоты.

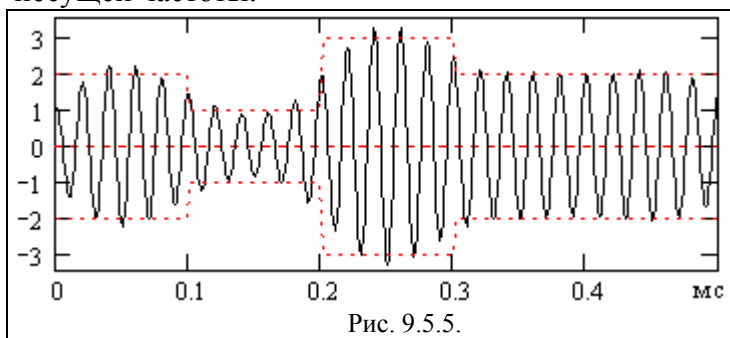


Рис. 9.5.5.

Естественно, что при передаче данных частотный диапазон канала передачи данных ограничивается значимой частью спектра, ширина которого устанавливается по допустимой степени искажения приемных сигналов. Степень искажения сигналов существенно зависит от длительности посылок. Пример искажения вышеприведенного сигнала при ограничении спектра интервалом 40-60 кГц приведен на рис. 9.5.5.

Угловая манипуляция, как правило, использует частотные методы модулирования, в которых каждому возможному значению передаваемого символа сопоставляется индивидуальное значение частоты гармонической несущей. При этом в точках сопряжения интервалов посылок могут происходить скачки напряжения, с соответствующим усложнением спектра модулированного сигнала. Самый простой способ – синусоидальное начало несущей на каждом интервале с кратным количеством периодов несущей в посылке. При более сложных способах, независимых от точного сопряжения несущих частот с интервалами посылок, осуществляется управление скоростью изменения фазы несущих на границах посылок.

Демодуляция сигналов осуществляется корреляционными методами. Сущность методов – вычисление взаимной корреляции между принимаемым сигналом и набором опорных частот, используемых при модулировании, с идентификацией символов по максимумам вза-

имной корреляции.

Для повышения помехоустойчивости передачи данных желательно, чтобы разносимвольные посылки были некоррелированы. Если для бинарных символов 0 и 1 принять частоты посылок равными

$$s_0(t) = \cos \omega_0(t), \quad s_1(t) = \cos \omega_1(t),$$

то их ВКФ при нулевом временном сдвиге определится выражением:

$$B_{01}(0) = \int_0^T s_0(t) s_1(t) dt = \frac{1}{2} (\sin (\omega_1 + \omega_0)T) / (\omega_1 + \omega_0) + \frac{1}{2} (\sin (\omega_1 - \omega_0)T) / (\omega_1 - \omega_0).$$

При $(\omega_1 + \omega_0)T \gg 1$ первым слагаемым можно пренебречь, оно много меньше второго. А второе слагаемое обращается в нуль при $(\omega_1 + \omega_0)T = \pi k$, где $k = 1, 2, \dots$ – целое число. Отсюда, минимальное значение между частотами манипуляции для некоррелированных посылок определяется выражениями:

$$\Delta\omega_{\min} = \pi/T, \quad \Delta f_{\min} = 1/2T = f_T/2,$$

где f_T – символьная скорость.

Фазовая манипуляция применяется значительно реже, в связи со значительными сложностями измерения абсолютных значений начальных фаз в посылках. Проще определяется относительный фазовый сдвиг в соседних посылках, поэтому обычно используется фазоразностная манипуляция.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов.- М.: Высшая школа, 1988.
25. Сергиенко А.Б. Цифровая обработка сигналов. – СПб.: Питер, 2003. – 608 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 10: ПРЕОБРАЗОВАНИЕ ГИЛЬБЕРТА

То, что не может произойти, никогда не может быть, а если произошло, то не должно нас удивлять.

Марк Туллий Цицерон. О девиации.
Римский философ и политик, I в.д.н.э.

Однако пока не создано строгой математической теории чудес, приходится наоборот не удивляться, когда они не осуществляются, и удивляться, когда они осуществляются.

Николай Пятин. О чудесах.
Воронежский геофизик Уральской школы, XX в.

Содержание: Введение. 10.1. Сущность преобразования Гильберта. Определение преобразования Гильберта. Спектральная характеристика преобразования Гильберта. Изменение спектра сигналов при выполнении преобразования Гильберта. Спектры каузальных функций. 10.2. Свойства преобразования Гильберта. Линейность. Сдвиг. Преобразование константы. Свойство четности и нечетности. Последовательное двойное преобразование. Обратное преобразование Гильберта. Подобие. Энергетическая эквивалентность. Свойство ортогональности. Свойство свертки. Свойство модуляции. 10.3. Вычисление преобразования Гильберта. Преобразование Гильберта аналоговых сигналов. Оператор дискретного преобразования. Литература.

ВВЕДЕНИЕ.

Преобразование Гильберта для любого произвольного сигнала представляет собой идеальный широкополосный фазовращатель, который осуществляет поворот начальных фаз всех частотных составляющих сигнала на угол, равный 90° (сдвиг на $\pi/2$). Применение преобразования Гильберта позволяет выполнять квадратурную модуляцию сигналов, в каждой текущей координате модулированных сигналов производить определение огибающей и мгновенной фазы (частоты) сигналов, выполнять анализ каузальных систем обработки сигналов.

10.1. СУЩНОСТЬ ПРЕОБРАЗОВАНИЯ ГИЛЬБЕРТА [1, 2, 21].

Определение преобразования. Прямое преобразование Гильберта произвольной действительной функции $x(t)$, $-\infty < t < \infty$, результат которого будем отображать знаком тильды над индексом исходной функции, задается сверткой $x(t)$ с функцией $hb(t) = 1/(\pi t)$:

$$\tilde{x}(t) = \text{TH}[x(t)] = x(t) * (1/\pi t),$$

$$\tilde{x}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau. \quad (10.1.1)$$

Функция $1/(t-\tau)$ называется *ядром преобразования* Гильберта. Обратное преобразование Гильберта определяется выражением:

$$x(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tilde{x}(\tau)}{t-\tau} d\tau. \quad (10.1.1')$$

Интегралы преобразования имеет особую точку при $a = t-\tau \Rightarrow 0$ и при вычислении используется их главное значение по Коши:

$$\lim_{a \rightarrow 0} \left[\int_{-\infty}^{t-a} \dots + \int_{t+a}^{\infty} \dots \right].$$

Оператор Гильберта определен по аргументу от $-\infty$ до ∞ и имеет полюс в точке $t=0$ с разрывом значений от $-\infty$ до ∞ . Основной участок формы оператора Гильберта и пример преобразования сигнала приведены на рис. 10.1.1.

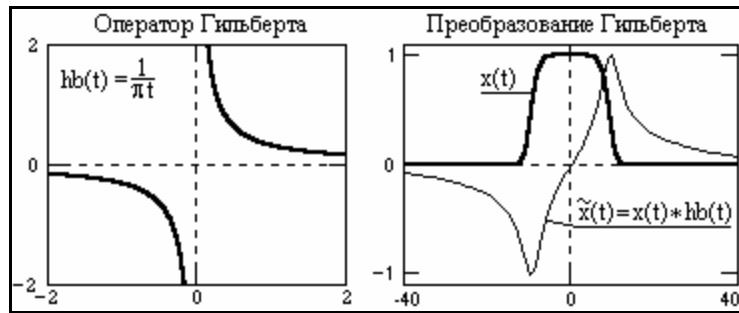


Рис. 10.1.1.

Спектральная характеристика преобразования. Выполним преобразование Фурье функции (10.1.1). В общей форме:

$$\tilde{X}(f) = \text{TF}[\tilde{x}(t)] = X(f) \cdot \text{Hb}(f), \quad (10.1.2)$$

$$\tilde{X}(f) = \int_{-\infty}^{\infty} \tilde{x}(t) \exp(-j2\pi ft) dt. \quad (10.1.2')$$

Заметим, что произведение $X(f) \cdot \text{Hb}(f)$ не является преобразованием Гильберта спектральной функции $X(f)$. Это не более чем преобразование Фурье свертки функций: $x(t) * \text{hb}(t) \Leftrightarrow X(f) \cdot \text{Hb}(f)$, которое позволяет вычислить результат преобразования Гильберта во временной области через частотную область:

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

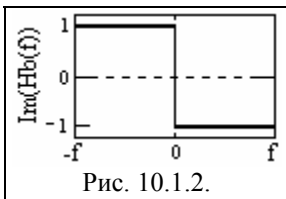


Рис. 10.1.2.

Функция $\text{hb}(t) = 1/\pi t$ является нечетной, а спектр этой функции, представленный только мнимой частью, является обратной сигнатурной функцией (рис. 10.1.2):

$$\text{Hb}(f) = \text{TF}[1/\pi t] = -j \cdot \text{sgn}(f) = \begin{cases} -j, & f > 0 \\ 0, & f = 0 \\ j, & f < 0 \end{cases} \quad (10.1.3)$$

Соответственно, формулы (10.1.1) задают преобразование сигнала $x(t)$ системой, частотная передаточная характеристика которой отображается функцией $-j \cdot \text{sgn}(f)$. Фурье-образ функции $\tilde{x}(t)$:

$$\tilde{X}(f) = -j \text{sgn}(f) \cdot X(f). \quad (10.1.2'')$$

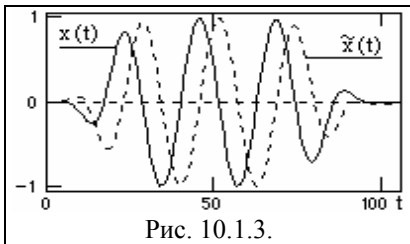
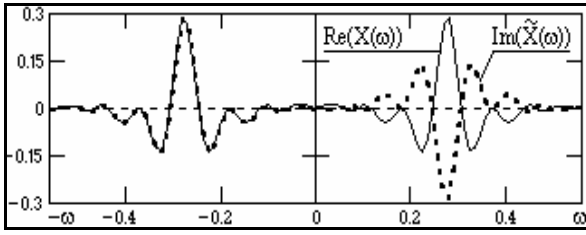
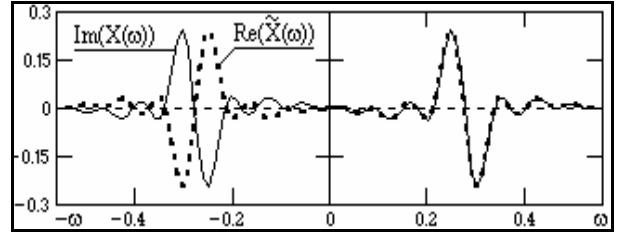


Рис. 10.1.3.

Изменение спектра сигналов при выполнении преобразования Гильберта. На рис. 10.1.3 приведено преобразование радиоимпульсного сигнала $x(t) = a(t) \cdot \cos(\omega_0 t)$ с несущей частотой ω_0 в сигнал $\tilde{x}(t)$ во временной области непосредственно через операцию свертки по (10.1.1). Сигнал $x(t)$ является односторонним каузальным. Спектр сигнала содержит реальную и мнимую составляющие, т.е. может быть записан в

виде $X(\omega) = \text{Re}(X(\omega)) + j \cdot \text{Im}(X(\omega))$. Эти составляющие для сигнала $x(t)$ на рис. 10.1.3 показаны непрерывными кривыми на рис. 10.1.4 и 10.1.5.

При выполнении преобразования (10.1.2'') реальная и мнимая части спектра $X(\omega)$ умножаются на $-j \cdot \text{sgn}(\omega)$. Функция $\text{Re}(X(\omega))$ (рис. 10.1.4) умножается на 1 при $\omega < 0$, на 0 при $\omega = 0$ и на -1 при $\omega > 0$, и тем самым превращается в нечетную мнимую часть $\text{Im}(\tilde{X}(\omega))$ спектра $\tilde{X}(\omega)$ функции $\tilde{x}(t)$, показанную пунктиром. Это означает, что все косинусные гармоники сигнала, которым соответствует реальная часть спектра сигнала, превращаются в синусные гармоники.

Рис. 10.1.4. Преобразование $\text{Re}(X) \Rightarrow \text{Im}(\tilde{X})$ Рис. 10.1.5. Преобразование $\text{Im}(X) \Rightarrow \text{Re}(\tilde{X})$

Аналогично на функцию $-j \cdot \text{sgn}(\omega)$ умножается и мнимая функция $j \cdot \text{Im}(X(\omega))$, при этом сигнатурная функция инвертируется ($-j \cdot j = 1$), что меняет знак левой части функции $\text{Im}(X(\omega))$ – области отрицательных частот, и превращает ее в реальную четную часть $\text{Re}(\tilde{X}(\omega))$ спектра $\tilde{X}(\omega)$ (рис. 10.1.5). Тем самым синусные гармоники спектра сигнала превращаются в косинусные гармоники.

Угол, на который изменяется фаза гармоник, можно определить из следующих соотношений. Частотную характеристику $H_b(f)$ (10.1.3) можно записать в следующем виде:

$$H_b(f) = |H_b(f)| \cdot \exp(j\varphi_b(f)), \quad \text{где } |H_b(f)| = 1.$$

$$H_b(f) = -j \cdot \text{sgn}(f) = \begin{cases} \exp(j(-\pi/2)), & f > 0 \\ 0, & f = 0 \\ \exp(j(\pi/2)), & f < 0 \end{cases}, \quad (10.1.3')$$

Если спектр функции $x(t)$ также представить в виде

$$X(f) = |X(f)| \cdot \exp(j\varphi_x(f)),$$

то выражение (10.1.2) преобразуется к следующей форме:

$$\tilde{X}(f) = |X(f)| \cdot \exp(j\varphi_x(f)) \cdot \exp(j\varphi_b(f)) = |X(f)| \cdot \exp[j(\varphi_x(f) + \varphi_b(f))], \quad (10.1.2'')$$

т.е. амплитудный спектр сигнала $\tilde{x}(t)$ – как результат преобразования Гильберта сигнала $x(t)$, не изменяется и остается равным амплитудному спектру сигнала $x(t)$. Фазовый спектр сигнала $\tilde{x}(t)$ (начальные фазовые углы всех гармонических составляющих сигнала) сдвигается на -90° при $f > 0$ и на 90° при $f < 0$ относительно фазового спектра сигнала $x(t)$. Но такой фазовый сдвиг и означает не что иное, как превращение косинусных гармоник в синусные, а синусных в косинусные. Последнее не трудно проверить на единичной гармонике.

Если $x(t) = \cos(2\pi f_0 t)$, то имеем следующее преобразование Гильберта через частотную область:

$$\tilde{x}(t) = H[x(t)] \Leftrightarrow \text{TF}[H[x(t)]] = -j \text{sgn}(f) \cdot [\delta(f+f_0) + \delta(f-f_0)]/2. \quad (10.1.4)$$

$$\tilde{X}(f) = -j \cdot [-\delta(f+f_0) + \delta(f-f_0)]/2 = j \cdot [\delta(f+f_0) - \delta(f-f_0)]/2. \quad (10.1.5)$$

Но последнее уравнение – спектр синусоиды. При обратном преобразовании Фурье:

$$\tilde{x}(t) = \text{TF}^{-1}[\tilde{X}(f)] = \sin(2\pi f_0 t). \quad (10.1.6)$$

При $x(t) = \sin(2\pi f_0 t)$ аналогичная операция дает $\tilde{x}(t) = -\cos(2\pi f_0 t)$. Знак минус демонстрирует отставание (запаздывание) выходного сигнала преобразования, как операции свертки, от входного сигнала. Для гармонических сигналов любой частоты с любой начальной фазой это запаздывание составляет четверть периода колебаний. На рис. 10.1.3 этот сдвиг на четверть периода для единичной гармонической составляющей (несущей частоты радиоимпульса) виден достаточно наглядно. Таким образом, преобразование Гильберта, по существу, представляет собой идеальный фазовращатель, осуществляющий фазовый сдвиг на 90° всех частотных составляющих сигналов одновременно.

Сдвиг фазы спектров сигналов $x(t)$ на $\pi/2$ определяет изменение четности и самих сигналов: четный $x(t) \Leftrightarrow$ нечетный $\tilde{x}(t)$, и наоборот.

Спектры каузальных функций. Каузальная (физически осуществимая) линейная система (равно как и произвольная причинно обусловленная функция) задается односторонним импульсным откликом (выражением) $h(t)$, $t \geq 0$, и имеет частотную характеристику $H(f)$:

$$H(f) = A(f) - j \cdot B(f),$$

где $A(f)$ и $B(f)$ - действительная (четная) и мнимая (нечетная) части частотной характеристики. Осуществим обратное преобразование Фурье для всех частей этого выражения:

$$h(t) = a(t) + b(t),$$

$$a(t) = \int_{-\infty}^{\infty} A(f) \cos(2\pi ft) df, \quad b(t) = \int_{-\infty}^{\infty} B(f) \sin(2\pi ft) df,$$

где $a(t)$ и $b(t)$ - соответственно четная и нечетная части импульсного отклика $h(t)$. Условие каузальности для импульсного отклика ($h(t) = 0$ при $t < 0$) будет выполнено, если при $t < 0$ функции $a(t)$ и $b(t)$ компенсируют друг друга. Тогда общее условие каузальности, как можно наглядно видеть на рис. 10.1.6, с учетом нечетности функции $b(t)$, запишется в следующем виде:

$$\begin{aligned} b(t) &= -a(t), \quad b(t) = -h(t)/2, \quad a(t) = h(t)/2, \quad t < 0, \\ b(t) &= 0, \quad a(t) = a(0) = h(0), \quad t = 0, \\ b(t) &= a(t) = h(t)/2, \quad t > 0. \end{aligned} \quad (10.1.7)$$

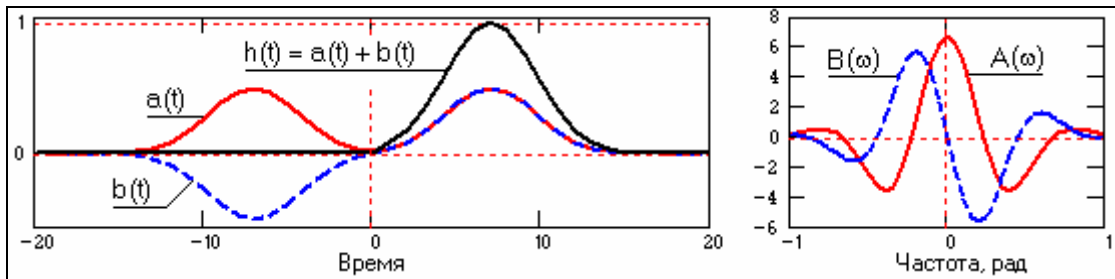


Рис. 10.1.6. Параметры каузальной функции.

Из этих условий следует, что нечетная функция $b(t)$ в каузальной системе однозначно связана с четной функцией:

$$b(t) = \text{sgn}(t) \cdot a(t), \quad (10.1.8)$$

Осуществляя преобразование Фурье обеих частей данного равенства при известном преобразовании сигнатурной функции ($\text{sgn}(t) \leftrightarrow j/\pi f$), получаем:

$$\text{Im}(H(f)) = (j/\pi f) * A(f),$$

или, с учетом знака мнимой части:

$$B(f) = -(1/\pi f) * A(f) = -(1/\pi) \int_{-\infty}^{\infty} [A(v)/(f-v)] dv. \quad (10.1.9)$$

Аналогично определяется и действительная компонента спектра по мнимой части:

$$A(f) = (1/\pi f) * B(f) = (1/\pi) \int_{-\infty}^{\infty} [B(v)/(f-v)] dv. \quad (10.1.10)$$

Таким образом, реальная и мнимая части спектра физически осуществимых (односторонних) систем, а равно и произвольных каузальных сигналов, связаны парой преобразований Гильберта. Они позволяют производить определение любой, действительной или мнимой, части частотной характеристики каузальной функции путем свертки другой ее части с функцией $1/\pi f$.

10.2. СВОЙСТВА ПРЕОБРАЗОВАНИЯ ГИЛЬБЕРТА [1, 2].

Для любых произвольных функций $x(t)$ и $y(t)$, имеющих Фурье – образы $X(\omega)$, $Y(\omega)$ и преобразования Гильберта $\tilde{x}(t) = \text{ТН}[x(t)]$ и $\tilde{y}(t) = \text{ТН}[y(t)]$, действительны следующие свойства:

Линейность. $\text{ТН}[a \cdot x(t) + b \cdot y(t)] = a \cdot \tilde{x}(t) + b \cdot \tilde{y}(t)$ при любых постоянных значениях коэффициентов a и b для любых произвольных функций $x(t)$ и $y(t)$.

Сдвиг. $\text{ТН}[x(t-a)] = \tilde{x}(t-a)$.

Преобразование константы, а в силу линейности преобразования, и постоянной составляющей сигнала, равно нулю. Это прямо следует из нечетности ядра преобразования Гильберта. Отсюда следует, что при преобразовании Гильберта из квадратурной составляющей исключается постоянная составляющая.

Свойство четности и нечетности определяется сдвигом всех гармоник сигнала на $\pi/2$, при этом четные сигналы $x(t)$ дают нечетные сигналы $\tilde{x}(t)$, и наоборот. Это действительно и для произвольных сигналов относительно их четных и нечетных частей.

Последовательное двойное преобразование Гильберта возвращает исходную функцию с обратным знаком $\text{TH}[\text{TH}[x(t)]] = \text{TH}[\tilde{x}(t)] = -x(t)$. Это определяется тем, что при двойном преобразовании фазы всех гармоники сигнала сдвигаются на π , что изменяет их знак гармоник. Однако в силу исключения из сигнала при первом преобразовании постоянной составляющей, при двойном преобразовании сигнал $x(t)$ восстанавливается с исключенным средним значением по интервалу задания.

Обратное преобразование Гильберта, по существу, это второе преобразование в последовательном двойном преобразовании Гильберта с изменением знака результата:

$$x(t) = \text{TH}^{-1}[\tilde{x}(t)] = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tilde{x}(\tau)}{t-\tau} d\tau = \tilde{x}(t) * (-1/\pi t). \quad (10.2.1)$$

Альтернативная форма вычисления $x(t)$ из $\tilde{x}(t)$:

$$x(t) = \text{TF}^{-1}[j \text{sgn}(f) \cdot \text{TF}[\tilde{x}(t)]]. \quad (10.2.1')$$

Подобие при изменении масштаба аргумента: $\text{TH}[x(at)] = \tilde{x}(at)$.

Энергетическая эквивалентность:

$$\int_{-\infty}^{\infty} x^2(t) dt = \int_{-\infty}^{\infty} \tilde{x}^2(t) dt. \quad (10.2.2)$$

Это следует из теоремы Парсеваля (энергия сигнала равна сумме энергии всех частотных составляющих сигнала) и равенства модулей спектров сигналов $x(t)$ и $\tilde{x}(t)$ (энергия сигнала не зависит от его фазочастотной характеристики).

Свойство ортогональности:

$$\int_{-\infty}^{\infty} x(t) \cdot \tilde{x}(t) dt = 0. \quad (10.2.3)$$

Если все косинусные составляющие сигнала $x(t)$ превращаются в ортогональные им синусные составляющие сигнала $\tilde{x}(t)$, а синусные – в ортогональные им косинусные, то и сигналы $x(t)$ и $\tilde{x}(t)$ должны быть ортогональны. Из теоремы Парсеваля следует:

$$\int_{-\infty}^{\infty} x(t) \cdot \tilde{x}(t) dt = \int_{-\infty}^{\infty} X^*(f) \cdot \tilde{X}(f) df.$$

Функция $X^*(f) \cdot \tilde{X}(f) = -X^* \cdot j \text{sgn}(f) \cdot X(f) = -j \text{sgn}(f) \cdot |X(f)|^2$ является нечетной, а поэтому определенный интеграл от этой функции по симметричным относительно нуля пределам равен нулю. Ортогональность сигналов наглядно видна на рис. 10.1.1.

Свойство свертки:

$$\text{TH}[x(t) * y(t)] = \tilde{x}(t) * y(t) = x(t) * \tilde{y}(t). \quad (10.2.4)$$

Это вытекает из следующих соображений. Примем $z(t) = x(t) * y(t)$, при этом:

$$Z(f) = X(f) \cdot Y(f), \quad \tilde{Z}(f) = -j \text{sgn}(f) \cdot Z(f) = -j \text{sgn}(f) X(f) \cdot Y(f).$$

$$\tilde{Z}(f) = [-j \text{sgn}(f) X(f)] \cdot Y(f) = \tilde{X}(f) \cdot Y(f) \Leftrightarrow \tilde{x}(t) * y(t).$$

$$\tilde{Z}(f) = X(f) \cdot [-j \text{sgn}(f) Y(f)] = X(f) \cdot \tilde{Y}(f) \Leftrightarrow x(t) * \tilde{y}(t).$$

Отсутствие коммутативности с преобразованием Фурье:

$$\text{TF}[\text{TH}[x(t)]] \neq \text{TH}[\text{TF}[x(t)]]. \quad (10.2.5)$$

Свойство модуляции: Модулирующие сигналы $u(t)$, как правило, имеют ограниченный спектр, максимальные частоты которого Ω много меньше значения несущей частоты ω_0 , при этом:

$$\text{TH}[u(t) \cdot \cos(\omega_0 t)] = u(t) \cdot \sin(\omega_0 t). \quad (10.2.6)$$

Для четных функций $u(t)$ это свойство очевидно. При переходе в частотную область:

$$\text{TH}[u(t) \cdot \cos(\omega_0 t)] \Leftrightarrow -j \cdot \text{sgn}(\omega) \cdot [U(\omega) * (\delta(\omega + \omega_0) + \delta(\omega - \omega_0))].$$

Множитель $-j \cdot \text{sgn}(\omega)$ является знаковой константой по ω и может быть внесен под интеграл свертки и умножен на $(\delta(\omega + \omega_0) + \delta(\omega - \omega_0))$, что, как уже рассматривалось ранее (см.

10.1.4 – 10.1.6), при обратном преобразовании Фурье дает $u(t) \cdot \sin(\omega_0 t)$.

Аналогично можно показать, что

$$\text{TH}[u(t) \cdot \sin(\omega_0 t)] = -u(t) \cdot \cos(\omega_0 t). \quad (10.2.7)$$

10.3. ВЫЧИСЛЕНИЕ ПРЕОБРАЗОВАНИЯ ГИЛЬБЕРТА [1,2,21].

Преобразование Гильберта аналоговых сигналов целесообразно выполнять не по формулам линейной свертки с оператором $1/\pi t$, который стремится к ∞ при $t \Rightarrow 0$, а через спектр аналитической функции:

$$z(t) = x(t) + j \cdot \tilde{x}(t) \Leftrightarrow X(f) + j \cdot \tilde{X}(f) = Z(f). \quad (10.3.1)$$

Заменяя в этом выражении функцию $\tilde{X}(f) = -j \operatorname{sgn}(f) \cdot X(f)$, получаем:

$$Z(f) = [1 + \operatorname{sgn}(f)] \cdot X(f), \quad (10.3.2)$$

где функция $1 + \operatorname{sgn}(f)$ равна 0 при $f < 0$, 1 при $f = 0$ и 2 при $f > 0$, при этом:

$$Z(f) = \begin{cases} 2 X(f), & f > 0, \\ X(0), & f = 0, \\ 0, & f < 0. \end{cases} \quad (10.3.2')$$

т.е. спектр функции $z(t)$ является односторонним и устанавливается непосредственно по спектру функции $x(t)$ при $f \geq 0$ (см. также (10.1.13)). Обратное преобразование Фурье функции $Z(f)$ должно давать комплексную функцию $z(t)$, при этом из (10.3.2') следует:

$$x(t) = \operatorname{Re} \left[2 \int_0^{\infty} X(f) \exp(j2\pi ft) df \right], \quad (10.3.3)$$

$$\tilde{x}(t) = \operatorname{Im} \left[2 \int_0^{\infty} X(f) \exp(j2\pi ft) df \right]. \quad (10.3.3')$$

В дискретной форме, при общем числе N отсчетов функции $x(t)$ с шагом Δt , с шагом по частоте $\Delta f = 1/(N\Delta t)$:

$$X(n\Delta f) = \Delta t \sum_{k=0}^{N-1} x(k\Delta t) \cdot \exp(-j2\pi kn/N), \quad n = 0, 1, \dots, N/2. \quad (10.3.4)$$

$$x(k\Delta t) = \Delta f \cdot \operatorname{Re} \left[X_0 + 2 \sum_{n=1}^{N/2} X(n\Delta f) \cdot \exp(j2\pi kn/N) \right]. \quad (10.3.5')$$

$$\tilde{x}(k\Delta t) = 2\Delta f \cdot \operatorname{Im} \left[\sum_{n=1}^{N/2} X(n\Delta f) \cdot \exp(j2\pi kn/N) \right]. \quad (10.3.5)$$

На рис. 10.3.1 приведен пример преобразования Гильберта, выполненный через частотную область. Естественно, что при реальном использовании преобразования Гильберта выполнять вычисления по (10.3.5') не требуется.

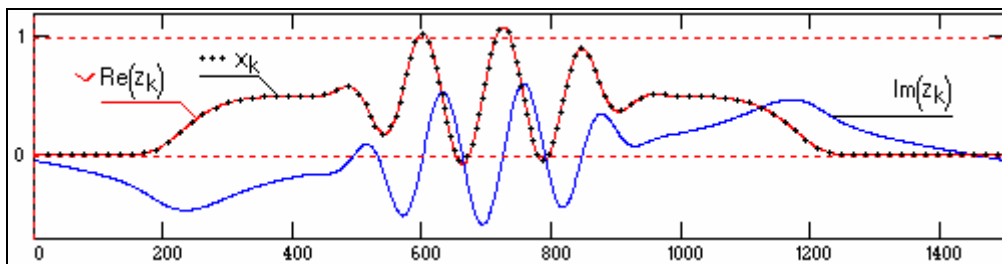


Рис. 10.3.1.

Оператор дискретного преобразования Гильберта $hb(k\Delta t) \Leftrightarrow 1/\pi t$ на интервале от $-T$ до T с шагом Δt можно получить обратным преобразованием Фурье частотной характеристики $Hb(f)$ (выражение 10.1.3) в интервале от $-f_N$ до f_N ($f_N = 1/2\Delta t$). При $\Delta t = 1$:

$$\begin{aligned} hb(k\Delta t) &= \int_{-f_N}^{f_N} Hb(f) \exp(j2\pi fk\Delta t) df = \int_{-1/2}^0 j \exp(j2\pi fk\Delta t) df - \int_0^{1/2} j \exp(j2\pi fk\Delta t) df = \\ &= [1/(2\pi k\Delta t)] \cdot [1 - \exp(-j\pi k\Delta t) - \exp(j\pi k\Delta t) + 1] = [1/(\pi k\Delta t)] \cdot [1 - (\exp(-j\pi k\Delta t) + \exp(j\pi k\Delta t))/2] = \\ &= [1/(\pi k\Delta t)] \cdot (1 - \cos(\pi k\Delta t)) = [2/(\pi k\Delta t)] \sin^2(\pi k\Delta t/2). \end{aligned} \quad (10.3.6)$$

$$\begin{aligned} \text{hb}(k\Delta t) &= 2/(\pi k\Delta t), \quad k = \pm 1, \pm 3, \pm 5, \dots, \\ \text{hb}(k\Delta t) &= 0, \quad k = \pm 0, \pm 2, \pm 4, \dots \end{aligned} \quad (10.3.6')$$

Нетрудно убедиться, что коэффициент усиления постоянной составляющей оператора равен нулю, а коэффициент усиления дисперсии помех равен 1.

В частотной области при выполнении преобразования Гильберта спектральных функций оператор свертки $\text{hb}(k\Delta f) \Leftarrow 1/\pi f$ не отличается от приведенного для временной области.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
2. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. - М.: Мир, 1989. - 540 с.
21. Рапопорт М.Б. Вычислительная техника в полевой геофизике: Учебник для вузов. - М.: Недра, 1993. - 350 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 10/6: АНАЛИТИЧЕСКИЕ СИГНАЛЫ

Нуль и единица - от Бога, все остальное - дело рук человеческих.

Леопольд Кронекер. Немецкий математик, XIX в.

И в этом остальном немалый вклад математиков. У них есть мания упрощения вычислений путем изобретения все новых и новых величин. Оно и понятно. Когда вводишь новое понятие, чувствуешь себя наравне с Богом.

Николай Пятин. Геофизик Уральской школы, XX в.

Содержание: 10.1. Понятие аналитического сигнала. Комплексное представление вещественных сигналов. Аналитический сигнал. Спектральная плотность аналитического сигнала. 10.2. Примеры применения аналитических сигналов. Огибающая и мгновенная фаза сигналов. Мгновенная частота. Огибающие модулированных сигналов. Анализ каузальных систем. Литература.

ВВЕДЕНИЕ

Аналитический сигнал – это один из способов комплексного представления сигнала, который применяется при анализе сигналов и систем их обработки. Он позволяет ввести в анализ понятия огибающей и мгновенной частоты сигнала.

10.1. ПОНЯТИЕ АНАЛИТИЧЕСКОГО СИГНАЛА [1,25].

Комплексное представление вещественных сигналов. При математическом анализе очень часто вместо вещественных сигналов с целью упрощения математического аппарата преобразований данных удобно использовать эквивалентное комплексное представление сигналов. Так, например, в теории электрических цепей вещественная запись синусоидального напряжения

$$u(t) = U_0 \cos(\omega_0 t + \varphi)$$

заменяется комплексной формой записи:

$$\dot{u}(t) = \dot{U}_0 \exp(j\omega_0 t), \quad \dot{U}_0 = U_0 \exp(j\varphi), \quad \text{при этом } u(t) = \operatorname{Re} \dot{u}(t).$$

В общем случае, произвольный динамический сигнал $s(t)$, заданный на определенном участке временной оси (как конечном, так и бесконечном) имеет комплексную двустороннюю спектральную плотность $S(\omega)$. При раздельном обратном преобразовании Фурье реальной и мнимой части спектра $S(\omega)$ сигнал $s(t)$ разделяется на четную и нечетную составляющие, которые являются двусторонними относительно $t = 0$, и суммирование которых полностью восстанавливает исходный сигнал. На рис. 10.1.1 приведен пример сигнала (А), его комплексного спектра (В) и получения четной и нечетной части сигнала из реальной и мнимой части спектра (С).

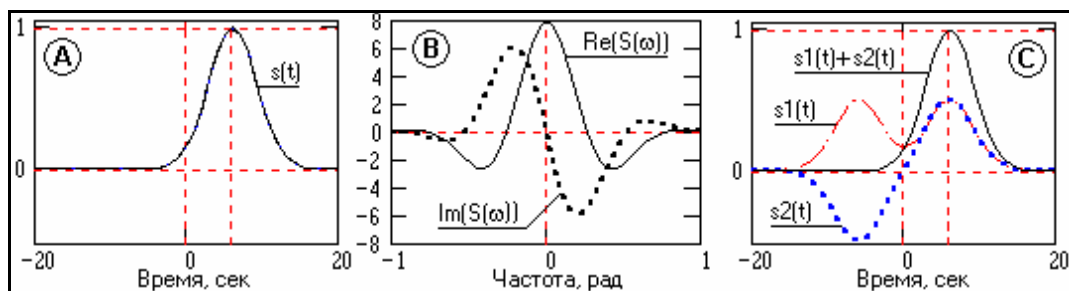


Рис. 10.1.1. Сигнал, спектральная плотность сигнала, четная и нечетная составляющие.

Аналитический сигнал. Можно выполнить обратное преобразование Фурье и в другой форме - раздельно для положительных и отрицательных частот спектра:

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^0 S(\omega) \cdot \exp(j\omega t) d\omega + \frac{1}{2\pi} \int_0^{\infty} S(\omega) \cdot \exp(j\omega t) d\omega. \quad (10.1.1)$$

Информация в комплексном спектре сигнала является избыточной. В силу комплексной сопряженности полную информацию о сигнале $s(t)$ содержит как левая (отрицательные частоты), так и правая (положительные частоты) часть спектра $S(\omega)$. Аналитическим сигна-

лом, отображающим вещественный сигнал $s(t)$, называют второй интеграл выражения (10.1.1), нормированный на π , т.е. обратное преобразование Фурье спектра сигнала $s(t)$ по положительным частотам:

$$z_s(t) = \frac{1}{\pi} \int_0^{\infty} S(\omega) \cdot \exp(j\omega t) d\omega. \quad (10.1.2)$$

Дуальность свойств преобразования Фурье определяет, что аналитический сигнал $z_s(t)$, полученный из односторонней спектральной функции, всегда является комплексным и может быть представлен в виде:

$$z_s(t) = \text{Re } z(t) + j \cdot \text{Im } z(t).$$

Аналогичное преобразование первого интеграла выражения (10.1.1) дает сигнал $z_s^*(t)$, комплексно сопряженный с сигналом $z(t)$:

$$z_s^*(t) = \text{Re } z(t) - j \cdot \text{Im } z(t),$$

что наглядно видно на рис. 10.1.2 при восстановлении сигналов по односторонним частям спектра, приведенного на рис. 10.1.1-В.

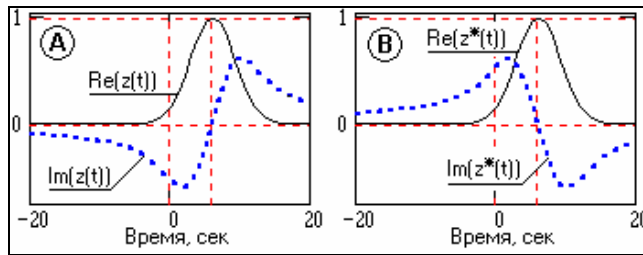


Рис. 10.1.2. Сигналы $z(t)$ и $z^*(t)$.

При сложении функций $z_s(t)$ и $z_s^*(t)$ с учетом нормировки в (10.1.2) только на $1/\pi$, а не на $1/2\pi$, как в (10.1.1), мы обязаны получить полный исходный сигнал $s(t)$:

$$s(t) = [z_s(t) + z_s^*(t)]/2 = \text{Re } z(t).$$

Отсюда следует, что реальная часть аналитического сигнала $z_s(t)$ равна самому сигналу $s(t)$.

Реальная и мнимая части спектра произвольных каузальных сигналов связаны преобразованием Гильберта. Оно позволяет производить определение любой части частотной характеристики каузальной функции, действительной или мнимой, путем свертки другой ее части с оператором Гильберта $1/\pi f$. Аналогично, мнимая часть аналитического сигнала $z_s(t)$ является аналитически сопряженной с его действительной частью $\text{Re } z(t) = s(t)$ через преобразование Гильберта, и называется *квадратурным дополнением* сигнала $s(t)$:

$$\text{Im}(z(t)) = \tilde{s}(t) = \text{TH}[s(t)] = s(t) * \text{hb}(t), \quad (10.1.3)$$

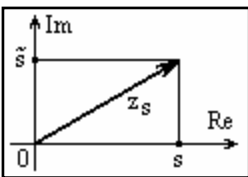
$$\text{hb}(t) = 1/(\pi t),$$

$$z_s(t) = s(t) + j \cdot \tilde{s}(t). \quad (10.1.4)$$

где индексом $\tilde{s}(t)$ обозначен сигнал, *аналитически сопряженный* с сигналом $s(t)$, $\text{hb}(t)$ – оператор Гильберта.

Таким образом, квадратурное дополнение сигнала $s(t)$ представляет собой свертку сигнала $s(t)$ с оператором $1/(\pi t)$ и может быть выполнено линейной системой с постоянными параметрами:

$$\tilde{s}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau, \quad (10.1.3')$$



Аналитический сигнал зависит от действительного аргумента, является однозначным и дифференцируемым. На комплексной плоскости он отображается вектором, модуль и фазовый угол которого изменяются от аргумента, а проекция сигнала на вещественную ось для любого значения аргумента равна значению исходного сигнала $s(t)$. Какой-либо новой ин-

формации аналитический сигнал не несет, так как получен линейным преобразованием из исходного сигнала и представляет собой его новую математическую модель.

Почему именно оператор Гильберта применяется для получения квадратурного дополнения сигнала? Какую физическую операцию он выполняет? Ответ на этот вопрос может быть получен при рассмотрении спектра аналитического сигнала.

Спектральная плотность аналитического сигнала, если он сформирован непосредственно во временной области, определяется обычным преобразованием Фурье:

$$Z_s(\omega) = \int_{-\infty}^{\infty} z_s(t) \exp(-j\omega t) dt.$$

Эта функция, с учетом определения аналитического сигнала по выражению (10.1.2), должна быть отлична от нуля только в области положительных частот, где ее значения (в силу нормировки на π , а не на 2π) должны быть равны удвоенным значениям спектральной плотности сигнала $s(t)$:

$$Z_s(\omega) = \begin{cases} 0, & \omega < 0, \\ 2 \cdot S(\omega), & \omega > 0. \end{cases} \quad (10.1.5)$$

С другой стороны, при непосредственном преобразовании Фурье левой и правой части формулы (10.1.4) аналитического сигнала $z_s(t)$, получаем:

$$Z_s(\omega) = S(\omega) + j\tilde{S}(\omega). \quad (10.1.6)$$

Данное выражение действительно для всей частотной оси (от $-\infty$ до $+\infty$) и должно быть равно выражению (10.1.5). А это означает, что левая часть спектра (отрицательные частоты ω) сигнала (10.1.6) должна быть обращена в ноль, аналогично формированию каузальной функции из ее четной и нечетной части. Это может быть выполнено следующим образом.

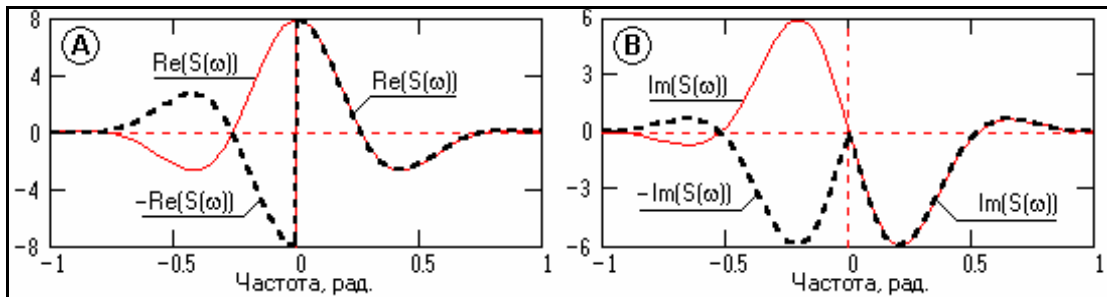
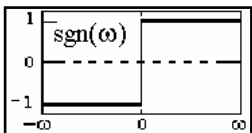


Рис. 10.1.3.

Если левые части спектра сигнала $S(\omega)$ умножить на -1 , обнулить реальную часть на частоте $\omega=0$ и оставить без изменения правые части спектра, то будут получены функции, показанные пунктиром на рис. 10.1.3), которые дают нули в левой части спектра при сложении с исходной функцией $S(\omega)$ и увеличивают в 2 раза правые части спектра. Такая операция может быть выполнена умножением спектра $S(\omega)$ на сигнатурную функцию $\text{sgn}(\omega)$:



$$\text{sgn}(\omega) = \begin{cases} -1, & \omega < 0, \\ 0, & \omega = 0, \\ 1, & \omega > 0. \end{cases} \quad (10.1.7)$$

Однако при этом реальная часть новой функции $\text{sgn}(\omega) \cdot S(\omega)$, как это можно видеть на рис. 10.1.3, становится нечетной, а мнимая часть четной, что не соответствует статусу спектральных функций. Для восстановления статуса полученный результат нужно дополнительно умножить на $-j$. Применяя для левой и правой части частотных аргументов индексирование соответственно ω_l и ω_r , можно записать подробные выражения для спектров:

$$S(\omega) = \text{Re } S(\omega_l) + j \cdot \text{Im}(\omega_l) + \text{Re } S(\omega_r) + j \cdot \text{Im}(\omega_r),$$

$$\tilde{S}(\omega) = j \cdot \text{Re } S(\omega_l) - \text{Im}(\omega_l) - j \cdot \text{Re } S(\omega_r) + \text{Im}(\omega_r).$$

При умножении квадратурной функции $\tilde{S}(\omega)$ на j (для выражения в (10.1.6)):

$$j \cdot \tilde{S}(\omega) = -\operatorname{Re} S(\omega_l) - j \cdot \operatorname{Im}(\omega_l) + \operatorname{Re} S(\omega_r) + j \cdot \operatorname{Im}(\omega_r).$$

Отсюда нетрудно видеть результат:

$$Z_s(\omega) = S(\omega) + j \tilde{S}(\omega) = 2 \cdot \operatorname{Re} S(\omega_r) + j \cdot 2 \cdot \operatorname{Im}(\omega_r) = 2 \cdot S(\omega_r),$$

что полностью соответствует выражению (10.1.5). В краткой форме:

$$\tilde{S}(\omega) = \begin{cases} jS(\omega), & \omega < 0, \\ 0, & \omega = 0, \\ -jS(\omega), & \omega > 0. \end{cases} \quad \tilde{S}(\omega) = -j \cdot \operatorname{sgn}(\omega) \cdot S(\omega), \quad (10.1.8)$$



$$Hb(\omega) = -j \cdot \operatorname{sgn}(\omega) = \begin{cases} j, & \omega < 0 \\ 0, & \omega = 0 \\ -j, & \omega > 0 \end{cases} \quad (10.1.9)$$

Таким образом, спектральная плотность $\tilde{S}(\omega)$ аналитически сопряженного сигнала $\tilde{s}(t)$ образуется из спектра $S(\omega)$ исходного сигнала $s(t)$ умножением на функцию $-j \cdot \operatorname{sgn}(\omega)$. Это обеспечивает при суммировании $S(\omega) + j \tilde{S}(\omega)$ удвоение амплитуд частотных составляющих в области положительных частот и их взаимную компенсацию в области отрицательных частот.

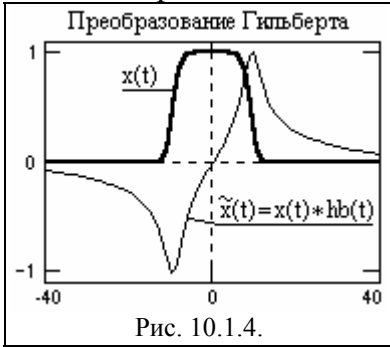


Рис. 10.1.4.

Из выражения (10.1.8) в спектральной области непосредственно следует соответствующая связь функций $s(t)$ и $\tilde{s}(t)$ во временной области:

$$\tilde{s}(t) = s(t) * hb(t), \quad (10.1.10)$$

$$s(t) = -\tilde{s}(t) * hb(t). \quad (10.1.11)$$

где $hb(t) = \operatorname{TF}[-j \cdot \operatorname{sgn}(\omega)] = 1/(\pi t)$ – обратное преобразование Фурье функции $-j \cdot \operatorname{sgn}(\omega)$.

Пример преобразования сигнала $x(t)$ оператором Гильберта для формирования аналитического сигнала $z_x(t) = x(t) + j \cdot \tilde{x}(t)$ приведен на рис. 10.1.4.

Частотную характеристику оператора Гильберта (10.1.9) можно записать и в следующем виде:

$$Hb(\omega) = |Hb(\omega)| \cdot \exp(j\varphi_h(\omega)), \quad \text{где } |Hb(\omega)| = 1.$$

$$Hb(\omega) = -j \cdot \operatorname{sgn}(\omega) = \begin{cases} \exp(j(-\pi/2)), & \omega > 0 \\ 0, & \omega = 0, \\ \exp(j(\pi/2)), & \omega < 0 \end{cases} \quad (10.1.12)$$

Если спектр функции $x(t)$ также представить в форме

$$S(\omega) = |S(\omega)| \cdot \exp(j\varphi_s(\omega)),$$

то выражение (10.1.8) преобразуется к следующей форме:

$$\tilde{S}(\omega) = |S(\omega)| \cdot \exp(j\varphi_s(\omega)) \cdot \exp(j\varphi_h(\omega)) = |S(\omega)| \cdot \exp[j(\varphi_s(\omega) + \varphi_h(\omega))], \quad (10.1.8')$$

т.е. модуль $|S(\omega)|$ – амплитудный спектр сигнала $\tilde{s}(t)$ как результат преобразования Гильберта сигнала $s(t)$, не изменяется и остается равным амплитудному спектру сигнала $s(t)$. Фазовый спектр сигнала $\tilde{s}(t)$ (начальные фазовые углы всех гармонических составляющих сигнала) сдвигается на -90° при $\omega > 0$ и на 90° при $\omega < 0$ относительно фазового спектра сигнала $s(t)$. Но такой фазовый сдвиг означает не что иное, как превращение косинусных гармоник в синусные, а синусных в косинусные. Это можно наглядно видеть на единичной гармонике. Так, если $x(t) = \cos(2\pi f_0 t)$, то имеем следующее преобразование Гильберта через частотную область:

$$\begin{aligned} \tilde{x}(t) &= \operatorname{TH}[x(t)] \Leftrightarrow \operatorname{TF}[\operatorname{TH}[x(t)]] = -j \operatorname{sgn}(f) \cdot [\delta(f+f_0) + \delta(f-f_0)]/2. \\ \tilde{X}(f) &= -j \cdot [-\delta(f+f_0) + \delta(f-f_0)]/2 = j \cdot [\delta(f+f_0) - \delta(f-f_0)]/2. \end{aligned}$$

Но последнее уравнение - спектр синусоиды. При обратном преобразовании Фурье:

$$\tilde{x}(t) = \text{TF}^{-1}[\tilde{X}(f)] = \sin(2\pi f_0 t).$$

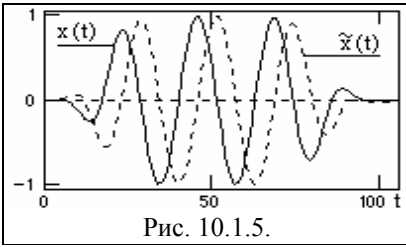


Рис. 10.1.5.

При $x(t) = \sin(2\pi f_0 t)$ аналогичная операция дает $\tilde{x}(t) = -\cos(2\pi f_0 t)$. Знак минус демонстрирует отставание (запаздывание) выходного сигнала преобразования, как операции свертки, от входного сигнала. Для гармонических сигналов любой частоты с любой начальной фазой это запаздывание составляет четверть периода колебаний. На рис. 10.1.5 этот сдвиг на четверть периода для единичной гармонической составляющей (несущей частоты радиоимпульса) виден достаточно наглядно.

Таким образом, аналитический сигнал, по существу, представляет собой двух ортогональных сигналов, все гармонические составляющие которых сдвинуты по фазе на 90° друг относительно друга.

10.2. ПРИМЕРЫ ПРИМЕНЕНИЯ АНАЛИТИЧЕСКИХ СИГНАЛОВ [1,2].

Огибающая и мгновенная фаза сигналов. Допустим, что имеем зарегистрированный радиоимпульсный сигнал $x(t)$ с несущей частотой ω_0 , который содержит определенную информацию, заключенную в огибающей сигнала $u(t)$ и его фазе $\varphi(t)$:

$$x(t) = u(t) \cos(\omega_0 t + \varphi(t)). \quad (10.2.1)$$

Требуется выделить информационные составляющие сигнала. Запишем выражение (10.2.1) в другой форме:

$$x(t) = a(t) \cdot \cos(\omega_0 t) + b(t) \cdot \sin(\omega_0 t), \quad (10.2.2)$$

где функции $a(t)$ и $b(t)$ называются низкочастотными квадратурными составляющими сигнала $x(t)$:

$$a(t) = u(t) \cos \varphi(t), \quad b(t) = u(t) \sin \varphi(t). \\ u(t) = \sqrt{a^2(t) + b^2(t)}, \quad \text{tg } \varphi(t) = b(t)/a(t).$$

С использованием преобразования Гильберта из сигнала $x(t)$ можно сформировать аналитически сопряженный сигнал $\tilde{x}(t)$. Математическую форму сигнала $\tilde{x}(t)$ получим из выражения (10.2.2) с учетом свойства модуляции преобразования Гильберта:

$$\tilde{x}(t) = a(t) \cdot \sin(\omega_0 t) - b(t) \cdot \cos(\omega_0 t).$$

$$z(t) = x(t) + j \cdot \tilde{x}(t).$$

Квадрат модуля сигнала $z(t)$:

$$|z(t)|^2 = x^2(t) + \tilde{x}^2(t) = a^2(t)[\cos^2(\omega_0 t) + \sin^2(\omega_0 t)] + b^2(t)[\cos^2(\omega_0 t) + \sin^2(\omega_0 t)] = u^2(t).$$

Отсюда, огибающая $u(t)$ и мгновенная фаза $\phi(t)$ сигнала $x(t)$:

$$u(t) = \sqrt{x^2(t) + \tilde{x}^2(t)}. \quad (10.2.3)$$

$$\phi(t) = \omega_0 t + \varphi(t) = \text{arctg}[\tilde{x}(t)/x(t)]. \quad (10.2.4)$$

$$\varphi(t) = \phi(t) - m_0 t.$$

Мгновенная частота сигнала определяется по скорости изменения мгновенной фазы:

$$d\phi(t)/dt = \frac{\tilde{x}'(t) \cdot x(t) - x'(t) \cdot \tilde{x}(t)}{x^2(t) + \tilde{x}^2(t)}. \quad (10.2.5)$$

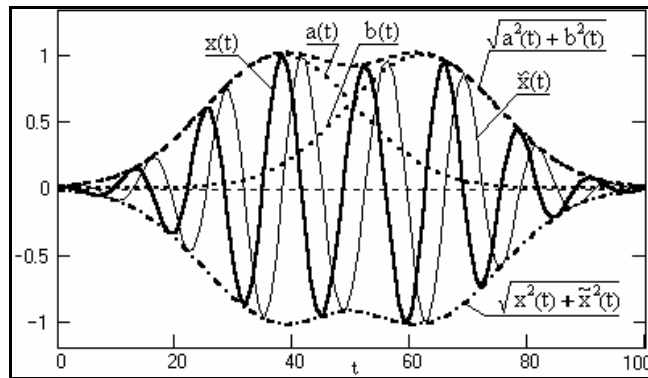


Рис. 10.2.1.

Для амплитудно-модулированных сигналов с одной несущей частотой эти результаты достаточно очевидны (см. рис. 10.2.1). Но выражения (10.2.3-10.2.5), полученные из общих соображений, остаются действительными и для любых произвольных сигналов.

На рис. 10.2.2. представлен сигнал, сложенный двумя гармониками:

$$x(t) = a(t) \cdot \cos(\omega_1 t) + b(t) \cdot \cos(\omega_2 t).$$

Квадратурное дополнение и аналитический сигнал:

$$\tilde{x}(t) = a(t) \cdot \sin(\omega_1 t) + b(t) \cdot \sin(\omega_2 t).$$

$$z(t) = x(t) + j \cdot \tilde{x}(t).$$

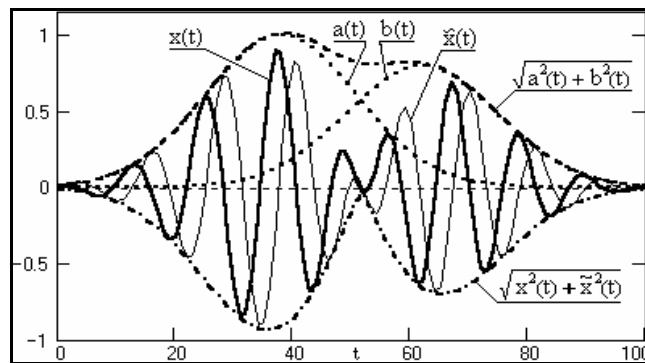


Рис. 10.2.2.

Огибающая такого сигнала, как это можно видеть на рисунке 10.2.2, должна вычисляться по формуле (10.2.3). При этом для данного сигнала получаем:

$$u(t) = \sqrt{a^2(t) + b^2(t) + 2 \cdot a(t) \cdot b(t) \cdot \cos((\omega_2 - \omega_1)t)},$$

что может существенно отличаться от функции $\sqrt{a^2(t) + b^2(t)}$.

Мгновенная фаза сигнала, график которой приведен на рис. 10.2.3, зависит от времени нелинейно:

$$\phi(t) = \arg \operatorname{tg} \frac{a(t) \cdot \sin(\omega_1 t) + b(t) \cdot \sin(\omega_2 t)}{a(t) \cdot \cos(\omega_1 t) + b(t) \cdot \cos(\omega_2 t)}.$$

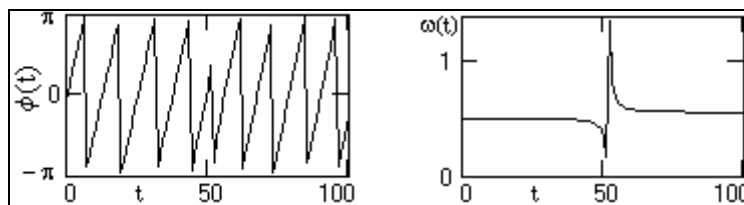


Рис. 10.2.3.

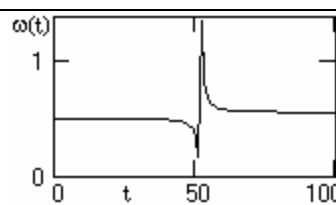


Рис. 10.2.4.

Мгновенная частота сигнала (рис. 10.2.4) также имеет нелинейную зависимость от времени, причем ее значения могут существенно превышать даже суммарное значение частот, составляющих сигнал:

$$\omega(t) = \frac{\omega_1 a^2(t) + \omega_2 b^2(t) + a(t)b(t) \cdot (\omega_1 + \omega_2) \cdot \cos((\omega_2 - \omega_1)t)}{a^2(t) + b^2(t) + 2 \cdot a(t)b(t) \cdot \cos((\omega_2 - \omega_1)t)}$$

Аналогичная методика определения огибающих, мгновенных значений фазы и частоты применяется и для анализа случайных процессов.

Огибающие модулированных сигналов. В качестве примера применения огибающих рассмотрим связь форм относительно узкополосных радиосигналов с формой модулирующих сообщений.

Амплитудная модуляция. Уравнение модулированного сигнала:

$$x(t) = U_0 \cdot [1 + m \cdot s(t)] \cdot \cos \omega_0 t, \quad s(t) \leq 1, \quad m \leq 1$$

Квадратурное дополнение и аналитический сигнал:

$$\tilde{x}(t) = U_0 \cdot [1 + m \cdot s(t)] \cdot \sin \omega_0 t, \quad z_x(t) = x(t) + j\tilde{x}(t).$$

Огибающая сигнала $x(t)$:

$$u(t) = |z_x(t)| = U_0 \cdot [1 + m \cdot s(t)],$$

т.е. точно повторяет форму модулирующего сообщения (см. рис. 10.2.5)

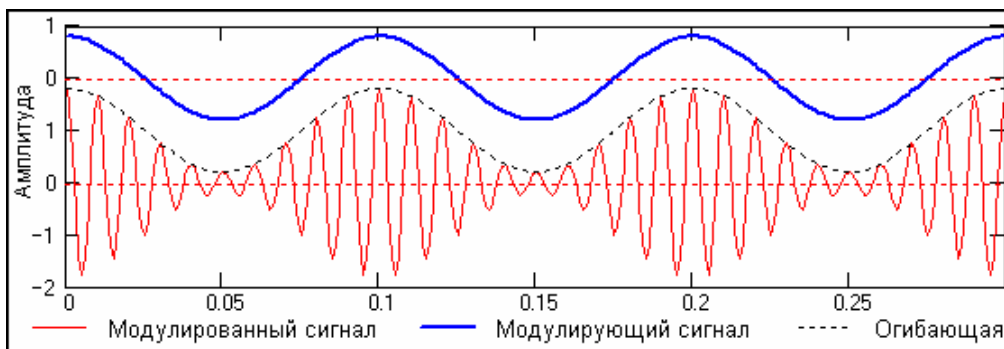


Рис. 10.2.5. Амплитудная модуляция.

Балансная модуляция. Уравнение модулированного сигнала, приведенного на рис. 10.2.6:

$$x(t) = U_0 \cdot s(t) \cdot \cos \omega_0 t,$$

Квадратурное дополнение, аналитический сигнал, огибающая сигнала $x(t)$:

$$\tilde{x}(t) = U_0 \cdot s(t) \cdot \sin \omega_0 t, \quad z_x(t) = x(t) + j\tilde{x}(t), \quad u(t) = |z_x(t)| = U_0 \cdot |s(t)|.$$

Огибающая сигнала $x(t)$ существенно отличается от модулирующего сообщения, но связана с ним простым соотношением.

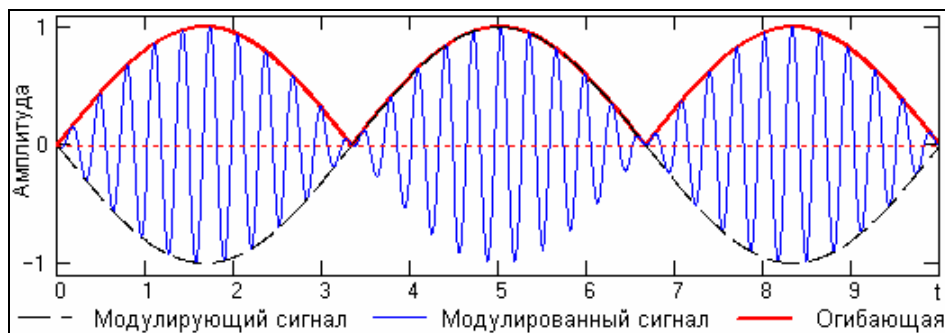


Рис. 10.2.6. Балансная модуляция.

Анализ каузальных систем. Каузальная (физически осуществимая) линейная система задается односторонним импульсным откликом $h(t)$, $t \geq 0$, и имеет частотную характеристику $H(f)$:

$$H(f) = X(f) - jY(f),$$

Осуществим обратное преобразование Фурье для всех частей выражения отдельно:

$$h(t) = x(t) + y(t),$$

$$x(t) = \int_{-\infty}^{\infty} X(f) \cos(2\pi ft) df,$$

$$y(t) = \int_{-\infty}^{\infty} Y(f) \sin(2\pi ft) df,$$

где $x(t)$ и $y(t)$ - четная и нечетная части функции $h(t)$. Нечетная функция $y(t)$ в каузальной системе однозначно связана с четной функцией $x(t)$:

$$y(t) = \operatorname{sgn}(t) \cdot x(t). \quad (10.2.6)$$

Осуществляя обратное преобразование Фурье обеих частей равенства (10.2.6) при известном преобразовании сигнатурной функции ($\operatorname{sgn}(t) \Leftrightarrow -j/(\pi f)$), получаем:

$$\operatorname{TF}[y(t)] = (-j/\pi f) * X(f) = (-j/\pi) \int_{-\infty}^{\infty} [X(u)/(f-u)] du.$$

Отсюда:

$$Y(f) = (1/\pi) \int_{-\infty}^{\infty} [X(u)/(f-u)] du = \operatorname{TH}[X(f)],$$

т.е. мнимая часть спектра импульсного отклика каузальной системы (и любой каузальной функции) является преобразованием Гильберта действительной части спектра. Соответственно, уравнение для определения действительной компоненты спектра по мнимой части:

$$X(f) = -\operatorname{TH}[Y(f)] = -(1/\pi) \int_{-\infty}^{\infty} [Y(u)/(f-u)] dv.$$

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
2. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. - М.: Мир, 1989. - 540 с.
25. Сергиенко А.Б. Цифровая обработка сигналов. - СПб.: Питер, 2003. - 608 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 11: СТАЦИОНАРНЫЕ ЛИНЕЙНЫЕ СИСТЕМЫ

Такова взаимосвязь: раз существует гадание, значит должны быть и боги; а раз существуют боги, значит должно быть и гадание.

Марк Туллий Цицерон. О гадании.
Римский философ и политик, I в.д.н.э.

Геология, наряду с медициной и богословием, относится к точным наукам. Есть геологический объект - должен быть сигнал. Есть сигнал - должен быть объект.

Владимир Мартынов. Основы гадания в точных науках.
Мордовский геофизик Уральской школы, XX в.

Содержание: Введение. 11.1. Линейные системы. Общие понятия систем. Линейные системы. Основные системные операции. Инвариантность систем к сдвигу. Математическая модель системы. Нерекурсивные цифровые системы. Рекурсивные цифровые системы. Стационарные и нестационарные системы. 11.2. Импульсная характеристика системы. Импульсный отклик системы. Реакция системы на произвольный сигнал. Усиление постоянной составляющей сигнала. Усиление шумов. Определение импульсной реакции. 11.3. Передаточные функции цифровых систем. Z-преобразование. Устойчивость систем. 11.4. Частотные характеристики систем. Основные свойства. 11.5. Реакция систем на случайные сигналы. Математическое ожидание. Корреляционные соотношения. Спектральные соотношения. Дисперсия выходного сигнала. Функция когерентности. 11.6. Структурные схемы систем. Структурные схемы. Графы систем. Соединения систем. Схемы реализации систем. Обращенные формы. Литература.

ВВЕДЕНИЕ

Преобразование и обработка сигналов осуществляется в системах. Понятия сигнала и системы неразрывны, так как любой сигнал существует в какой-либо системе его обращения. Система обработки сигналов может быть реализована как в материальной форме (специальное устройство, измерительный прибор и т.п.), так и программно на ЭВМ или на любом другом вычислительном устройстве. Существуют и комплексные измерительно-вычислительные системы (ИВС), которые выполняют как регистрацию и первичную обработку сигналов непосредственно в материальной форме их представления, так и преобразование сигналов в цифровую форму, и последующую программную обработку. Форма реализации систем существенного значения не имеет и определяет только их возможности при анализе и обработке сигналов. Основное внимание при рассмотрении данной темы будем уделять цифровым системам и дискретной математике их отображения и анализа, применяя аналитическую математику при рассмотрении общих вопросов, если последнее упрощает изложение и понимание теоретического материала.

11.1. ЛИНЕЙНЫЕ СИСТЕМЫ [1,2,9,14,18,20]

Общие понятия систем. Безотносительно к назначению и исполнению система всегда имеет *вход*, на который подается входной сигнал или входное воздействие, в общем случае многомерное, и *выход*, с которого снимается обработанный выходной сигнал. Если устройство системы и внутренние операции преобразований принципиального значения не имеют, то система в целом может восприниматься как “черный ящик”, в формализованном виде. Формализованная система представляет собой определенный *системный оператор* (алгоритм) преобразования входного сигнала – *воздействия* $s(t)$, в сигнал на выходе системы $y(t)$ – *отклик* или *выходную реакцию* системы. Символическое обозначение операции преобразования (трансформации):

$$y(t) = T[s(t)].$$

Системный оператор T - это правило (набор правил, алгоритм) преобразования сигнала $s(t)$ в сигнал $y(t)$. Для общеизвестных операций преобразования сигналов применяются также расширенные символы операторов трансформации, где вторым символом и специальными индексами обозначается конкретный вид операции (как, например, TF - преобразование Фурье, TF^{-1} - обратное преобразование Фурье).

Входной сигнал системы может представлять собой m - мерный вектор (m входных сигналов), а выходной сигнал n - мерный вектор, при этом система будет иметь m входов и n выходов. Пример такой системы в геофизике: трехканальный гамма-спектрометр, на три

входа решающего блока которого поступают сигналы от калиевого, радиевого и ториевого каналов спектрометра, а на три выхода выводятся сигналы содержания калия, урана и тория, при этом системный оператор реализует алгоритм решения системы трех линейных уравнений с тремя неизвестными.

Для детерминированных входных сигналов соотношение между выходными и входными сигналами однозначно задается системным оператором. В случае реализации на входе системы случайного входного процесса также существует однозначное соответствие процессов на выходе и входе системы, однако при этом одновременно происходит изменение статистических характеристик выходного сигнала (математического ожидания, дисперсии, корреляционной функции и пр.), которое также определяется системным оператором.

Для определения системы необходимо задать характер, тип и области допустимых величин входных и выходных сигналов. Как правило, системы выполняются на сигналы одного типа по входу/выходу и подразделяются на системы непрерывного времени (аналоговые или дискретные сигналы на входе и выходе) и цифровые системы. Совокупность системного оператора T и пространства сигналов образует математическую модель системы.

Линейные системы. Любые преобразования сигналов сопровождаются изменением их спектра и по характеру этих изменений разделяются на два вида: линейные и нелинейные. К нелинейным относят изменения, при которых в составе спектра сигналов появляются новые гармонические составляющие. При линейных изменениях сигналов изменяются амплитуды и/или начальные фазы гармонических составляющих спектра. Оба вида изменений могут происходить как с сохранением полезной информации, так и с ее искажением. Это зависит не только от характера изменения спектра сигналов, но и от спектрального состава самой полезной информации.

Линейные системы составляют основной класс систем обработки сигналов. Термин линейности означает, что система преобразования сигналов должна иметь произвольную, но в обязательном порядке линейную связь между входным сигналом (возбуждением) и выходным сигналом (откликом). В нелинейных системах связь между входным и выходным сигналом определяется произвольным нелинейным законом.

Система считается линейной, если в пределах установленной области входных и выходных сигналов ее реакция на входные сигналы аддитивна (выполняется принцип суперпозиции сигналов) и однородна (выполняется принцип пропорционального подобия).

Принцип *аддитивности* требует, чтобы реакция на сумму двух входных сигналов была равна сумме реакций на каждый сигнал в отдельности:

$$T[a(t)+b(t)] = T[a(t)]+T[b(t)].$$

Принцип *однородности* или пропорционального подобия требует сохранения однозначности масштаба преобразования при любой амплитуде входного сигнала:

$$T[c \times a(t)] = c \times T[a(t)].$$

Другими словами, отклик линейной системы на взвешенную сумму входных сигналов должен быть равен взвешенной сумме откликов на отдельные входные сигналы независимо от их количества и для любых весовых коэффициентов, в том числе комплексных.

Примеры.

1. Система $y(t) = a^2t$. $y(t_1) = a^2t_1$, $y(t_2) = a^2t_2$, $y(ct) = a^2ct$.
 $y(t_1+t_2) = a^2(t_1+t_2) = a^2t_1+a^2t_2 = y(t_1)+y(t_2)$. Система аддитивна.
 $cy(t) = ca^2t = a^2ct = y(ct)$. Система однородна. И в целом линейна.
2. Система $y(t) = at^2$. $y(t_1) = at_1^2$, $y(t_2) = at_2^2$, $y(ct) = a(ct)^2 = ac^2t^2$.
 $y(t_1+t_2) = a(t_1+t_2)^2 \neq y(t_1)+y(t_2) = at_1^2+at_2^2$. Система не аддитивна.
 $cy(t) = c at^2 \neq y(ct) = ac^2t^2$. Система неоднородна. И в целом нелинейна.

При программной реализации линейных систем на ЭВМ особых затруднений с обеспечением линейности в разумных пределах значений входных и выходных сигналов, как правило, не возникает. При физической (аппаратной) реализации систем обработки данных диапазон входных и/или выходных сигналов, в котором обеспечивается линейность преобра-

зования сигналов, всегда ограничен и должен быть специально оговорен в технической документации или методической инструкции.

Основные системные операции. К базовым линейным операциям, из которых могут быть сформированы любые линейные операторы преобразования, относятся операции скалярного умножения, сдвига и сложения сигналов:



$$y(t) = b \times x(t),$$

$$y(t) = x(t-\Delta t),$$

$$y(t) = a(t)+b(t).$$

Графическое отображение операций (цифровая форма) приведено на рис. 11.1.1.

Отметим, что операции сложения и умножения являются линейными только для аналоговых и дискретных сигналов. В случае цифровых сигналов они линейны относительно самих цифровых сигналов, но если последние - результат операции амплитудно-цифрового преобразования, то сложение и умножение не может считаться линейным абсолютно точно по отношению к исходным сигналам.

Для систем, с размерностью 2 и более существует также еще одна базовая операция, которая называется операцией *пространственного маскирования*, которая может рассматриваться как обобщение скалярного умножения. Так, для двумерных систем:

$$z(x,y) = c(x,y) \cdot u(x,y),$$

где $u(x,y)$ – двумерный входной сигнал, $c(x,y)$ – пространственная маска постоянных (весовых) коэффициентов. Пространственное маскирование представляет собой поэлементное произведение значений сигнала с коэффициентами маски.

Инвариантность систем к сдвигу. Система называется инвариантной к сдвигу (инвариантной во времени, а равно и по любым другим аргументам), если сдвиг входного сигнала по аргументам вызывает соответствующий сдвиг выходного сигнала:

$$s(x,t) = T[a(x,t)], \quad T[a(x-\Delta x,t-\Delta t)] = s(x-\Delta x,t-\Delta t).$$

Линейность и инвариантность к сдвигу являются независимыми свойствами систем и не определяют друг друга. Так, например, операция квадратирования сигнала (возведения в квадрат всех значений сигнала) инвариантна к сдвигу, но нелинейна.

В теории анализа и обработки данных основное место занимают системы, линейные и инвариантные к сдвигу (ЛИС - системы). Они обладают достаточно широкими практически возможными возможностями при относительной простоте математического аппарата. В дальнейшем, если это специально не оговаривается, будем иметь в виду именно такие системы.

Преимущество, которое отдается ЛИС - системам в методах обработки информации, базируется на возможности разложения входного сигнала любой, сколь угодно сложной формы, на составляющие простейших форм, отклик системы на которые известен и хорошо изучен, с последующим вычислением выходного сигнала в виде суммы откликов на все составляющие входного сигнала. В качестве простейших форм разложения сигналов используются, как правило, единичные импульсы и гармонические составляющие. Первая применяется при представлении сигнала в динамической форме и использует преобразование свертки, вторая - частотное представление сигнала и преобразование Фурье.

Другой важной особенностью ЛИС - систем является то, что любые их комбинации также являются ЛИС - системами, а любую сложную ЛИС - систему можно разложить на комбинации простых систем. Так, например, при последовательном (каскадном) соединении систем, когда выходной сигнал одной системы служит входным сигналом для второй и т.д., образуемая система в целом также является ЛИС - системой, если линейны и инвариантны к сдвигу все системы, в нее входящие, при этом по отношению к общей системной операции преобразования порядок соединения входящих в нее систем значения не имеет.

Математическая модель системы задается связью между сигналами входа и выхода и в аналоговой одномерной линейной системе обычно выражается линейным дифференциальным уравнением:

$$\sum_{m=0}^M a_m \frac{d^m y(t)}{dt^m} = \sum_{n=0}^N b_n \frac{d^n s(t)}{dt^n}. \quad (11.1.1)$$

При нормировке к $a_0 = 1$, отсюда следует:

$$y(t) = \sum_{n=0}^N b_n \frac{d^n s(t)}{dt^n} - \sum_{m=1}^M a_m \frac{d^m y(t)}{dt^m}. \quad (11.1.1')$$

По существу, правой частью этого выражения в самой общей математической форме отображается содержание операции преобразования входного сигнала, т.е. задается оператор трансформации входного сигнала в выходной.

Аналогичная связь в цифровой системы описывается разностными уравнениями:

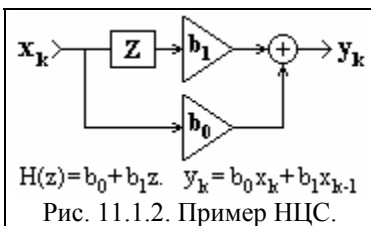
$$\sum_{m=0}^M a_m y((k-m)\Delta t) = \sum_{n=0}^N b_n s((k-n)\Delta t). \quad (11.1.2)$$

$$y(k\Delta t) = \sum_{n=0}^N b_n s((k-n)\Delta t) - \sum_{m=1}^M a_m y((k-m)\Delta t). \quad (11.1.2')$$

Последнее уравнение можно рассматривать как алгоритм последовательного вычисления значений $y(k\Delta t)$, $k = 0, 1, 2, \dots$, по значениям входного сигнала $s(k\Delta t)$ и предыдущих вычисленных значений $y(k\Delta t)$ при известных значениях коэффициентов a_m , b_n и с учетом задании определенных начальных условий - значений $s(k\Delta t)$ и $y(k\Delta t)$ при $k < 0$. Интервал дискретизации в цифровых последовательностях отсчетов обычно принимается равным 1, т.к. выполняет только роль масштабного множителя.

Нерекурсивные цифровые системы. При нулевых значениях коэффициентов a_m уравнение (11.1.2') переходит в уравнение дискретной свертки $x(k)$ с оператором b_n :

$$y(k) = \sum_{n=0}^N b_n x(k-n). \quad (11.1.3)$$



При установленных значениях коэффициентов b_n значения выходных отсчетов свертки для любого аргумента k определяются текущим и "прошлыми" значениями входных отсчетов. Такая система называется нерекурсивной цифровой системой (НЦС). Пример простейшей НЦС приведен на рис. 11.1.2. Интервал суммирования по n получил название "окна" системы.

Окно системы (11.1.3) составляет $N+1$ точку, система является односторонней каузальной, причинно обусловленной текущими и "прошлыми" значениями входного сигнала, выходной сигнал не опережает входного. Каузальная система может быть реализована аппаратно в реальном масштабе времени. При $k < n$ проведение обработки входных данных возможно только при задании определенных начальных условий для точек $x(-k)$, $k=1, 2, \dots, N$. Как правило, в качестве начальных условий принимаются нулевые значения или значения отсчета $x(0)$. Применяется также четное или нечетное продление функции $x(k)$ на интервал отрицательных значений k . Если при обработке данных начальные интервалы массивов $x(k)$ существенного значения не имеют, то обработку можно начинать с отсчета $k=N$.

При обработке данных на ЭВМ ограничение по каузальности системного оператора снимается. В программном распоряжении системы могут находиться как "прошлые", так и "будущие" значения входных отсчетов, при этом уравнение (11.1.3) будет иметь вид:

$$y(k) = \sum_{n=-N'}^N b_n x(k-n). \quad (11.1.3')$$

При $N' = N$ система называется двусторонней симметричной. Симметричные системы, в отличие от каузальных, не изменяют фазы обрабатываемых сигналов.

Техника выполнения свертки в координатной области не отличается от техники выполнения обычной дискретной свертки двух массивов данных.

Представим, что на одной полоске бумаги выписаны по порядку сверху вниз значения данных $x(k)$. На второй полоске бумаги находятся записанные в обратном порядке значения

коэффициентов системы b_n . Для вычисления $y(k)$ располагаем вторую полосу против первой таким образом, чтобы значение b_0 совпало со значением $x(k)$, перемножаем все значения b_n с расположенными против них значениями $x(k-n)$ и суммируем результаты перемножения. Результат суммирования является выходным значением сигнала $y(k)$. Сдвигаем окно системы - полосу коэффициентов b_k , на один отсчет последовательности $x(k)$ вниз (по порядку возрастания номеров k) или массив $x(k)$ сдвигаем на отсчет вверх и вычисляем аналогично следующее значение, и т.д.

Описанный процесс свертки в вещественной области массива данных $x(k)$ с нерекурсивным оператором системы b_n (массивом весовых коэффициентов системы) обычно называют нерекурсивной цифровой фильтрацией данных, а саму систему, если она выполняет только данную операцию, нерекурсивным цифровым фильтром (НЦФ).

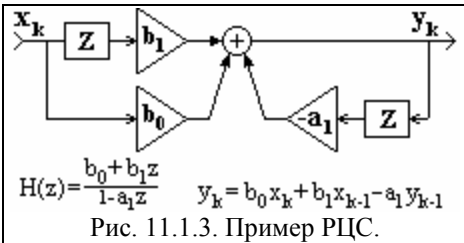
Пример. Сглаживание данных скользящим симметричным П-окном размером 5 отсчетов.

Уравнение НЦФ: $y_k = \sum_{n=-2}^2 b_n x_{k-n}$, $b_n = 0,2$. Начальные условия - нулевые.

Входной сигнал: $x_k = \{0,0,0,0,10,0,10,0,0,0\}$. Выходной сигнал: $y_k = \{0,0,2,2,4,4,4,2,2,0,0\}$.

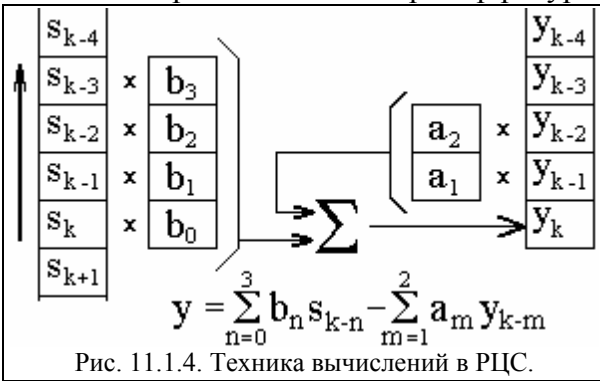
Проверьте результат (выполните фильтрацию сигнала).

Заметим: сумма коэффициентов b_n сглаживающего НЦФ равна 1, при этом сумма значений выходного сигнала равна сумме значений входного сигнала. Сигнал "размазался" по аргументу.



Рекурсивные цифровые системы. Системы, которые описываются полным разностным уравнением (11.1.2), принято называть рекурсивными цифровыми системами (РЦС) или рекурсивными цифровыми фильтрами (РЦФ), так как в вычислении текущих значений выходного сигнала участвует не только входной сигнал, но и значения выходного сигнала, вычисленные в предшествующих циклах расчетов.

С учетом последнего фактора рекурсивные системы называют системами с обратной связью. Пример рекурсивной системы приведен на рис. 11.1.3.



Полное окно рекурсивной системы состоит из двух составляющих: нерекурсивной части b_n , аналогичной окну нерекурсивной системы и ограниченной в работе текущими и "прошлыми" значениями входного сигнала (при реализации на ЭВМ возможно использование и "будущих" отсчетов сигнала), и рекурсивной части a_m , которая работает только с "прошлыми", ранее вычисленными значениями выходного сигнала. Техника вычислений для РЦС приведена на рис. 11.1.4.

Пример. Уравнение РЦС: $y_k = b_0 x_k + a_1 y_{k-1}$, при $b_0 = a_1 = 0,5$, $y_{-1} = 0$. Входной сигнал: $x_k = \{0, 10, 0, 0, 0, \dots\}$.

Расчет выходного сигнала:

$y_0 = 0,5x_0 + 0,5y_{-1} = 0$ $y_1 = 0,5x_1 + 0,5y_0 = 5$ $y_2 = 0,5x_2 + 0,5y_1 = 2,5$ $y_3 = 0,5x_3 + 0,5y_2 = 1,25$

$y_4 = 0,5x_4 + 0,5y_3 = 0,625$ $y_5 = 0,5x_5 + 0,5y_4 = 0,3125$ и т.д.

Выходной сигнал: $y_k = \{0, 5, 2,5, 1,25, 0,625, 0,3125, 0,15625, \dots\}$

Из примера можно видеть, что реакция РЦС на конечный входной сигнал, в принципе, может иметь бесконечную длительность, в отличие от реакции НЦС, которая всегда ограничена количеством членов b_k (окном системы).

Пример. Уравнение РЦС: $y_k = b_0 x_k + a_1 y_{k-1} + a_2 y_{k-2}$, при $b_0 = a_1 = a_2 = 1$. $x_k = \{0, 1, 0, 0, 0, \dots\}$, $y_{-1} = y_{-2} = 0$.

Расчет выходного сигнала:

$y_0 = x_0 + y_{-1} + y_{-2} = 0$ $y_1 = x_1 + y_0 + y_{-1} = 1$ $y_2 = x_2 + y_1 + y_0 = 1$ $y_3 = x_3 + y_2 + y_1 = 2$

$y_4 = x_4 + y_3 + y_2 = 3$ $y_5 = x_5 + y_4 + y_3 = 5$ $y_6 = x_6 + y_5 + y_4 = 8$ $y_7 = x_7 + y_6 + y_5 = 13$ и т.д.

Заметим: сумма коэффициентов обратной связи a_i больше 1 и выходной сигнал идет "в разнос".

Пример. Интегрирование данных по формуле трапеций.

Уравнение РЦФ: $y_k = (x_k + x_{k-1})/2 + y_{k-1}$, начальные условия - нулевые.

Входной сигнал: $x_k = \{0, 0, 2, 2, 8, 0, 4, 4, 0, \dots\}$.

Выполните фильтрацию. Контроль: $y_k = \{0, 0, 1, 3, 8, 12, 14, 18, 20, 20, \dots\}$.

Стационарные и нестационарные системы. Система считается *стационарной* и имеет постоянные параметры, если ее свойства (математический алгоритм оператора преобразования) в пределах заданной точности не зависят от входного и выходного сигналов и не изменяются ни во времени, ни от каких-либо других внешних факторов. Математически это означает задание системы уравнениями типа (11.1.1-2) с постоянными значениями коэффициентов a_j и b_i и реакция системы на какое-либо воздействие не зависит от времени (координат) его приложения. В противном случае система является нестационарной или *параметрической* (системой с переменными параметрами). Среди последних большое значение имеют так называемые адаптивные системы обработки данных. В этих системах производится, например, оценивание определенных параметров входных и выходных сигналов, по результатам сравнения которых осуществляется подстройка параметров преобразования (переходной характеристики системы) таким образом, чтобы обеспечить оптимальные по производительности условия обработки сигналов или минимизировать погрешность обработки.

11.2. ИМПУЛЬСНАЯ ХАРАКТЕРИСТИКА СИСТЕМЫ [1,5]

Импульсный отклик системы. По определению, импульсными характеристиками систем (второй широко используемый термин - импульсный отклик систем) называются функции $h(t)$ для аналоговых и $h(k\Delta t)$ для цифровых систем, которые являются реакцией (откликом) систем на единичные входные сигналы: дельта-функцию $\delta(t)$ для аналоговых и импульс Кронекера $\delta(k\Delta t)$ для цифровых систем, поступающие на вход систем соответственно при $t=0$ и $k=0$. Эта реакция однозначно определяется оператором преобразования:

$$y(t) = T[\delta(t)] \equiv h(t). \quad (11.2.1)$$

$$y(k\Delta t) = T[\delta(k\Delta t)] \equiv h(k\Delta t). \quad (11.2.1')$$

Импульсный отклик аналоговой системы, как результат операции над дельта-функцией, в определенной степени представляет собой математическую абстракцию идеального преобразования. С практической точки зрения под импульсным откликом аналоговой системы можно понимать математическое отображение реакции системы на импульсный входной сигнал произвольной формы с площадью, равной 1, если длительность сигнала пренебрежимо мала по сравнению со временем реакции системы или с периодом ее собственных колебаний. Под временем (длиной) реакции системы обычно понимают интервал, на котором значения функции $h(t)$ существенно отличаются от нуля после прекращения действия единичного сигнала на ее входе.

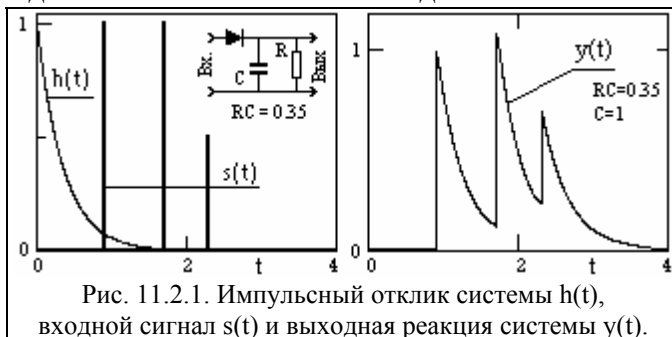


Рис. 11.2.1. Импульсный отклик системы $h(t)$, входной сигнал $s(t)$ и выходная реакция системы $y(t)$.

Для цифровых систем импульсный отклик однозначно определяется реакцией системы на импульс Кронекера $\delta(k\Delta t) = 1$ при $k=0$.

Функцию импульсного отклика называют также *весовой функцией* системы.

На рисунке 11.2.1 приведен пример импульсного отклика $h(t)$ интегрирующей RC-цепи. При подаче на вход RC-цепи импульса заряда Δq емкость C заряжается до

напряжения $V_0 = \Delta q/C$ и начинает разряжаться через сопротивление R , при этом напряжение на ней изменяется по закону $v(t) = V_0 \cdot e^{-t/RC} = (\Delta q/C) \cdot e^{-t/RC}$. Отсюда, отклик RC-цепи по выходному напряжению на входной сигнал $\Delta q = 1$: $h(t) = (1/C) \cdot e^{-t/RC}$. По существу, импульсным откликом системы $h(t)$ определяется доля входного сигнала, которая действует на выходе системы по истечении времени t после поступления сигнала на вход (запаздывающая реак-

ция системы).

Реакция системы на произвольный сигнал. Если функция импульсного отклика системы известна, то, с учетом принципа суперпозиции сигналов в линейной системе, можно выполнить расчет реакции системы в любой произвольный момент времени на любое количество входных сигналов с любыми моментами времени их прихода путем суммирования запаздывающих реакций системы на эти входные сигналы, как это показано на рис. 11.2.1 для трех входных импульсов. В общем случае произвольный сигнал на входе системы может быть разложен в линейную последовательность взвешенных единичных импульсов:

$$y(t) = T[s(t)] \equiv T\left[\int_{-\infty}^{\infty} s(\tau)\delta(t-\tau) d\tau\right]. \quad (11.2.2)$$

На основании принципа суперпозиции линейный оператор T может быть внесен под знак интеграла, т.к. последний представляет собой предельное значение суммы. При этом операция преобразования действует только по переменной t :

$$y(t) = \int_{-\infty}^{\infty} s(\tau) T[\delta(t-\tau)] d\tau = \int_{-\infty}^{\infty} s(\tau) h(t-\tau) d\tau. \quad (11.2.3)$$

Это выражение представляет собой интеграл Дюамеля или свертку (конволюцию) входного сигнала с импульсной характеристикой системы. Заменой переменных $t-\tau = \tau$ можно убедиться в том, что эта операция, как и положено свертке, коммутативна:

$$\int_{-\infty}^{\infty} s(\tau) h(t-\tau) d\tau \equiv \int_{-\infty}^{\infty} h(\tau) s(t-\tau) d\tau.$$

Аналогично, для дискретных сигналов:

$$y(k\Delta t) = \sum_{n=-\infty}^{\infty} s(n\Delta t) h(k\Delta t - n\Delta t) \equiv \sum_{n=-\infty}^{\infty} h(n\Delta t) s(k\Delta t - n\Delta t). \quad (11.2.3')$$

В символической форме математического представления:

$$y(t) = s(\tau) * h(t-\tau) \equiv s(t-\tau) * h(\tau) \equiv s(t) * h(t).$$

В реальных физических системах импульсный отклик $h(t)$ равен нулю при $t < 0$ (реакция на выходе системы не может опережать входной сигнал) и, как правило, отличен от нуля только на определенном интервале τ , по которому и ведется интегрирование или суммирование в выражениях свертки. При обработке данных на ЭВМ требований по односторонности импульсного отклика не предъявляется, равно как и по его размерам вперед и назад от нуля по координатам.

Усиление постоянной составляющей сигнала. Подадим на вход системы постоянный сигнал $s(t) = A$. При этом сигнал на выходе системы:

$$y(t) = \int_{-\infty}^{\infty} h(\tau) s(t-\tau) d\tau = A \int_{-\infty}^{\infty} h(\tau) d\tau = A \cdot K_{пс}, \quad (11.2.4)$$

т.е. площадь импульсного отклика (для цифровой системы соответственно сумма коэффициентов импульсного отклика) является коэффициентом $K_{пс}$ усиления постоянной составляющей входного сигнала. Если при обработке сигналов должны изменяться только динамические характеристики их формы без изменения постоянной составляющей, а равно и различных постоянных уровней (фона, пьедесталов, региональных трендов и т.п.), то площадь импульсного отклика (сумма коэффициентов) должна нормироваться к единице.

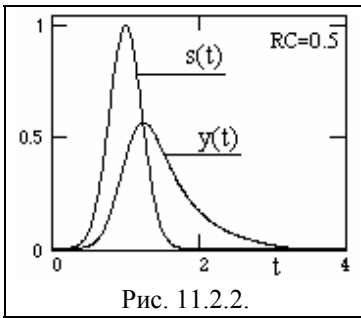


Рис. 11.2.2.

На рис. 11.2.2 приведен пример выполнения свертки рассмотренной нами выше RC-цепью при нормированной к 1 площади импульсного отклика $h(\tau)$. Входной сигнал $s(t)$ находится на постоянном фоновом значении, в данном случае - нулевом, при этом, как и следовало ожидать, площадь выходного сигнала $y(t)$ равна площади входного сигнала.

Усиление шумов. Критерием качества системы при использовании любого метода обработки информации можно считать выполнение целевого назначения с минимальным усилением шумов (максимальным их подавлением). Допустим, что система имеет нормированный к 1 импульсный отклик $h(k)$. Обозначим через $\varepsilon(k)$ аддитивный шум с математическим ожиданием $M\{\varepsilon(k)\} = \bar{\varepsilon} = 0$ и дисперсией $D_\varepsilon = \sigma^2$, который в сумме с сигналом поступает на вход системы. Значения $\varepsilon(k)$ статистически независимы и некоррелированы с сигналом. С учетом помехи во входном сигнале значение сигнала на выходе системы:

$$y(k) = \sum_n h(n)[x(k-n) + \varepsilon(k-n)].$$

Математическое ожидание значений выходного сигнала:

$$M\{y(k)\} = \sum_n h(n)[x(k-n) + M\{\varepsilon(k-n)\}] = \sum_n h(n) x(k-n).$$

Вычислим дисперсию распределения отсчетов выходного сигнала:

$$D\{y(k)\} = M\left\{\left[\sum_n h(n)[x(k-n) + \varepsilon(k-n)] - M\{y(k)\}\right]^2\right\} = M\left\{\left[\sum_n h(n) \varepsilon(k-n)\right]^2\right\}.$$

Если правую часть последнего выражения представить в виде

$$M\left\{\left[\sum_n h(n) \varepsilon(k-n)\right] \cdot \left[\sum_m h(m) \varepsilon(k-m)\right]\right\},$$

то в этом выражении математические ожидания всех членов произведения с сомножителями $\varepsilon(n)\varepsilon(m)$ при $n \neq m$ равны 0 в силу статистической независимости значений шума. Остаются только члены с $n = m$, т.е.:

$$M\left\{\sum_n h^2(n) \varepsilon^2(n)\right\} = \sum_n h^2(n) M\{\varepsilon^2(n)\} = D_\varepsilon \sum_n h^2(n) = \sigma^2 \sum_n h^2(n). \quad (11.2.5)$$

Отсюда следует, что сумма квадратов значений нормированного импульсного отклика системы представляет собой коэффициент усиления аддитивных шумов во входном сигнале.

Пример. Сглаживающий фильтр: $y(k) = 0.2 \sum_{n=-2}^2 x(k-n)$.

Коэффициент усиления шумов: $5(0.2^2) = 0.2$. Дисперсия шумов уменьшается в $1/0.2 = 5$ раз.

Для систем с m входами и n выходами аналогично определяются парциальные импульсные отклики $h_{ij}(t)$, $i = \{1, 2, \dots, n\}$, $j = \{1, 2, \dots, m\}$, каждым из которых отображается сигнал на i -м выходе при поступлении сигнала $\delta(t)$ на j -й вход. Полная совокупность импульсных откликов образует матрицу:

$$\bar{h}(t) = \begin{vmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,m} \\ h_{2,1} & h_{2,2} & \dots & h_{2,m} \\ \cdot & \cdot & \cdot & \cdot \\ h_{n,1} & h_{n,2} & \dots & h_{n,m} \end{vmatrix},$$

а выражение свертки приобретает вид:

$$\bar{U}_{\text{вых}}(t) = \int_{-\infty}^{\infty} \bar{h}(t-\tau) \bar{u}_{\text{вх}}(\tau) dt.$$

Определение импульсной реакции требуется, как правило, для рекурсивных систем, так как импульсная реакция для НЦС специального определения не требует:

$$h(k) = \sum_{n=-N}^N b(n)\delta(k-n) \equiv b(k).$$

Если выражение для системы известно в общей форме (11.1.2), определение импульсной реакции производится подстановкой в уравнение системы импульса Кронекера с координатой $k = 0$ при нулевых начальных условиях, при этом сигнал на выходе системы будет представлять собой импульсную реакцию системы: $y(k) \equiv h(k)$.

Пример. Уравнение РЦС: $y_k = x_k + 0.5y_{k-1}$. Входной сигнал: $x_k = \delta_0 = \{1, 0, 0, 0, \dots\}$.

Расчет выходного сигнала при нулевых начальных условиях:

$$y_0 = x_0 + 0.5 y_{-1} = 1 + 0 = 1 = h_0. \quad y_1 = x_1 + 0.5 y_0 = 0 + 0.5 = 0.5 = h_1. \quad y_2 = x_2 + 0.5 y_1 = 0 + 0.25 = 0.25 = h_2$$

$$y_3 = x_3 + 0.5 y_2 = 0.125 = h_3. \quad y_4 = x_4 + 0.5 y_3 = 0.0625 = h_4, \quad \text{т.д.}$$

Импульсный отклик системы: $h_k = \{1, 0.5, 0.25, 0.125, \dots\} \equiv (0.5)^k, \quad k = 0, 1, 2, \dots$

Определение импульсной реакции физических систем обычно производится подачей на вход систем ступенчатой функции $u_0(k) = 1$ при $k \geq 0$, и $u_0(k) = 0$ при $k < 0$:

$$g(k) = \sum_{n=0}^k h(n) u_0(k-n) = \sum_{n=0}^k h(n).$$

$$h(k) = g(k) - g(k-1), \quad k=0, 1, 2, \dots$$

Функция $g(k)$ получила название переходной характеристики системы (перехода из одного статического состояния в другое).

11.3. ПЕРЕДАТОЧНЫЕ ФУНКЦИИ ЦИФРОВЫХ СИСТЕМ [5,18].

Z-преобразование является удобным методом решения разностных уравнений линейных систем. Применяя z-преобразование к обеим частям равенства (11.1.2), получаем:

$$Y(z) \sum_{m=0}^M a_m z^m = X(z) \sum_{n=0}^N b_n z^n, \quad (11.3.1)$$

где $X(z), Y(z)$ - соответствующие z-образы входного и выходного сигнала. Из этого выражения, полагая $a_0 = 1$, получаем в общей форме функцию связи входа и выхода системы - уравнение *передаточной* функции системы (или *системной* функции) в z-области:

$$H(z) = Y(z)/X(z) = \sum_{n=0}^N b_n z^n / (1 + \sum_{m=1}^M a_m z^m). \quad (11.3.2)$$

Для нерекурсивных систем при $a_m = 0$:

$$H(z) = \sum_{n=0}^N b_n z^n. \quad (11.3.3)$$

При подаче на вход системы единичного импульса Кронекера δ_0 , имеющего z-образ $\delta(z) = z^0 = 1$, сигнал на выходе системы будет представлять собой импульсную реакцию системы $y(k) = h(k)$, при этом:

$$H(z) \equiv Y(z) = Y(z)/\delta(z) = TZ[y(k)] = \sum_{k=0}^{\infty} h(k)z^k, \quad (11.3.4)$$

т.е. передаточная функция системы является z-образом ее импульсной реакции. При обратном z-преобразовании передаточной функции соответственно получаем импульсную характеристику системы:

$$h(k) \Leftrightarrow H(z). \quad (11.3.5)$$

Если функция $H(z)$ представлена конечным степенным полиномом, как это обычно имеет место для НЦС, то обратное z-преобразование элементарно. Передаточная функция РЦС также может быть представлена степенным полиномом прямым делением числителя на знаменатель правой части выражения (11.3.2), однако результат при этом может оказаться как конечным, так и бесконечным, т.е. система может иметь либо конечную, либо бесконечную импульсную характеристику. Системы с бесконечной импульсной характеристикой получили название БИХ-систем, с конечной импульсной характеристикой соответственно КИХ-систем. Нерекурсивные системы всегда имеют конечную импульсную характеристику, т.к. длительность импульсной реакции НЦС определяется окном фильтра.

Пример. Передаточная функция рекурсивной системы: $H(z) = (1-z^5)/(1-z)$.

Делением числителя на знаменатель получаем: $H(z) = 1+z+z^2+z^3+z^4$. $H(z) \Leftrightarrow h(n) = \{1,1,1,1,1\}$.

Система имеет конечную импульсную характеристику.

Пример. Передаточная функция: $H(z)=1/(1-2z)$. Методом обратного преобразования: $h(n) = 2^n$.

Система имеет бесконечную импульсную характеристику.

Устойчивость систем. Любая практическая система должна быть *устойчивой*, т.е. для сигналов, конечных по энергии или средней мощности, выходные сигналы также должны быть конечными по этим параметрам. Система называется устойчивой, если при любых начальных условиях реакция системы на любое ограниченное воздействие также ограничена.

Для конечного по энергии входного сигнала, можно записать:

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

Отсюда следует условие, при котором выходной сигнал системы также будет ограниченным:

$$\int_{-\infty}^{\infty} |h(\tau)| d\tau < \infty, \quad (11.3.6)$$

т.е. необходимым и достаточным условием устойчивости системы является абсолютная сходимость ее импульсной характеристики, или, для цифровых систем, абсолютная суммируемость импульсного отклика:

$$\sum_n |h(n)| < \infty. \quad (11.3.6')$$

Анализ устойчивости может быть проведен по передаточной функции. В устойчивой системе значение $H(z)$ должно быть конечным во всех точках z -плоскости, где $|z| \leq 1$, а, следовательно, передаточная функция не должна иметь особых точек (полюсов) при $z \leq 1$ (на и внутри единичного круга на z -плоскости). Полюсы определяются корнями многочлена знаменателя передаточной функции $H(z)$.

Пример.

Передаточная функция РЦФ: $H(z) = 1/(1+1.25z)$. Корни знаменателя: $z = -0,8$. $|z| < 1$.

Фильтр неустойчив.

Передаточная функция: $H(z) = (1-z^3)/(1-0.6z+0.25z^2)$. Корни: $z_{1,2} = 1.2+1.6j$. $|z| = 2 > 1$.

Фильтр устойчив.

Приведенный критерий устойчивости относится к несократимой дроби, т.к. в противном случае возможна компенсация полюса нулем передаточной функции и следует проверить наличие однозначных нулей и полюсов.

Проверка на устойчивость требуется только для рекурсивных цифровых фильтров (систем с обратной связью), нерекурсивные системы всегда устойчивы.

11.4. ЧАСТОТНЫЕ ХАРАКТЕРИСТИКИ СИСТЕМ [2,5].

Для линейных систем, принимая в качестве сигнала на входе системы собственную функцию $x(k\Delta t) = B(\omega)\exp(j\omega k\Delta t)$, мы вправе ожидать на выходе системы сигнал $y(k\Delta t) = A(\omega)\exp(j\omega k\Delta t)$. Подставляя эти выражения в разностное уравнение системы (11.1.2), получаем:

$$\begin{aligned} \sum_{m=0}^M a_m A(\omega)\exp(j\omega k\Delta t - j\omega m\Delta t) &= \sum_{n=0}^N b_n B(\omega)\exp(j\omega k\Delta t - j\omega n\Delta t). \\ A(\omega)\exp(j\omega k\Delta t) \sum_{m=0}^M a_m \exp(-j\omega m\Delta t) &= B(\omega)\exp(j\omega k\Delta t) \sum_{n=0}^N b_n \exp(-j\omega n\Delta t). \\ A(\omega) \sum_{m=0}^M a_m \exp(-j\omega m\Delta t) &= B(\omega) \sum_{n=0}^N b_n \exp(-j\omega n\Delta t). \end{aligned} \quad (11.4.1)$$

Отсюда, частотная передаточная функция системы (частотная характеристика при нормировке к $a_0=1$):

$$H(\omega) = A(\omega)/B(\omega) = \sum_{n=0}^N b_n \exp(-j\omega n\Delta t) / \left[1 + \sum_{m=1}^M a_m \exp(-j\omega m\Delta t) \right]. \quad (11.4.2)$$

Нетрудно убедиться, что подстановкой $z = \exp(-j\omega\Delta t)$ в выражение передаточной функции $H(z)$ (11.3.2) может быть получено абсолютно такое же выражение для частотной характеристики, т.е.:

$$H(\omega) \equiv H(z) \text{ при } z = \exp(-j\omega\Delta t).$$

При обратном преобразовании $H(z)$ во временную область с использованием выражений (11.3.4-5) отсюда следует также, что частотная характеристика системы представляет собой Фурье-образ ее импульсной реакции, и наоборот. При $\Delta t = 1$:

$$H(\omega) = \sum_{n=-\infty}^{\infty} h(n) \exp(-j\omega n), \quad (11.4.3)$$

$$h(n) = (1/2\pi) \int_{-\pi}^{\pi} H(\omega) \exp(j\omega n) d\omega. \quad (11.4.4)$$

В общем случае $H(\omega)$ является комплексной функцией, модуль которой $R(\omega)$ называется амплитудно-частотной характеристикой системы (АЧХ), а аргумент $\varphi(\omega)$ - фазочастотной характеристикой (ФЧХ).

$$A(\omega) = |H(\omega)| = \sqrt{\operatorname{Re}^2 H(\omega) + \operatorname{Im}^2 H(\omega)}.$$

$$\varphi(\omega) = \operatorname{arctg}(\operatorname{Im} H(\omega)/\operatorname{Re} H(\omega)).$$

Физический смысл частотной характеристики системы достаточно прост. Произвольный сигнал на входе системы может рассматриваться в виде суммы гармонических составляющих с различным набором амплитуд и начальных фазовых углов. Амплитудно-частотной характеристикой системы устанавливаются коэффициенты усиления системой (коэффициенты передачи) этих частотных составляющих, а фазочастотной характеристикой - сдвиг фаз этих частотных составляющих в выходном сигнале относительно начальных фаз во входном сигнале.

Основные свойства частотных характеристик систем:

1. Частотные характеристики являются непрерывными функциями частоты.
2. При дискретизации данных по интервалам Δt функция $H(\omega)$ является периодической. Период функции $H(\omega)$ равен частоте дискретизации входных данных $F = 1/\Delta t$. Первый низкочастотный период (по аргументу ω от $-\pi/\Delta t$ до $\pi/\Delta t$, по f от $-1/2\Delta t$ до $1/2\Delta t$) называется главным частотным диапазоном передачи сигнала. Граничные частоты главного частотного диапазона соответствуют частоте Найквиста $\pm\omega_N$, $\omega_N = \pi/\Delta t$. Частота Найквиста определяет предел частотной разрешающей способности системы по обработке данных.
3. Для систем с вещественными коэффициентами импульсной реакции $h(n\Delta t)$ функция АЧХ является четной, а функция ФЧХ - нечетной. С учетом этого частотные характеристики систем обычно задаются только на интервале положительных частот $0-\omega_N$ главного частотного диапазона. Значения функций на интервале отрицательных частот являются комплексно сопряженными со значениями на интервале положительных частот.

11.5. РЕАКЦИЯ СИСТЕМ НА СЛУЧАЙНЫЕ СИГНАЛЫ [2,14].

Если сигнал на входе линейной системы является детерминированным, то, при известных параметрах системы, его соотношение с выходным сигналом является однозначным. Таким же однозначным является соотношение процессов на входе и выходе и для случайных сигналов, однако в силу природы последних явное представление, как входного сигнала, так и отклика системы, не представляется возможным. Для описания отклика системы необходимо использовать статистический подход. При рассмотрении данной темы ограничимся только физически реализуемыми системами с односторонним импульсным откликом $h(t)$ ($h(t)=0$ при $t<0$) и соответствующей частотной характеристикой $H(f)$. Если параметры входного сигнала специально не оговариваются, то по умолчанию принимается, что на вход системы поступает реализация случайного стационарного процесса $x(t)$ с нулевым средним и

вызывает сигнал $y(t)$ на выходе системы.

Квазидетерминированный сигнал в какой-то мере позволяет оценить сохранение однозначности преобразования системой случайных сигналов.

Допустим, что система имеет импульсный отклик $h(t) = \exp(-at)$, $t \geq 0$. Квазидетерминированный случайный сигнал стационарен, не обладает свойством эргодичности, но может быть описан в явной математической форме. Зададим сигнал на входе системы следующего вида:

$$x(t) = A + \cos(2t+\varphi),$$

где A и φ - взаимно независимые случайные величины, причем φ равномерно распределена в интервале $[0, 2\pi]$. Выходной сигнал определится выражением:

$$y(t) = h(\tau) * x(t-\tau) \equiv \int_0^{\infty} h(\tau) \cdot x(t-\tau) d\tau = A/3 + [3 \cos(2t+\varphi) + 2 \sin(2t+\varphi)]/13.$$

Из выражения следует, что выходной сигнал системы также является случайным процессом и содержит те же самые случайные параметры, что и входной сигнал, а, следовательно, для него также могут быть определены статистические характеристики.

Математическое ожидание произвольного случайного стационарного сигнала $x(t)$ на выходе линейной системы определится выражением:

$$\bar{y} = M\{y(t)\} = \int_0^{\infty} M\{x(t-\tau)\} \cdot h(\tau) d\tau = \bar{x} \int_0^{\infty} h(\tau) d\tau = \bar{x} K_{\text{пс}}. \quad (11.5.1)$$

Отсюда следует, что математическое ожидание выходных сигналов системы равно математическому ожиданию входных сигналов, умноженному на коэффициент усиления системой постоянной составляющей. Если система не пропускает постоянную составляющую сигналов (площадь или сумма коэффициентов импульсного отклика системы равна нулю), то случайный выходной сигнал всегда будет иметь нулевое математическое ожидание.

Корреляционные соотношения. Для произведения выходных сигналов $y(t)$ и $y(t+\tau)$ линейной системы можно записать:

$$y(t) \cdot y(t+\tau) = \int_0^{\infty} \int_0^{\infty} h(\alpha) h(\beta) x(t-\alpha) x(t+\tau-\beta) d\alpha d\beta.$$

Если взять математические ожидания от обеих частей этого равенства, то, с учетом соотношения в подынтегральном выражении

$$M\{x(t-\alpha) x(t+\tau-\beta)\} = -R_x(t-\alpha-t+\tau+\beta) = R_x(\tau+\alpha-\beta),$$

получим:

$$R_y(\tau) = \int_0^{\infty} \int_0^{\infty} h(\alpha) h(\beta) R_x(\tau+\alpha-\beta) d\alpha d\beta \equiv R_x(\tau) * h(\tau+\alpha) * h(\tau-\beta). \quad (11.5.2)$$

Таким образом, функция автокорреляции (АКФ) выходного сигнала равна АКФ входного сигнала, свернутой дважды, в прямом и обратном направлении, с импульсным откликом системы, что сохраняет четность АКФ выходного сигнала. Для нецентрированных процессов аналогичное заключение действительно и для ковариационных функций.

Заметим, что для свертки импульсных откликов, производя замену $\tau-\beta = t$, мы имеем равенство:

$$h(\tau+\alpha) * h(\tau-\beta) = h(t+\alpha+\beta) * h(t) = h(t) * h(t+\gamma) = K_h(t),$$

где $K_h(t)$ - функция ковариации импульсного отклика системы. Отсюда:

$$R_y(\tau) = R_x(\tau) * K_h(\tau). \quad (11.5.2')$$

Это означает появление в случайном сигнале на выходе системы определенной корреляционной зависимости, вызванной инерционностью системы, причем радиус корреляции выходного сигнала обратно пропорционален верхней частоте, пропускаемой системой.

Для взаимной корреляционной функции (ВКФ) R_{xy} входного и выходного сигналов соответственно имеем:

$$x(t) \cdot y(t+\tau) = \int_0^{\infty} h(\alpha) x(t)y(t+\tau-\alpha) d\alpha.$$

$$R_{xy}(\tau) = \int_0^{\infty} h(\alpha) R_x(\tau-\alpha) d\alpha \equiv R_x(\tau) * h(\tau-\alpha). \quad (11.5.3)$$

т.е. функция взаимной корреляции входного и выходного сигналов равна свертке АКФ входного сигнала с функцией импульсного отклика системы - фильтрации АКФ сигнала этим же фильтром. Заключение действительно и для функций ковариации.

Другая взаимно корреляционная функция R_{yx} может быть получена из соотношения:

$$R_{yx}(\tau) = R_{xy}(-\tau) \equiv R_x(\tau) * h(\tau+\alpha). \quad (11.5.4)$$

Отметим, что для статистически независимых случайных величин при одностороннем импульсном отклике $h(\tau)=0$ при $\tau < 0$ функция $R_{xy}(\tau)$ также является односторонней и равна 0 при $\tau < 0$, а функция R_{yx} соответственно равна 0 при $\tau > 0$.

Спектральные соотношения, которые характеризуют систему в целом по отношению к преобразованию случайных сигналов, это соотношения частотных плотностей распределения мощности случайных процессов на входе и выходе, которые для краткости обычно называют спектральными плотностями процессов (сигналов) или спектрами мощности.

Применяя преобразование Фурье к выражениям (11.5.2), для спектра мощности выходного сигнала получаем:

$$W_y(f) = W_x(f) |H(f)|^2. \quad (11.5.5)$$

Спектр мощности сигнала на выходе системы равен спектру мощности входного сигнала, умноженному на квадрат модуля частотной характеристики фильтра. С учетом четности корреляционных функций спектр мощности выходного сигнала также является четной действительной функцией и не имеет фазовой характеристики процесса.

Аналогично, для взаимного спектра мощности на основе выражений (11.5.3-4):

$$W_{xy}(f) = W_x(f) H(f). \quad (11.5.6)$$

$$W_{yx}(f) = W_x(f) H(-f).$$

Взаимный спектр мощности при одностороннем импульсном отклике является комплексным и содержит как амплитудную, так и фазовую характеристику процесса.

Отметим, что с использованием выражения (11.5.6) можно производить определение частотной характеристики и импульсного отклика системы:

$$H(f) = W_{xy}/W_x \Leftrightarrow h(t).$$

Дисперсия выходного сигнала может быть определена с использованием формул (11.5.2,5):

$$\sigma_y^2 = R_y(0) = \int_{-\infty}^{\infty} W_x(f) |H(f)|^2 df \equiv R_x(0) \int_0^{\infty} h^2(t) dt = \sigma_x^2 \int_0^{\infty} h^2(t) dt, \quad (11.5.7)$$

что полностью соответствует полученной ранее формуле (11.2.5) для цифровой системы.

Если сигнал нецентрированный и значение дисперсии входного сигнала неизвестно, то по аналогичным формулам вычисляется сначала *средний квадрат* выходного сигнала или так называемая *средняя мощность сигнала*:

$$\overline{y^2} = \overline{y^2(t)} = R_y(0) \equiv \overline{x^2} \int_0^{\infty} h^2(t) dt \equiv \int_0^{\infty} W_x(f) |H(f)|^2 df, \quad (11.5.8)$$

Вывод: средняя мощность выходного сигнала равна средней мощности входного сигнала, умноженной на квадрат площади импульсной реакции системы (для цифровых систем - сумму квадратов коэффициентов импульсного отклика). Для центрированных случайных сигналов средняя мощность равна дисперсии сигналов. Для нецентрированных выходных сигналов:

$$\sigma_y^2 = \overline{y^2} - \bar{y}^2 \equiv (\overline{x^2} - \bar{x}^2) \int_0^{\infty} h^2(t) dt. \quad (11.5.9)$$

Функция когерентности дает оценку точности принятой линейной модели системы. Когерентность входного и выходного сигналов системы оценивается по формуле:

$$\gamma_{xy}^2(f) = |W_{xy}(f)|^2 / [W_x(f) \cdot W_y(f)]. \quad (11.5.10)$$

Если функции $W_x(f)$ и $W_y(f)$ отличны от нуля и не содержат дельта-функций, то для всех частот f значения функции когерентности заключены в интервале:

$$0 \leq \gamma_{xy}^2(f) \leq 1.$$

Для исключения дельта-функций на нулевой частоте (постоянная составляющая сигнала) определение функции когерентности производится по центрированным сигналам. Для линейных систем с постоянными параметрами функция когерентности равна 1, в чем нетрудно убедиться, если в формулу (11.5.10) подставить выражения W_{xy} и W_y , определенные через W_x в формулах (11.5.5-6). Для совершенно не связанных сигналов функция когерентности равна нулю. Промежуточные между 0 и 1 значения могут соответствовать трем ситуациям:

1. Система осуществляет преобразование $x(t) \Rightarrow y(t)$, но в измерениях этих сигналов или одного из них присутствует внешний шум. Так, например, в сигналах, зарегистрированных с ограничением по разрядности, появляется шум квантования (округления значений).
2. Система не является строго линейной. Это может наблюдаться, например, при определенном ограничении по разрядности вычислений в цифровых системах, при накоплении ошибки в рекурсивных системах и т.п.
3. Выходной сигнал $y(t)$ помимо $x(t)$ зависит еще от каких-то входных или внутренних системных процессов.

Величина $1 - \gamma_{xy}^2(f)$ задает долю среднего квадрата сигнала $y(t)$ на частоте f , не связанную с сигналом $x(t)$.

Использование функций когерентности в практических методах анализа случайных данных подробно рассмотрено в работе /2/.

11.6. СТРУКТУРНЫЕ СХЕМЫ СИСТЕМ [5,20].

Структурные схемы. Алгоритмы обработки сигналов в системах наглядно отображаются в виде структурных схем. Базовые элементы схем и примеры построения структурных схем приводились ранее на рис. 11.1.1-3. Как правило, структурные схемы соответствуют программной реализации систем и будут рассматриваться ниже применительно к цифровым системам, но не определяет аппаратной реализации в специальных радиотехнических устройствах, которая может существенно отличаться от программной реализации, особенно для аналоговых систем.

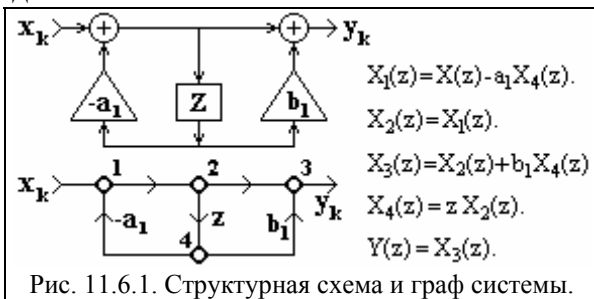


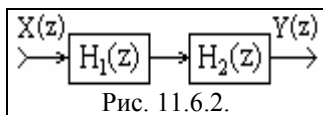
Рис. 11.6.1. Структурная схема и граф системы.

Графы систем. Наряду со структурной схемой система может быть представлена в виде графа, который отображает диаграмму прохождения сигналов и состоит из направленных ветвей и узлов.

Пример структурной схемы с передаточной функцией $H(z) = (1+b_1z)/(1+a_1z)$ и графа, ей соответствующего, приведен на рисунке 11.6.1.

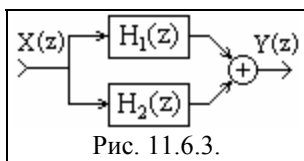
С каждым i узлом графа связано значение сигнала $x_i(k)$ или его образа $X_i(z)$, которые определяются суммой всех сигналов или их z -образов входящих в узел ветвей. В каждой ij - ветви (из узла i в узел j) происходит преобразование сигнала в соответствии с передаточной функцией ветви, например, задержка сигнала или умножение на коэффициент.

Соединения систем. Различают следующие соединения систем (рис. 11.6.2-4).



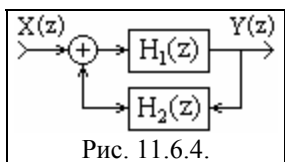
1. **Последовательное соединение.** Выходной сигнал предыдущей системы является входным для последующей. Эквивалентная передаточная функция общей системы равна произведению передаточных функций систем, в нее входящих:

$$H(z) = H_1(z) \cdot H_2(z) \cdot \dots \cdot H_N(z).$$



2. **Параллельное соединение.** Сигнал подается на входы всех параллельно соединенных систем одновременно, выходные сигналы систем суммируются. Эквивалентная передаточная функция общей системы равна сумме передаточных функций систем, в нее входящих:

$$H(z) = H_1(z) + H_2(z) + \dots + H_N(z).$$

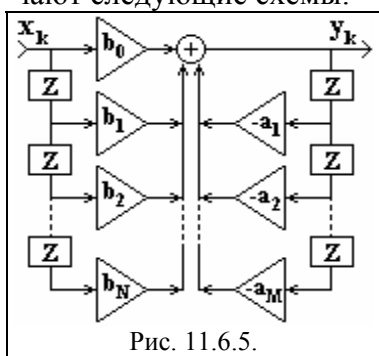


3. **Соединение обратной связи.** Выходной сигнал первой системы подается на выход системы и одновременно на вход системы обратной связи, выходной сигнал которой суммируется, со знаком плюс или минус в зависимости от вида связи (отрицательной или положительной), с входным сигналом первой системы. Эквивалентная передаточная функция общей системы:

$$H(z) = H_1(z) / (1 \pm H_1(z)H_2(z)).$$

Схемы реализации систем.

По принципам структурной реализации систем различают следующие схемы:



1. **Прямая форма.** Реализуется непосредственно по разностному уравнению

$$y_k = \sum_{n=0}^N b_n x_{k-n} - \sum_{m=1}^M a_m y_{k-m},$$

или по передаточной функции

$$H(z) = \sum_{n=0}^N b_n z^n / (1 + \sum_{m=1}^M a_m z^m).$$

Пример прямой системы приведен на рис. 11.6.5.

2. **Прямая каноническая форма** содержит минимальное

число элементов задержки. Передаточную функцию РЦС можно представить в виде:

$$H(z) = Y(z)/X(z) = H_1(z)H_2(z),$$

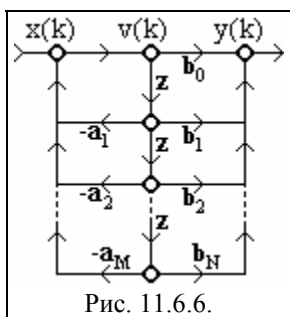
$$H_1(z) = V(z)/X(z) = 1 / (1 + \sum_{m=1}^M a_m z^m), \quad H_2(z) = Y(z)/V(z) = \sum_{n=0}^N b_n z^n.$$

Отсюда:

$$v(k) = x(k) - \sum_{m=1}^M a_m v(k-m), \tag{11.6.1}$$

$$y(k) = \sum_{n=0}^N b_n v(k-n). \tag{11.6.2}$$

В разностных уравнениях (11.6.1-2) осуществляется только задержка сигналов $v(k)$. Граф реализации РЦС приведен на рисунке 11.6.6.



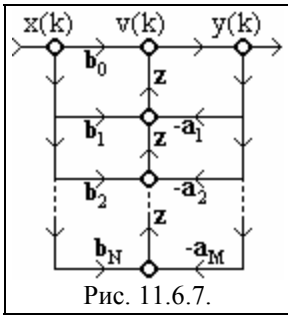
3. **Каскадная (последовательная) форма.** Соответствует представлению передаточной функции в виде произведения:

$$H(z) = \prod_{i=1}^k H_i(z).$$

$H_i(z)$ - составляющие функции типа $(1-r_i z)/(1-p_i z)$ при представлении $H(z)$ в факторизованной форме, где r_i и p_i - нули и полюсы функции $H(z)$. В качестве функций $H_i(z)$ обычно используются передаточные функции биквадратных блоков - фильтров второго порядка:

$$H_i(z) = (b_{0i} + b_{1i} \cdot z + b_{2i} \cdot z^2) / (1 + a_{1i} \cdot z + a_{2i} \cdot z^2).$$

4. **Параллельная форма.** Используется много реже и соответствует представлению передаточной функции в виде суммы биквадратных блоков или более простых функций.



Обращенные формы. В теории линейных сигнальных графов существуют процедуры преобразования исходных графов с сохранением передаточных функций. Одна из таких процедур - обращение графов, которое выполняется путем изменения направления всех ветвей цепи, при этом вход и выход графа также меняются местами. Для ряда систем такая транспозиция позволяет реализовать более эффективные алгоритмы обработки данных. Пример обращения графа прямой канонической формы рекурсивной системы (с перестроением на расположение входа с левой стороны) приведен на рис. 11.6.7.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
2. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. - М.: Мир, 1989. - 540 с.
5. Гольденберг Л.М. и др. Цифровая обработка сигналов: Учебное пособие для вузов. - М.: Радио и связь, 1990.- 256 с.
9. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. - М.: Мир, 1988. - 488 с.
14. Купер Дж., Макгиллем А. Вероятностные методы анализа сигналов и систем. - М.: Мир, 1989.
18. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. - М.: Связь, 1979. - 416 с.
20. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. - М.: Мир, 1978.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 12: ПЕРЕДАЧА СИГНАЛОВ ПО КАБЕЛЯМ

Все должно быть изложено так просто, как только возможно, но не проще.

Альберт Эйнштейн. Немецкий физик, XX в.

Осталось установить границу возможного и выкинуть излишние подробности. Для теоретиков не составит труда отнести эти подробности к известным по умолчанию. А практики хорошо знают принцип лишней детали - никогда не известно, для чего она нужна, пока не выбросишь.

Владислав Микшевич. Уральский геофизик, XX в.

Содержание: Введение. 12.1. Основное уравнение кабельной линии. 12.2. Волновое сопротивление кабельной линии. 12.3. Режимы передачи сигналов кабельной линией. Режим бегущей волны. Режим стоячей волны. Режим несогласованной нагрузки. Задержка сигналов в кабеле. Литература.

ВВЕДЕНИЕ

Передача электрических сигналов по кабельным линиям связи обычно рассматривается в рамках общей теории однородных длинных видеолиний – симметричных и коаксиальных кабельных линий передачи аналоговых, дискретных и цифровых сигналов в спектре частот 0...10 МГц. Современное промышленное производство практически невозможно без стационарных и широко разветвленных внешних и внутренних линий связи.

Внутренние линии связи, управления, сбора и обработки данных в геологоразведочных и горнопромышленных отраслях производства в силу его специфики имеют свои особенности. Обычно они относятся к пассивным каналам связи и работают в условиях жестких климатических, механических и химических дестабилизирующих факторов на высоком уровне внешних электромагнитных помех. Как правило, эти каналы связи являются многожильными и для защиты от внешних факторов имеют общий стальной экран (оплетку), что существенно ограничивает частотные параметры передачи сигналов.

12.1. ОСНОВНОЕ УРАВНЕНИЕ КАБЕЛЬНОЙ ЛИНИИ [7,23]

Однородная кабельная линия, эквивалентная электрическая схема которой приведена на рис. 12.1.1, определяется первичными электрическими параметрами: погонными значениями активного сопротивления R , индуктивности L , емкости C и проводимости G на единицу длины линии (как правило, на 1 км). На вход линии подключается источник сигналов (генератор, передатчик) с выходным сопротивлением Z_0 , на выход линии – приемник сигналов с входным сопротивлением Z_H (нагрузка линии).

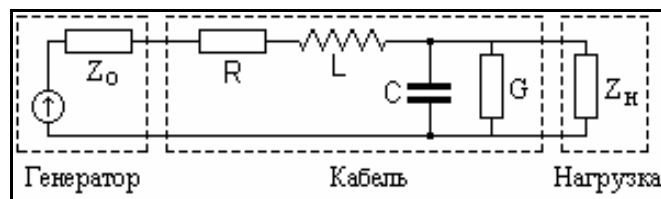


Рис. 12.1.1. Кабельная линия передачи сигнала.

По своей физической природе первичные электрические параметры кабеля аналогичны параметрам колебательных контуров, но в отличие от них они являются не сосредоточенными, а распределены по всей длине кабеля. Этим объясняется определенная зависимость первичных параметров кабеля от частоты сигнала и от конструкции кабеля.

Сигнал на входе линии задается в виде временной функции напряжения \dot{U}_0 и тока \dot{I}_0 . На выходе линии (на нагрузке) соответственно имеем \dot{U}_H и \dot{I}_H . Падение напряжения и утечка тока на произвольном участке dx линии определяются уравнениями:

$$-d \dot{U} / dx = \dot{I} (R + j\omega L),$$

$$-d \dot{I} / dx = \dot{U} (G + j\omega C).$$

Решение данных уравнений для напряжения и тока в произвольной точке x линии дает следующие выражения [2]:

$$\dot{U}_X = \dot{U}_0 \operatorname{ch} \gamma x - \dot{I}_0 \dot{Z}_B \operatorname{sh} \gamma x, \quad (12.1.1)$$

$$\dot{I}_X = \dot{I}_0 \operatorname{ch} \gamma x - (\dot{U}_0 / \dot{Z}_B) \operatorname{sh} \gamma x, \quad (12.1.2)$$

$$\dot{Z}_B = (R + j\omega L) / \gamma = \sqrt{\frac{R + j\omega L}{G + j\omega C}}, \quad (12.1.3)$$

где \dot{Z}_B – волновое сопротивление кабеля, γ – коэффициент (постоянная) распространения линии (сигнала в кабеле):

$$\gamma = \beta + j\alpha = \sqrt{(R + j\omega L)(G + j\omega C)}, \quad (12.1.4)$$

В выражениях (12.1.1)-(12.1.2) первые члены правой части представляют собой уравнения падающих волн напряжения и тока, распространяющихся по кабелю от генератора к нагрузке, а вторые члены – уравнения волн, отраженных от конца кабеля, энергия которых не поглотилась в нагрузке. Коэффициенты \dot{Z}_B и γ относят к вторичным параметрам кабеля. Выражения действительны для любой точки кабеля, в том числе и на нагрузке кабеля при $x = \Lambda$, где Λ – длина кабеля. Коэффициенты β и α являются собственными коэффициентами (собственными постоянными), соответственно, амплитудного затухания и фазового сдвига волны напряжения, проходящей через кабель. Их величины обычно задаются в значениях на 1 км кабеля и в этом случае называются километрическими (хотя последнее часто опускается и подразумевается по умолчанию). Численное значение километрического коэффициента β определяет коэффициент затухания волны напряжения, проходящей через кабель длиной 1 км. Соответственно, численным значением километрического коэффициента α задается величина сдвига фазы волны напряжения, проходящей через однокilометровый кабель.

$$\beta = \sqrt{\frac{1}{2} \left[RG - \omega^2 LC + \sqrt{(R^2 + \omega^2 L^2)(G^2 + \omega^2 C^2)} \right]}, \quad (12.1.5)$$

$$\alpha = \sqrt{\frac{1}{2} \left[\omega^2 LC - RG + \sqrt{(R^2 + \omega^2 L^2)(G^2 + \omega^2 C^2)} \right]}. \quad (12.1.6)$$

Кабель, удовлетворяющий уравнениям 12.1.1-12.1.6, является идеальным кабелем передачи сигналов. Реальный кабель может существенно отличаться от идеального. Но основной характер зависимостей вторичных электрических параметров кабелей от первичных и качественную картину передачи сигналов по кабелю в различных условиях согласования с источником сигналов и нагрузкой целесообразно выяснить сначала на идеальном кабеле.

12.2. ВОЛНОВОЕ СОПРОТИВЛЕНИЕ КАБЕЛЬНОЙ ЛИНИИ [7,23]

Волновое сопротивление – это сопротивление линии электромагнитной волне при отсутствии отражений от концов линии. Оно зависит от первичных электрических параметров кабеля и частоты сигнала. Если электромагнитную волну представить в виде отдельных волн напряжения и тока, то соотношение между ними и представляет собой волновое сопротивление цепи: $\dot{Z}_B = \dot{U} / \dot{I}$.

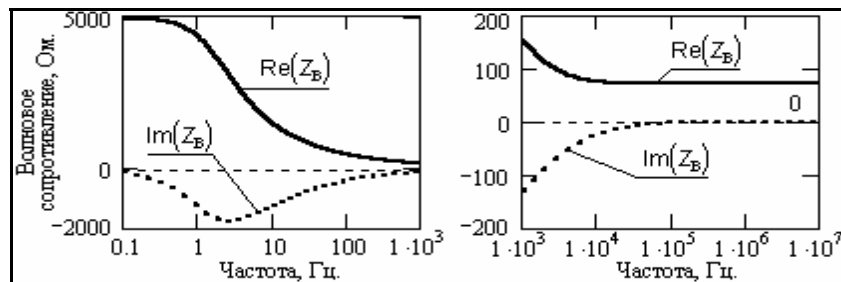


Рис. 12.2.1.

Волновое сопротивление является комплексной величиной и состоит из активной и

реактивной части, частотная зависимость которых показана на рис 12.2.1. Расчет графиков проведен при условно постоянных частотно-независимых значениях электрических параметров кабеля: $R = 25 \text{ Ом}$, $L = 0.5 \text{ мГн}$, $C = 0.1 \text{ мкФ}$, $G = 0.1 \text{ мкСм}$. Эти значения, типичные для каротажных кабелей, будем использовать и в дальнейшем без дополнительных пояснений. В действительности эти параметры являются частотно - зависимыми и определяются конструкцией кабеля, но они широко используются при сравнении кабелей по электрическим параметрам, при этом значение R измеряется на постоянном токе, а значения L , C и G – на определенной частоте в диапазоне 10-50 кГц.

Как следует из рисунка, зависимость волнового сопротивления от частоты наиболее существенна в области низких частот (менее 10 кГц) и имеет емкостной характер. В области частот более 10-20 кГц имеет место $\omega L > R$, $\omega C \gg G$ и значение волнового сопротивления стремится к постоянной величине $\sqrt{L/C}$. Эту величину называют номинальным (характеристическим) волновым сопротивлением кабеля. В дальнейшем индексом $Z_B = \sqrt{L/C} \equiv R_B$ будем обозначать постоянное характеристическое сопротивление кабеля (на частотах более 50-100 кГц). Для частотной функции волнового сопротивления будем применять обозначение с аргументом по частоте $Z_B(\omega)$ или индекс \dot{Z}_B .

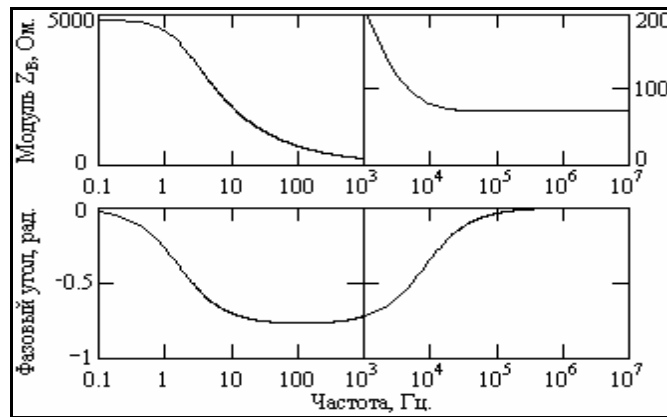


Рис. 12.2.2. Модуль и фаза волнового сопротивления.

Как комплексную величину, волновое сопротивление можно представить в форме:

$$Z_B(\omega) = z_B(\omega) \exp(j\varphi_B(\omega)),$$

где: $z_B(\omega)$ – частотная функция модуля волнового сопротивления (абсолютная величина отношения амплитудных значений напряжения и тока по аргументу – частоте ω) в любой точке линии, φ - частотная функция угловых значений, равных разности фаз волн напряжения и тока. Частотная зависимость значений модуля и фазового угла волнового сопротивления приведена на рис. 12.2.2. Как следует из графиков, волна тока в области низких частот опережает волну напряжения в максимуме на 45° .

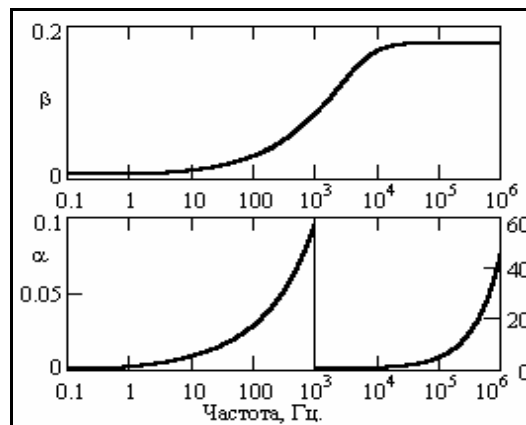
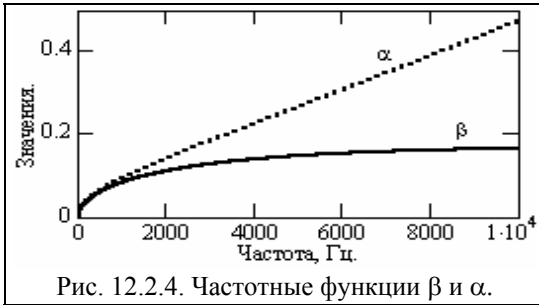


Рис. 12.2.3. Частотные функции β и α .

На рис. 12.2.3-4 приведены графики зависимости коэффициентов затухания β и фазо-

вого сдвига α (в относительных единицах) от частоты. В технической документации значение коэффициента β обычно приводится в неперах на километр. По мере нарастания частоты коэффициент затухания сначала плавно увеличивается с постепенным уменьшением степени увеличения, а затем, начиная с частоты порядка 10-20 кГц, практически постоянен и равен:

$$\beta = 0.5 \cdot (R\sqrt{C/L} + G\sqrt{L/C}).$$



Коэффициент фазового сдвига на низких частотах увеличивается синхронно с коэффициентом затухания (при $f < 1$ кГц $\beta \approx \alpha$), а затем, начиная с частоты порядка 1 кГц, нарастает линейно и пропорционально частоте ($\alpha = \omega\sqrt{LC}$). Это обеспечивает формирование фронтальной волны распространения сигнала по кабелю с постоянной скоростью для всех частотных составляющих сигнала, за исключением

низких частот, и сохранение формы сигналов на нагрузке. Отсюда следует, что кабель является оптимальной линией передачи высокочастотных и радиоимпульсных сигналов, энергия частотного спектра которых минимальна в области низких частот.

При совместном рассмотрении рисунков 12.2.1-12.2.4 нетрудно сделать вывод, что, начиная с частот порядка 5-10 кГц, кабельные линии связи имеют практически постоянные параметры. Именно в этой частотной области обеспечиваются минимальные искажения формы частотного спектра сигналов, а соответственно и формы самих сигналов при их передаче по линии связи.

Коэффициент передачи сигнала по напряжению по кабельной линии в общем виде может быть определен из выражения (12.1.1):

$$\dot{K} = \dot{U}_H / \dot{U}_0 = \text{ch } \gamma\Lambda - (\dot{I}_0 \dot{Z}_B / \dot{U}_0) \text{sh } \gamma\Lambda, \quad (12.2.1)$$

$$\dot{K} = \text{ch } \gamma\Lambda - (\dot{Z}_B / \dot{Z}_{BX}) \text{sh } \gamma\Lambda, \quad (12.2.1')$$

где Λ — длина кабеля, Z_{BX} — входное сопротивление кабеля, которое также является комплексной величиной и зависит от частоты:

$$\dot{Z}_{BX} = \dot{Z}_B \cdot (Z_H \text{ch } \gamma\Lambda + \dot{Z}_B \text{sh } \gamma\Lambda) / (\dot{Z}_B \text{ch } \gamma\Lambda + Z_H \text{sh } \gamma\Lambda). \quad (12.2.2)$$

12.3. РЕЖИМЫ ПЕРЕДАЧИ СИГНАЛОВ КАБЕЛЬНОЙ ЛИНИЕЙ [7,23]

В зависимости от величины нагрузки Z_H на выходе линии различают три режима передачи сигналов:

1. Режим бегущей волны сигнала при $Z_H = \dot{Z}_B$. В этом (согласованном) режиме $\dot{U}_1 / \dot{I}_1 = \dot{U}_0 / \dot{I}_0 = \dot{Z}_B$, входное сопротивление кабеля также равно волновому сопротивлению $\dot{Z}_{BX} = \dot{Z}_B$, отражения сигнала от концов линии отсутствуют и выражения (12.1.1), (12.1.2), (12.2.1) упрощаются:

$$\begin{aligned} \dot{U}_1 &= \dot{U}_0 \exp(-\gamma\Lambda), & \dot{I}_1 &= \dot{I}_0 \exp(-\gamma\Lambda), \\ \dot{K} &= \exp(-\gamma\Lambda). \end{aligned} \quad (12.3.1)$$

Коэффициент передачи сигнала представить в форме:

$$\begin{aligned} \dot{K} &= e^{-(\beta+j\alpha)\Lambda} = e^{-\beta\Lambda} e^{-j\alpha\Lambda} = |\dot{K}| e^{-j\varphi}, \\ |\dot{K}| &= e^{-\beta\Lambda}, & e^{-j\alpha\Lambda} &= e^{-j\varphi}, \end{aligned} \quad (12.3.1')$$

Эти выражения достаточно ясно показывают, почему коэффициенты β и α называют коэффициентами затухания β (коэффициент, характеризующий уменьшение абсолютных значений напряжения или тока на выходе линии по отношению к входному сигналу) и фазового сдвига α (изменение угла векторов тока или напряжения на выходе линии относительно

входного сигнала) при прохождении сигнала через единичный отрезок кабеля.

На практике коэффициент затухания амплитудных значений сигнала при его передаче по кабелю обычно измеряют в логарифмических единицах отношения амплитуды сигнала на входе кабеля к амплитуде сигнала на его выходе в неперах на 1 км, т.е.:

$$\ln(U_{\text{вх}}/U_{\text{вых}}) = \ln(1/\exp(-\beta)) = \beta,$$

при этом численные значения коэффициента затухания сигнала в неперах совпадают со значениями коэффициента β в относительных единицах.

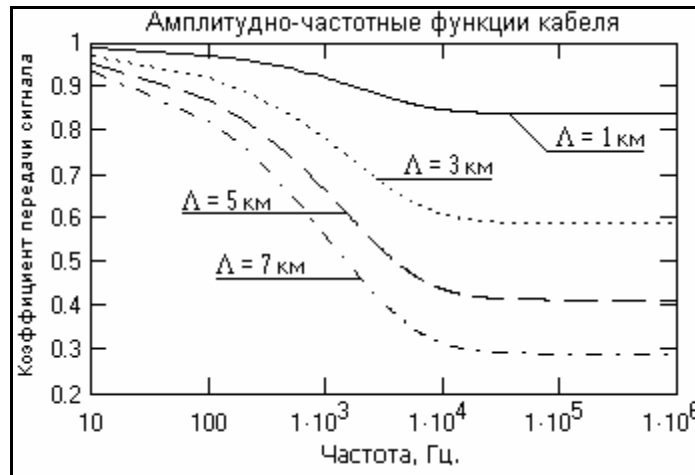


Рис. 12.3.1. Передаточные характеристики кабелей.

На рис. 12.3.1 приведены графики модуля коэффициента передачи сигнала, вычисленные по (12.3.1) при разных значениях длины кабеля. По существу, эти графики представляют собой передаточные амплитудно – частотные характеристики (АЧХ) идеальных кабелей, согласованных по нагрузке во всем частотном диапазоне.

Затухание сигналов в кабеле определяется потерями его энергии. Потери энергии в кабеле подразделяются на два вида: потери в активном сопротивлении кабеля R (нагревание токопроводящих проводников), и потери в изоляции кабеля, определяемые значением G ее проводимости. В общем случае проводимость G зависит от прямой утечки тока через диэлектрик и затрат энергии на его поляризацию:

$$G = (1/R_{\text{и}}) + \omega \cdot C \cdot \text{tg}(\delta), \quad (12.3.2)$$

где: $R_{\text{и}}$ – омическое сопротивление изоляции постоянному току, C – емкость кабеля, $\text{tg}(\delta)$ – угол диэлектрических потерь. Для современных изоляционных материалов значение угла диэлектрических потерь обычно не превышает десятитысячных долей и начинает сказываться только на очень высоких частотах (десятки МГц). Отсюда следует, что изменение АЧХ в области низких и средних частот обусловлено, в основном, изменением соотношения сопротивлений R и $Z_{\text{н}} = \dot{Z}_{\text{в}}$ и стабилизируется при частотах выше 10 кГц, где $Z_{\text{н}} = \dot{Z}_{\text{в}} \Rightarrow R_{\text{в}} = \text{const}$ (рис.12.2.2).

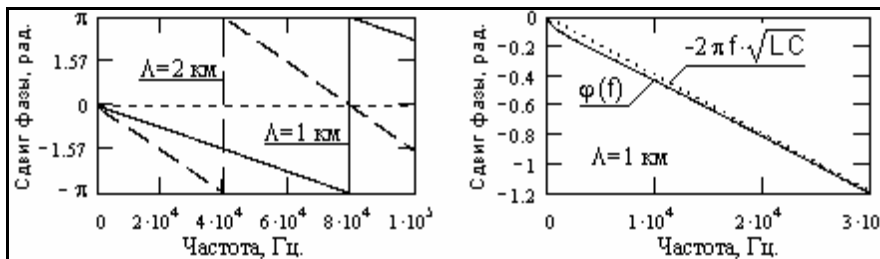


Рис. 12.3.2. Фазочастотные характеристики согласованных кабелей.

Что касается фазочастотных характеристик (ФЧХ) кабелей, приведенных на рис. 12.3.2, то, как это следует из величины α (см. рис. 12.2.4), значение фазового угла $\varphi = -\alpha L$ увеличивается с увеличением частоты, а на частотах выше 20 кГц равно $-\omega L \sqrt{LC}$, т.е. прямо пропорционально частоте и длине кабеля. Из выражения (12.3.1') следует также, что значе-

ния коэффициента сдвига фазы α в относительных единицах без учета знаков численно равны значениям угла сдвига фазы волны напряжения (тока) в радианах при прохождении через единичный отрезок кабеля. Соответственно, с учетом знаков, $\varphi = -\alpha\Delta$ в радианах. Отклонение от линейности (в сторону больших углов сдвига) наблюдается только на низких частотах (менее 20 кГц).

Режим согласованной нагрузки наиболее эффективен при передаче сигналов по кабелю. Однако в пассивных линиях связи обеспечить такой режим при передаче импульсных широкополосных сигналов практически невозможно без применения систем формирования специальной формы частотной зависимости выходного сопротивления генератора и входного сопротивления приемника, соответствующей волновому сопротивлению кабеля. Как правило, для решения данной проблемы используются альтернативные решения: применение частотно-зависимых корректоров формы сигналов на выходе кабеля или фильтров частичной деконволюции импульсного отклика кабеля (фильтры направленного сжатия формы импульсного отклика).

2. Режим стоячей волны устанавливается в кабеле при $Z_{\text{н}} = 0$ (короткозамкнутая линия) или $Z_{\text{н}} = \infty$ (режим холостого хода). Эти режимы применяются при измерениях волнового сопротивления кабеля.

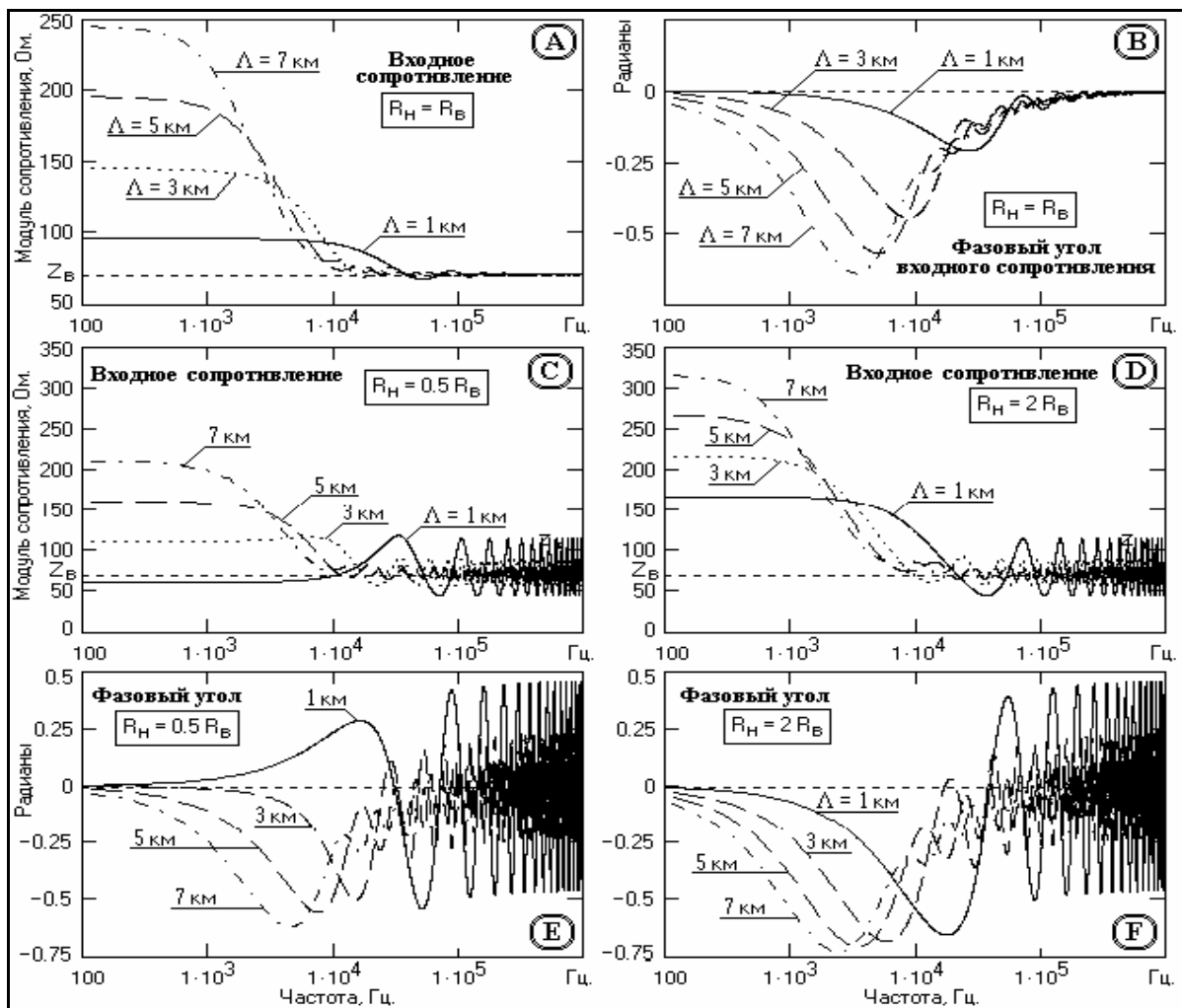


Рис. 12.3.3. Частотные характеристики входного сопротивления кабелей в зависимости от длины кабеля и значения сопротивления нагрузки.

3. Режим несогласованной нагрузки при $Z_{\text{н}} \neq Z_{\text{в}}$. Как правило, сопротивление нагрузки представляет собой постоянную величину $Z_{\text{н}} \Rightarrow R_{\text{н}}$, независимую или слабо зависящую от частоты сигнала. Но согласованность кабеля с нагрузкой является частотно-зависимой даже при $Z_{\text{н}} = R_{\text{в}}$, что определяет зависимость от частоты и входного сопротивле-

ния кабеля. На рис. 12.3.3(А,В) приведены частотные зависимости модулей и фазовых углов входного сопротивления кабелей различной длины при сопротивлении нагрузки, равной характеристическому волновому сопротивлению.

Как следует из графиков, входное сопротивление для низкочастотного диапазона (менее 1-20 кГц в зависимости от длины кабеля) отличается по своей величине и характеру от диапазона выше 10-50 кГц, где оно практически постоянно и равно волновому сопротивлению кабеля. В диапазоне менее 1 кГц входное сопротивление выше волнового на величину, примерно равную сопротивлению жилы постоянному току, и также является преимущественно активным. Между этими двумя диапазонами выделяется переходная зона, где входное сопротивление имеет реактивную составляющую емкостного характера.

При $R_H \neq R_B$ характер входного сопротивления кабеля и его фазового угла усложняется, что можно наглядно видеть на рис. 12.3.3(С, D, E, F). Из этих рисунков очевидно, что при невозможности согласования кабеля по всему частотному диапазону необходимо, по крайней мере, стремиться выполнить условие $R_H = R_B$.

Неполная согласованность кабеля с нагрузкой создает отраженные волны, которые достигают начала кабеля и при $Z_0 \neq Z_B$ снова отражается назад в кабель, что приводит к искажению сигналов. Условие $R_0 = R_B$ является оптимальным и для согласования источника сигнала с кабелем. Затухание линии при фиксированных значениях сопротивлений нагрузки R_H и источника сигнала R_0 называют рабочим затуханием кабеля и вычисляют по формуле:

$$\beta_p = \beta \cdot \Lambda + \ln \left| \frac{Z_0 + Z_B(\omega)}{2\sqrt{Z_0 Z_B(\omega)}} \right| + \ln \left| \frac{Z_H + Z_B(\omega)}{2\sqrt{Z_H Z_B(\omega)}} \right| + \ln \left| 1 - \frac{Z_0 - Z_B(\omega)}{Z_0 + Z_B(\omega)} \cdot \frac{Z_H - Z_B(\omega)}{Z_H + Z_B(\omega)} \cdot e^{-2\gamma \Lambda} \right|, \quad (12.3.3)$$

где первый член $\beta \cdot \Lambda$ в правой части уравнения – собственное затухание кабеля, второй и третий члены – дополнительное затухание вследствие неполного согласования с нагрузкой и с генератором, а последний член – влияние многократных отражений от концов кабеля.

При постоянных сопротивлениях источника сигнала и нагрузки, равных волновому сопротивлению, кабель остается существенно рассогласован на низких частотах, при этом километрический коэффициент рабочего затухания, вычисленный по (12.3.3) с приведением к 1 км, на низких частотах зависит от длины кабеля, что видно на рис. 12.3.4(А). Это объясняется тем, что при больших коэффициентах отражения электромагнитных волн от концов кабеля и их многократной пульсации по кабелю общие потери энергии на кабеле существенно зависят от его длины. Этот факт необходимо учитывать при использовании наземных приборов с коррекцией частотных искажений сигнала.

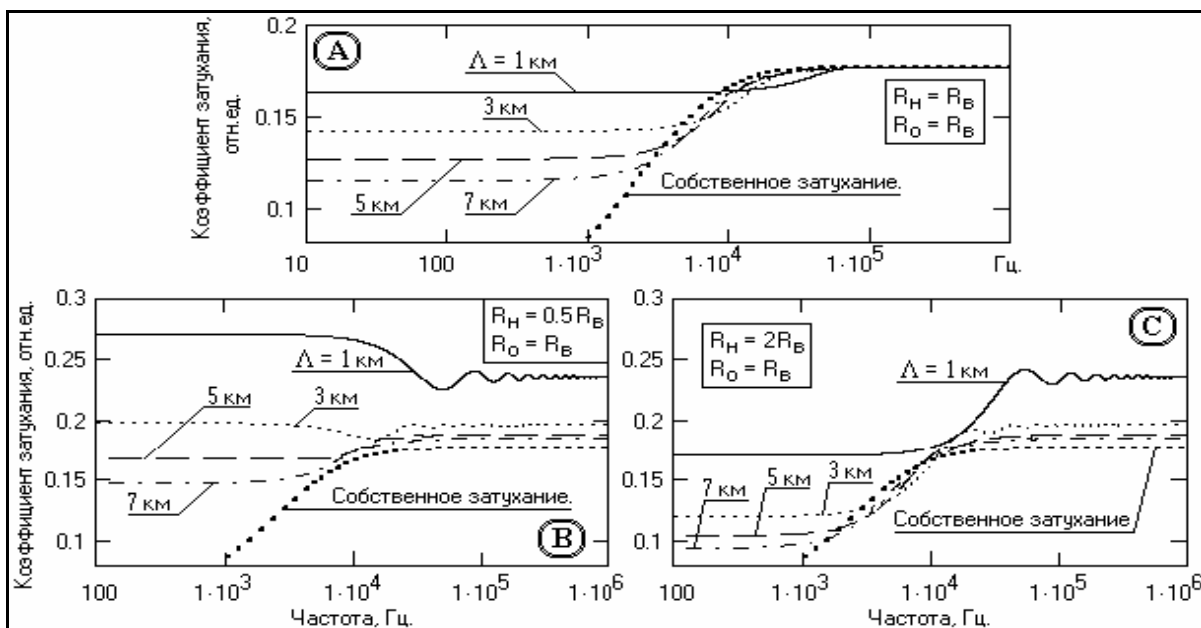
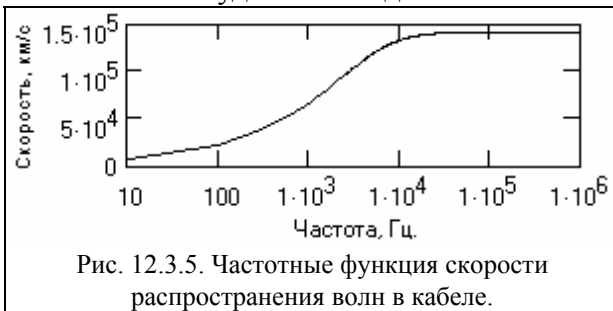


Рис. 12.3.4. Частотные функции рабочих километрических коэффициентов затухания

сигнала в зависимости от длины кабеля и согласования с нагрузкой

На рис. 12.3.4(В,С) дополнительно приведены графики частотной зависимости километрического коэффициента рабочего затухания β_p , вычисленные по (12.3.3) при $Z_0 = Z_B$ и разных значениях нагрузки кабеля по отношению к его номинальному волновому сопротивлению (Z_B на высоких частотах). Как следует из графиков, при рассогласовании кабеля с нагрузкой затухание сигнала на частотах более 50 кГц увеличивается, в основном, за счет отражения сигнала от нагрузки, причем в тем большей степени, чем меньше длина кабеля (и, соответственно, больше абсолютная доля отраженной энергии сигнала и ее поглощение в кабеле). Характер изменения затухания на частотах ниже 50 кГц еще более сложен и зависит как от знака изменения нагрузки относительно волнового сопротивления, так и от длины кабеля, причем при $R_H < R_B$ затухание увеличивается, а при $R_H > R_B$ уменьшается для кабелей большой длины. Аналогичная картина наблюдается и при изменении сопротивления генератора относительно номинального волнового при постоянном сопротивлении нагрузки.

Таким образом, при сопротивлении нагрузки близкой к номинальному волновому сопротивлению, километрический коэффициент затухания имеет два разных уровня с переходной границей в области средних частот. Высокочастотный уровень достаточно слабо зависит от длины кабеля и сопротивлений нагрузки и генератора, а низкочастотный уровень может существенно изменяться при изменении сопротивления нагрузки или генератора. Это позволяет использовать режим согласованной с номинальным волновым сопротивлением нагрузки кабеля в качестве основного режима передачи сигналов по кабелю, при этом небольшим направленным рассогласованием кабеля с нагрузкой или генератором коэффициент затухания сигнала может быть сделан практически равномерным по всему частотному диапазону. Вместе с тем графики еще раз свидетельствуют о целесообразности передачи информации сигналами, имеющими минимальную энергию в области низких частот (с нулевым средним значением амплитудной последовательности сигналов).



Задержка сигналов в кабеле. Если коэффициент α определяет сдвиг по фазе колебания с частотой f на единице длины, то длина волны λ в единицах длины кабеля будет равна длине кабеля, при которой сдвиг по фазе достигает величины 2π , т.е. $\lambda\alpha = 2\pi$. С учетом этого скорость распространения электромагнитных волн в кабеле, график зависимости которой от частоты колебаний приведен на рис.

12.3.5, определяется выражением:

$$v(f) = f \cdot \lambda, \quad v(\omega) = \omega / \alpha. \quad (12.3.4)$$

Максимальная задержка сигнала соответствует низким частотам. На частотах выше 10 кГц при $\alpha = \omega \sqrt{LC}$ значение скорости распространения волны стремится к постоянной величине $v = 1 / \sqrt{LC}$.

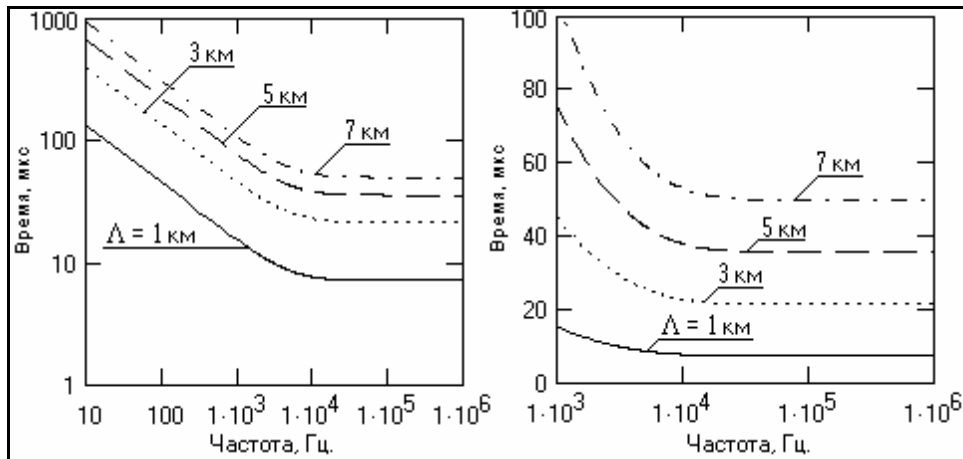


Рис. 12.3.6. Функции временной задержки волн в кабеле.

На рис. 12.3.6 приведены функции временной задержки частотных составляющих ($t_s(\omega) = \Lambda/v(\omega)$) в кабеле.

В целом, из рассмотрения основных электрических характеристик кабеля следуют два, во многом очевидных для практиков вывода:

1. Оптимальная величина сопротивления нагрузки кабеля и выходного сопротивления источника сигналов должна быть равна характеристическому сопротивлению кабеля.
2. Энергия сигналов должна быть минимальной в области низких частот.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.
6. Горбенко Л.А., Месенжник Я.З. Кабели и провода для геофизических работ. - М.: Энергия, 1977.
7. Гроднев И.И., Фролов Н.А. Коаксиальные кабели связи. - М.: Радио и связь, 1983. - 209 с.
23. Стрижевский Н.З. Коаксиальные видеопроводы. - М.: Радио и связь, 1988. - 200 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 13: КАРОТАЖНЫЕ КАБЕЛИ

Нет смысла добиваться точности там, где неизвестно, о чем идет речь.

Джон фон Нойман.
Американский математик и физик, XX в.

Каротажный кабель, это сначала трос, а уже потом кабель. Зачем же требовать от него высокой точности передачи данных. Лучше довольствоваться малым, но с умом.

Владимир Кузьмин.
Новосибирский геофизик Уральской школы, XX в.

Содержание: Введение. 13.1. Первичные электрические параметры кабелей. Активное сопротивление. Проводимость изоляции жил кабеля и электромагнитные потери. Емкость токопроводящих жил. Индуктивность кабеля. 13.2. Вторичные электрические параметры кабеля. Волновое сопротивление. Частотные характеристики жил кабеля. 13.3. Импульсный отклик кабеля. Литература.

ВВЕДЕНИЕ

Начиная с 80-90 годов прошлого века в методах и технологиях геофизических исследований скважин (ГИС) наблюдаются существенные качественные изменения, а именно: переход на комплексные и многопараметровые измерения с применением комбинированных и/или многофункциональных скважинных приборов с достаточно глубокой обработкой первичных данных в реальном масштабе времени непосредственно в каротажных лабораториях (станциях). Такой переход требует как повышения качества первичных данных, так и передачи этих данных, как правило, в цифровой (кодовой) форме, с достаточно высокой скоростью в наземные обрабатывающие (измерительно-вычислительные) и регистрирующие устройства. И если выполнение первого требования достаточно успешно базируется на высоком уровне развития современной электронной техники, то практически единственной телеметрической линией передачи данных ГИС остается традиционный каротажный кабель.

Каротажный геофизический кабель относится к типу универсальных аналоговых кабельных линий передачи информации от скважинных приборов к каротажной станции и передачи управляющих сигналов на скважинные приборы. Пропускная информационная способность каротажного кабеля определяет скорость каротажа, особенно в комплексных методах ГИС. Однако каротажный кабель является не только электрической линией передачи информации, но и тросом с достаточно большим (до нескольких тонн) разрывным усилием, несущим скважинные приборы в химически- и механически агрессивной среде скважин. По существу, это кабель-трос специального технологического назначения, работающий в широком диапазоне температур (от минусовых на поверхности до 100-150 и более градусов на больших глубинах), что накладывает определенные ограничения на его характеристики, как линии связи. Реальная скорость передачи информации (бит/с) современных кабелей в зависимости от их длины ограничиваются диапазоном до 10-100 кГц, что начинает существенно сдерживать развитие и совершенствование технологий ГИС.

По числу токопроводящих жил (ТПЖ) каротажные кабели разделяются на три основных вида: одно-, трех- и семижильные. Как правило, кабельные каналы связи рассматриваются в рамках теории однородных длинных линий с постоянными электрическими параметрами кабеля по всей его длине. Однако в процессе каротажа различные части кабеля находятся в неодинаковых условиях давления и температуры, что приводит к изменению первичных параметров, как во времени, так и по длине кабеля. Однако, как показали исследования, изменение первичных параметров бронированных кабелей хотя и имеет место, но не столь значительно, чтобы отказаться от использования теории однородных линий.

Одножильные бронированные каротажные кабели, в принципе, относятся к разновидности коаксиальных кабелей с концентрическим расположением жилы (прямого провода) внутри брони (обратного провода). Взаимодействие электромагнитных полей прямого и обратного проводника в идеальном коаксиальном кабеле при равных значениях тока и разных его направлениях создает нулевое значение электромагнитного поля за пределами кабеля, т.е. электромагнитное поле сигналов сосредоточено внутри кабеля, что и обеспечивает эффек-

тивную передачу электромагнитной энергии с минимальными потерями. Центральная жила и оплетка коаксиальных кабелей выполняются из немагнитных материалов (медь), что также не создает потерь на перемагничивание магнитных материалов.

Каротажный кабель, в отличие от коаксиального, в качестве оплетки имеют стальную броню без поверхностной изоляции, а, следовательно, локализованного обратного тока в этой броне не существует как для одножильного, так и для многожильного кабеля. Это действительно как для кабеля на барабане лебедки, где броня представляет в какой-то мере сплошной металлический монолит, так и для кабеля в скважине, где броня – линейный заземленный электрод. Следовательно, в каротажных кабелях появляется весьма существенный источник потерь электромагнитной энергии сигналов – на перемагничивание стальной брони электромагнитными полями токопроводящих жил и межпроводниковыми электромагнитными полями (при двухпроводной передаче сигналов), а также на потери электромагнитной энергии обратного тока в окружающей среде. Эти потери нарастают с увеличением частоты тока и приводят к существенному частотному ограничению импульсной пропускной способности кабеля. Их место в математической модели кабеля подлежат уточнению.

13.1. ПЕРВИЧНЫЕ ЭЛЕКТРИЧЕСКИЕ ПАРАМЕТРЫ КАБЕЛЕЙ [6, 7]

Активное сопротивление.

Полное активное сопротивление линии передачи сигналов кабеля состоит из суммы сопротивлений прямого и обратного проводников. Для одножильного бронированного кабеля (ОБК) и многожильного бронированного кабеля (МБК) с однопроводной линией передачи информации (МБК-1п, броня в качестве обратного проводника) сопротивление постоянному току R_0 определяется сопротивлением токопроводящей жилы (ТПЖ): длиной и конструкцией жилы, диаметром и материалом проволок жилы и температурой окружающей среды. Сопротивление обратного проводника (брони) много меньше величины сопротивления жил и его значением можно пренебречь. Для двухпроводной линии передачи (МБК-2п) полное сопротивление R_0 складывается из сопротивления двух ТПЖ.

Жилы кабелей свивают, в общем случае, из нескольких стренг, каждая из которых свивается из нескольких проволок. В настоящее время распространены кабели с жилами из одной стренги. Сопротивление жилы постоянному току определяется уравнением:

$$R_0 = (\rho \cdot \Lambda / s) \cdot (m_1 \cdot m_2),$$

где: ρ - удельное сопротивление материала проволоки при 20⁰С в Ом·мм²/км, Λ - длина жилы в км, s - суммарная площадь всех проволок жилы в мм, m - коэффициенты скрутки проволок в стренгу и стренг в жилу. Для новых кабелей значения коэффициентов m близки к 1. С течением времени переходное сопротивление между проволоками возрастает (окисление их поверхности) и значения коэффициентов приближаются к величине $1/\cos(\phi)$, где углом скрутки ϕ учитывается длина проволок в скрутках (обычно больше длины кабеля на 10-15%). Если жилы скручиваются из медных и стальных проволок, то сопротивления для них вычисляются отдельно и затем объединяются как параллельные электрические цепи. Типовые значения ρ для меди и стали принимаются равными соответственно 17.5 и 98 Ом·мм²/км.

Сопротивление жил, особенно комбинированных из стали и меди, существенно зависит от температуры. В принципе, оно может вычисляться по известному уравнению:

$$R_t = R_{20}[1 + \mu(t - 20)],$$

где μ - температурный коэффициент (0.0039 для меди, 0.0062 для стали). Расчеты с учетом геотермического градиента по стволу скважины показывают, что при работе в скважинах с температурой до 150⁰С на забое сопротивление жил кабеля может увеличиваться на 10-20%. Отсюда следует, что нормальным явлением можно считать изменение сопротивления жил кабелей в процессе каротажа в пределах до 10%. Соответственно, этой величиной может ограничиваться и точность математической модели активного сопротивления жил кабеля.

Активное сопротивление кабелей на высоких частотах практически прямо пропорционально корню квадратному из частоты. Это определяется так называемым поверхностным эффектом - вытеснением тока к поверхности проводов, которое возрастает с увеличени-

ем частоты, и взаимодействием полей проволок кабеля. На частотах свыше 100 кГц активное сопротивление определяется, в основном, поверхностным эффектом и в несколько раз больше сопротивления постоянному току. На частотах ниже 100 кГц эта зависимость от частоты несколько уменьшается по величине, но сохраняется по характеру. Для жил каротажного кабеля зависимость активного сопротивления от частоты дополнительно осложняется влиянием скрутки проволок в стренгу (стренг в жилу) и зависит от материала проволок (медь, сталь). Строго обоснованные аналитические методы расчетов электрических параметров многопроволочных проводников каротажных кабелей отсутствуют. С использованием известных экспериментальных данных полное активное сопротивление жилы каротажного кабеля в первом приближении может быть аппроксимировано следующими формулами:

$$R(f) = R_0 \cdot [1 + K(f) \cdot P(f)], \quad (13.1.1)$$

$$K(f) = 0.0078 \cdot \left(\frac{3.7 \cdot r}{3.7 \cdot s_{Cu} + s_{Fe}} \right) \cdot \sqrt{f}, \quad (13.1.1')$$

$$P(f) = 1 - \exp(-a \cdot f \cdot 10^{-4}), \quad (13.1.1'')$$

где: $K(f)$ - усредненный коэффициент поверхностного эффекта, $P(f)$ - поправочный коэффициент для каротажного кабеля на конструкцию жилы и кабеля, r - радиус жилы в мм, s - относительные площади сечения медных и стальных проволок в жиле ($s_{Cu} + s_{Fe} = 1$), a - частотная постоянная кабеля, значение которой порядка (0.1-0.3) для ОБК, $\sim(0.3 \div 0.5)$ для МБК-1п, $\sim(0.5 \div 0.8)$ для МБК-2п, и приближается к 1 для коаксиальных видеолиний.

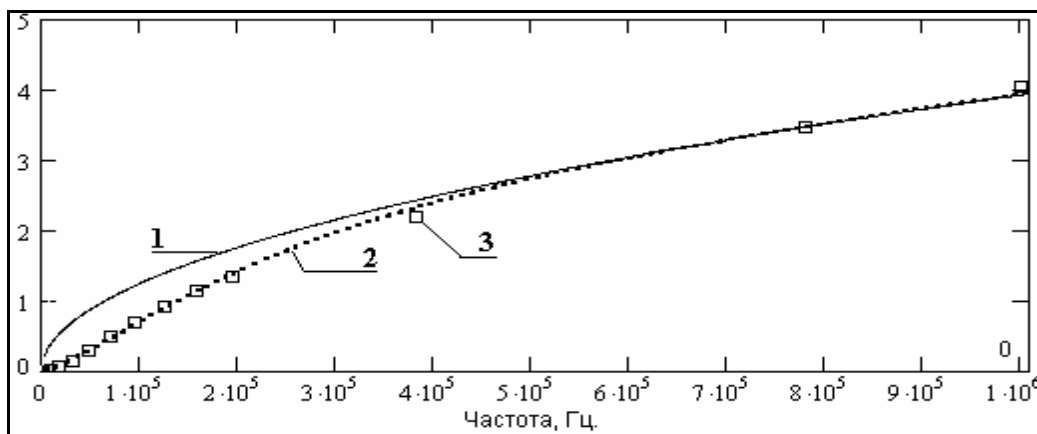


Рис. 13.1.1. Коэффициент $Q(f)$ кабеля с медной жилой диаметром 1.05 мм (1- коаксиальной линии, 2- расчет по формулам аппроксимации, 3- экспериментальные данные)

На рис. 13.1.1 приведено сопоставление экспериментальных данных по значению полного коэффициента $Q(f) = K(f) \cdot P(f)$ с расчетами по формулам (13.1.1).

Проводимость изоляции жил кабеля и электромагнитные потери.

Проводимость изоляции жил кабеля G определяется уравнением

$$G = (1/R_{и}) + \omega \cdot C \cdot \operatorname{tg}(\delta),$$

где значение сопротивления изоляции $R_{и}$ бронированных каротажных кабелей не менее 1000 МОм/км, а параметр δ потерь на поляризацию изоляции не более 0.0005. В условиях эксплуатации значение $R_{и}$ обычно поддерживается не ниже 10 МОм. Эти величины определяют проводимость изоляции жил не менее 0.0003 См/км на частотах до 1 МГц и потерями в изоляции кабелей можно пренебречь.

Однако в бронированном каротажном кабеле имеют место существенные потери на перемагничивание брони и магнитных элементов окружающей среды, точное количественное значение которых предусмотреть невозможно. Влияние этих потерь на расчеты коэффициентов затухания сигнала эквивалентно потерям на поляризацию изоляции кабеля (частотно зависимая потеря энергии). Учитывая последнее, в простейшем случае можно считать возможным введение в формулу расчета проводимости дополнительного коэффициента χ -коэффициента приведения электромагнитных потерь к потерям в изоляции кабеля:

$$G = (1/R_n) + \omega \cdot C \cdot \operatorname{tg}(\delta + \chi),$$

где значение χ для одножильных бронированных кабелей на частотах до 200 кГц порядка 0.1-0.2. Уравнение в этом случае можно считать уравнением эквивалентной проводимости кабеля, что позволяет сохранить без изменений общую теорию однородных длинных линий.

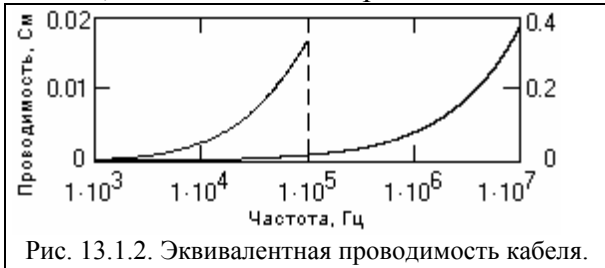


Рис. 13.1.2. Эквивалентная проводимость кабеля.

Параметр эквивалентной проводимости кабеля существенно влияет на передачу сигналов в области высоких частот. По результатам сопоставления расчетных и экспериментальных данных затухания сигнала в каротажных кабелях в диапазоне до 5 МГц может быть предложена для использования более простая формула аппроксимации эквивалентной проводимости

кабеля:

$$G = 2\pi f \cdot C \cdot \eta / R(f), \quad (13.1.2)$$

где значение η порядка 12-14 для одножильных и 10-13 для многожильных кабелей. Пример расчета проводимости кабеля КГ 1х0.75-55-150 по данной формуле приведен на рис. 13.1.2.

Емкость токопроводящих жил.

Емкость жил определяется конструкцией кабеля и зависит от материала и толщины изоляции ТПЖ. Для одножильных кабелей емкость ТПЖ может оцениваться по уравнению (в мкФ/км): $C = \varepsilon / (18 \cdot \ln(D/d))$, где ε - диэлектрическая проницаемость изоляции, D - внешний диаметр жилы по изоляции, d - диаметр токопроводника жилы. Значение емкости, как правило, находится в диапазоне 0.1-0.12 мкФ/км. Емкость между жилой и броней в многожильных кабелях имеет примерно такие же значения, а емкость жила-жила в 1.6-1.8 раз меньше. Зависимость емкостей жил от частоты практического значения не имеет. При повышении температуры емкость жил несколько уменьшается за счет уменьшения диэлектрической проницаемости изоляции, а при повышении давления - увеличивается. В скважинных условиях эти два процесса практически компенсируют друг друга и общее изменение емкости незначительно.

Индуктивность кабеля.

Собственная индуктивность коаксиальных линий с увеличением частоты уменьшается вследствие поверхностного эффекта в проводниках. Во внешних проводниках ток вытесняется к центру кабеля, во внутренних - от центра. Соответственно, индуктивность внешних проводов уменьшается, а внутренних возрастает, но их меньший диаметр не создает полной компенсации уменьшения индуктивности внешних проводов.

Общая индуктивность пары жила-броня складывается из собственной индуктивности жилы и межпроводниковой индуктивности жила-броня. В первом приближении, для пары жила-броня из разных металлов, она может определяться с помощью уравнения, которое используется для расчетов индуктивности коаксиальных кабелей (в Гн/км):

$$L(f) = \left(\frac{L_c}{\sqrt{f}} + L_m \right) \cdot 10^{-4}, \quad (13.1.3)$$

$$L_c = K_c \cdot \left(\frac{\sqrt{\mu_1 \rho_1}}{d} + \frac{\sqrt{\mu_2 \rho_2}}{D} \right), \quad K_c = 100/\pi \quad (13.1.4)$$

$$L_m = b \cdot \ln(D/d), \quad (13.1.5)$$

где: f - частота тока (Гц), μ - магнитные проницаемости жилы и брони (медь $\mu_1=1$, сталь $\mu_2=100 \div 120$), ρ - удельные сопротивления ($\rho_1=17.5$ Ом·мм²/км, $\rho_2=130$ Ом·мм²/км), d - диаметр жилы (мм), D - внутренний диаметр брони кабеля, (мм), b - коэффициент учета конструкции кабеля ($b \approx 2$ для ОБК, $b \approx 3$ для МБК). Значение L_c/\sqrt{f} определяет собственную индуктивность жил, L_m - межпроводниковую индуктивность. Для жил, содержащих стальные проволоки в своем составе, значения μ_1 и $1/\rho_1$ вычисляются с учетом весовых коэффициентов площади сечений медных и стальных составляющих в общей площади сечения жилы.

Сопоставление расчетов индуктивности кабелей по формулам (13.1.3-13.1.5) с результатами экспериментальных измерений показало, что расчетные формулы завышают значения индуктивностей практически на всех частотах выше 100 Гц. По видимому, здесь играет роль, в первых, тот же фактор разницы значений прямого и обратного токов в жиле и броне (по существу – отсутствие локализованного обратного тока в броне), а во вторых, конструкция брони. Достаточно толстый двойной слой из стальных проволок с перекрестным повивом слоев не может быть эквивалентным оплетке коаксиального кабеля.

Сходимость расчетных и измеренных значений индуктивностей на частотах выше 200 Гц обеспечивается установлением в формуле (13.1.4) значения "эффективного" диаметра $D_3 = k \cdot D$ кабеля вместо внутреннего диаметра брони D . Величина коэффициента k по результатам расчетных и экспериментальных значений параметров жил находится в пределах (1-1.2) для МБК-2п, (1.2-1.4) для МБК-1п и (1.4-1.6) для ОБК. В последнем случае это практически соответствует диаметру по среднему слою брони, что объяснимо по своей физической сущности. При известных значениях индуктивности жил на нескольких частотах более точная аппроксимация может производиться дополнительным изменением коэффициента K_c .

Как следует из формулы (13.1.3), зависимость индуктивности от частоты определяется собственной индуктивностью жил. Она же определяет зависимость индуктивности от температуры (изменение ρ). В области высоких частот индуктивность определяется, в основном, межпроводниковой индуктивностью. При намотке кабеля на лебедку его индуктивность может увеличиваться на 1-3% в зависимости от конструкции лебедки и состояния (степени окисления) поверхности брони кабеля.

Межпроводниковая индуктивность многожильных кабелей для пары жила-жила увеличивается за счет индуктивности пар и влияния соседних жил. В первом приближении она может оцениваться по формуле аппроксимации экспериментальных данных:

$$L_m \approx b \cdot \ln((D_3/d) + s \cdot (2c/d)), \quad (13.1.6)$$

где c – расстояние между центрами жил, $s \approx (0.4-0.5)$ в зависимости от конструкции кабеля и определяется по измерениям индуктивности на высоких частотах.

По измеренному значению индуктивности на частоте f_1 частотная функция индуктивности может вычисляться по формуле:

$$L(f) \approx [L(f_1) - L_m] \sqrt{\frac{f}{f_1}} + L_m. \quad (13.1.7)$$

13.2. ВТОРИЧНЫЕ ЭЛЕКТРИЧЕСКИЕ ПАРАМЕТРЫ КАБЕЛЕЙ [6,7,23]

Волновое сопротивление.

В принципе, понятие волнового сопротивления применимо только к однородным линиям, как постоянное сопротивление электромагнитной волне в любой точке кабеля при ее распространении вдоль линии любой длины. Волновое сопротивление – комплексная величина, зависящая от частоты и первичных параметров линии. В общей форме:

$$Z(\omega) = \sqrt{\frac{R(\omega) + j\omega L(\omega)}{G(\omega) + j\omega C}}, \quad (13.2.1)$$

Если при достаточно высоком значении сопротивлении изоляции $R_{и}$ в значении G пренебречь величиной $1/R_{и}$, то это выражение можно представить в виде:

$$Z(\omega) = q \cdot \sqrt{\frac{R(\omega)}{j\omega C} + \frac{L(\omega)}{C}}, \quad q = \sqrt{1/(1 + \tan(\delta + \chi))} \approx 0.93.$$

На частотах более 0.1 МГц отношение $R(\omega)/\omega C$ становится много меньше значений $L(\omega)/C$ и формула упрощается:

$$Z(\omega) \cong q \cdot \sqrt{\frac{L(\omega)}{C}},$$

при этом индуктивность жил определяется, в основном, независимой от частоты межпроводниковой индуктивностью и волновое сопротивление кабеля имеет преимущественно рези-

стивный характер, практически не зависит от частоты, и считается номинальным волновым сопротивлением кабеля Z_B .

По мере уменьшения частоты отношение $R(\omega)/\omega C$ увеличивается, а на частотах менее 1 кГц становится преобладающим

$$Z(\omega) \cong q \cdot \sqrt{\frac{R(\omega)}{j\omega C}},$$

т.е. содержит соизмеримые действительную и мнимую (емкостную) составляющие.

Каротажные кабели не в полной мере удовлетворяют требованиям однородных линий. Это не мешает использованию выражения (13.2.1) как в теоретическом плане, так и при чисто практических расчетах с погрешностью не более 3-5%, если специфика каротажных кабелей учитывается в функциях частотного распределения первичных параметров.

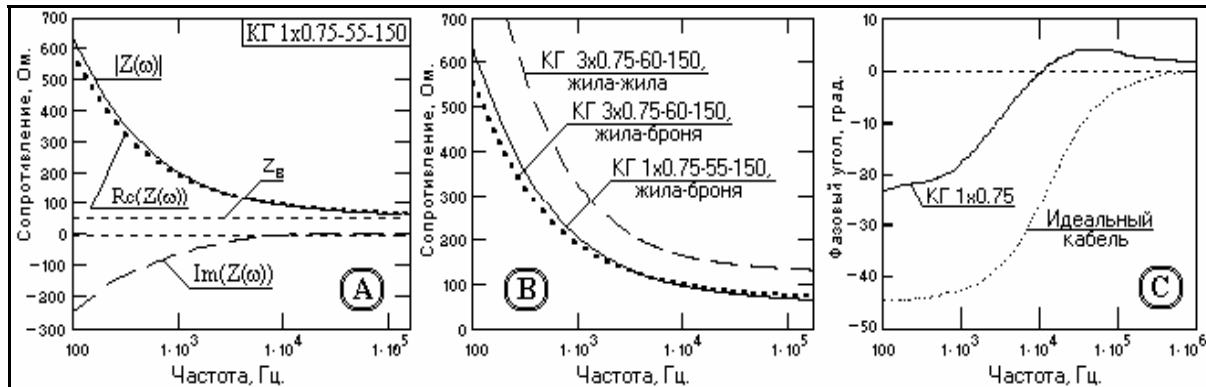


Рис. 13.2.1. Частотные характеристики волнового сопротивления реальных кабелей.

На рис. 13.2.1(А) приведены расчетные данные по волновому сопротивлению кабеля. При сравнении с идеальным кабелем можно видеть, что зависимость волнового сопротивления от частоты остается без изменения, несколько изменяются только численные значения зависимости. Волновое сопротивление жила-броня многожильных кабелей при однотипных параметрах жил и брони практически не отличается от одножильных, а для пар жила-жила отличается только увеличением числовых значений сопротивления (рис. 13.2.1(В)).

На рис. 13.2.1(С) приведено сопоставление расчетных фазочастотных характеристик волнового сопротивления реального (сплошная кривая) и идеального (нанесена пунктиром) кабеля. В целом характеристики сходны, но имеются и достаточно существенные отличия. За счет существенного возрастания значений индуктивности на низких частотах, значения и начальный интервал емкостной составляющей реального кабеля сокращаются, а на частотах выше 10 кГц в волновом сопротивлении кабеля появляется индуктивная составляющая, которая дает ощутимый фазовый сдвиг высокочастотных составляющих сигнала.

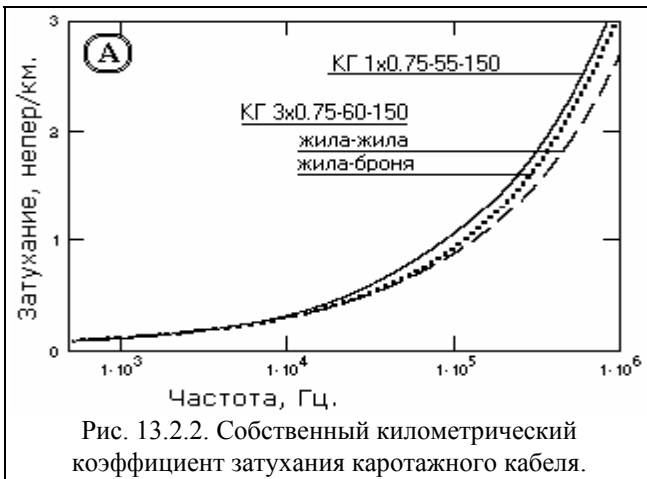
Следует ожидать, что значительная зависимость волнового сопротивления от частоты будет затруднять согласование кабеля как с наземной, так и со скважинной аппаратурой.

Частотные характеристики жил кабеля.

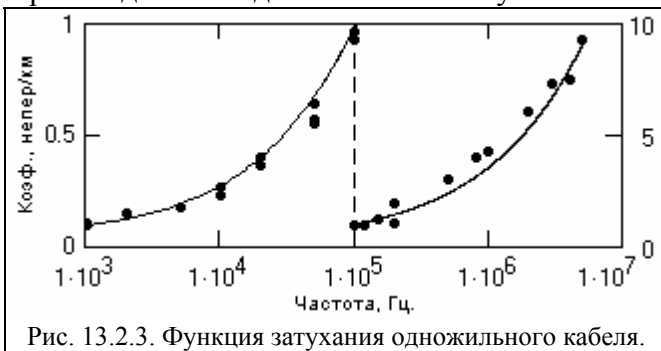
Характеристики коэффициента передачи сигналов по жилам кабеля, согласованным с нагрузкой по всему частотному диапазону, определяются передаточной функцией:

$$K(\omega) = \exp[-\gamma(\omega)\Lambda] = \exp[-(\beta(\omega) + j\alpha(\omega))\Lambda], \quad (13.2.2)$$

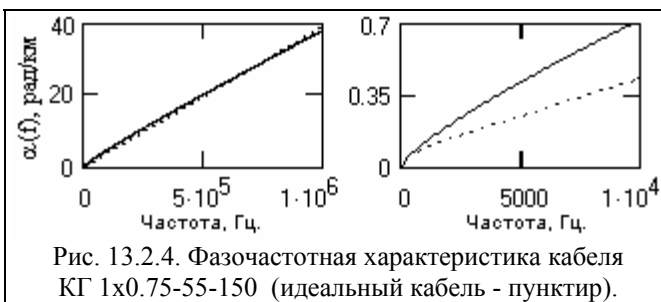
где: Λ – длина кабеля, $\gamma(\omega)\Lambda$ – комплексный коэффициент распространения кабеля, $\beta(\omega)\Lambda$ – частотная характеристика затухания кабеля (собственная частотная функция затухания), $\alpha(\omega)\Lambda$ – фазочастотная функция жилы кабеля. Передаточные функции являются системными функциями кабеля, обобщающими электрические параметры его токопроводящих жил.



На рис. 13.2.2 приведены функции километрического затухания близких по конструкции каротажных кабелей, вычисленные с учетом реальных частотно-зависимых первичных параметров жил. Нетрудно убедиться в весьма существенном отличии затухания реального кабеля от затухания идеального кабеля, что определяется, в основном, электромагнитными потерями высокочастотной энергии в реальных кабелях. Характер частотных зависимостей на низких частотах для одно- и многожильных кабелей практически однотипный. На частотах более 10 кГц затухание одножильных кабелей несколько выше затухания многожильных, причем для последних меньшее затухание имеют пары жила-жила.



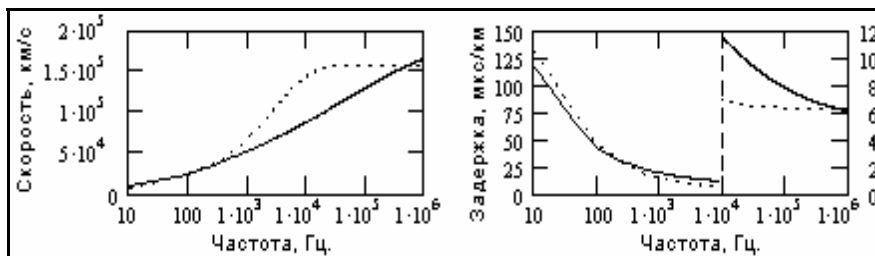
О достаточной достоверности расчетных данных можно судить по рис. 13.2.3, где приведено сопоставление расчетных значений собственного километрического коэффициента затухания одножильного бронированного кабеля с экспериментальными данными измерений коэффициента затухания сигнала на кабеле, как это принято – в неперах на 1 км, на разных частотах (нанесены точки).



Что касается фазочастотной характеристики реального кабеля, то она практически не отличается от характеристики идеального кабеля по форме и имеет небольшие изменения числовых значений. Коэффициент фазового сдвига частотных составляющих сигнала, приведенный на рис. 13.2.4, за исключением начальной части до частоты порядка 10 кГц, остается

прямо пропорциональным частоте.

Однотипность частотных характеристик одно- и многожильных кабелей позволяет в дальнейшем при рассмотрении зависимостей передаточных функций кабелей от значений различных параметров ограничиваться типичной средней частотной характеристикой жил, которая соответствует паре жила-броня многожильного кабеля.



Изменения фазочастотной характеристики кабеля приводят к соответствующим изменениям скорости передачи и временной задержки гармоник по кабелю без существенного изменения характера их частотной зависимости, что можно видеть на рис. 13.2.5.

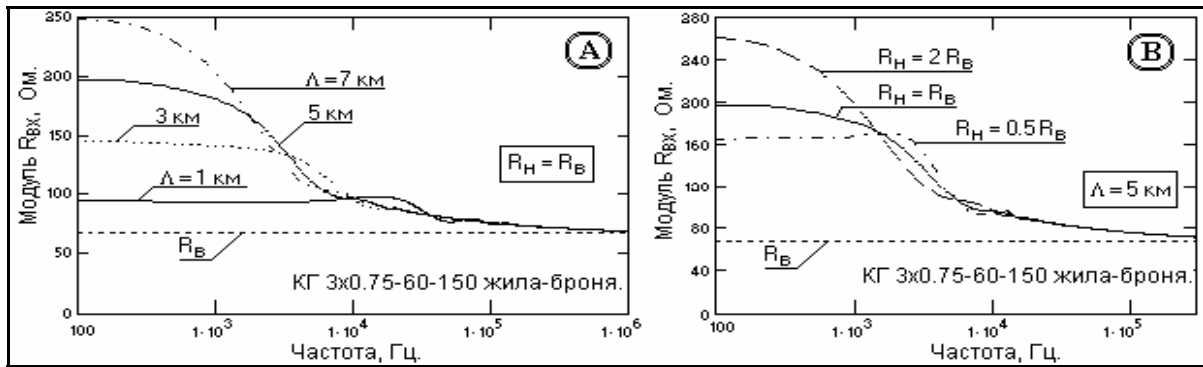


Рис. 13.2.6. Входное сопротивление кабеля.

На рис. 13.2.6 приведены расчеты по формуле (12.2.2) зависимости входного сопротивления жил от частоты при различной длине кабеля и сопротивлении нагрузки. Частотные функции модуля входного сопротивления и по форме, и по значениям, мало отличается от функций идеального кабеля (рис. 12.3.3). Изменение величины нагрузки сказывается только на изменении входного сопротивления в области частот ниже 10 кГц. Несогласованность с нагрузкой в реальном кабеле не вызывает осцилляций функций в области высоких частот, что объясняется пониженной добротностью реального кабеля.

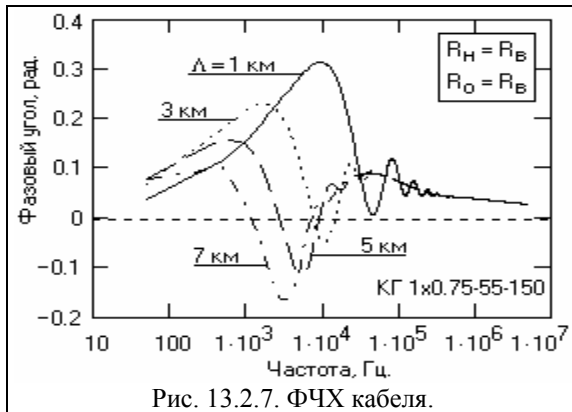


Рис. 13.2.7. ФЧХ кабеля.

Что касается фазочастотной характеристики кабелей, пример которой приведен на рис. 13.2.7, то она по сравнению с идеальным кабелем изменяется существенно и приобретает преимущественно индуктивный характер, особенно для

коротких кабелей в области средних и низких частот. Это может быть объяснено довольно существенным увеличением индуктивности реального кабеля по мере уменьшения частоты.

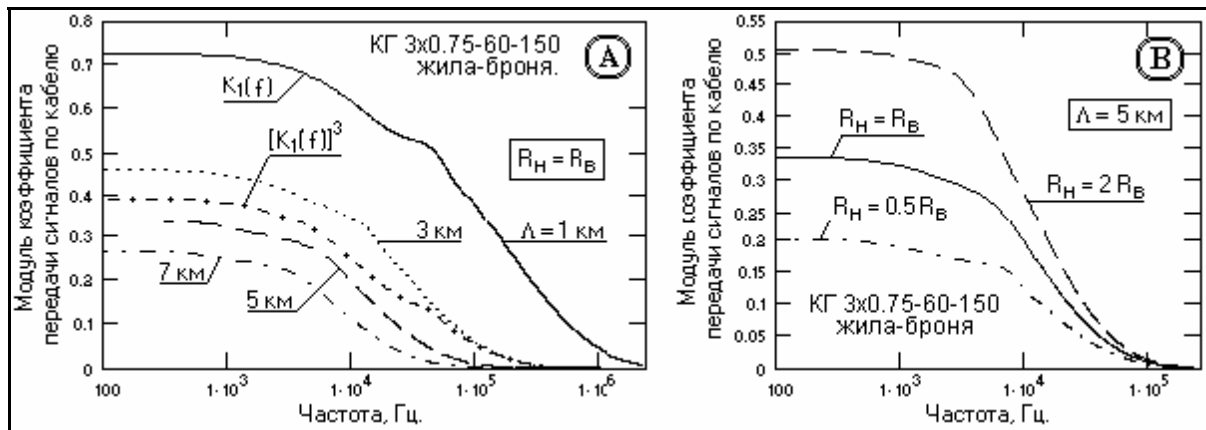


Рис. 13.2.8. Амплитудно-частотные характеристики ТПЖ кабеля.

На рис. 13.2.8 приведены графики модулей рабочих коэффициентов $K_p(f)$ передачи сигналов по кабелю, вычисленные по формуле (12.2.1'). Сравнением с рис. 12.3.1 можно убедиться в существенном отличии амплитудных характеристик реального кабеля от аналогичных функций идеального, особенно в области высоких частот. Электромагнитные потери энергии в реальном кабеле существенно ограничивают его частотный диапазон. Влияние согласования кабеля с нагрузкой также достаточно существенно по всему частотному диапазону передачи (рис. 13.2.8(B)) и требует стабилизации ее величины при влиянии дестабилизирующих факторов.

Фазочастотная характеристика кабеля практически не отличается от ФЧХ идеального кабеля и остается прямо пропорциональной значению $\omega\lambda$, т.е. частоте сигнала и длине кабе-

ля, при этом коэффициент пропорциональности по своему значению несколько зависит от величины нагрузки кабеля только в области низких частот (менее 20 кГц) и только для коротких кабелей. При длине кабеля более 1.5 км его величина при изменении сопротивления нагрузки в пределах $(0.5 \div 2) R_B$ остается практически постоянной.

Неполная согласованность жил кабеля с нагрузкой по частотному диапазону даже при выполнении равенства $R_H = R_B$ проявляется в том, что не существует кабельного дискрета, который мог бы использоваться для моделирования кабеля любой длины, как линейной системы, последовательным соединением дискретов. При таком моделировании результирующий импульсный отклик кабеля на единичный входной сигнал получается путем последовательной свертки импульсных откликов кабельных дискретов, что отображается в частотной области перемножением спектров импульсных откликов дискретов. Однако, как это можно видеть на рис. 13.2.8 (А), даже при использовании достаточно длинного дискрета длиной 1 км коэффициент передачи сигнала для кабеля длиной 3 км, полученный перемножением спектров трех дискретов $- [K_1(f)]^3$, существенно отличается от коэффициента передачи, вычисленного по формуле (12.2.1'). Следовательно, реальный кабель определенной длины даже при оптимальном согласовании с нагрузкой ($R_H = R_B$) должен рассматриваться в виде самостоятельной системы передачи сигналов с индивидуальными для него частотными характеристиками.

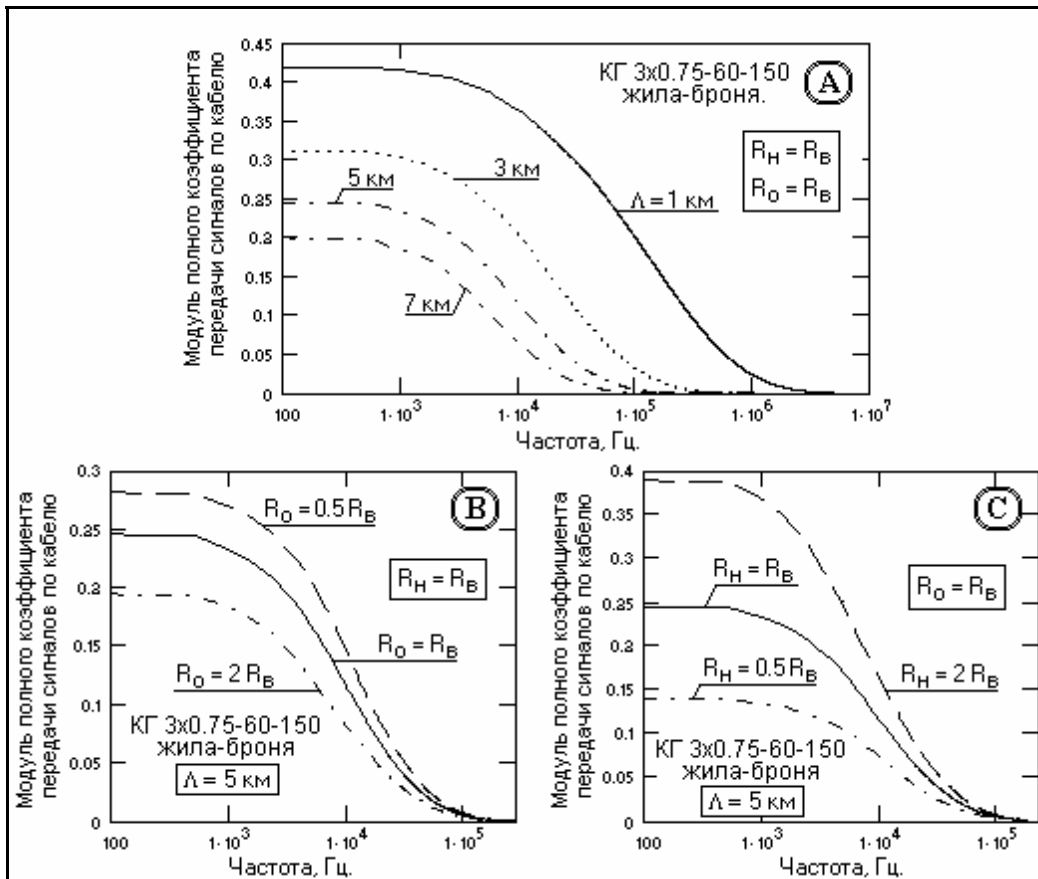


Рис. 13.2.9. Полный коэффициент передачи сигналов кабельной линией.

Формулы (12.2.1) не учитывают выходного сопротивления источника сигналов, т.е. отражают коэффициент передачи сигналов с входа линии на нагрузку при формировании (и измерении) амплитудных значений сигнала непосредственно на входе кабеля (сопротивление источника сигнала много меньше входного сопротивления кабеля на всех частотах). Расчет полного коэффициента передачи сигналов при формировании сигнала на входе источника сигналов, т.е. с учетом падения напряжения на внутреннем сопротивлении источника сигналов, выполняется по той же формуле (12.2.1) с соответствующим уточнением значения входного напряжения U_0 :

$$U_o(f) = U_c(f) Z_{вх}(f)/(Z_{вх}(f) + Z_o), \quad (13.2.3)$$

где U_c – сигнал на входе источника сигналов с внутренним сопротивлением Z_o .

На рис. 13.2.9 приведены графики частотной зависимости модуля полного коэффициента передачи сигналов кабельной линией от значений длины кабеля и сопротивления источников сигнала и нагрузки. Как следует из графиков, согласование кабеля с источником сигналов улучшает форму частотной характеристики и практически не изменяет относительную зависимость коэффициента передачи сигналов от частоты. Фазочастотные характеристики жил также практически не изменяются. Вместе с тем увеличивается относительная зависимость коэффициента передачи от согласования с нагрузкой. Очевидно, что только определенная стабилизация значений сопротивления нагрузки и источника сигналов от влияния дестабилизирующих факторов может гарантировать стабильность коэффициента передачи сигналов.

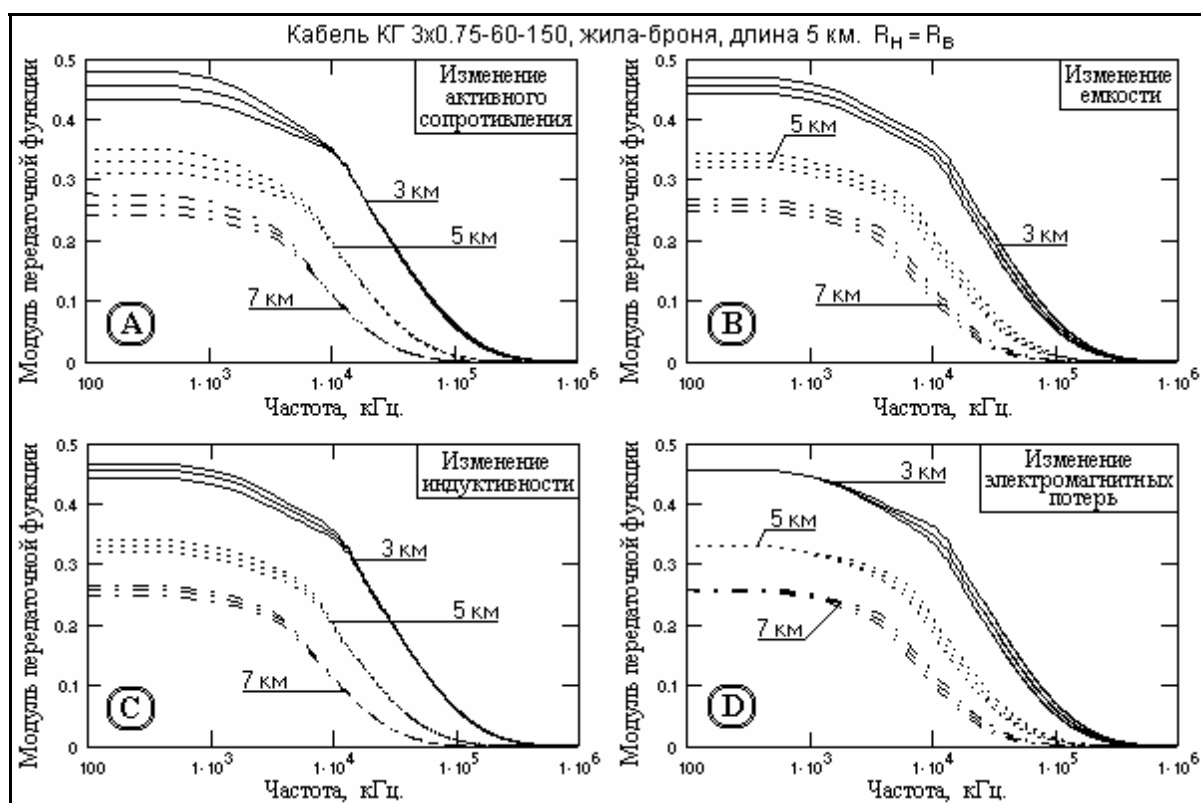


Рис. 13.2.10. Изменения амплитудно-частотной характеристики ТПЖ при изменении ее первичных параметров в пределах $\pm 10\%$.

Передаточная функция кабеля позволяет дать интегральную оценку влияния изменения первичных параметров жил на амплитудно-частотную характеристику кабеля. На рис. 13.2.10 приведены графики передаточных функций при возможных в процессе эксплуатации изменениях первичных электрических параметров жил в пределах $\pm 10\%$ от номинальных.

Как следует из этих графиков, на низких частотах степень чувствительности АЧХ жил кабеля порядка 0.5% на 1% изменения активного сопротивления жил, $\sim 0.2\%$ на 1% изменения емкости и индуктивности, и практически отсутствует при изменении электромагнитных потерь. На высоких частотах характеристики жил практически нечувствительны к изменениям активного сопротивления и индуктивности при повышении степени чувствительности к изменениям емкости и электромагнитным потерям. Учитывая, что в процессе эксплуатации изменения емкости и электромагнитных потерь много меньше изменений активного сопротивления и индуктивности (температурных и пр.) и носят, в основном, долговременный характер (старение кабеля), особое внимание следует обратить на снижение влияния активного сопротивления и индуктивности на передачу сигналов, тем более что это может достигаться уже рекомендованным методом в разделе 2: энергия сигналов должна быть минимальной в

области низких частот.

Что касается устойчивости математической модели передаточных функций жил кабеля к изменению постоянных коэффициентов во всех расчетных формулах (частотная постоянная «а» в (13.1.1"), коэффициент учета конструкции кабеля «b» в (13.1.5) и пр.), то степень чувствительности передаточных функций к их изменению на порядок ниже, что свидетельствует о достаточной точности модели (не хуже 3-5% по рекомендованным средним значениям). Исключением является коэффициент электромагнитных потерь η прямого действия, значение которого устанавливается непосредственно по передаточной функции (затуханию сигнала на высоких частотах) и приводит модель к фактическим (или паспортным) данным реального кабеля, фиксируя тем самым все возможные факторы отклонения характеристик реального кабеля от идеальной линии передачи сигналов.

13.3. ИМПУЛЬСНЫЙ ОТКЛИК КАБЕЛЯ

Полной математической моделью кабеля в реальном масштабе времени, которой обобщаются все электрические параметры жил, является импульсный отклик жил – форма сигнала на нагрузке жилы при подаче на ее вход единичного дельта-импульса. Импульсный отклик находится обратным преобразованием Фурье частотной передаточной функции кабеля: $h(t) \leftrightarrow K_p(f)$. Точность математической модели передаточной функции определяет соответствующую точность модели импульсного отклика жилы кабеля.

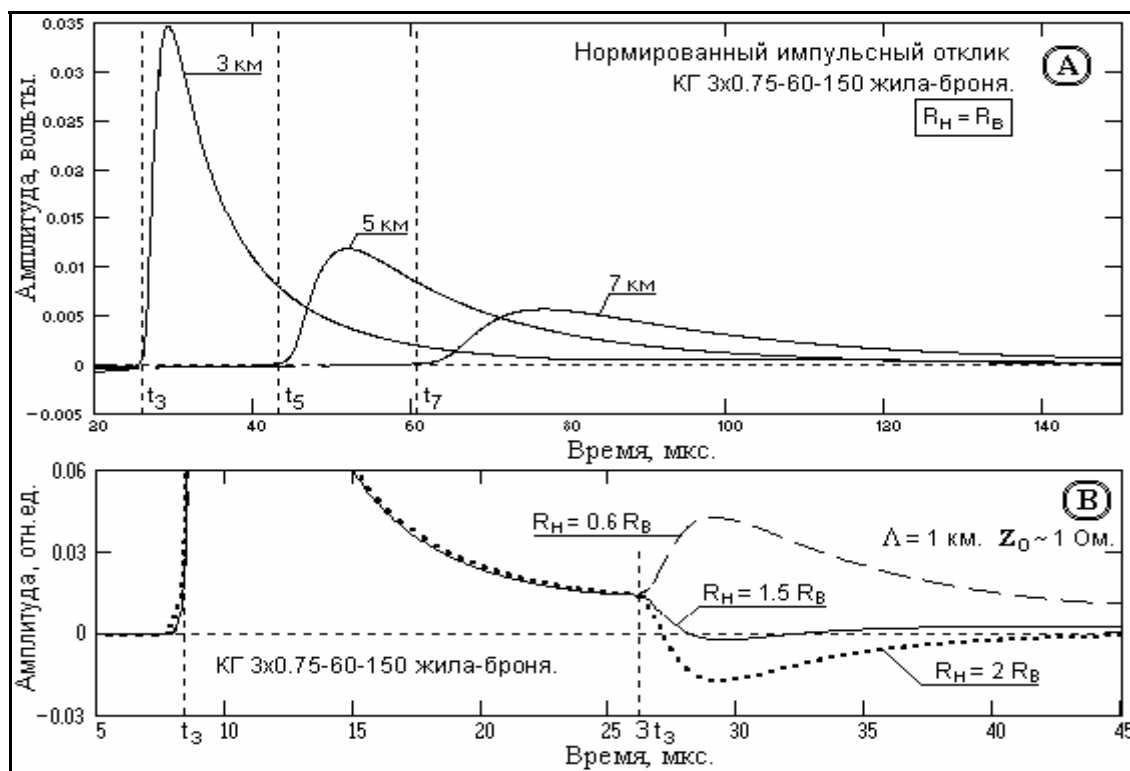


Рис. 13.3.1. Нормированные импульсные отклики кабеля.

При работе в частотном диапазоне до 1 МГц временной масштаб импульсного отклика целесообразно установить в микросекундах и вычислять отклик кабеля на единичный входной импульс $U \cdot \Delta t = 1$ (вольт·микросекунда), при этом масштаб значений интегрального импульсного отклика кабеля на микросекундной временной оси также будет измеряться в вольтах, а площадь импульсного отклика (в вольт·мкс), как и положено при прохождении импульса в пассивной линейной системе с потерями энергии, будет меньше 1.

На рис. 13.3.1(A) приведены импульсные отклики кабеля, вычисленные преобразованием Фурье по функции $K_p(f)$. Задержка t_3 фронта откликов (показана пунктиром) соответствует расчетному времени задержки на высоких частотах. Форма импульсного отклика жилы кабеля складывается из двух примерно экспоненциальных функций: короткой "зарядной" функции, определяющей фронт отклика, и длинной "разрядной" реакции жилы. При полном

согласовании модели с электрическими параметрами жилы на интервале $0-t_3$, не должно быть никакой реакции, что позволяет производить уточнение основных коэффициентов модели (K_c , D и η) приведением отклика на данном временном интервале к мало значимым значениям и их обнулением (при условии, что значение обнуляемой площади отклика не превышает 1-2% его полной площади). В связи с зависимостью рабочего коэффициента передачи сигнала от длины кабеля вышеописанное уточнение импульсного отклика для кабелей определенной длины, если имеется необходимость в повышении точности модели, может производиться индивидуально. Появление при расчете существенного отрицательного выброса перед фронтом импульсного отклика свидетельствует о заниженном значении коэффициентов η или эффективного диаметра D и сопровождается, как правило, занижением расчетных значений волнового сопротивления R_B относительно фактических (измеренных или среднестатистических). Ликвидацию выброса целесообразно выполнять подбором сначала коэффициента η (грубо), а затем диаметра D (точно).

При $Z_0 \Rightarrow 0$ форма собственного импульсного отклика кабеля существенно зависит от согласования с нагрузкой. Это определяется тем, что при несогласованной нагрузке отраженная от нагрузки волна полностью отражается от источника сигнала и через утроенное время задержки снова появляется на выходе кабеля, создавая второй пик на спаде отклика, положительный при $R_H < R_B$ или отрицательный при $R_H > R_B$, что можно видеть на рис. 13.3.1(B) в относительных единицах от максимума отклика. Это позволяет подбором нагрузки $R_H > R_B$ использовать повторный отрицательный пик для компенсации спада импульсного отклика и уменьшать его длительность (пример приведен на рисунке). Оптимальное согласование достигается при значении R_H порядка $(1.2-1.5)R_B$ в зависимости от длины кабеля (при увеличении длины кабеля этот эффект уменьшается в связи с большим затуханием отраженных волн при двойном прохождении по кабелю). Для кабелей, согласованных с источником сигнала ($Z_0 \sim R_B$), эффект несогласования с нагрузкой уменьшается практически на порядок.

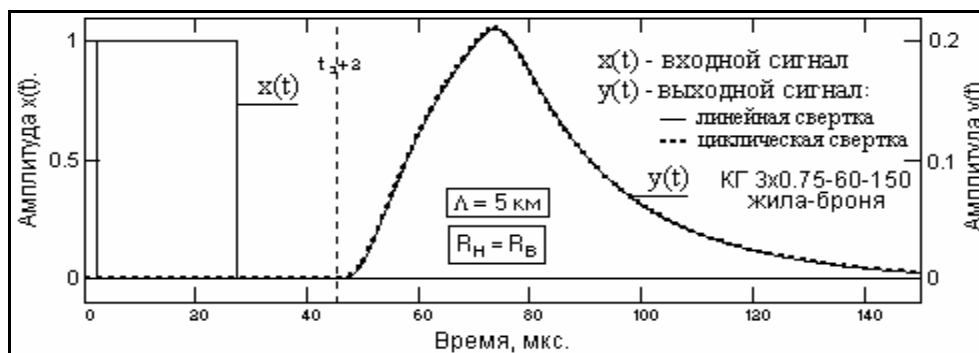


Рис. 13.3.2. Форма сигналов.

На рис. 13.3.2 приводятся результаты сопоставления формы импульса на нагрузке жилы кабеля при двух вариантах расчета:

1. Линейная свертка входного сигнала $x(t)$ с импульсным откликом кабеля $h(t)$ (свертка во временной области в дискретном варианте с интервалом дискретизации данных через 0.1 мкс):

$$y(t) = h(t) * x(t-\tau).$$

2. Через спектральную область (с использованием быстрого преобразования Фурье) умножением спектра сигнала на передаточную функцию кабеля (так называемая циклическая свертка):

$$x(t) \Rightarrow X(f). Y(f) = X(f) \cdot K_p(f). Y(f) \Rightarrow y(t).$$

Форма сигналов практически идентична, что позволяет использовать для дальнейшего анализа и вычислений как циклическую свертку, так и оператор импульсного отклика.

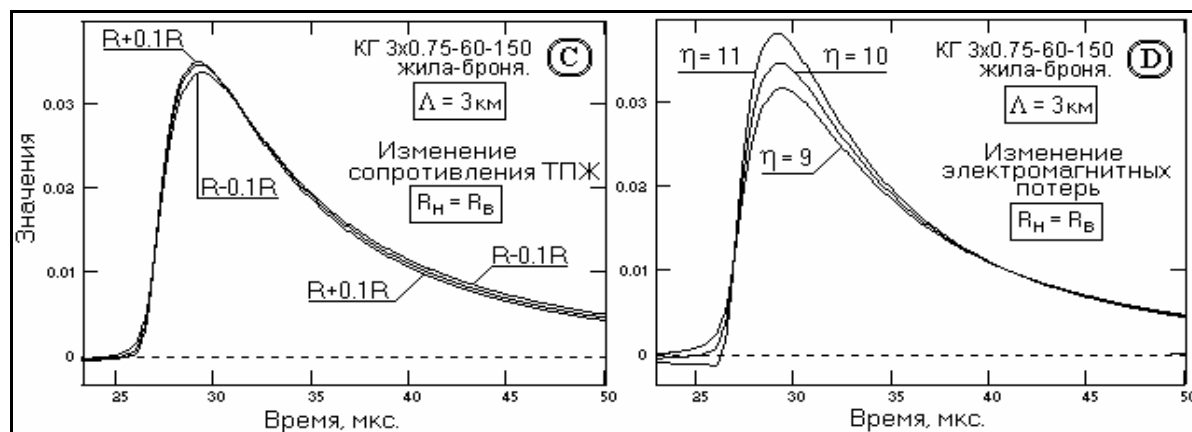


Рис. 13.3.3. Зависимость импульсных откликов ТПЖ кабеля от изменения первичных электрических параметров.

Влияние изменения первичных электрических параметров жил на форму импульсного отклика приведено на рис. 13.3.3. Активное сопротивление жилы практически не изменяет форму отклика и несколько изменяет постоянную времени его «хвоста» за счет изменения низкочастотного состава спектра передаточной функции. Аналогично влияние изменения индуктивности жил. Более существенно изменение амплитудных значений импульсного отклика при изменении электромагнитных потерь, относительное значение изменения которых приводит к практически такому же относительному изменению амплитуды отклика (обратному по знаку), так как пиковое значение отклика формируется высокочастотными составляющими спектра передаточной функции жил. Влияние изменения емкости жил и по содержанию, и по числовым значениям практически аналогично влиянию изменения электромагнитных потерь.

Таким образом, реальный кабель представляет собой линию связи с пониженной добротностью и существенной нелинейной зависимостью передачи сигналов от длины кабеля. Каждый тип каротажного кабеля определенной длины представляет собой самостоятельную и индивидуальную систему передачи сигналов. Фактор пониженной добротности каротажных кабелей в рамках общей теории однородных длинных линий можно учитывать введением в расчетные формулы вторичных электрических параметров жил дополнительного коэффициента электромагнитных потерь.

Математические модели передаточной функции и импульсного отклика кабеля с учетом его фактических частотно-зависимых электрических параметров и пониженной добротностью отображают реальные электрические параметры кабеля с точностью не хуже 5%.

ЛИТЕРАТУРА

6. Горбенко Л.А., Месенжник Я.З. Кабели и провода для геофизических работ. - М.: Энергия, 1977.
7. Гроднев И.И., Фролов Н.А. Коаксиальные кабели связи. - М.: Радио и связь, 1983. - 209 с.
23. Стрижевский Н.З. Коаксиальные видеопровода. - М.: Радио и связь, 1988. - 200 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 14: ИМПУЛЬСНЫЕ ПАРАМЕТРЫ КАРОТАЖНОГО КАБЕЛЯ

Когда дело доходит до разбора конкретных случаев, все оказывается сложнее.

Альбер Камю. Физик, XX в.

Сложности – дело житейское. Сами создаем, сами преодолеваем. В этом весь смысл исконно российской героики.

Евгений Кучурин. Уральский геофизик, XX в.

Содержание: 14.1. Индекс неопределенности кабеля. Принцип неопределенности. Ограничения передачи сигналов. 14.2. Оптимальная форма сигналов. 14.3. Кодовые сигналы. Литература.

ВВЕДЕНИЕ

Импульсная пропускная способность кабеля, равно как и возникающие при этом искажения передаваемых сигналов, полностью определяются передаточной функцией жил. Однако передаточная функция, как индивидуальная характеристика жилы определенной длины, мало пригодна для сопоставления эффективности различных методов кодирования данных и сравнения различных типов кабелей по скорости передачи данных. Для этих целей желательно иметь нормированную числовую характеристику.

14.1. ИНДЕКС НЕОПРЕДЕЛЕННОСТИ КАБЕЛЯ [23]

Все дальнейшие вычисления проводятся на кабеле, согласованном с нагрузкой, если отклонения от согласованности специально не оговорены.

Принцип неопределенности. Эффективная длительность произвольного импульсного сигнала (или импульсного отклика линейной системы) и эффективная ширина его спектра (передаточной функции системы) связаны определенным постоянным соотношением. Чем меньше длительность сигнала (отклика), тем больше ширина его спектра. Согласно принципу неопределенности, для любого сигнала и системы должно выполняться соотношение:

$$\Delta T \Delta \Omega \geq 1/\pi, \quad (14.1.1)$$

где ΔT – длительность сигнала (или импульсного отклика системы), $\Delta \Omega$ - ширина спектра сигнала (полоса пропускания системы), при условии, что мерой эффективной длительности сигнала и эффективной ширины спектра служат вторые моменты функций распределения сигнала и спектра, т.е. средние квадратические отклонения от центра временного распределения сигнала и от центра спектрального распределения (последний, при двустороннем распределении относительно нуля частот, равен нулю). Для произвольной системной функции $h(t) \Leftrightarrow H(\omega)$, заданной на интервале $0 - \infty$, значения ΔT и $\Delta \Omega$ определяются по формулам:

$$(\Delta T)^2 = 4 \cdot \left[\frac{\int_0^{\infty} t^2 h^2(t) dt}{\int_0^{\infty} h^2(t) dt} - \left(\frac{\int_0^{\infty} t \cdot h^2(t) dt}{\int_0^{\infty} h^2(t) dt} \right)^2 \right], \quad (14.1.2)$$

$$(\Delta \Omega)^2 = 4 \cdot \frac{\int_{-\infty}^{\infty} f^2 |H(f)|^2 df}{\int_{-\infty}^{\infty} |H(f)|^2 df}. \quad (14.1.3)$$

В таблице 14.1.1 приведены расчетные значения параметров кабеля, вычисленные по вышеприведенным формулам. Полная физическая длительность импульсного отклика в первом приближении должна приниматься равной $2\Delta T$, а ширина спектра (с учетом отрицательных частот) равной $2\Delta \Omega$. В области реальных (только положительных) частот параметр $\Delta \Omega$ может отождествляться с реальной полосой пропускания системы (реальной шириной спектра сигнала).

Таблица 14.1.1. Импульсные параметры кабелей.

Тип кабеля и ТПЖ	ОБК			МБК, жила-броня			МБК, жила-жила		
	3	5	7	3	5	7	3	5	7
Длина кабеля, км.	3	5	7	3	5	7	3	5	7
ΔT_k отклика, мкс.	19.2	32.9	48.3	14.1	26.1	40.9	8.21	17.4	29.7
$\Delta \Omega_k$ отклика, кГц.	67.6	26.2	15.2	83.6	34.4	19.4	102	40.1	22.1
Индекс $\Delta T_k \Delta \Omega_k$	1.30	0.86	0.73	1.18	0.89	0.79	0.84	0.70	0.66

Ограничения передачи сигналов. В принципе, индекс неопределенности кабеля $\Delta T_k \Delta \Omega_k$, равно как и входящие в него параметры ΔT_k и $\Delta \Omega_k$, не устанавливают каких-либо жестких ограничений по импульсной пропускной способности кабеля, особенно для импульсных сигналов, так как она существенно зависит как от системы формирования импульсных последовательностей и, соответственно, определенных априорных данных для идентификации сигналов на выходе кабеля, так и от уровня статистических шумов линии передачи сигналов. Однако следует учитывать, что при передаче импульсов с $\Delta \Omega > \Delta \Omega_k$ система приводит (в той или иной мере) индекс неопределенности сигналов к собственному индексу неопределенности, а при установленном значении $\Delta \Omega_k$ системы это накладывает ограничения на длительность импульсов. При этом различают полнометрическую, параметрическую и кодовую передачу сигналов.

При полнометрической передаче на выходе кабеля требуется сохранить форму входных сигналов, что, естественно, возможно только с определенной погрешностью и только для сигналов, эффективная ширина спектра которых минимум на порядок ниже ширины спектра передаточной функции кабеля и достаточно быстро затухает. В общем случае, возможность восстановления сигналов на выходе кабеля методами частотной коррекции передаточной функции кабеля, а равно и погрешность восстановления, определяется безвозвратными потерями высокочастотных информационных составляющих сигналов. С учетом этого импульсный сигнал произвольной формы без скачков и угловых изломов может быть восстановлен на выходе кабеля, если эффективная ширина его спектра не превышает эффективной ширины спектра кабеля, а сам сигнал задается на интервале не менее $2\Delta T_k$ с выходом на нулевые значения по его краям и длительность его фронтов не меньше длительности фронта импульсного отклика кабеля (по уровням 0.1 и 0.9 амплитудных значений).

При параметрической передаче сигналов кабель должен обеспечивать линейную передачу определенных параметров сигналов и достаточно точное измерение их значений на выходе кабеля. К ним обычно относятся: амплитудные значения, значения площади (энергии) сигналов или временные интервалы между сигналами. По существу, это частный и достаточно упрощенный вариант полнометрической передачи произвольных сигналов. Без восстановления формы сигналов минимальный временной интервал следования сигналов при определениях амплитуд и площадей (энергии) определяется длительностью спада импульсного отклика кабеля (до требуемых значений остаточной реакции на предыдущий импульс). При использовании систем коррекции (восстановления) сигналов форма и входных, и выходных сигналов (после коррекции) может быть произвольной. При формировании входных сигналов на интервале не менее $2\Delta T_k$ точность измерений может быть не менее 1%.

Идентификация кодовых сигналов гарантирована при любой форме приема сигналов (с коррекцией и без коррекции) и при любой форме входных сигналов, задаваемых на интервале $2\Delta T_k$. При использовании корректоров сигналов битовый интервал может быть сокращен до ΔT_k , а при априорно известном интервале битовых посылок и жесткой структуре кодовых слов – и еще более.

Как видно из данных, приведенных в таблице 14.1.1, каротажные кабели имеют довольно низкие параметры пропускания импульсных сигналов. Кроме того, все вышеизложенное не учитывает уровня статистических кабельных помех, которые могут значительно усложнить прием сигналов.

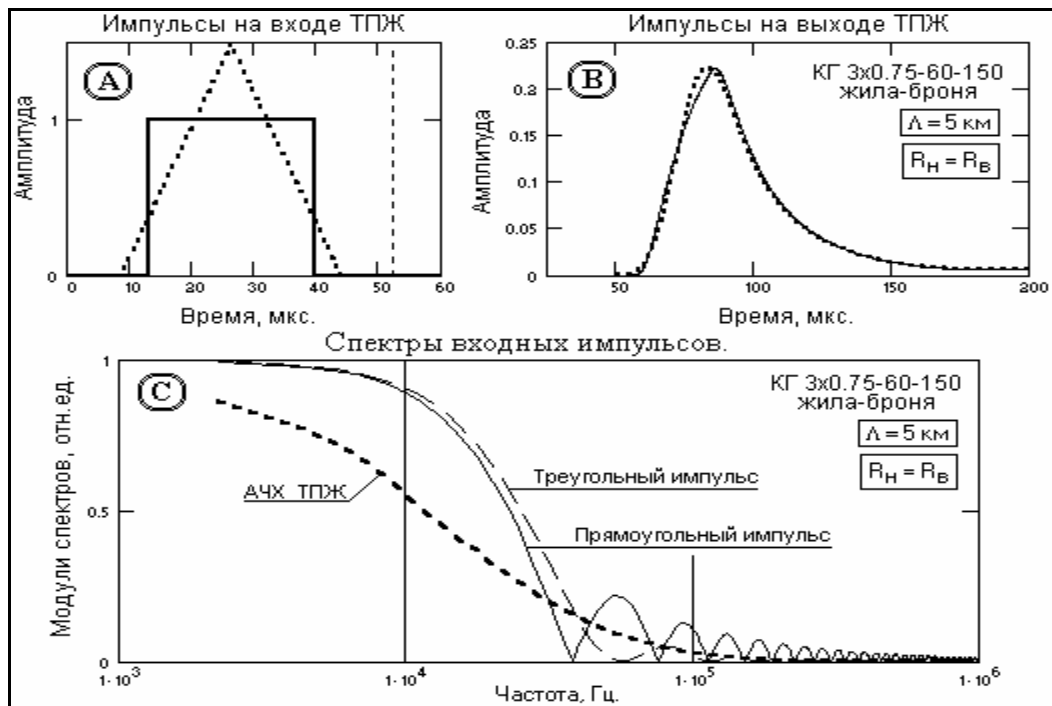


Рис. 14.1.1. Форма и спектры однополярных импульсов.

На рис. 14.1.1(А, В) для жил кабеля КГ 3x0.75-60-150 длиной 5 км приведена форма двух типов импульсов, полярных по своим индексам неопределенности, в границах интервала $2\Delta T_k$ кабеля (граница интервала выделена вертикальным пунктиром). Прямоугольный импульс из всех типов простых импульсов имеет наибольший индекс неопределенности. У треугольного сигнала индекс неопределенности только на 10% больше индекса гауссовских сигналов, имеющих предельно возможный минимальный индекс $1/\pi$. Эффективная длительность сигналов установлена такой, чтобы при свертке с импульсным откликом жилы амплитуды импульсов на выходе кабеля были примерно равными, что позволяет наглядно сравнить качество передачи данных сигналов по кабелю. На рис. 14.1.1(С) приведены модули спектров сигналов в высокочастотной области в сравнении с модулем передаточной функции жилы. Для наглядности сравнения спектры нормированы по максимальным значениям к 1.

Таблица 14.1.2. Параметры однополярных сигналов на нагрузке кабеля КГ 3x0.75-60-150.

Тактовый интервал	$T = 2\Delta T_k$				$T = \Delta T_k$			
	Входные		Выходные		Входные		Выходные	
Сигналы	П	Л	П	Л	П	Л	П	Л
Параметры	П	Л	П	Л	П	Л	П	Л
ΔT импульса, мкс.	15.3	11.2	30.2	29.7	7.8	5.5	27.3	27.1
$\Delta \Omega$ импульса, кГц.	326	31.2	16.2	16.6	456	63.6	22.7	23.8
Индекс $\Delta T \Delta \Omega$	4.98	0.35	0.49	0.49	3.56	0.35	0.62	0.64

Как следует из этих рисунков, оба типа сигналов имеют существенные потери в области частот выше $\Delta \Omega_k$. Форма сигналов на выходе жил становится практически идентичной. Последнее наглядно показывают импульсные характеристики сигналов, приведенные в таблице 14.1.2. При сокращении интервала задания импульсов до величины ΔT_k импульсные характеристики сигналов практически приводятся к импульсным характеристикам жил кабеля. Отсюда следует, что на предельных частотах передачи сигналов форма входных импульсов значения не имеет. В относительных единицах ΔT_k динамика временных процессов в кабелях различных типов и различной длины также идентична, что позволяет индекс неопределенности $\Delta T \Delta \Omega$ и входящие в него параметры ΔT и $\Delta \Omega$ считать основными импульсными числовыми характеристиками систем и использовать их в качестве относительных еди-

ниц сравнительного анализа. С учетом этого дальнейший анализ будем проводить, как правило, на один кабель (типа КГ 3x0.75-60-150) длиной 5 км с временными параметрами сигналов в единицах (относительных долях) параметра ΔT_k .

Простейший кодовый сигнал двоичной разрядности, это конечная последовательность однополярных битовых сигналов 0 или 1 с определенной тактовой частотой. В пределе, на одном тактовом периоде сигнал 1 может представлять собой импульс с единичной площадью и с длительностью, много меньшей длительности тактового интервала, т.е. физическое приближение дельта-импульса. Выходным сигналом кабеля в этом случае будет импульсный отклик кабеля. Соответственно, кодовое слово будет представлять собой сумму импульсных откликов на последовательность битовых сигналов.

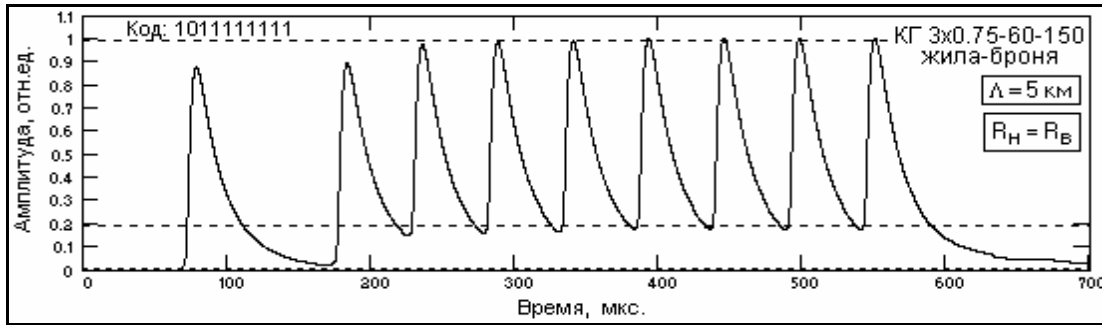


Рис. 14.1.2. Последовательность импульсных откликов на выходе кабеля (такт $2\Delta T_k$)

Пример такой выходной последовательности с длительностью тактового интервала, равной $2\Delta T_k$, кабеля, приведен на рис. 14.1.2. Предельное максимальное смещение нулевой линии последовательности, определяемое асимметрией импульсного отклика, не превышает $1/5$ амплитудного значения отклика и является максимально возможным для кабеля при любой последовательности кодовых импульсов и при любой их длительности в пределах тактового интервала. Выход на практически нулевые значения – 3 тактовых интервала после импульса. Сокращение тактового интервала до величины ΔT_k приводит к смещению нулевой линии до $2/3$ от амплитуды импульсного отклика и практически исключает возможность идентификации импульсов на выходе кабеля, особенно в присутствии шумов.

Таким образом, при передаче по кабелю двоичных кодовых сигналов без применения на выходе кабеля частотных корректоров максимальная частота гарантированной передачи однополярных сигналов по кабелю соответствует тактовым интервалам, равным удвоенному значению эффективной ширины импульсного отклика кабеля, при этом сигналы с гладкой формой и минимальной шириной спектральной характеристики не имеют преимуществ перед прямоугольными сигналами.

14.2. ОПТИМАЛЬНАЯ ФОРМА СИГНАЛОВ

Однополярные импульсы, рассмотренные в предыдущем разделе, по отношению к электрическим параметрам кабеля и его частотным характеристикам не являются оптимальными. Практически все первичные и вторичные электрические характеристики кабеля, за исключением емкости, существенно нелинейны в области низких частот, где сосредоточена значительная часть энергии однополярных сигналов. Соответственно, кабель оказывается рассогласованным на этих частотах с источником сигнала и нагрузкой, а стабильность передачи сигналов зависит от изменения степени рассогласования при влиянии на кабель дестабилизирующих факторов (температуры, условий заземления оплетки в скважине и пр.). При однополярных сигналах на ограничение импульсной пропускной способности кабеля оказывает свое влияние и существенное смещение нулевой линии, определяемое резкой асимметрией импульсного отклика кабеля и, соответственно, зависимое от конкретной импульсной нагрузки кодового слова. По этой же причине для однополярных кодов не может быть выполнена система надежной автоматической синхронизации тактовых частот передатчика и приемника, весьма существенная для каротажных станций.

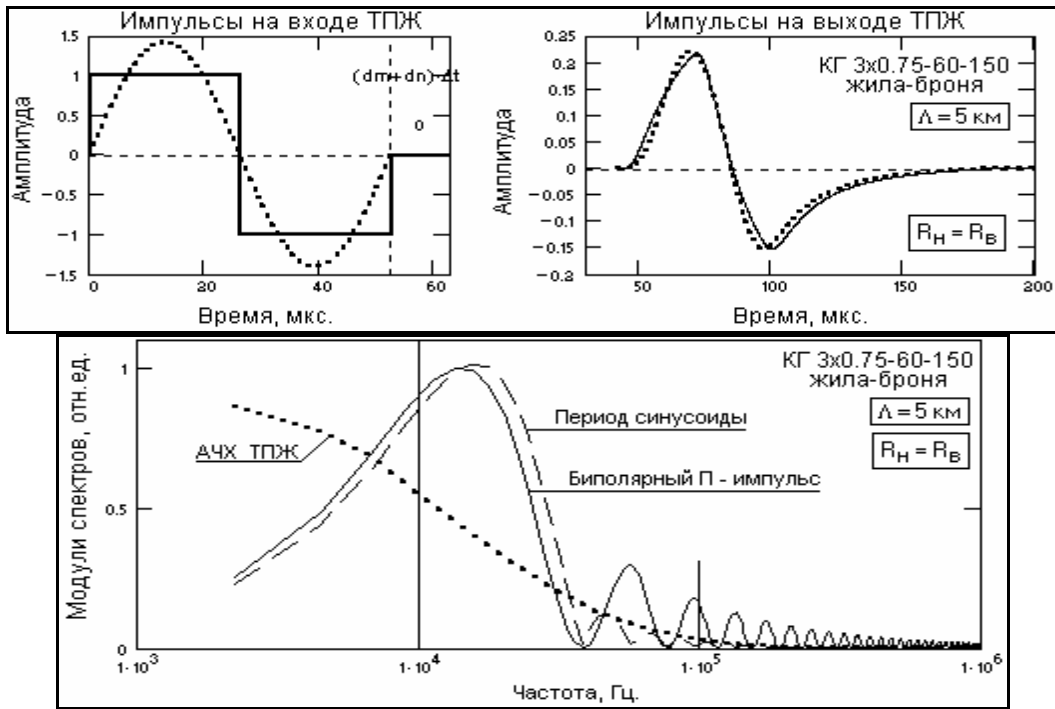


Рис. 14.2.1. Форма и спектры биполярных импульсов.

Решение данной проблемы в радиотехнике известно: частотное перераспределение энергии сигналов со смещением в средне- и высокочастотную область спектра кабеля, т.е. применение биполярных сигналов с нулевым средним значением по тактовым интервалам.

На рис. 14.2.1 приведена форма двух входных и выходных биполярных сигналов равной энергии пределах одного тактового периода (сплошная линия – один период меандра, пунктир – период синусоиды). Как и в предыдущем разделе, в качестве примера используется кабель КГ 3x0.75-55-150 длиной 5 км, а параметры сигналов и период их следования определены с привязкой на параметры кабеля значением $T = 2\Delta T_k$. Среднее значение сигналов в пределах тактового периода равно нулю. Модули спектров сигналов приведены в нормированном виде в сопоставлении с передаточной функцией кабеля (пунктирная кривая). Импульсные параметры сигналов на входе и выходе кабеля приведены в таблице 14.2.1.

Таблица 14.2.1. Параметры биполярных сигналов на кабеле КГ 3x0.75-60-150.

Тактовый интервал	$T = 2\Delta T_k$				$T = \Delta T_k$			
	Входные		Выходные		Входные		Выходные	
Сигналы	Би-П	Син.	Би-П	Син.	Би-П	Син.	Би-П	Син.
Параметры								
ΔT импульса, мкс.	30.3	27.9	39.3	36.2	15.2	13.9	24.9	23.5
$\Delta \Omega$ импульса, кГц.	400	38.2	26.3	28.2	566	76.4	44.1	48.0
Индекс $\Delta T \Delta \Omega$	12.1	1.06	1.03	1.02	8.06	1.07	1.10	1.13

Естественно, что для биполярных сигналов значение $\Delta \Omega$ увеличилось, но при этом существенно увеличились и значения $\Delta \Omega$ выходных сигналов и амплитудные характеристики выходных сигналов изменились несущественно. Гладкая форма входных сигналов на предельных тактовых частотах не имеет преимуществ перед прямоугольными импульсами, по крайней мере, для устройств без частотной коррекции передаточных функций жил.

Пример двоичного кодового сигнала, переданного биполярными П-импульсами с тактовым интервалом $T = 2\Delta T_k$, приведен на рис. 14.2.2(В). Пунктиром на рисунке приведена форма единичного (первого) импульса. При сравнении с рис. 14.1.2 можно видеть преимущества биполярной передачи сигналов.

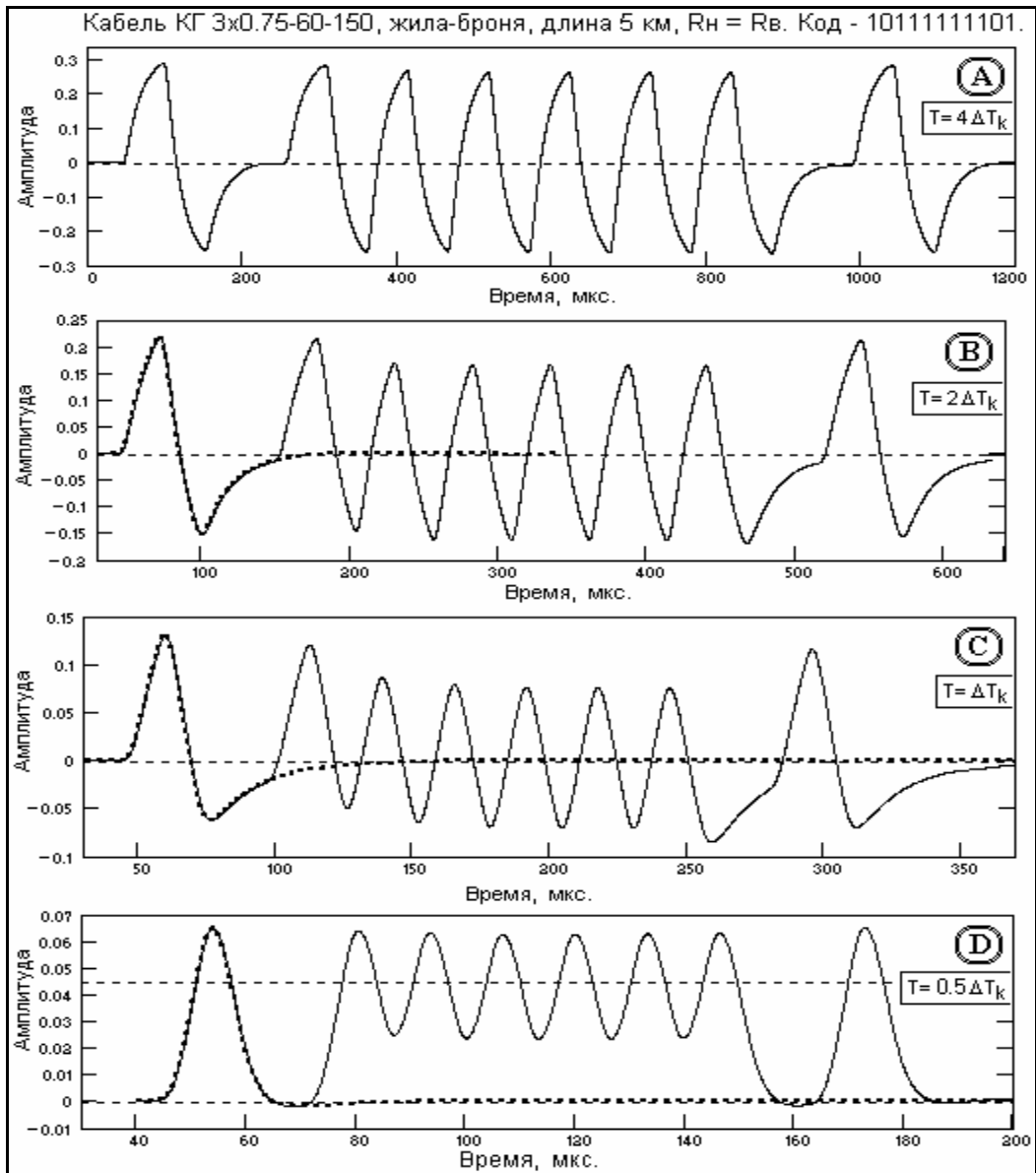


Рис. 14.2.2. Кодовый сигнал на выходе кабеля при различной тактовой частоте.

В первых, биполярные импульсы не дают смещения нулевой линии в пределах кодового слова, при этом по внутритактовому пересечению нуля может выполняться автоматическая синхронизация тактовых частот передатчика и приемника.

Во вторых, импульсы отрицательной полярности представляют собой дублирующую кодовую информацию, сдвинутую на половину такта. Она также может быть использована для повышения достоверности идентификации импульсов и допустимого уровня шумов (детектирование сигналов по полярности на два канала, задержка канала положительных импульсов на полтакта, инверсия полярности канала отрицательных импульсов и суммирование каналов, при этом амплитуды импульсов увеличиваются в 2 раза, а среднеквадратический уровень статистических шумов уменьшается в $\sqrt{2}$ раз).

На рис. 14.2.2(А,С) приведены выходные сигналы с тактовыми интервалами передачи $4\Delta T_k$ и ΔT_k при единичных амплитудах входных биполярных сигналов. На тактовых интервалах $4\Delta T_k$ взаимное влияние битовых импульсов исключается практически полностью. На интервалах ΔT_k взаимное влияние битовых сигналов становится достаточно существенным, но идентификация сигналов по синхроимпульсам сомнений не вызывает. Без специальных устройств декодирования сигналов этот интервал может считаться предельным для передачи данных с использованием биполярных кодов.

Попутно заметим, что первые импульсы в кодовых комбинациях при $T = \Delta T_k$, а также

импульсы после предыдущих нулевых битов, несколько больше по своим значениям последующих единичных импульсов в непрерывной последовательности. При необходимости этот эффект может быть устранен уменьшением амплитуды импульсов входного сигнала, следующих за нулевыми интервалами.

Следовательно, для биполярных сигналов значение ΔT_k можно считать тактовым интервалом гарантированной идентификации кода при любой форме сигнала.

В принципе, для биполярных сигналов интервал ΔT_k не является предельным. Биполярная форма импульсов позволяет подобрать соотношение длительности положительных и отрицательных частей импульсов (или их амплитудных значений) таким образом, чтобы частично компенсировать асимметрию импульсного отклика кабеля. На рис. 14.2.2(D) приведен пример формы сигналов на выходе кабеля для биполярных импульсов на входе кабеля с тактовым интервалом $0.5 \cdot \Delta T_k$ при равных значениях длительности импульсов и амплитуде отрицательного импульса порядка $\frac{3}{4}$ амплитуды положительного. Амплитудная асимметрия биполярного импульса создает "разрядный" импульс на входе кабеля, обратный ток которого, сдвинутый на полтакта, вычитается из тока среза "зарядного" импульса и компенсирует тем самым обратную по знаку асимметрию импульсного отклика кабеля. Результатом является практически симметричная однополярная форма единичного выходного сигнала (показана пунктиром) и одноуровневая форма кодового слова, которая хорошо декодируется.

Таким образом, оптимальной формой импульсов для каротажного кабеля при кодовой передаче информации, обеспечивающей максимальную скорость передачи данных, можно считать биполярный импульс типа одного периода меандра.

14.3. КОДОВЫЕ СИГНАЛЫ

Скорость передачи информации (бит/с) зависит не только от тактовой частоты передачи данных, но и от сетевого протокола кодирования информации. Хотя каротажную линию передачи данных с трудом можно отнести к информационной магистрали, скорее это специальная технологическая линия управляемого сбора данных, тем не менее, протокол передачи каротажных данных целесообразно заимствовать из стандартов открытых систем обмена информацией OSI (Open System Interconnect).

С учетом специфики каротажных систем и условий эксплуатации (один приемник, большое разнообразие и частая смена передатчиков – скважинных приборов) для каротажных систем используются, как правило, только двух и трехуровневые коды RZ (Return to Zero), NRZ (Non Return to Zero) и Манчестер-II.

Двухуровневый код NRZ (ноль – нижний уровень, 1 – верхний) и его вариант обратной полярности NRZI имеет информационные переходы на границе битов. Максимальная частота кода – при чередовании единиц и нулей, минимальная (нулевая) – при передаче последовательности одинаковых битов. Достоинство кода – максимальная простота. Основным недостатком – отсутствие надежной синхронизации и, соответственно, необходимость для синхронизации специальных (пакетных старт-стоповых) служебных битов.

Трехуровневый код RZ обеспечивает возврат к нулевому уровню после каждого бита информации. Логическому нулю соответствует положительный импульс, единице – отрицательный, Информационный переход осуществляется в начале бита, возврат к нулю – в середине бита. Код синхронизируется по битам переходом по его центру.

Код Манчестер-II является наиболее распространенным двухуровневым самосинхронизирующимся кодом. Логическому нулю соответствует переход на верхний уровень в центре битового интервала с возвратом на нижний по концу битового интервала, если следующий бит также нулевой. Соответственно, логической единице – переход на нижний уровень с возвратом на верхний по концу интервала, если следующий бит также 1. Бит обозначен переходом в центре бита, по которому выделяется синхросигнал. Максимальная частота кода – при передаче последовательности нулей или единиц. При чередовании нулей и единиц частота кода уменьшается в два раза. Несомненное достоинство кода – отсутствие постоянной составляющей при передачах длинных последовательностей нулей или единиц.

На рис. 14.3.1(A) приведен пример кодировки битовой последовательности тремя данными кодами по тактовым интервалам $2 \cdot \Delta T_k$ жилы кабеля. Выходные сигналы кабеля приведены на рис. 14.3.1(B). Масштаб выходных сигналов (утолщенные линии) – в относительных единицах амплитуды входных сигналов. Тонкими линиями показаны входные сигналы, приведенные к выходу кабеля с учетом безвозвратных потерь энергии умножением на площадь импульсного отклика и смещенные на время задержки.

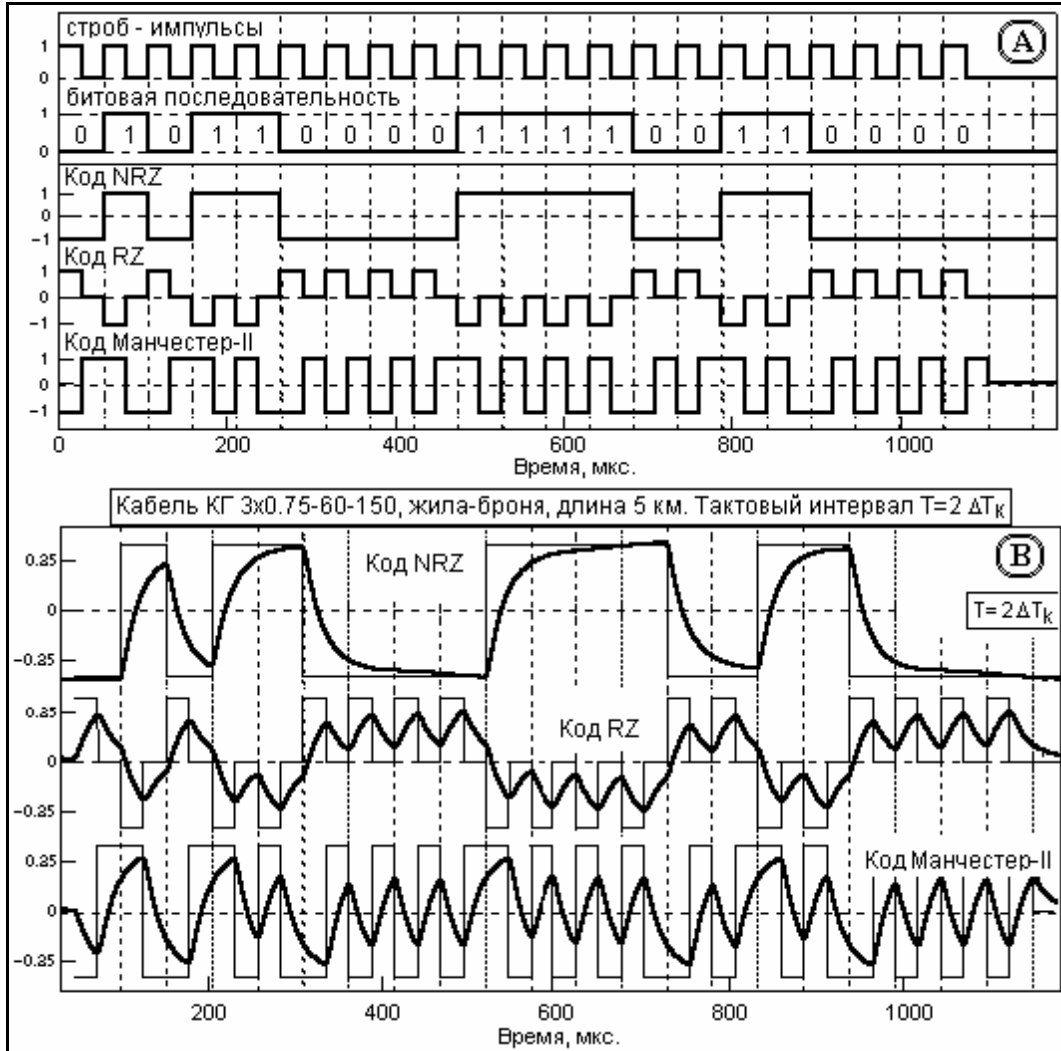


Рис. 14.3.1. Примеры кодировки сигналов и форма сигналов на выходе кабеля.

Сравнение кодов достаточно наглядно. Скорость передачи данных $f_T = 1/(2\Delta T_k)$ бит/с можно считать предельной для всех трех типов кодов. Код NRZ имеет преимущество и запас "прочности" по амплитудным параметрам выходных сигналов, но без "жесткой" синхронизации приемника надежное декодирование сигналов не гарантируется. Амплитудные характеристики кода Манчестер-II практически в 2 раза лучше кода RZ и, соответственно, выше помехозащищенность кода и надежность автосинхронизации.

Заметим также, что частотный спектр кода Манчестер-II содержит только две несущих частоты: f_T при передаче нулей и единиц, и $f_T/2$ при чередовании нулей и единиц (см. рис. 14.3.2). Это позволяет при приеме сигналов применять полосовые фильтры и тем самым повышать помехозащищенность линии передачи данных.

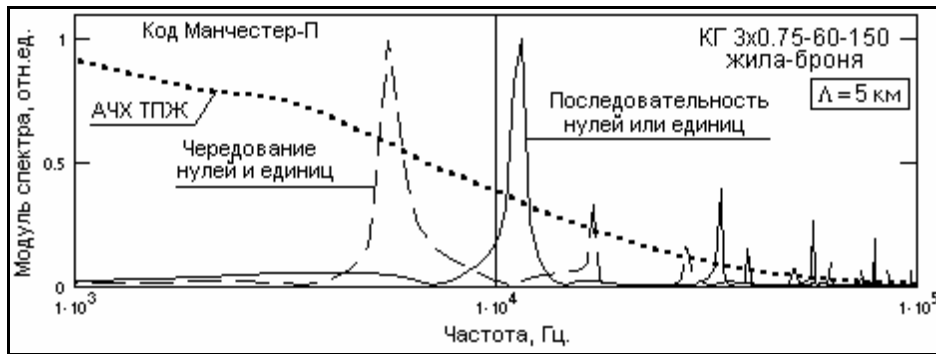


Рис. 14.3.2. Частотные характеристики кода Манчестер-П.

Выводы:

1. Основными числовыми характеристиками системы передачи данных, полностью определяющими скорость и качество передачи данных, являются эффективная ширина частотного спектра пропускания системы $\Delta\Omega$, эффективная ширина импульсного отклика системы ΔT и индекс неопределенности $\Delta T\Delta\Omega$.

2. Максимальная скорость передачи однополярных импульсных сигналов по кабелю соответствует тактовым интервалам, равным удвоенному значению эффективной ширины импульсного отклика кабеля. Максимальная скорость передачи биполярных сигналов соответствует тактовым интервалам, равным значению эффективной ширины импульсного отклика кабеля.

3. На предельной частоте передачи сигналы с гладкой формой и минимальной шириной спектральной характеристики (минимальным значением индекса неопределенности) не имеют преимуществ перед прямоугольными сигналами.

4. Оптимальной формой импульсов для каротажного кабеля при кодовой передаче сигналов, обеспечивающей максимальную скорость передачи данных, являются биполярные импульсы. Основным кодом передачи каротажных данных, максимально использующим возможности каротажного кабеля, можно считать код Манчестер-П.

ЛИТЕРАТУРА

22. Сиберт У.М. Цепи, сигналы, системы. – М.: Мир, 1988. – 336 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davydro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 15: ПОВЫШЕНИЕ СКОРОСТИ ПЕРЕДАЧИ СИГНАЛОВ ПО КАБЕЛЮ

Гораздо труднее увидеть проблему, чем найти ее решение. Для первого требуется воображение, для второго - только умение.

Джон Бернал. Английский физик, XX в.

Ученые склонны преувеличивать ценность своего воображения. Между тем, в основе постановки любой проблемы лежит благоприятный прогноз "погоды" на ее решение. При этом, с учетом принципа ветеринара – лекарство не должно быть дороже пациента, которому оно назначается.

Александр Рукавичников. Уральский геофизик, XX в.

Содержание: Введение. 15.1. Фильтр восстановления формы сигналов. Общая методика синтеза фильтров. Точность воспроизводства формы импульсов. Усиление шумов. 15.2. Фильтр восстановления кодовых сигналов. Оператор фильтра. Предельная частота передачи данных. Кодовые сигналы. 15.3. Принципы реализации фильтров частичной деконволюции. Литература.

ВВЕДЕНИЕ

Принципы частотной коррекции сигналов при их искажениях в линиях связи известны [23]. Отметим сразу, что говорить о метрологически точном восстановлении сигналов с определенной погрешностью, задаваемой метрикой выходного и входного сигналов, имеет смысл только в том случае, если эффективная ширина спектра сигналов много меньше эффективной ширины спектра импульсного отклика кабеля и затухает достаточно быстро. Если эти условия выполняются, то полная частотная коррекция передачи сигналов может быть выполнена инверсией импульсного отклика кабеля (без учета временной задержки сигналов). Такая инверсия, в принципе, возможна, но практического интереса не представляет. Это определяется тем, что спектральная плотность импульсного отклика жил кабелей меньше 1 по всему частотному диапазону с затуханием, близким к монотонному, также по всему частотному диапазону, начиная с частот 1-10 кГц. Фильтры деконволюции таких импульсных откликов имеют значительные (до нескольких тысяч) коэффициенты усиления дисперсии шумов, а потому рассматривать этот вариант не имеет смысла. Соответственно, для широкополосных сигналов, передаваемых по кабелю, возможна только частичная деконволюция импульсного отклика жил, т.е. применение фильтров сжатия импульсных откликов жил до определенной формы. Оптимальной с позиции минимального значения коэффициента усиления дисперсии шумов в этом случае считается симметричная гауссовская форма выходных импульсов фильтров частичной деконволюции (ЧД) импульсных откликов.

Методику расчета фильтров ЧД будем рассматривать в общей форме, возможности технической реализации в реальном масштабе времени в аналоговой форме (трансверсальные фильтры на линиях задержки) и в чисто цифровой форме с использованием амплитудно-цифровых преобразователей (АЦП) при дискретизации сигналов на выходе кабеля и микропроцессоров обработки цифровых данных в реальном масштабе времени.

15.1. ФИЛЬТР ВОССТАНОВЛЕНИЯ ФОРМЫ СИГНАЛОВ

Коррекция формы выходных сигналов обычно применяется только для кабелей небольшой длины при прямой передаче сигналов геофизических детекторов на измерительные приборы. С учетом этого, все нижеследующие расчеты выполняются в качестве примера для кабеля КГ 1x0.75-55-150 длиной 2 км.

Общая методика синтеза фильтров ЧД (которая в следующем разделе будет использоваться и для расчетов фильтров ЧД кодовых импульсных сигналов) включает следующие операции:

1. Задание длины кабеля, определение его импульсного отклика и сдвиг отклика (по началу значимых значений фронта) в начало координат. Сдвинутую функцию будем считать амплитудным откликом $h(t)$ кабеля (без учета задержки сигнала) и выполняем преобразование Фурье $h(t) \Rightarrow H(\omega)$.

2. Задание формы выходного импульса $z(t)$ фильтра ЧД в виде гауссовской функции (или любой другой с ограниченным и быстро затухающим спектром) и определение его

спектра $z(t) \Leftrightarrow Z(\omega)$. Временное расположение импульса $z(t)$ должно быть таким, чтобы площадь импульса практически полностью располагалась за пределами фронта импульсного отклика кабеля $h(t)$. Максимум $z(t)$ должен располагаться за максимумом отклика кабеля. В принципе, он может задаваться на максимуме и до максимума импульсного отклика, но это ухудшает параметры оператора ЧД. Площадь $z(t)$ при метрологической коррекции сигналов должна быть равна площади импульсного отклика. При восстановлении кодовых сигналов этот параметр не нормируется.

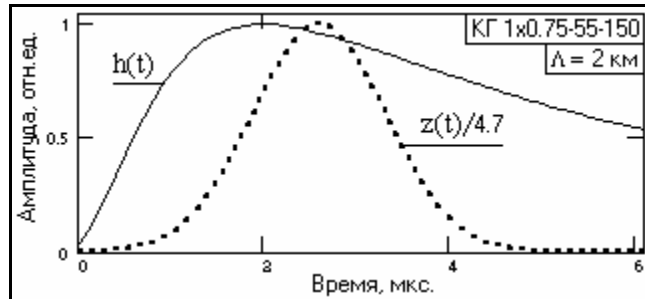


Рис. 15.1.1. Форма импульсов.

На рис. 15.1.1 приведены примеры формы сигналов, нормированные по амплитуде к 1 для наглядности сравнения. Ширина гауссовского импульса (значение стандарта – среднеквадратического отклонения от центра импульса) подбирается по допустимому коэффициенту усиления дисперсии шумов (после расчета оператора ЧД). Для приведенных далее расчетов площади импульсов установлены равными, ширина импульса $z(t)$ установлена такой, чтобы коэффициент усиления дисперсии шумов был не более 4 (амплитудное усиление шумов не более 2), при этом амплитудное усиление сигналов порядка 4.7, что позволяет повысить отношение сигнал/шум не менее чем в 2 раза.

3. Вычисление спектральной плотности передаточной функции фильтра ЧД, которое выполняется по формуле:

$$H_d(\omega) = Z(\omega)/H(\omega). \quad (15.1.1)$$

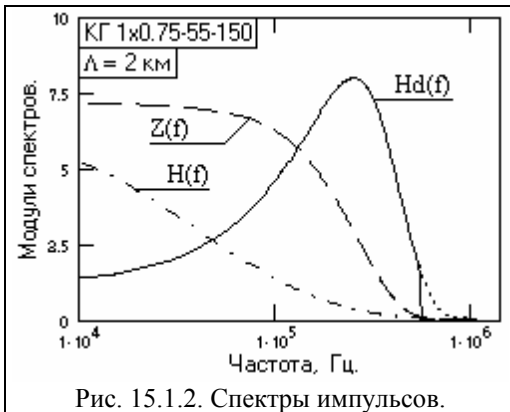


Рис. 15.1.2. Спектры импульсов.

При равных площадях импульсного отклика кабеля и функции $z(t)$ значения спектров $Z(\omega)$ и $H(\omega)$ при $\omega=0$ равны и, соответственно, их отношение равно 1, т.е. коэффициент усиления фильтром ЧД постоянной составляющей равен 1. Для фильтров ЧД кодовых импульсов этот параметр не регламентируется. Более того, при коэффициенте усиления постоянной составляющей менее 1 фильтр ЧД в определенной мере стабилизирует нулевую линию сигналов.

Функция $H_d(\omega)$ имеет смысл в области значимых значений $H(\omega)$ и за пределами этой области обнуляется, что исключит усиление высокочастотных шумов кабеля. Усечение, начиная с определенной частоты ω_c , целесообразно выполнять весовой гауссовской функцией:

$$H_d(\omega) = H_d(\omega) \cdot \exp(-a(\omega_c + \omega)^2), \quad \omega \geq \omega_c. \quad (15.1.2)$$

Пример формы спектральной плотности фильтра ЧД приведен на рис. 15.1.2 в сопоставлении со спектрами исходных импульсов (масштаб спектров импульсов для наглядности увеличен). Пунктиром на рисунке показана усеченная часть функции $H_d(\omega)$.

Коэффициент усиления фильтром дисперсии входных статистических шумов при равномерном распределении шумов в диапазоне от 0 до Ω :

$$K_\varepsilon = (1/\Omega) \int_{-\Omega}^{\Omega} |H_d(\omega)|^2 d\omega. \quad (15.1.3)$$

По рисунку 15.1.2 нетрудно сделать заключение, что коэффициент усиления фильт-

ром ЧД дисперсии статистических шумов больше 1 и существенно зависит от эффективной ширины спектра импульса $z(t)$. Чем меньше значение $\Delta\Omega$ импульса $z(t)$ (больше эффективная ширина ΔT импульса), тем меньше коэффициент усиления дисперсии шумов.

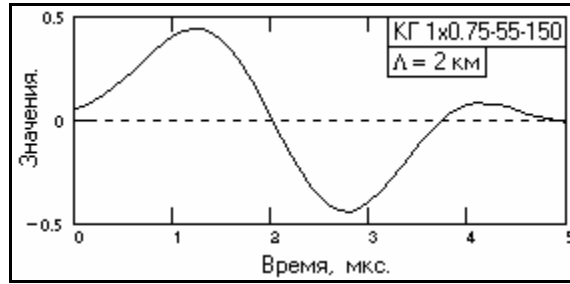


Рис. 15.1.3. Оператор ЧД.

4. Обратным преобразованием Фурье функции $H_d(\omega)$ вычисляется оператор фильтра ЧД: $H_d(\omega) \Rightarrow h_d(t)$. Пример оператора, нормированного на интервал дискретизации данных $\Delta t = 0.1$ мкс для цифровой обработки сигналов, приведен на рис. 15.1.3 и получен по исходной функции $H_d(\omega)$ рис. 15.1.2. Коэффициент усиления дисперсии шумов данного оператора равен 3,4. При заданном на рис. 15.1.1 коэффициенте усиления амплитуды импульса $z(t)$ порядка 4,7 это дает улучшение отношения сигнал/шум в $4.7/\sqrt{3.4} \approx 2.5$ раза при равномерном распределении шумов в интервале от 0 до $\Omega = 1/(2\Delta t)$.

При подаче на вход кабеля импульса Кронекера δ_0 на выходе фильтра будем иметь:

$$s(t) = \delta_0 * h(t) * h_d(t) \Leftrightarrow 1 \cdot H(\omega) \cdot H_d(\omega) = Z(\omega) \Leftrightarrow z(t),$$

т.е. выходной единичный импульс вместо асимметричной формы импульсного отклика кабеля будет иметь симметричную форму гауссовского импульса. Аналогичным является и результат непосредственной свертки импульсного отклика с оператором фильтра ЧД, что может использоваться для контроля расчетов операторов фильтров ЧД.

Точность воспроизводства формы заданной функции выходного импульса определяется размером оператора ЧД, который имеет существенное значение при его технической реализации в виде трансверсальных фильтров на линиях задержки. Основная часть энергии оператора (более 99%) сосредоточена в пределах интервала, примерно равного (2-3) значениям фронта импульсного отклика, т.е. начальная часть импульсного отклика содержит практически всю информацию о форме сигналов на входе кабеля. Однако, вследствие достаточно длинного спада отклика с постепенным уменьшением скорости спада, ограничение размера оператора ЧД короткой значимой частью приводит к появлению на его выходе послеимпульсных «выбросов» с амплитудным значением до 2% от максимальной амплитуды выходного сигнала (чем меньше размер оператора ЧД, тем больше амплитуда выбросов) и с длительностью, равной соответствующей длительности значимой части импульсного отклика (с постепенным затуханием). Положение послеимпульсного "хвоста" (включая первый отрицательный выброс) относительно нулевой линии может смещаться изменением конечного значения оператора ЧД.

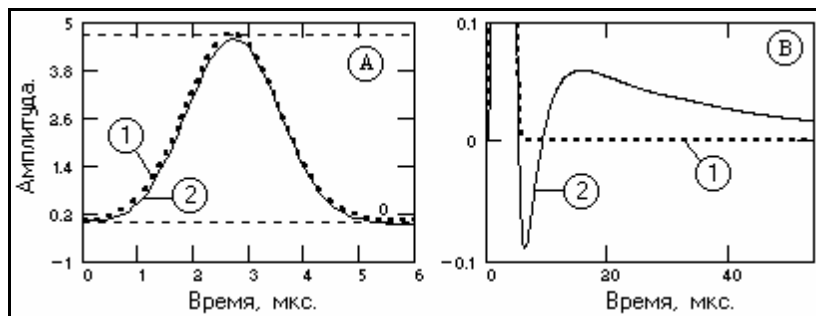


Рис. 15.1.4. Свертка импульсного отклика с оператором ЧД.

На рис. 15.1.4 приведены результаты свертки (кривая 2) импульсного отклика кабеля с

оператором ЧД размером 4.2 мкс при временной длине значимой части импульсного отклика кабеля (более 0.01 максимального значения) порядка 70 мкс. Кривая 1 – заданная функция $z(t)$. Как видно на графике А, форма заданного импульса сжатия реализуется достаточно точно. Среднее квадратическое расхождение с заданной формой $z(t)$ – 0.024, отрицательный выброс не более 2% от амплитуды при центрировании "хвоста" относительно нулевой линии (график В).

На рис. 15.1.5 приведено сопоставление формы импульсов на входе кабеля, заданных на интервалах $2\Delta T_k$, на выходе кабеля (на входе фильтра ЧД) и на выходе фильтра ЧД. Для фильтрации использован оператор ЧД, приведенный на рис. 15.1.3, размером 4.2 мкс. Амплитудные значения сигналов на входе кабеля умножены на площадь импульсного отклика (учет безвозвратных электромагнитных потерь), и сдвинуты во времени для сопоставления формы к сигналам на выходе фильтра ЧД.

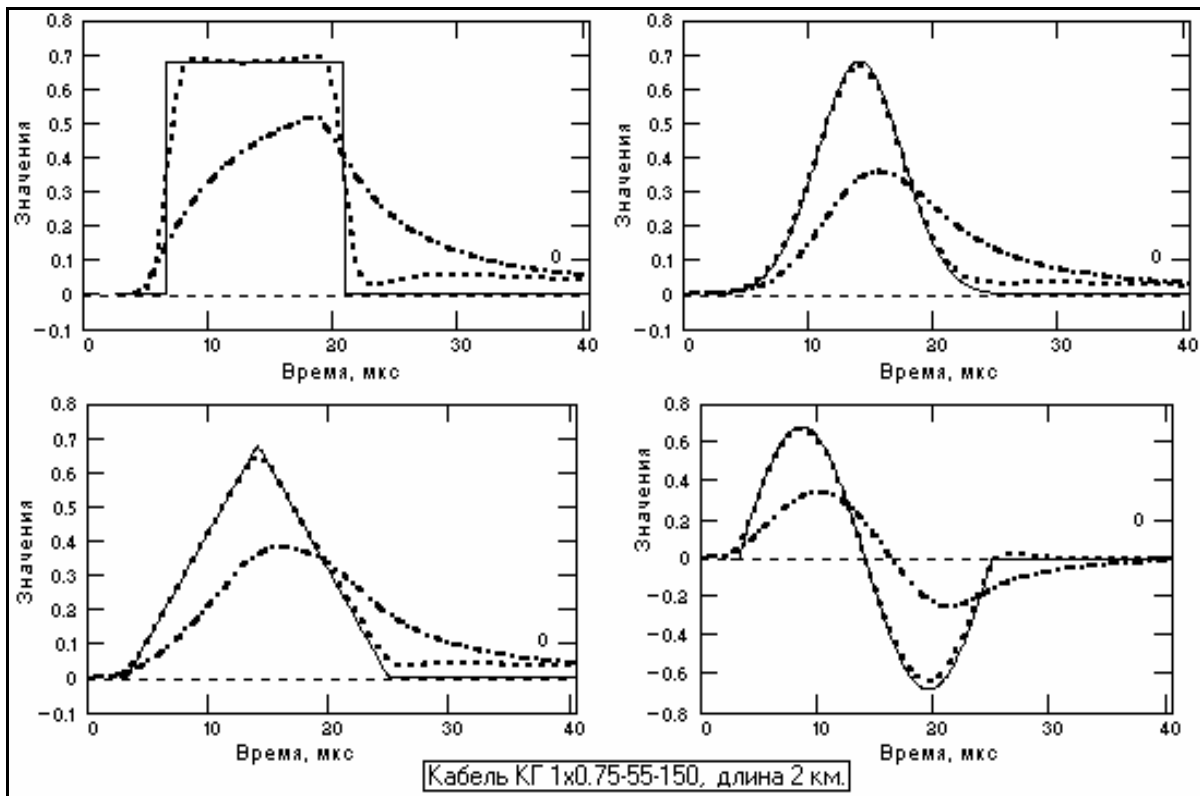


Рис. 15.1.5. Формы сигналов на входе (сплошные линии), на выходе кабеля (штрих-пунктир) и на выходе фильтра ЧД (пунктир). Входные сигналы смещены во времени.

Как видно на графиках, форма сигналов со значением $\Delta\Omega$, соизмеримом и меньшим значения $\Delta\Omega_k$, восстанавливается достаточно хорошо (со среднеквадратической погрешностью в пределах сигнала не более 1% относительно амплитудного значения), но для однополярных сигналов значение накопленного послеимпульсного «хвоста» возрастает до 5% от амплитуды сигнала. Если «хвост» послеимпульса должен быть ограничен по амплитуде определенной величиной, то это выполняется увеличением размера оператора фильтра ЧД. При длине оператора, равной длине импульсного отклика кабеля, амплитуда послеимпульса уменьшается более чем на порядок. Для прямоугольных импульсов при $\Delta\Omega > \Delta\Omega_k$ погрешность, естественно, возрастает за счет сглаживания фронтов, но остается в пределах не более (2-3)% по амплитудным значениям.

Усиление шумов. На рис. 15.1.6 приведены графики модельного статистического шума на входе (сплошная линия) и на выходе (пунктир) фильтра ЧД с параметрами, описанными выше. По результатам моделирования коэффициент усиления дисперсии шумов полностью соответствует расчетному значению (расчетное – 3.4, модельное на односекундном интервале – 3.38). Фильтр ЧД сглаживает шумы в соответствии со своей частотной характери-

стикой, и форма выходных шумов становится похожей на форму произвольных сигналов.

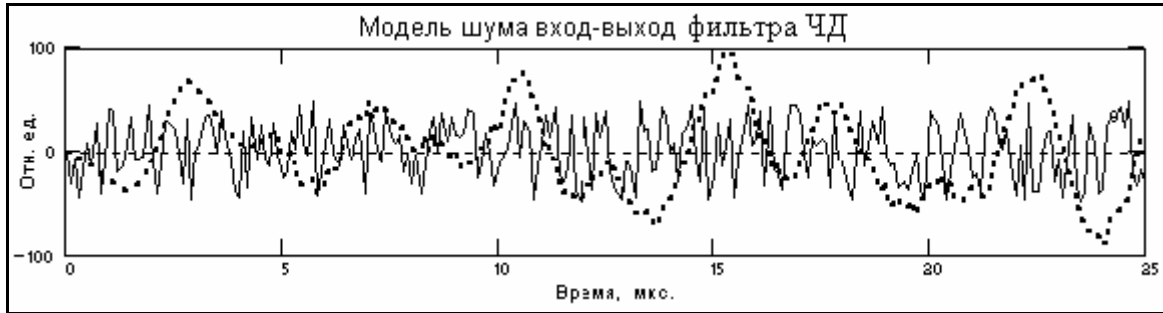


Рис. 15.1.6. Форма шумовых сигналов на входе и на выходе фильтра ЧД (пунктир).

Усиление шумов может существенно усложнить метрологическое восстановление произвольных сигналов. На рис. 15.1.7 приведен пример фильтрации зашумленного сигнала при среднеквадратическом уровне шума порядка 10% от амплитудных значений входного сигнала. Заметим, что хотя частотный диапазон спектра статистических шумов на выходе кабеля значительно превышает частотный диапазон сигналов, любые линейные методы фильтрации сигналов (подавления шумов) в данном случае не имеют смысла. Оператор ЧД, как это следует из (15.1.2) и можно видеть на рис. 15.1.2, самостоятельно подавляет все высокочастотные шумы.

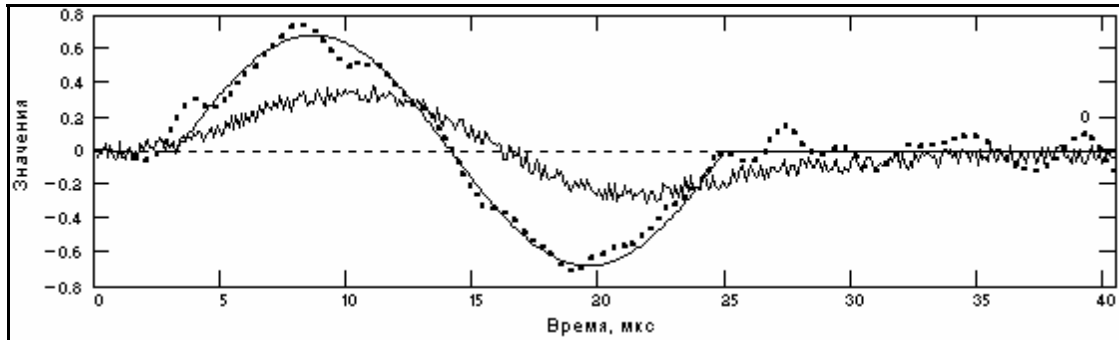


Рис. 15.1.7. Форма сигнала с шумом на входе и на выходе фильтра ЧД (пунктир).

При использовании цифровых методов обработки данных плавная форма выходных информационных сигналов с априорно известными параметрами их динамики позволяет использовать до подачи сигналов на фильтр ЧД нелинейные методы фильтрации, и в частности – адаптивные фильтры, не изменяющие динамики информационных сигналов при подавлении шумов, в том числе перекрывающихся по частотному диапазону с частотным диапазоном сигналов.

15.2. ФИЛЬТР ВОССТАНОВЛЕНИЯ КОДОВЫХ СИГНАЛОВ

Методика синтеза фильтров ЧД на кодовые сигналы практически полностью повторяет методику расчетов при восстановлении формы сигналов. Все нижеследующие расчеты выполняются в качестве примера для кабеля КГ 3х0.75-60-150 длиной 5 км для цифрового фильтра ЧД при интервале дискретизации данных 0.1 мкс. Импульсные параметры кабеля: $\Delta T_k \approx 26$ мкс, $\Delta \Omega_k \approx 34$ кГц.

Оператор фильтра. При приеме кодовых сигналов главное значение имеет простота и надежность выделения кода, которая определяется надежностью выделения единичных битовых импульсов. Для кода Манчестер-II она определяется интервалом минимум-максимум в непрерывной последовательности единиц (нулей). Для кодов NRZ и RZ имеет значение и сдвиг нулевой линии. Форма выходных импульсов особого значения не имеет, если не нарушается пространственное распределение импульсов. Естественно, что на предельной частоте передачи сигналов при достаточно высокой степени их амплитудного затухания существенное значение имеет увеличение (или, по крайней мере, сохранение на прежнем уровне) выходного отношения сигнал/шум фильтром ЧД.

При использовании фильтра ЧД можно считать, что предельная пропускная способность кабеля должна быть оценена с ограничением параметров оператора ЧД по определенному допустимому уровню коэффициента усиления дисперсии шумов. Примем этот уровень для дальнейших расчетов равным 1, т.е. оператор ЧД не должен усиливать среднеквадратичный уровень шумов. Это обеспечивается заданием соответствующей длительности импульса $z(t)$ на половине высоты гауссовского пика (в данном случае порядка 7-8 мкс).

Пример оператора, и результат его проверки сверткой с импульсным откликом жилы кабеля приведен на рис. 15.2.1. Длительность оператора 20 мкс, коэффициент усиления дисперсии шумов ~ 0.65 , коэффициент усиления амплитуды импульсного отклика жилы ~ 3.8 .

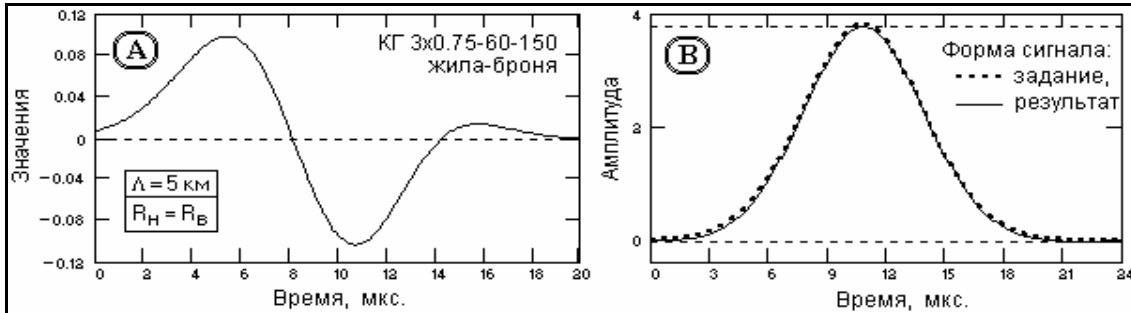


Рис. 15.2.1. Оператор фильтра ЧД кодов.

Предельная частота передачи данных. В качестве входных импульсов кабеля имеет смысл рассмотреть только прямоугольные импульсы, фронты которых на выходе фильтра ЧД будут соответственно сглажены до величины порядка ширины импульса $z(t)$. Эффект применения фильтра ЧД можно видеть на рис. 15.2.2 на примерах кодовой последовательности биполярных импульсов с единичной амплитудой на тактовых интервалах $T = 2 \cdot \Delta T_k$, $T = \Delta T_k$ и $T = 0.5 \cdot \Delta T_k$ мкс (тактовые частоты 19, 38 и 76 кГц). Пунктиром на рисунке показаны входные импульсы, приведенные к выходу кабеля с учетом безвозвратных потерь и задержки в кабеле.

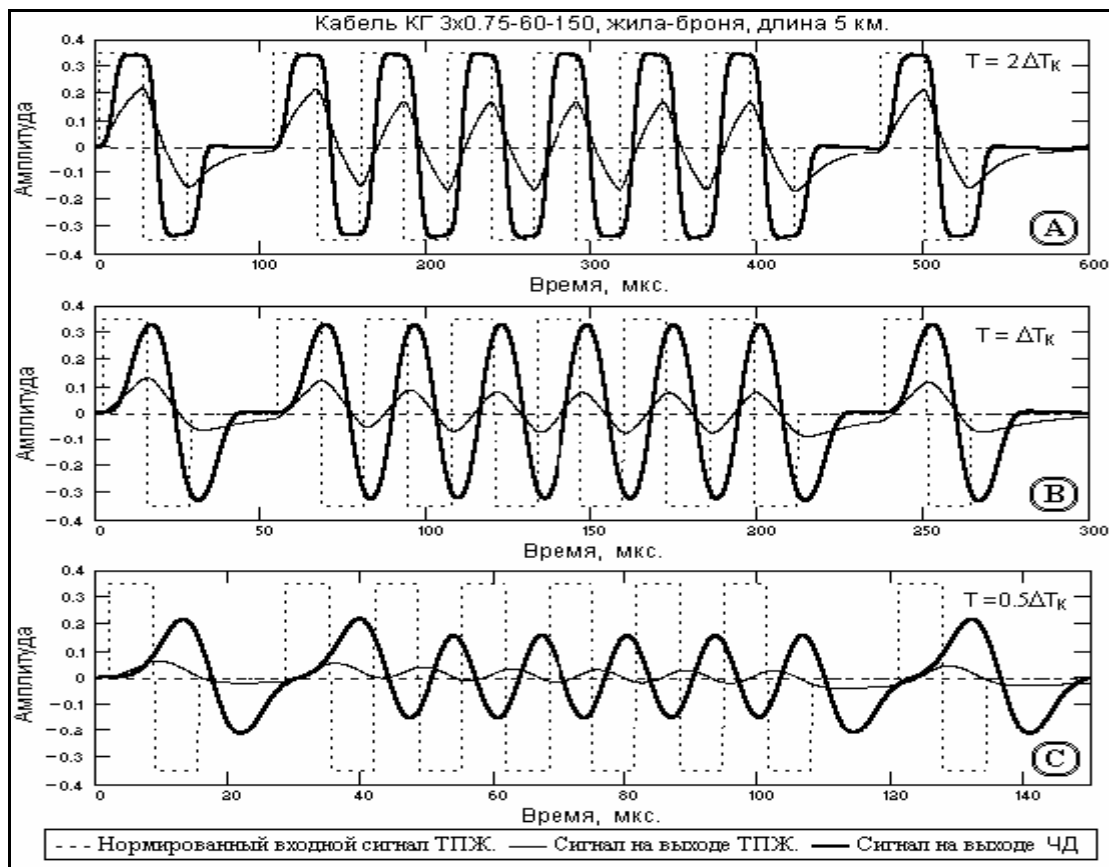


Рис. 15.2.2. Биполярные импульсы на входе и на выходе фильтра ЧД

Как следует из этого рисунка, применение фильтра ЧД позволяет уменьшить предельный тактовый интервал следования битовых импульсов при идентификации кода минимум в 2 раза, т.е. в 2 раза увеличить скорость передачи информации по кабелю.

Заметим, что при коэффициенте усиления дисперсии шумов данного оператора ЧД не более 1 отношение сигнал/шум на выходе фильтра ЧД при статистических шумах на входе улучшается практически в 2 раза, т.к. оператор ЧД в этом случае выполняет и роль низкочастотного сглаживающего фильтра. О последнем наглядно свидетельствует рис. 15.2.3(А).

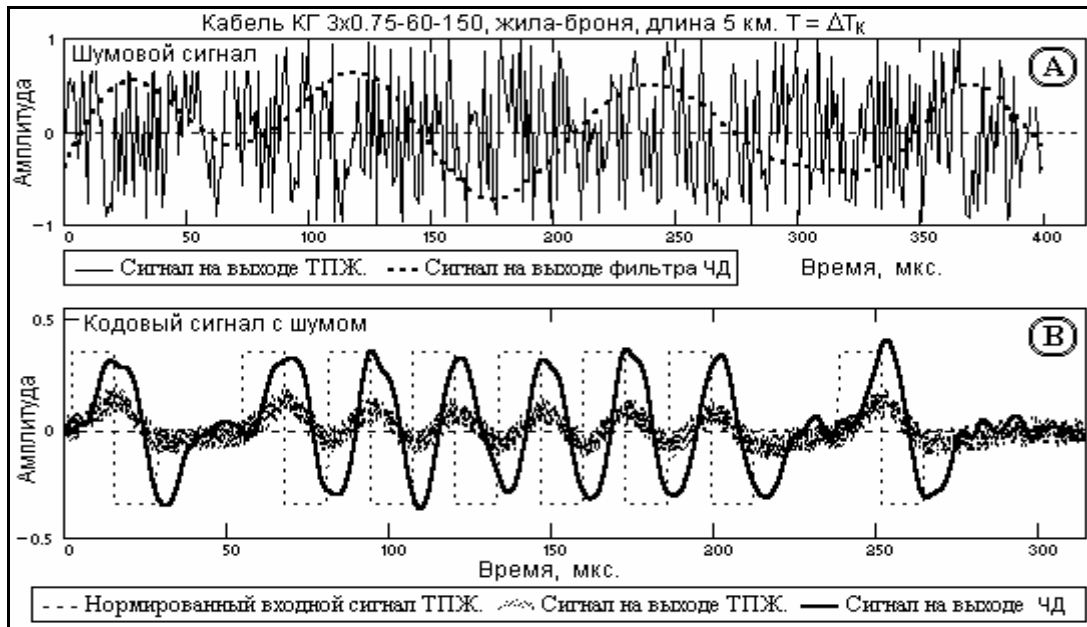


Рис. 15.2.3. Преобразование шумов и сигналов с шумом фильтром ЧД.

Пример фильтрации сигналов с наложенным статистическим шумом, средние квадратические флуктуации которого составляют порядка 30% от амплитуды информационных импульсов, приведен на рис. 15.2.3(В) и показывает уверенное сохранение информационной структуры сигнала.

Кодовые сигналы. На рис. 15.2.4(А) показана деконволюция кодовых сигналов с тактовой частотой 38 кГц, приведенных ранее на рис. 14.3.1 без деконволюции на предельной тактовой частоте 19 кГц. Как следует из рисунка, деконволюция позволяет практически полностью восстановить амплитудные значения сигналов (за вычетом безвозвратных потерь) на удвоенной тактовой частоте, т.е. импульсная пропускная способность кабеля повышается минимум в 2 раза для всех видов кодирования. Что касается предельной импульсной пропускной способности кабеля, то она повышается практически в 4 раза по сравнению с приемом сигналов без деконволюции, о чем достаточно наглядно свидетельствует пример формы тех же кодовых сигналов на тактовой частоте 76 кГц, приведенный на рис. 15.2.4(В), и сравнение этого рисунка с рисунком 14.3.1(В). Преимущество кода Манчестер-II перед кодами NRZ и RZ при использовании частичной деконволюции очевидно как на рис. 15.2.4(А), так и на рис. 15.2.4(В). С определенным запасом "прочности" можно считать, что при передаче информации биполярными кодами скорость передачи данных с использованием частичной деконволюции импульсного отклика кабеля может быть увеличена в 3 раза.

Краткие выводы по возможностям повышения импульсной пропускной способности коротких кабелей:

1. Качество приема и надежность идентификации кодовой информации на выходе кабеля могут быть существенно повышены при частичной деконволюции импульсного отклика кабеля до симметричной (гауссовской) формы.
2. Основная (значимая) часть энергии оператора частичной деконволюции импульсного отклика кабеля сосредоточена в пределах 2-3 значений фронта отклика.
3. Для финитных сигналов, задаваемых на интервале не менее $2\Delta T_k$ и имеющих эф-

фактивную ширину спектра не более $\Delta\Omega_k$, частичная деконволюция импульсного отклика жил кабеля обеспечивает восстановление формы сигналов с погрешностью по основным метрологическим параметрам (амплитуда, площадь, временная привязка) не более (1-2)%.

4. Скорость передачи кодовых данных при использовании частичной деконволюции импульсного отклика может быть увеличена минимум в 2 раза при любых методах кодирования и равна $1/\Delta T_k$. При передаче информации биполярными импульсами предельная скорость передачи данных может быть увеличена минимум в 3 раза.

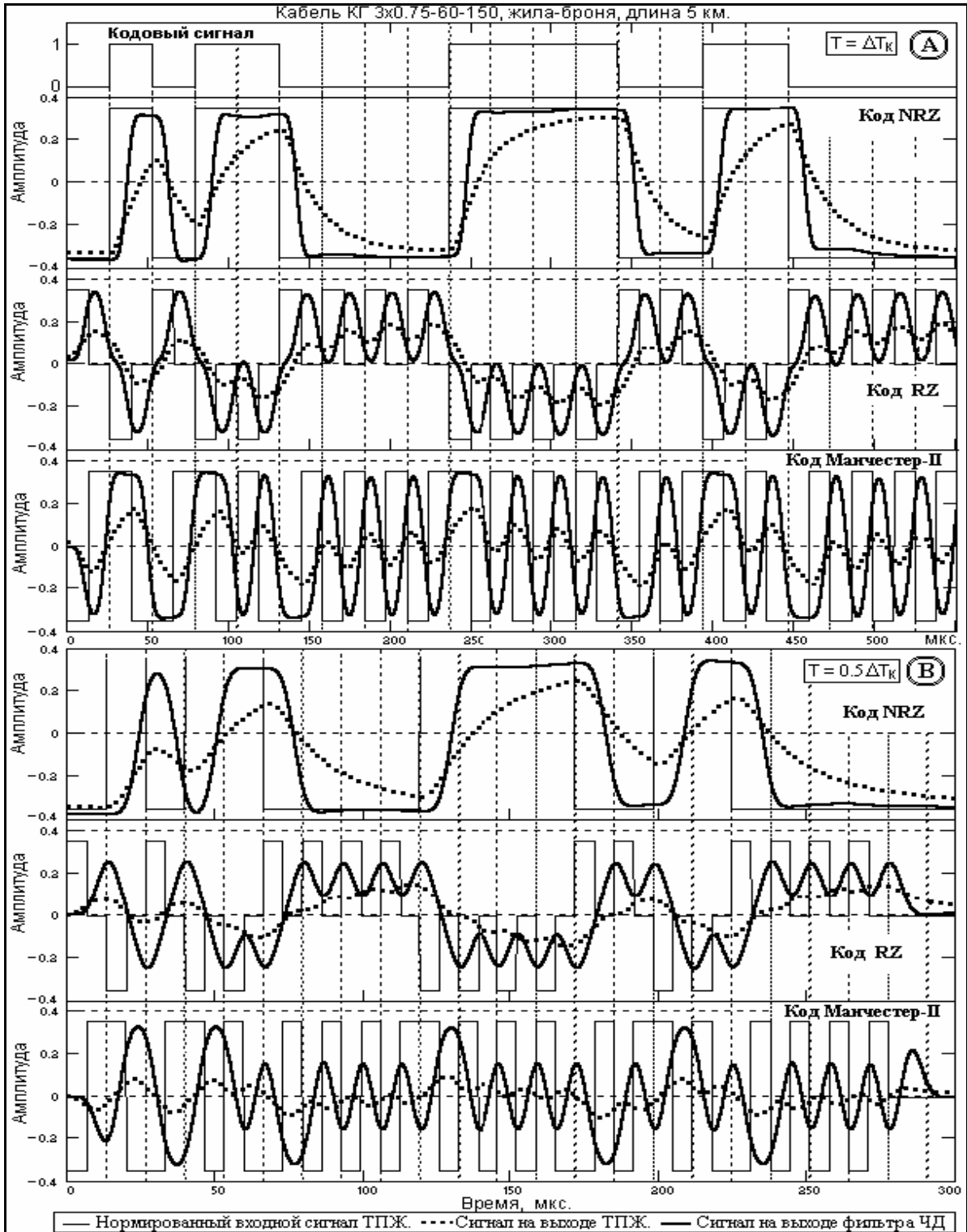


Рис. 15.2.4. Деконволюция кодовых выходных сигналов, тактовые частоты 38 и 76 кГц.

15.3. ПРИНЦИПЫ РЕАЛИЗАЦИИ ФИЛЬТРОВ ЧАСТИЧНОЙ ДЕКОНВОЛЮЦИИ

Основное условие технической реализации фильтров ЧД – работа в реальном мас-

штабе времени. Дополнительное и желательное условие – полная автономность от аппаратного блока каротажных станций, позволяющая включать фильтры ЧД на выход кабеля в состав любой станции, если в том появится необходимость.

Достаточно простая форма передаточной функции фильтра ЧД позволяет выполнить его реализацию непосредственно в виде аналогового нерекурсивного или рекурсивного фильтра. Следует отметить, что такие фильтры будут являться индивидуальными для кабелей различных типов и различной длины (в определенных пределах возможной подстройки) и могут отличаться по исполнению: нерекурсивные - для фильтров с короткой функцией отклика, и рекурсивные для кабелей длиной 3 и более км.

С учетом современных темпов развития электроники перспективным и универсальным направлением следует считать цифровые фильтры, т.е. микропроцессорное исполнение фильтров ЧД или программное выполнение фильтров в составе каротажных измерительно-вычислительных комплексов непосредственно на их входе. Заметим, что детерминированность кодовых сигналов позволяет выполнять микропроцессорные ЧД с автоматической адаптацией под конкретный кабель, стоящий на каротажной станции.

Своеобразной комбинацией дискретного синтеза с аналоговым процессом фильтрации являются трансверсальные фильтры на линиях задержки. Исходные условия для их реализации аналогичны цифровым фильтрам.

Возможность цифрового исполнения фильтров ЧД можно оценить по таблице 15.3.1. При определении частоты Найквиста цифрового фильтра ЧД по предельной частоте $f_{пр}$ передаточной функции кабеля (по уровню порядка 1% от коэффициента передачи на низких частотах) значение шага дискретизации данных на выходе кабеля находится в диапазоне 0.2-6 мкс в зависимости от длины кабеля. Этот диапазон достаточно хорошо согласуется с шагом дискретизации данных по предельной тактовой частоте передачи информации биполярными импульсами (частота Найквиста за третьей гармоникой тактовой частоты, что обеспечивает регистрацию более 98% энергии сигналов).

Таблица 15.3.1. Параметры передачи данных (кабель КГ 3х0.75-60-150, жила-броня).

Длина кабеля	км	1	2	3	4	5	6	7
Предельная частота $f_{пр}$ [по уровню 0.01 K_{max}]	кГц	2460	718	358	220	151	111	85
Шаг дискретизации на предельной частоте ($1/2f_{пр}$)	мкс	0.2	0.7	1.4	2.3	3.3	4.5	5.9
Эффективная ширина импульсного отклика ΔT_k	мкс	3.3	8.6	14.1	19.9	26.1	33.1	40.9
Тактовая частота передачи данных $f_T=1/\Delta T_k$	кГц	306	117	71	50.3	38.3	30.3	24.4
Шаг дискретизации по тактовой частоте ($1/4f_T$)	мкс	0.4	1.0	2.2	2.5	3.3	4.1	5.1
Размер оператора ЧД (>98% энергии оператора)	мкс	2.6	5.1	9	13.5	19	25	32
Оптимальный шаг дискретизации данных Δt	мкс	0.12	0.24	0.4	0.65	1	1.2	1.5

В принципе, эти значения шага дискретизации данных вполне достаточны для обработки сигналов, четкой автосинхронизации тактовой частоты приема данных и их кодовой идентификации, а при необходимости и полного восстановления аналоговой формы сигналов. Для исключения трансформации высокочастотных шумов в рабочий диапазон сигнала, фильтру ЧД должен предшествовать аналоговый низкочастотный фильтр.

Коэффициент усиления дисперсии статистических шумов оператором цифрового фильтра, равный сумме квадратов значений его коэффициентов, зависит от интервала дискретизации. В данном случае, при постоянной заданной форме импульса $z(t)$ сжатия импульсного отклика кабеля и увеличении значений Δt шага дискретизации (относительно $\Delta t \Rightarrow 0$ для аналоговой формы фильтра) количество коэффициентов оператора ЧД в пределах его импульсного отклика уменьшается, а их значения возрастают, что вызывает соответствующее возрастание коэффициента усиления дисперсии статистических помех. Для сохранения значения коэффициента усиления дисперсии помех на уровне, не большем 1, приходится увеличивать задаваемую ширину импульса $z(t)$, при этом уменьшается амплитуда им-

пульса $z(t)$ и качество деконволюции сигнала. Это можно наглядно видеть на рис. 15.3.2 и 15.3.3.

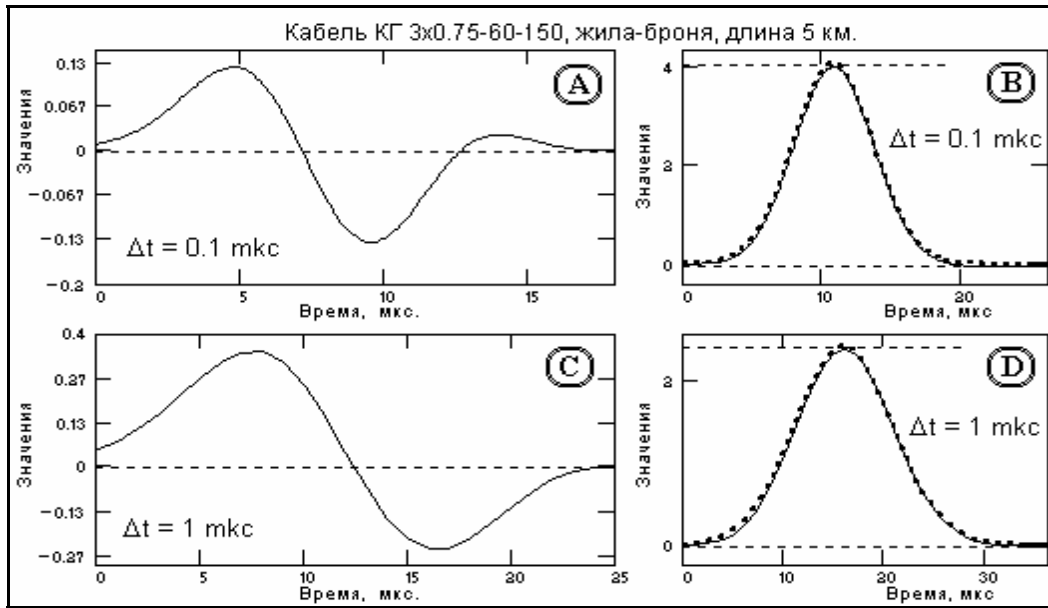


Рис. 15.3.2. Операторы ЧД (А, С) и форма сигналов сжатия импульсного отклика кабеля (В, D) при разном шаге дискретизации данных.

На рис. 15.3.2 (А, С) приведены два оператора ЧД, вычисленные с разным шагом дискретизации данных (0.1 и 1 мкс). Ширина гауссовских импульсов $z(t)$ (приведены пунктиром на рис. 15.3.2 (В, D)), была установлена такой, чтобы коэффициенты усиления дисперсии помех операторов ЧД были примерно равными в пределах 0.95-1.

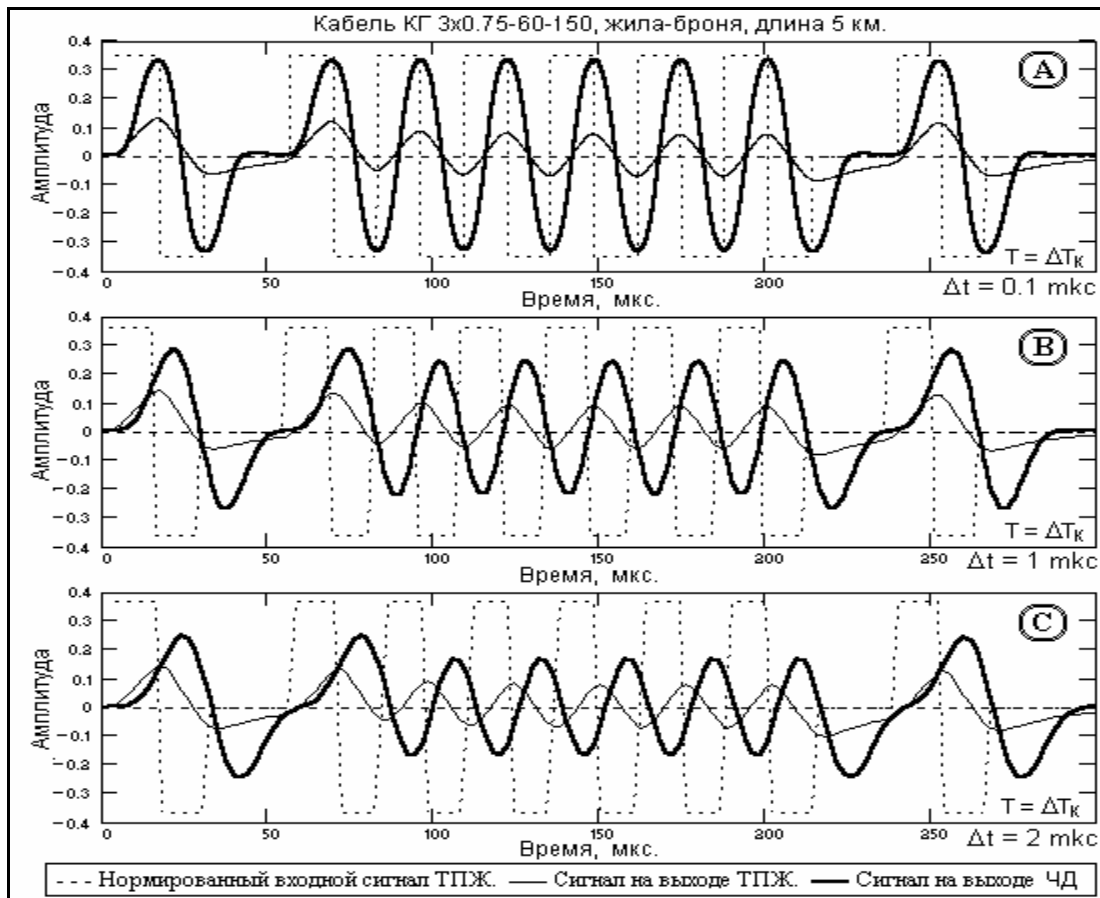


Рис. 15.3.3. Деконволюция выходных сигналов при разных интервалах дискретизации.

Как видно на рисунке, интервалу дискретизации 1 мкс соответствует в 1.5 раза боль-

шая ширина импульса $z(t)$, чем интервалу дискретизации 0.1 мкс, и, соответственно, меньшая амплитуда импульса. Качество работы операторов по реализации заданной формы импульса $z(t)$ при свертке с импульсным откликом кабеля практически одинаково (сплошные линии на рис. 15.3.2(B,D)), но при деконволюции импульсного кода оператор с большим шагом дискретизации данных соответственно имеет меньшую временную разрешающую способность и занижает амплитуды восстановленных импульсов, что можно видеть на рис. 15.3.3 (А, В). Дополнительно на рис. 15.3.3(С) приведен пример деконволюции сигнала с шагом дискретизации данных 2 мкс (оператор ЧД – 32 мкс, 16 точек, коэффициент усиления дисперсии шумов 0.98).

При сопоставлении графиков на рис. 15.3.3 можно сделать вывод, что оптимальный шаг дискретизации данных для цифрового фильтра ЧД соответствует 20-30 коэффициентам фильтра в пределах длительности значимой части оператора ЧД. Этот вывод подтверждают аналогичные вычисления и для кабелей других размеров. Соответственно, диапазон оптимальных значений шага дискретизации данных на входе фильтра ЧД, показанных в последней строке таблицы 15.3.1, установлен по длине операторов ЧД и составляет от 0.12 до 1.5 мкс. В принципе, такой тактовый диапазон при 20-30 операциях умножения и сложения вполне доступен для современных микропроцессорных систем, особенно для кабелей большой длины, для которых фильтр ЧД и необходим в максимальной степени.

В трансверсальных фильтрах сигнал с кабеля подается на последовательную цепочку линий задержки, в каждой из которых осуществляется задержка сигнала на интервал дискретизации данных. К выходам линий задержки подсоединена матрица резисторов, значения которых обратно пропорциональны значениям коэффициентов оператора ЧД. Токи через резисторы, пропорциональные положительным и отрицательным значениям коэффициентов оператора, суммируются раздельно (на входах двух операционных усилителей), после чего из "положительного" тока вычитается "отрицательный" и результат подается на вход аппаратуры станции, как выходной сигнал фильтра ЧД. Трансверсальный фильтр идеально приспособлен для исполнения в качестве автономного промежуточного блока между кабелем и станцией. При переменных сопротивлениях резисторной матрицы фильтр легко подстраивается под любой тип и любую жилу кабеля, а изменение длины кабеля с изменением интервала дискретизации данных выполняется заменой линии задержки. Некоторые технические трудности могут возникать только в наборе линий задержки для длинных кабелей (большое время задержки) с проявлением дополнительного затухания сигнала в самой ЛЗ, но последнее компенсируется соответствующим изменением коэффициентов резисторной матрицы.

Возможно и комбинированное цифро-аналоговое исполнение фильтра ЧД, в котором роль ЛЗ исполняет сдвиговый цифровой регистр с АЦП на входе (тактовая частота сдвига определяет шаг дискретизации входных данных), каждая цифровая ячейка которого имеет обратный резисторный ЦАП. Дальнейшая обработка токов ЦАП и формирование выходного сигнала аналогично трансверсальному фильтру. Для кодовых сигналов объем цифровых ячеек регистра может быть в пределах 5-7 двоичных разрядов. В таком исполнении фильтр ЧД становится автономным универсальным блоком с простой и гибкой настройкой под любой тип кабеля любой длины с изменением интервала дискретизации данных частотой тактового сдвига цифрового регистра.

Таким образом, каких-либо особых препятствий в технической реализации фильтров ЧД не имеется.

ЛИТЕРАТУРА

23. Стрижевский Н.З. Коаксиальные видеолнии. – М.: Радио и связь, 1988. – 200 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 16: СИНХРОНИЗАЦИЯ КОДА МАНЧЕСТЕР-II НА ВЫХОДЕ КАРОТАЖНОГО КАБЕЛЯ

Всякое предсказание зла только тогда доброе дело, когда сопровождается советом, как это зло отвести.

Марк Туллий Цицерон. Римский философ и политик, I в.д.н.э

Однако советы в технических вопросах могут приносить вред. Настоящее решение обычно всегда неожиданное и простое.

Валерий Самойлин. Геофизик и радиоинженер Уральской школы, XX в.

Содержание: Введение. 16.1. Исходные условия передачи данных. Код Манчестер-II. Спектры выходных сигналов. 16.2. Методика формирования выходных строб - импульсов. Амплитудное детектирование. Выделение несущих частот. 16.3. Декодирование выходных сигналов. Формирование последовательности строб – импульсов. Контроль фазовой синхронизации. Литература.

ВВЕДЕНИЕ.

Начиная с 80-90 годов прошлого века в технологиях геофизических исследований скважин (ГИС) наблюдается переход на комплексные многопараметровые измерения с применением многофункциональных скважинных приборов. Это требует передачи данных в наземные измерительно-вычислительные устройства, как правило, в цифровой форме и с высокой скоростью. В то же время линией передачи данных ГИС остается традиционный каротажный кабель со скоростью передачи информации (бит/с) не выше 10-100 кГц (в зависимости от длины), что сдерживает развитие и совершенствование технологий ГИС.

Каротажный кабель представляет собой линию связи с пониженной добротностью и нелинейной зависимостью передаточной функции от длины кабеля. Импульсную пропускную способность кабеля и качество передачи данных определяют эффективная ширина частотного спектра $\Delta\Omega_k$ и эффективная длительность импульсного отклика ΔT_k токопроводящих жил кабеля. Максимальная скорость передачи по кабелю кодовых импульсов без применения устройств частотной коррекции передаточной функции кабеля ограничивается тактовой частотой $f_T = 1/(2\Delta T_k)$ бит/с при эффективной длительности импульсов, не превышающей эффективной длительности импульсного отклика кабеля.

Скорость передачи информации зависит не только от тактовой частоты передачи данных, но и от протокола кодирования информации. Протокол передачи каротажных данных обычно заимствуется из стандартов открытых систем обмена информацией OSI (Open System Interconnect). С учетом эксплуатации скважинных приборов в условиях существенного влияния различных дестабилизирующих факторов в процессе каротажа, которые могут вызывать значительные вариации тактовой частоты передачи данных, для надежной передачи информации используются, в основном, коды с автосинхронизацией тактовых частот передатчика и приемника, среди которых наибольшее распространение получил код Манчестер-II.

16.1. ИСХОДНЫЕ УСЛОВИЯ ПЕРЕДАЧИ ДАННЫХ.

Код Манчестер-II является биполярным двухуровневым кодом. Логическому нулю соответствует переход на верхний уровень в центре тактового интервала с возвратом на нижний по концу тактового интервала, если следующий бит также нулевой. Соответственно, логической единице – переход на нижний уровень с возвратом на верхний по концу интервала, если следующий бит также 1. Бит обозначен переходом в центре тактового интервала, по которому и выделяется синхросигнал. Первая несущая частота кода соответствует чередованию нулей и единиц. Вторая несущая частота – последовательности нулей или единиц, и в 2 раза больше первой. При передаче произвольных последовательностей нулей и единиц более 50% энергии сигналов сосредоточено в области этих частот и между ними. Несомненное достоинство кода – отсутствие постоянной составляющей при передачах длинных последовательностей нулей или единиц.

Для большей конкретности и наглядности последующего текста все примеры и рисунки

будем приводить для трехжильного бронированного каротажного кабеля типа КГЗх0.75-60-150 длиной 5 км по линии передачи сигналов жила-броня. Электрическое сопротивление токопроводящих жил (ТПЖ) кабеля порядка 25 Ом/км, характеристическое (волновое) сопротивление 65 Ом, емкость ТПЖ не более 130 нФ/км, индуктивность на частоте 1 кГц порядка 2.5 мГн/км, коэффициент затухания на частоте 50 кГц не более 8 дБ/км. Эффективная ширина $\Delta\Omega_k$ спектральной передаточной функции ТПЖ порядка 35 кГц, эффективная длительность импульсного отклика ΔT_k порядка 25 мкс. Мера эффективной длительности - вторые моменты функций распределения отклика и спектра /22/. Сигналы будем моделировать со стороны скважинного прибора непосредственно на входе кабеля при условном выходном сопротивлении источника сигнала порядка 1 Ом. Входное сопротивление приемника сигнала на выходе кабеля примем равным его волновому сопротивлению. Все расчеты и моделирование выполняются в системе Mathcad в цифровой форме с шагом дискретизации временных данных 0.1 мкс и частотных спектров порядка 500 и 1000 Гц в зависимости от длины временных интервалов реализации кодовых последовательностей. Вычисления выполняются по методике, рассмотренной нами в предыдущих лекциях по каротажным кабелям. Амплитуда входных сигналов условно принята равной 1.

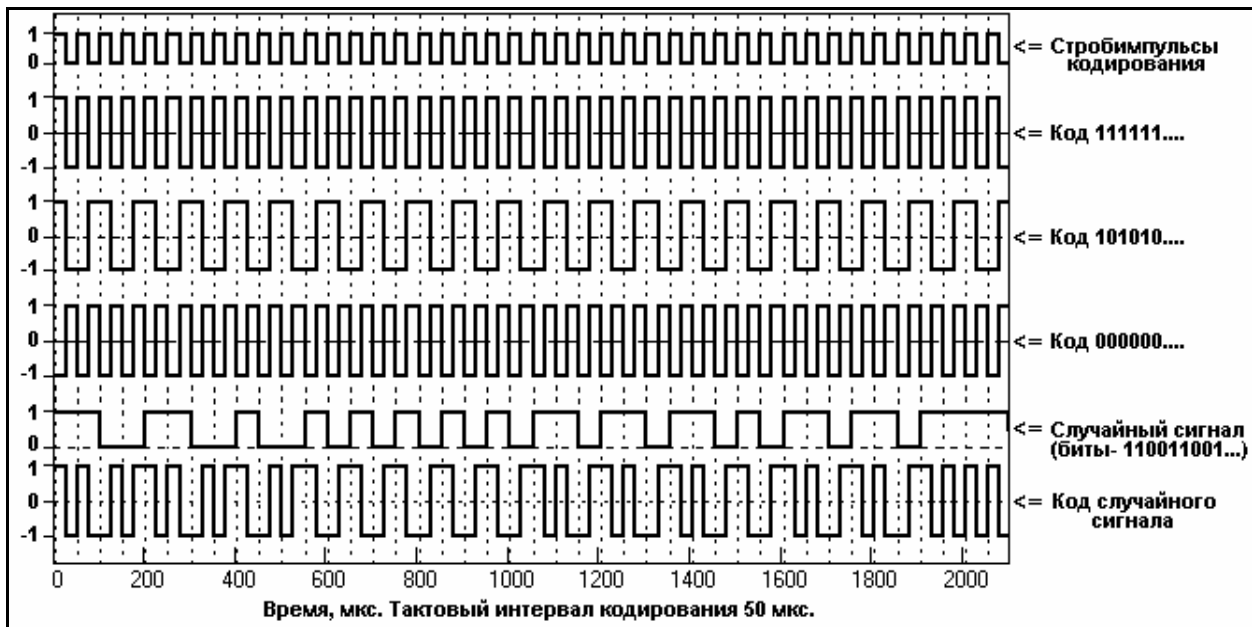


Рис. 16.1.1. Сигналы в кодовом протоколе Манчестер-II.

На рис. 16.1.1 приведен пример кодировки четырех вариантов битовых последовательностей (нулей, единиц, чередования единиц и нулей, и произвольного случайного сигнала) в протоколе Манчестер-II на входе кабеля. Основная несущая частота кодирования $f_T = 1/(2\Delta T_k)$ установлена равной 20 кГц (тактовый интервал кодирования 50 мкс).

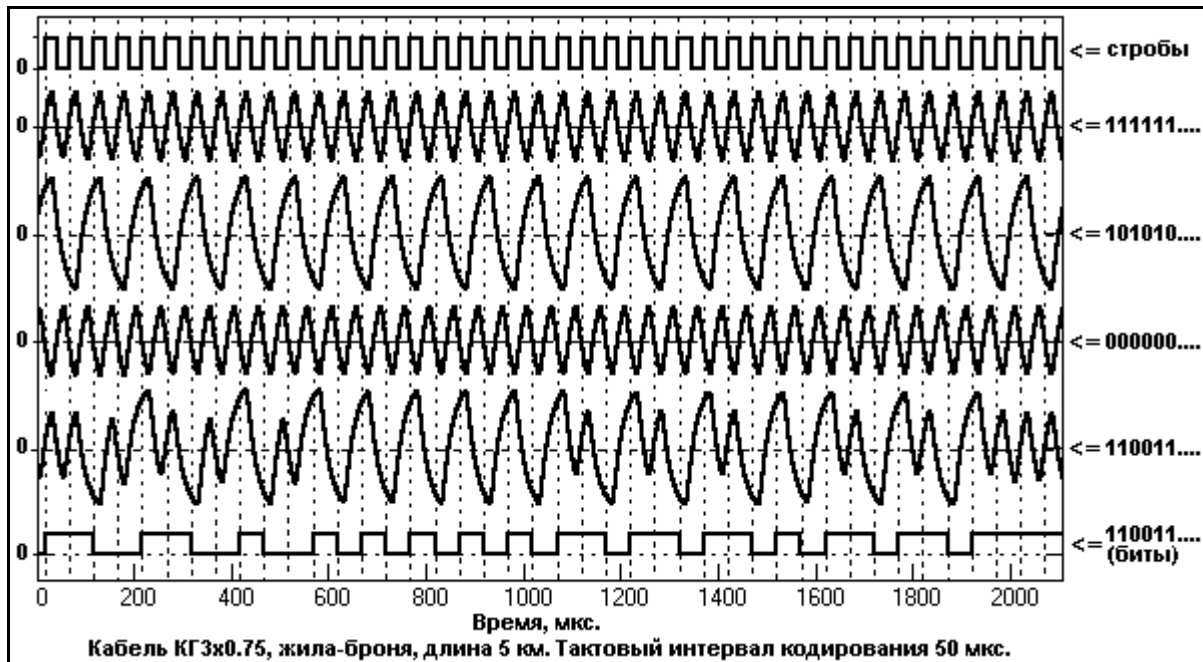


Рис. 16.1.2. Сигналы на выходе кабеля.

Для принятых первичных электрических параметрах кабеля коэффициент передачи с входа на выход амплитудных значений последовательностей единиц и нулей на данном кабеле порядка 0.15, чередования единиц и нулей - порядка 0.26. Форма сигналов на выходе кабеля для приведенного примера входных сигналов при отсутствии шумов на выходе кабеля приведена на рис. 16.1.2 (со сдвигом на временную задержку сигналов в кабеле).

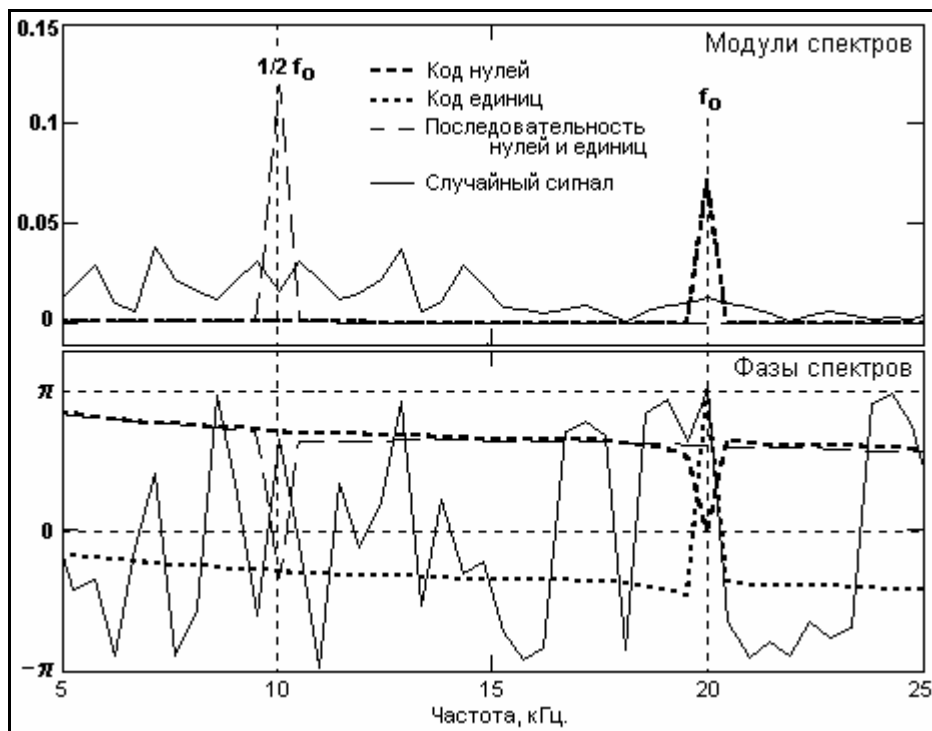


Рис. 16.1.3. Спектры выходных сигналов.

Спектры выходных сигналов приведены на рис. 16.1.3. Как следует из этого рисунка, основные несущие частоты сигнала хорошо выражены только для упорядоченных последовательностей. Основная энергия случайных импульсных последовательностей рассредотачивается по частотному диапазону от 0 до частоты порядка $1.5f_0$, т.е. практически по всей ширине эффективного частотного диапазона кабеля, что более наглядно видно на рис. 16.1.4, где приведены спектры для шести реализаций случайных импульсных последовательностей, зарегистрированных на 20-ти тактовых интервалах каждая.

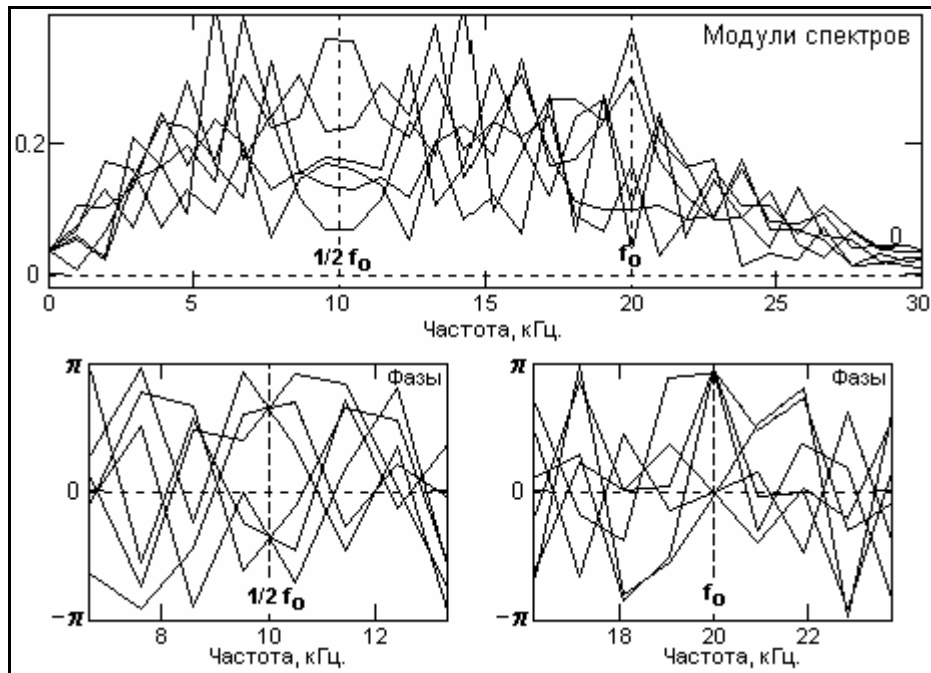


Рис. 16.1.4. Спектры случайных кодовых последовательностей.

16.2. МЕТОДИКА ФОРМИРОВАНИЯ СТРОБ-ИМПУЛЬСОВ.

Первой задачей приемника на выходе кабеля является формирование строб-импульсов, синхронных по частоте с тактовой частотой передачи кода передатчиком. Спектры случайных кодовых последовательностей, приведенные на рис. 16.1.4, достаточно наглядно показывают, что основные несущие частоты случайных кодов существенно варьируют по амплитуде и не имеют взаимной корреляции. Фазы гармоник на каждой из основных несущих частот имеют два постоянных значения, различающихся на 180° , что определяется противофазностью формирования кодов 0 и 1, и кодовых последовательностей 01 и 10, но какой-либо корреляции между этими четырьмя значениями фаз для случайных последовательностей также не имеется. Это означает, что устойчивую и достаточно надежную автосинхронизацию как по одной из основных частот, так и по их комбинации, выполнить невозможно. Практическое применение нашли фазовые методы синхронизации непосредственно по форме сигналов на выходе кабеля (например, по пересечениям нулевой линии) с учетом логики формирования кодовых сигналов.

Амплитудное детектирование. Различие фаз колебаний на каждой из основных частот кодовых последовательностей на 180° при нулевой постоянной составляющей позволяет простым преобразованием выходного сигнала, а именно – двухполупериодным амплитудным детектированием, получить сигнал амплитудных значений выходного сигнала кабеля, который имеет удвоенные значения основных частот с однозначным значением фаз любых комбинаций кодовых последовательностей на этих частотах. Пример амплитудного детектирования выходного сигнала для отрезка случайной кодовой последовательности приведен на рис. 16.2.1.

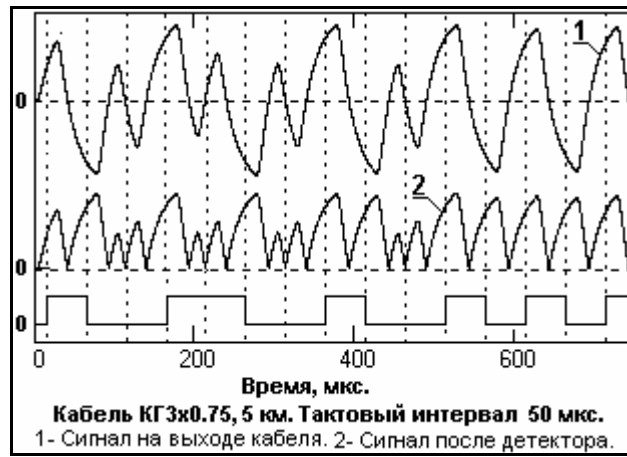


Рис. 16.2.1. Амплитудное детектирование

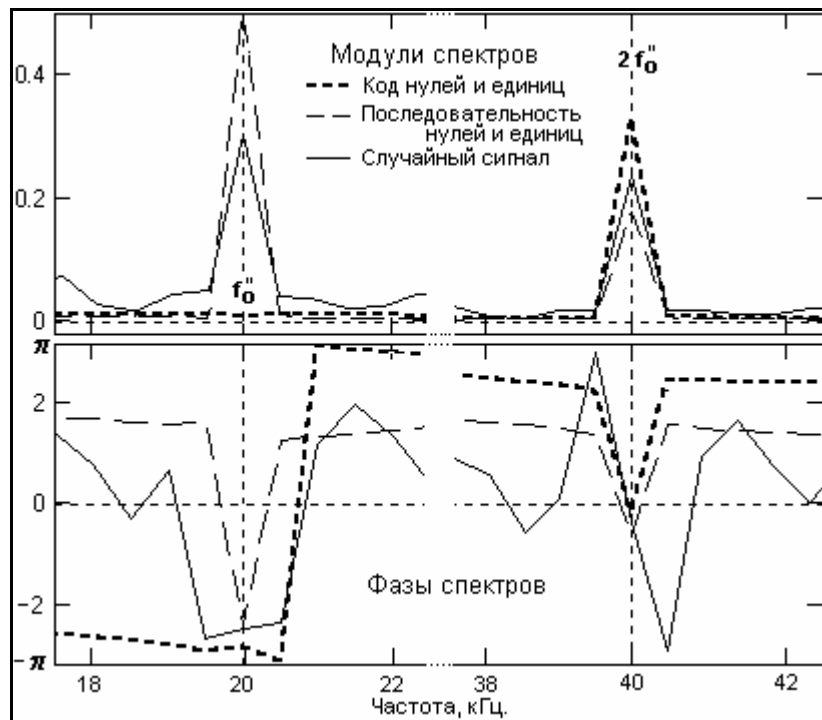


Рис. 16.2.2. Спектры сигналов после амплитудного детектора.

На рис. 16.2.2 приведены спектры сигналов после амплитудного детектирования. Как следует из этого рисунка, несущая гармоника последовательности чередования нулей и единиц и соответствующие гармоники кодов 01 и 10 в последовательностях случайных кодовых сигналов с частоты $0.5 \cdot f_0$ перемещаются на частоту $f'_0 = f_0$ практически с одинаковыми значениями фазы колебаний. Соответственно, несущие частоты f_0 последовательностей только нулей и единиц, и непрерывных отрезков кодов нулей и единиц любой длины (начиная с минимальных отрезков 00 и 11) в случайных кодовых последовательностях с практически одинаковой фазой переходят на частоту $2f'_0 = 2f_0$. Заметим также, что на частоте $2f'_0$ с той же фазой появляются и вторые гармоники несущей частоты f'_0 последовательностей чередования нулей и единиц, амплитуда которых достигает значений $1/3$ частоты f'_0 . так и вторые гармоники первой несущей частоты при чередования нулей и единиц. Эти два фактора приводят к тому, что для случайных кодовых последовательностей амплитудные значения колебаний на новых несущих частотах f'_0 и $2f'_0$ практически соизмеримы.

Таким образом, после двухполупериодного амплитудного детектирования сигналов кода Манчестер-II на выходе кабеля новые несущие частоты продетектированных сигналов, удвоенные по отношению к старым частотам, являются однофазовыми для любых кодовых последовательностей. Это позволяет применить частотные методы построения систем автосинхронизации передатчиков и приемников. Это может быть выполнено несколькими спосо-

бами.

Способ 1. По частоте f''_0 .



Рис. 16.2.3. Спектры кодов после амплитудного детектора.

На рис. 16.2.3 приведены спектры в диапазоне частоты f''_0 шести различных реализаций случайных последовательностей кодов на двадцати тактовых интервалах кодирования. Максимальные амплитуды гармоник соответствуют реализациям с большими весовыми долями кодов 01 и 10 в реализациях, минимальные (нулевые в пределе) - при полном отсутствии таких кодов в реализациях. В принципе, при непрерывной передаче информации последнее не имеет большого значения, если временная постоянная автосинхронизации задается больше длительности возможных пауз в передаче данных. При приеме данных в присутствии статистических шумов роль сигнала поддержки системы автостабилизации в рабочем состоянии в период пауз, как это будет показано ниже, могут выполнять непосредственно шумовые сигналы.



Рис. 16.2.4. Сигнал на выходе СФ на частоте f''_0 .

Выделение несущих частот. Фазовое постоянство частоты f''_0 любых кодовых последовательностей позволяет выделить колебания с частотой f''_0 узкополосным селекторным фильтром (СФ) и получить синхронизирующий сигнал (частоту синхронизации). Пример выделения частоты синхронизации селекторным фильтром приведен на рис. 16.2.4. Методы автоматической настройки селекторных фильтров на несущую частоту f''_0 , а равно и методы автоматического слежения за несущей частотой с определенной постоянной времени ее сохранения на интервалах отсутствия несущей частоты во входном сигнале (периоды пауз), в радиотехнике известны и хорошо отработаны. В простейшем случае, длительность временной постоянной автостабилизации может регулироваться непосредственно добротностью селекторного фильтра или шириной полосы его избирательности. Формирование стробимпульсов по выделенной частоте синхронизации также известно в самых различных вариантах. В простейшем случае это можно выполнить усилением синхронизирующего сигнала с ограничением до формы меандра.

Способ 2. По частоте $2f''_0$.

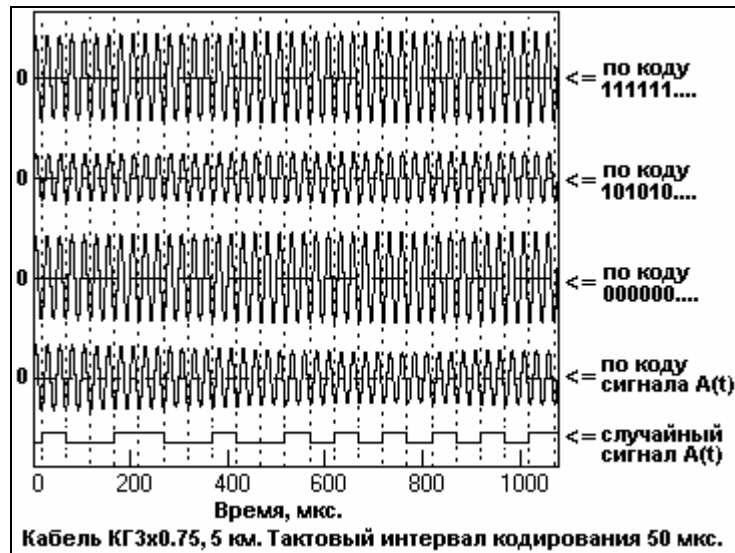


Рис. 16.2.5. Сигнал на выходе СФ на частоте $2f_0$.

Как следует из рисунка 16.2.2, на эту частоту при двухполупериодном амплитудном детектировании сигнала на выходе кабеля с практически одинаковой фазой переходят как несущие частоты непрерывных кодовых последовательностей нулей и единиц (и их отрезков любой длины в случайных последовательностях), так и вторые гармоники первой несущей частоты кода Манчестер-II при чередовании нулей и единиц (и кодовых пар 01 и 10 в произвольных кодовых последовательностях). Соответственно, частота $2f_0$ выделяется селективным фильтром в одной фазе для любой кодовой последовательности, что наглядно видно на рис. 16.2.5, в том числе и в периоды технологических пауз передачи данных (нулевых или единичных в зависимости от принятой системы передачи информационных данных).



Рис. 16.2.6. Спектры кодов после амплитудного детектора.

Стабильность фазовой синхронизации по частоте $2f_0$ можно видеть на рис. 16.2.6, где приведены спектры шести реализаций случайных кодовых последовательностей.

16.3. ДЕКОДИРОВАНИЕ ВЫХОДНЫХ СИГНАЛОВ.

Частота $2f_0$ в 2 раза выше требуемой частоты синхронизации для формирования строб-импульсов, но кратное понижение частоты в радиотехнике известно и может быть выполнено как в аналоговой форме, так и непосредственно при формировании строб-импульсов (усиление и ограничение частоты $2f_0$ до формы меандра с последующим триггерным понижением в 2 раза).

Формирование последовательности строб-импульсов целесообразно выполнять в виде меандра с положительной полярностью импульса строга в первой половине тактового интервала кодирования и отрицательной – во второй половине, т.е. строб-импульс в целом занимает полный такт и является двуполярным импульсом (один период меандра). В этом случае детектирование сигнала в двоичную (битовую) форму можно производить путем интегрирования сигнала в пределах тактовых интервалов с умножением на полный строб-импульс, т.е. значение интеграла сигнала первой половины тактового интервала суммировать

со значением интеграла сигнала второй половины тактового интервала со сменой его знака. С учетом протокола кодирования Манчестер-II и формы кодовых сигналов на выходе кабеля (рис. 16.1.2) это дает восстановление разнополярной битовой формы информации (1 – положительная полярность, 0 – отрицательная полярность).

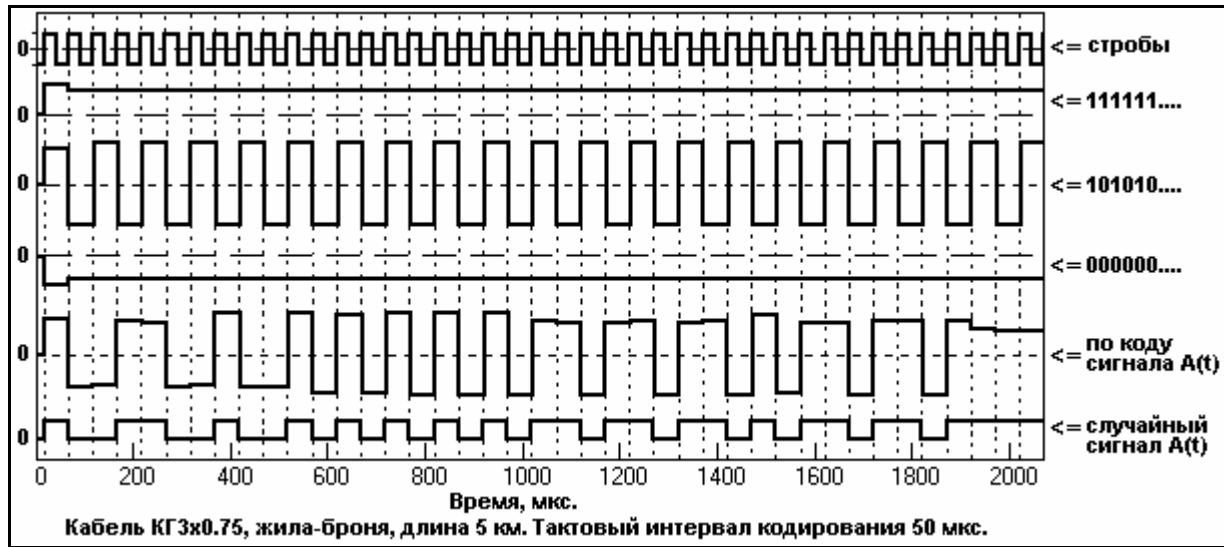


Рис. 16.3.1. Декодирование выходных сигналов.

Пример декодирования различных кодовых последовательностей выходного сигнала по вышеприведенному методу приведен на рис. 16.3.1. Амплитудные вариации сигналов, вызываемые различной амплитудой выходных сигналов для последовательностей нулей и единиц и кодов 01 и 10, на выходе декодера нормируются по амплитуде соответствующим пороговым ограничением или логической фиксацией полярности выходного сигнала декодера при помощи нуль-детектора. На рис. 16.3.1 и на дальнейших рисунках выходные сигналы декодера приводятся без нормировки для показа вариаций значений интеграла в пределах тактовых периодов при влиянии дестабилизирующих факторов.

Контроль фазовой синхронизации фронта положительного импульса строба с началом тактовых интервалов кодирования устанавливается следующим образом. Параллельно с интегрированием сигнала в пределах такта с учетом полярности строб-импульса проводится интегрирование сигнала в пределах такта без изменения знака во второй половине такта. Переход выходных сигналов кода Манчестер-II через ноль в середине такта приводит к тому, что выходные сигналы интегрирования без изменения знака существенно меньше выходных сигналов интегрирования с изменением знака (практически нулевые для непрерывных последовательностей нулей и единиц и в 2 и более раза меньше для кодов 01 и 10). Если синхронизация нарушается (первым в фактическом тактовом интервале идет импульс строба отрицательной полярности, т.е. меандр стробирования смещается на полтакта), то значения выходных сигналов интегрирования меняются прямо на противоположные как по значениям, так и по соотношению значений. Это может использоваться для постоянного контроля правильности автосинхронизации и для немедленного восстановления синхронизации и исправления кода при сбое синхронизации.

Метод автосинхронизации тактовой частоты приемника и передатчика при использовании достаточно узкополосного селективного фильтра выделения частоты $2f_0$ обеспечивает надежную синхронизацию на уровне статистических шумов на выходе кабеля, мощность которых может в несколько раз превышает среднюю мощность сигнала. Известные системы автосинхронизации в этих условиях полностью неработоспособны. Соответственно, интегрирование сигнала в пределах четко синхронизированных тактовых интервалов обеспечивает на высоком уровне шумов устойчивое декодирование сигналов. Предельное значение уровня шумов может определяться по устойчивости работы системы контроля за фазовой синхронизацией приемника и передатчика (разность интегралов со сменой и без смены знака на второй половине такта всегда должна иметь одну полярность).

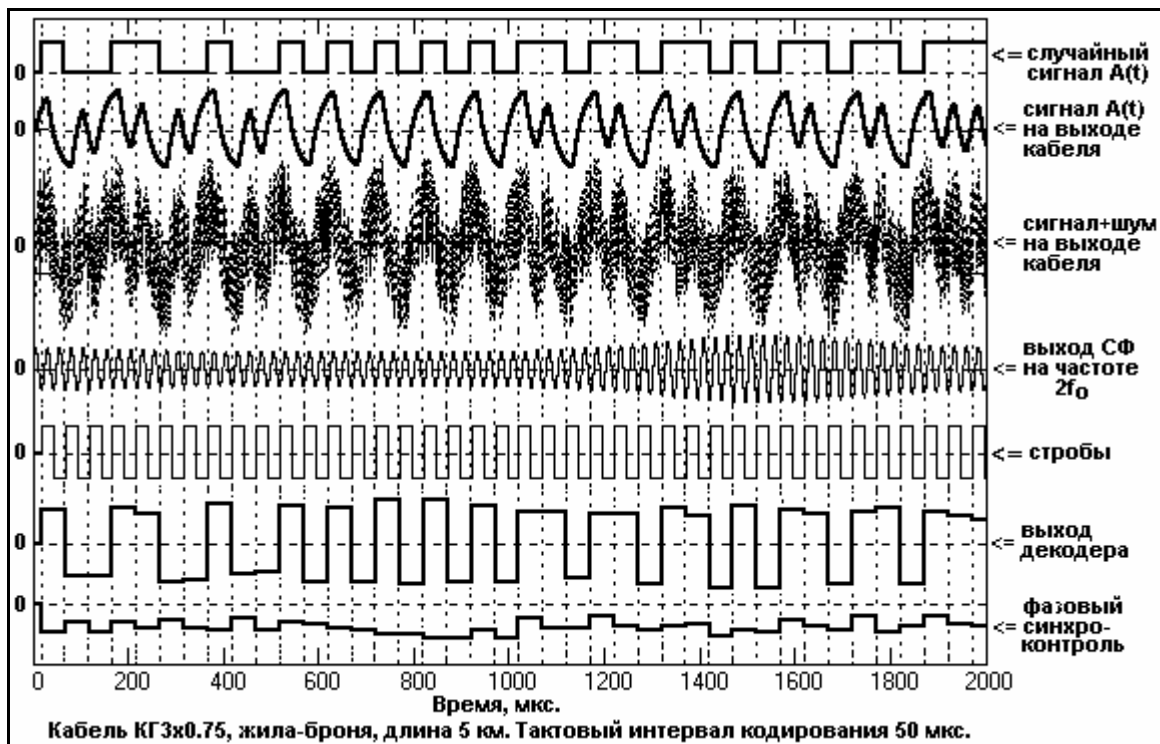


Рис. 16.3.2. Декодирование сигнала на высоком уровне шумов.

В качестве примера на рис. 16.3.2 приведены графики приема и декодирования произвольного сигнала на выходе кабеля, средняя мощность которого по интервалу 40 тактов в 2 раза меньше мощности шумов. Ширина полосы пропускания селективирующего фильтра на частоту $2f_0=40$ кГц была установлена равной 2 кГц на уровне 0.5 при постоянной времени установления фильтра порядка 10 тактовых интервалов (500 мкс).

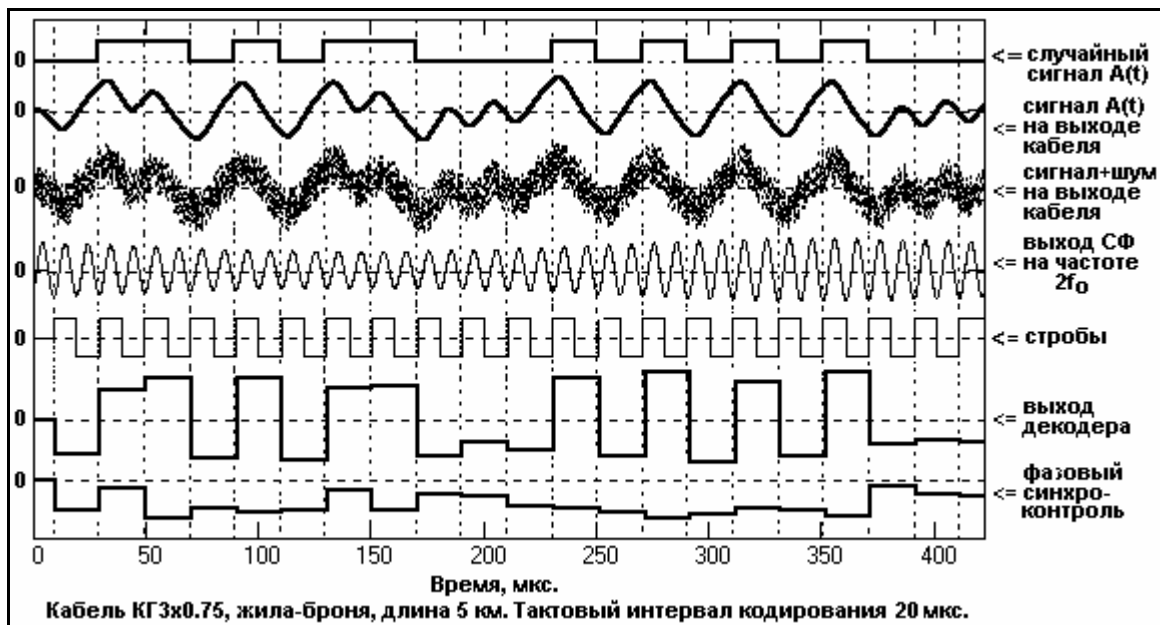


Рис. 16.3.3. Декодирование сигнала с повышенной тактовой частотой.

Возможность уверенного приема сигналов на высоком уровне шумов позволяет повысить тактовую частоту кодирования минимум в 2 раза. На рис. 16.3.3 приведен пример повышения тактовой частоты кодирования в 2.5 раза, до частоты 50 кГц. Амплитудные значения сигналов на такой частоте затухают для непрерывных последовательностей единиц и нулей в 20 раз, для чередования единиц и нулей в 8 раз. Мощность статистических шумов в пределах приведенного интервала равна средней мощности сигнала. Ширина полосы пропускания селективного фильтра, настроенного на частоту 100 кГц, в данном примере была

установлена равной 5 кГц на уровне 0.5.

Как можно видеть из приведенного примера, для кодов 01 и 10 в кодовых последовательностях сигналов без шума пересечения нулевой линии для второго бита вообще не наблюдается, что является следствием резкой асимметрии импульсного отклика кабеля. На работоспособности системы синхронизации это не отражается, и система декодирования обеспечивает уверенное выделение информационного сигнала.

ЛИТЕРАТУРА

22. Сиберт У.М. Цепи, сигналы, системы. – М.: Мир, 1988. – 336 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 17: СЛУЧАЙНЫЕ СИГНАЛЫ

Нет ничего более противного разуму и постоянству природы, чем случайность. Сам бог не может знать того, что произойдет случайно. Ибо если знает, то это определено произойдет, а если определено произойдет, то не случайно.

Марк Туллий Цицерон. О девиации.
Римский философ и политик, I в.д.н.э.

Случайность противна разуму, но не природе. Для проверки теории случайных процессов боги и создали мир. Швыряться яблоками они уже перестали, со времен Ньютона здесь ничего нового не наблюдалось. Но арбузные корки продолжают подсовывать - фиксируется непредсказуемая и зачастую очень даже интересная реакция.

Рудольф Гавшин. Случайность определенности.
Уральский геофизик, XX в.

Содержание: Введение. 17.1. Случайные процессы и функции. Случайный процесс. Функции математического ожидания и дисперсии. Корреляционная функция. Ковариационные функции. Свойства функций автоковариации и автокорреляции. Взаимные моменты случайных процессов. Классификация случайных процессов. 17.2. Функции спектральной плотности. Каноническое разложение случайных функций. Комплексные случайные функции. Финитное преобразование Фурье. Спектр функций случайных процессов. Взаимные спектральные функции. Теорема Винера-Хинчина. 17.3. Преобразования случайных функций. Системы преобразования случайных функций. Математическое ожидание выходного сигнала. Корреляционная функция выходного сигнала. Функция взаимной корреляции входного и выходного сигналов. Спектральные соотношения. Дисперсия выходного сигнала. Функция когерентности. Преобразования случайных функций. Преобразования стационарных случайных функций. 17.4. Модели случайных сигналов и помех. Телеграфный сигнал. Белый шум. Гауссовский шум. Гауссовские случайные процессы. Литература.

ВВЕДЕНИЕ.

Теория вероятностей рассматривает случайные величины и их характеристики в "статике". Задачи описания и изучения случайных сигналов "в динамике", как отображения случайных явлений, развивающихся во времени или по любой другой переменной, решает теория случайных процессов.

В качестве универсальной координаты для распределения случайных величин по независимой переменной будем использовать, как правило, переменную "t" и трактовать ее, чисто для удобства, как временную координату. Распределения случайных величин во времени, а равно и сигналов их отображающих в любой математической форме, обычно называют случайными процессами. В технической литературе термины "случайный сигнал" и "случайный процесс" используются как синонимы.

В отличие от детерминированных сигналов значения случайных сигналов в произвольные моменты времени не могут быть вычислены. Они могут быть только предсказаны в определенном диапазоне значений с определенной вероятностью, меньшей единицы. Количественные характеристики случайных сигналов, позволяющие производить их оценку и сравнение, называют *статистическими*.

В процессе обработки и анализа физико-технических данных обычно приходится иметь дело с тремя типами сигналов, описываемых методами статистики. Во-первых, это информационные сигналы, отображающие физические процессы, вероятностные по своей природе, как, например, акты регистрации частиц ионизирующих излучения при распаде радионуклидов. Во вторых, информационные сигналы, зависящие от определенных параметров физических процессов или объектов, значения которых заранее неизвестны, и которые обычно подлежат определению по данным информационным сигналам. И в третьих, это шумы и помехи, хаотически изменяющиеся во времени, которые сопутствуют информационным сигналам, но, как правило, статистически независимы от них как по своим значениям, так и по изменениям во времени.

17.1. СЛУЧАЙНЫЕ ПРОЦЕССЫ И ФУНКЦИИ [1, 2, 25].

Случайный процесс $X(t)$ представляет собой функцию, которая отличается тем, что принимаемые ею значения в любые произвольные моменты времени по координате t являются случайными. Строго с теоретических позиций, случайный процесс $X(t)$ следует рас-

считать как совокупность временных функций $x_k(t)$, имеющих определенную общую статистическую закономерность. При регистрации случайного процесса на определенном временном интервале осуществляется фиксирование единичной реализации $x_k(t)$ из бесчисленного числа возможных реализаций процесса $X(t)$. Эта единичная реализация называется *выборочной функцией* случайного процесса $X(t)$. Примеры выборочных функций модельного случайного процесса $X(t)$ приведены на рис. 17.1.1. В дальнейшем без дополнительных пояснений при рассмотрении различных параметров и характеристик случайных процессов для сопровождающих примеров будем использовать данную модель процесса.

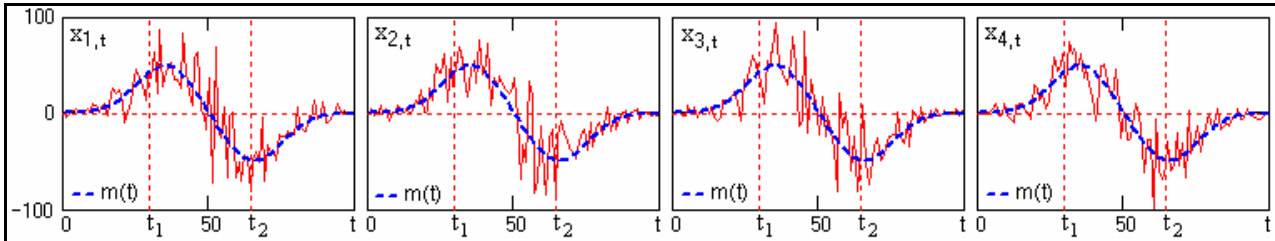


Рис. 17.1.1. Выборочные функции случайного процесса.

С практической точки зрения выборочная функция является результатом отдельного эксперимента, после которого данную реализацию $x_k(t)$ можно считать детерминированной функцией. Сам случайный процесс в целом должен анализироваться с позиции бесконечной совокупности таких реализаций, образующих *статистический ансамбль*. Полной статистической характеристикой такой системы является N -мерная плотность вероятностей $p(x_n, t_n)$. Однако, как экспериментальное определение N -мерных плотностей вероятностей процессов, так и их использование в математическом анализе представляет значительные математические трудности. Поэтому на практике обычно ограничиваются одно- и двумерной плотностью вероятностей процессов.

Функциональные характеристики случайного процесса.

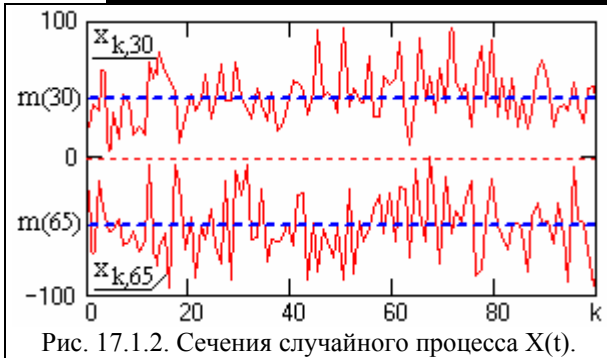


Рис. 17.1.2. Сечения случайного процесса $X(t)$.

Допустим, что случайный процесс $X(t)$ задан ансамблем реализаций $\{x_1(t), x_2(t), \dots, x_k(t), \dots\}$. В произвольный момент времени t_1 зафиксируем значения всех реализаций $\{x_1(t_1), x_2(t_1), \dots, x_k(t_1), \dots\}$. Совокупность этих значений представляет собой случайную величину $X(t_1)$ и является одномерным сечением случайного процесса $X(t)$. Примеры сечений по 100 выборкам случайного процесса $X(t)$ в точках t_1 и t_2 (рис. 17.1.1) приведены на

рис. 17.1.2.

Одномерная функция распределения вероятностей (x, t_i) определяет вероятность того, что в момент времени t_i значение случайной величины $X(t_i)$ не превысит значения x :

$$F(x, t_i) = P\{X(t_i) \leq x\}.$$

Очевидно, что в диапазоне значений вероятностей от 0 до 1 функция $F(x, t)$ является неубывающей с предельными значениями $F(-\infty, t) = 0$ и $F(\infty, t) = 1$. При известной функции $F(x, t)$ вероятность того, что значение $X(t_i)$ в выборках будет попадать в определенный интервал значений $[a, b]$ будет определяться выражением:

$$P\{a < X(t_i) \leq b\} = F(b, t_i) - F(a, t_i).$$

Одномерная плотность вероятностей $p(x, t)$ случайного процесса $X(t)$ характеризует распределение вероятностей реализации случайной величины $X(t_i)$ в произвольный момент времени t_i . Она представляет собой производную от функции распределения вероятностей:

$$p(x, t_i) = dF(x, t_i)/dx.$$

Моменты времени t_i являются сечениями случайного процесса $X(t)$ по пространству возможных состояний и плотность вероятностей $p(x, t_i)$ представляет собой плотность вероятностей случайных величин $X(t_i)$ данных сечений. Произведение $p(x, t_i) \cdot dx$ равно вероятности реализации случайной величины $X(t_i)$ в бесконечно малом интервале dx в окрестности значения x , откуда следует, что плотность вероятностей также является неотрицательной величиной.

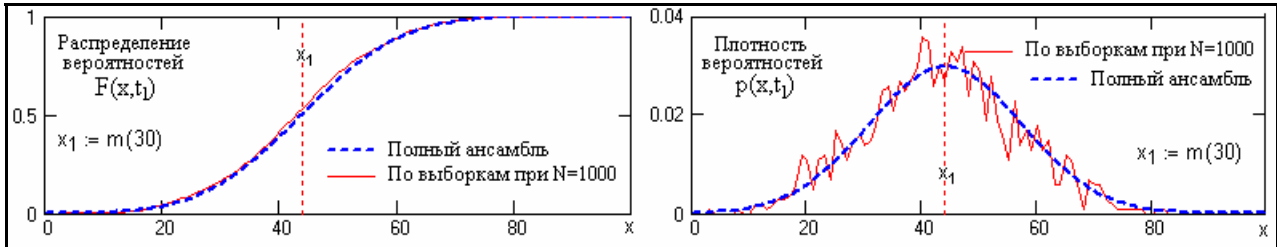


Рис. 17.1.3. Распределение вероятностей и плотность вероятностей сечения случайного процесса

На рис. 17.1.3 приведены примеры распределения вероятностей и плотности вероятностей сечения случайного процесса $X(t)$ в точке t_i (рис. 17.1.1). Функции вероятностей определены по $N=1000$ выборок дискретной модели случайного процесса и сопоставлены с теоретическими распределениями при $N \rightarrow \infty$.

При известной функции плотности вероятностей вероятность реализации значения $X(t_i)$ в произвольном интервале значений $[a, b]$ вычисляется по формуле:

$$P(a < X(t_i) \leq b) = \int_a^b p(x, t_i) dx.$$

Функция плотности вероятностей должна быть нормирована к 1, т.к. случайная величина обязана принимать какое-либо значение из числа возможных, образующих полное пространство случайных величин:

$$\int_{-\infty}^{\infty} p(x, t_i) dx = 1.$$

По известной плотности распределения вычисляется и функция распределения вероятностей:

$$F(x, t_i) = \int_{-\infty}^x p(x, t_i) dx.$$

Случайные процессы и их функции характеризуются неслучайными функциями математического ожидания (среднего значения), дисперсии и корреляции:

Математическое ожидание (mean value) представляет собой *статистическое усреднение* случайной величины $X(t_i)$, под которым понимают усреднение по ансамблю реализаций в каком либо фиксированном сечении t_i случайного процесса. Соответственно, функция математического ожидания является теоретической оценкой среднего взвешенного значения случайного процесса по временной оси:

$$m_x(t) \equiv M\{X(t)\} \equiv \overline{x(t)} = \int_{-\infty}^{\infty} x p(x; t) dx, \quad (17.1.1)$$

Математическое ожидание $m_x(t)$ представляет собой *неслучайную составляющую* случайного процесса $X(t)$. На рис. 17.1.1. и 17.1.2 неслучайные составляющие $m(t)$ модели случайного процесса $X(t)$ выделены пунктиром и соответствуют выборкам $N \rightarrow \infty$.

Функция дисперсии (variance) случайного процесса является теоретической оценкой среднего взвешенного значения разности $X(t) - m_x(t)$, которая называется *флуктуационной частью* процесса:

$$D_x(t) = M\{[X(t) - m_x(t)]^2\} = M\{X^2(t)\} - m_x^2(t) = \int_{-\infty}^{\infty} [x_o(t)]^2 p(x; t) dx, \quad (17.1.2)$$

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

Функция среднего квадратического отклонения (standard deviation) служит амплитудной мерой разброса значений случайного процесса по временной оси относительно математического ожидания процесса:

$$\sigma_x(t) = \sqrt{D_x(t)}. \quad (17.1.3)$$

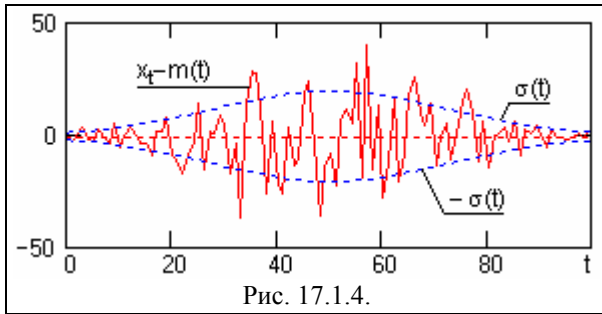


Рис. 17.1.4.

Учитывая последнее выражение, дисперсия случайной величины обычно обозначается индексом σ_x^2 .

На рис. 17.1.4 приведен пример флюктуационной составляющей процесса $X(t)$ (рис. 17.1.1) в одной из реализаций в сопоставлении со средним квадратическим отклонением $\pm\sigma$ случайных величин от математического ожидания $m(t)$.

Корреляционные функции случайных процессов. Одномерные законы плотности распределения вероятностей случайных процессов не несут каких-либо характеристик связи между значениями случайных величин для различных значений аргументов.

Двумерная плотность вероятностей $p(x_1, x_2; t_1, t_2)$ определяет вероятность совместной реализации значений случайных величин $X(t_1)$ и $X(t_2)$ в произвольные моменты времени t_1 и t_2 и в какой-то мере уже позволяет оценивать динамику развития процесса. Двумерная плотность вероятностей описывает двумерную случайную величину $\{X(t_i), X(t_j)\}$ в виде функции вероятности реализации случайной величины $X(t_i)$ в бесконечно малом интервале dx_i в окрестностях x_i в момент времени t_i при условии, что в момент времени t_j значение $X(t_j)$ будет реализовано в бесконечно малом интервале dx_j в окрестностях x_j :

$$p(x_i, x_j; t_i, t_j) dx_i dx_j = P\{|X(t_i - x_i)| \leq dx_i/2, |X(t_j - x_j)| \leq dx_j/2\}.$$

Характеристикой динамики изменения двумерной случайной величины $\{X(t_i), X(t_j)\}$ является корреляционная функция, которая описывает случайный процесс в целом:

$$R_X(t_i, t_j) = M\{X(t_i) X(t_j)\}.$$

Корреляционная функция представляет собой статистически усредненное произведение значений случайного процесса $X(t)$ в моменты времени t_i и t_j по всем значениям временных осей t_i и t_j , а следовательно тоже является двумерной функцией. В терминах теории вероятностей корреляционная функция является вторым начальным моментом случайного процесса.

На рис. 17.1.5 приведены примеры реализаций двух случайных процессов, которые характеризуются одной и той же функцией математического ожидания и дисперсии.

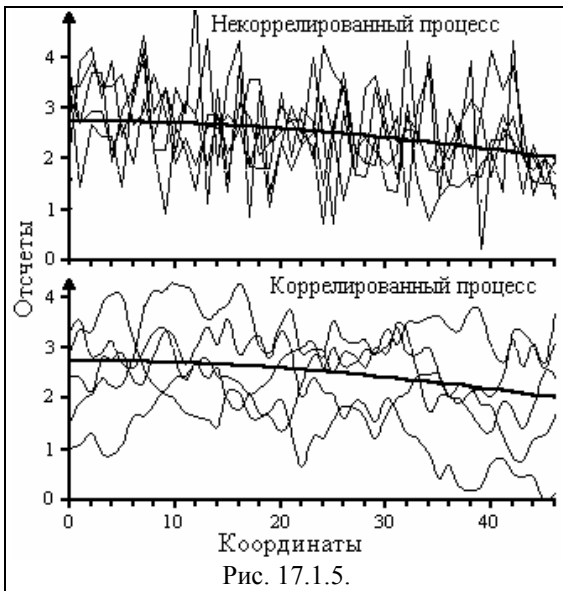
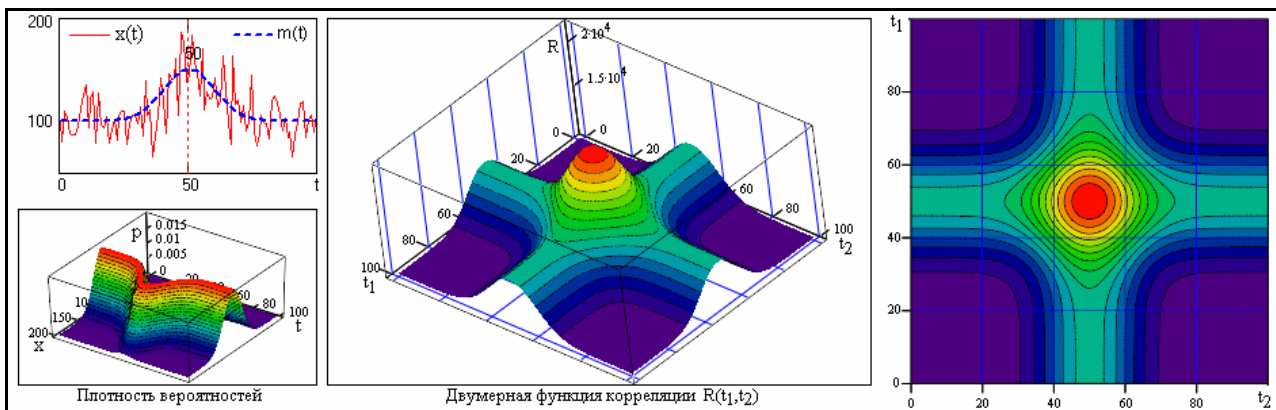


Рис. 17.1.5.

На рисунке видно, что хотя пространство состояний обоих процессов практически одно и то же, динамика развития процессов в реализациях существенно различается. Единичные реализации коррелированных процессов в произвольный момент времени могут быть такими же случайными, как и некоррелированных, а в пределе, во всех сечениях оба процесса могут иметь один и тот же закон распределения случайных величин. Однако динамика развития по координате t (или любой другой независимой переменной) единичной реализации коррелированного процесса по сравнению с некоррелированным является более плавной, а, следовательно, в коррелированном процессе имеется определенная связь между последовательными значениями случайных величин. Оценка степени

статистической зависимости мгновенных значений какого-либо процесса $X(t)$ в произвольные моменты времени t_1 и t_2 и производится функцией корреляции. По всему пространству значений случайного процесса $X(t)$ корреляционная функция определяется выражением:

$$R_X(t_i, t_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t_i)x(t_j) p(x_i, t_j; x_i, t_j) dx_i dx_j, \quad (17.1.4)$$

Рис. 17.1.6. Двумерная плотность вероятностей и корреляционная функция процесса $X(t)$.

На рис. 17.1.6 приведена форма модельного случайного процесса $X(t)$ в одной выборке со значительной и изменяющейся неслучайной составляющей. Модель задана на интервале $0-T$ ($T=100$) в дискретной форме с шагом $\Delta t=1$. Корреляционная функция вычислена по заданной плотности вероятностей модели

При анализе случайных процессов второй момент времени t_j удобно задавать величиной сдвига τ относительно первого момента, который при этом может быть задан в виде координатной переменной:

$$R_X(t, t+\tau) = M\{X(t)X(t+\tau)\}. \quad (17.1.4')$$

Функция, задаваемая этим выражением, обычно называется автокорреляционной функцией случайного процесса.

Ковариационные функции. Частным случаем корреляционной функции является функция автоковариации (ФАК), которая широко используется при анализе сигналов. Она представляет собой статистически усредненное произведение значений центрированной случайной функции $X(t)-m_X(t)$ в моменты времени t_i и t_j и характеризует флюктуационную составляющую процесса:

$$K_X(t_i, t_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x(t_i) - m_X(t_i)) (x(t_j) - m_X(t_j)) p(x_i, t_j; x_i, t_j) dx_i dx_j, \quad (17.1.5)$$

В терминах теории вероятностей ковариационная функция является вторым центральным моментом случайного процесса. Для центрированных случайных процессов ФАК тождественна функции корреляции. При произвольных значениях m_X ковариационные и корреляционные функции связаны соотношением:

$$K_X(t, t+\tau) = R_X(t, t+\tau) - m_X^2(t).$$

Нормированная функция автоковариации (функция корреляционных коэффициентов):

$$\rho_X(t, t+\tau) = K_X(t, t+\tau) / [\sigma(t)\sigma(t+\tau)]. \quad (17.1.6)$$

При $\tau = 0$ значение ρ_X равно 1, а ФАК вырождается в дисперсию случайного процесса:

$$K_X(t) = D_X(t).$$

Отсюда следует, что для случайных процессов и функций основными характеристиками являются функции математического ожидания и корреляции (ковариации). Особой необходимости в отдельной функции дисперсии не имеется.

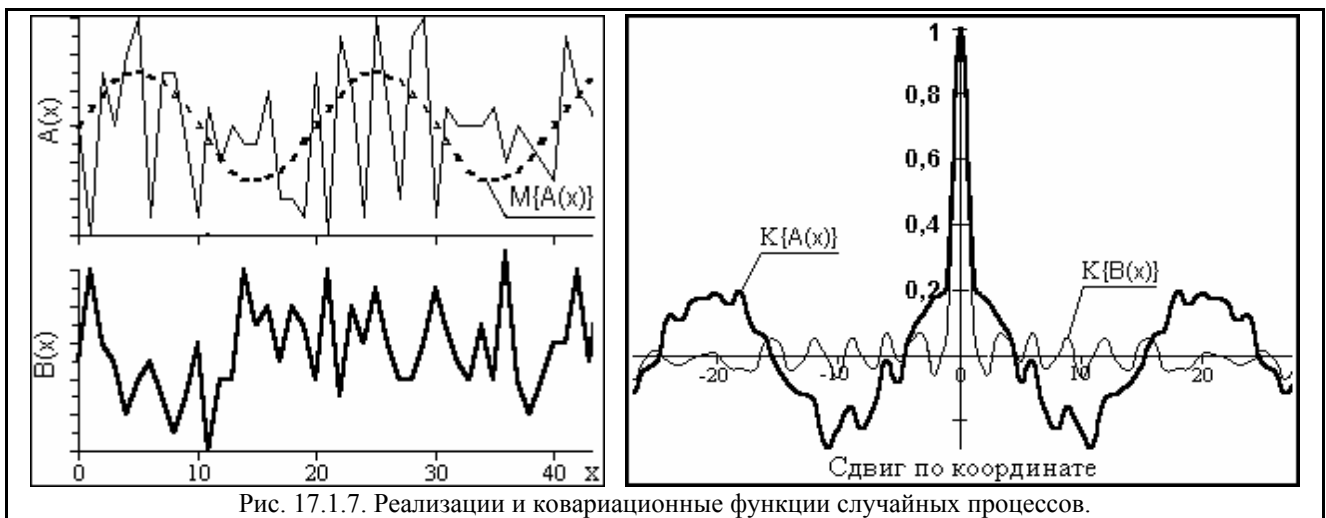


Рис. 17.1.7. Реализации и ковариационные функции случайных процессов.

Примеры реализаций двух различных случайных процессов и их нормированных ковариационных функций приведены на рис. 17.1.7.

Свойства функций автоковариации и автокорреляции.

1. Максимум функций наблюдается при $\tau = 0$. Это очевидно, т.к. при $\tau = 0$ вычисляется степень связи отсчетов с собой же, которая не может быть меньше связи разных отсчетов. Значение максимума функции корреляции равно средней мощности сигнала.

2. Функции автокорреляции и автоковариации являются четными: $R_X(\tau) = R_X(-\tau)$. Последнее также очевидно: $X(t)X(t+\tau) = X(t-\tau)X(t)$ при $t = t-\tau$. Говоря иначе, моменты двух случайных величин $X(t_1)$ и $X(t_2)$ не зависят от последовательности, в которой эти величины рассматриваются, и соответственно симметричны относительно своих аргументов: $R_X(t_1, t_2) = R_X(t_2, t_1)$, равно как и $K_X(t_1, t_2) = K_X(t_2, t_1)$.

3. При $\tau \Rightarrow \infty$ значения ФАК для сигналов, конечных по энергии, стремятся к нулю, что прямо следует из физического смысла ФАК. Это позволяет ограничивать длину ФАК определенным максимальным значением τ_{\max} - радиусом корреляции, за пределами которого отсчеты можно считать независимыми. Интегральной характеристикой времени корреляции случайных величин обычно считают *эффективный интервал корреляции*, определяемый по формуле:

$$T_k = 2 \int_0^{\infty} |\rho_X(\tau)| d\tau \equiv (2/K_X(0)) \int_0^{\infty} |K_X(\tau)| d\tau. \quad (17.1.7)$$

Отсчеты (сечения) случайных функций, отстоящие друг от друга на расстояние большее T_k , при инженерных расчетах считают некоррелированными.

Заметим, что для некоррелированных процессов при $t \Rightarrow \infty$ значение T_k стремится к 2,

что несколько противоречит физическому смыслу радиуса корреляции, который в этом случае должен был бы стремиться к 1. С учетом последнего эффективный интервал корреляции целесообразно определять по формуле:

$$T_k = 2 \int_0^{\infty} |\rho_x(\tau)| d\tau - 1 \equiv (2/K_x(0)) \int_0^{\infty} |K_x(\tau)| d\tau - 1. \quad (17.1.7')$$

4. Если к случайной функции $X(t)$ прибавить неслучайную функцию $f(t)$, то ковариационная функция не изменяется.

Обозначим новую случайную функцию как $Y(t)=X(t)+f(t)$. Функция математического ожидания новой величины: $\bar{y}(t) = \bar{x}(t) + f(t)$. Отсюда следует, что $Y(t) - \bar{y}(t) = X(t) - \bar{x}(t)$, и соответственно $K_y(t_1, t_2) = K_x(t_1, t_2)$.

5. Если случайную функцию $X(t)$ умножить на неслучайную функцию $f(t)$, то ее корреляционная функция $R_x(t_1, t_2)$ умножится на $f(t_1) \cdot f(t_2)$. Обоснование данного свойства проводится по методике, аналогичной предыдущему пункту.

6. При умножении функции случайного процесса на постоянное значение C значения ФАК увеличиваются в C^2 раз.

Взаимные моменты случайных процессов второго порядка дают возможность оценить совместные свойства двух случайных процессов $X(t)$ и $Y(t)$ путем анализа произвольной пары выборочных функций $x_k(t)$ и $y_k(t)$.

Мера связи между двумя случайными процессами $X(t)$ и $Y(t)$ также устанавливается корреляционными функциями, а именно - функциями взаимной корреляции и взаимной ковариации. В общем случае, для произвольных фиксированных моментов времени $t_1 = t$ и $t_2 = t+\tau$:

$$R_{XY}(t, t+\tau) = M\{X(t)Y(t+\tau)\}. \quad (17.1.8)$$

$$K_{XY}(t, t+\tau) = M\{(X(t)-m_x(t))(Y(t+\tau)-m_y(t+\tau))\}. \quad (17.1.9)$$

Взаимные функции являются произвольными функциями (не обладают свойствами четности или нечетности), и удовлетворяют следующим соотношениям:

$$\begin{aligned} R_{xy}(-\tau) &= R_{yx}(\tau), \\ |R_{xy}(\tau)|^2 &\leq R_x(0)R_y(0). \end{aligned} \quad (17.1.10)$$

Если один из процессов центрированный, то имеет место $R_{xy}(t) = K_{xy}(t)$.

Нормированная взаимная ковариационная функция (коэффициент корреляции двух процессов), которая характеризует степень линейной зависимости между случайными процессами при данном сдвиге τ одного процесса по отношению ко второму, определяется выражением:

$$\rho_{xy}(\tau) = K_{xy}(\tau)/(\sigma_x \sigma_y). \quad (17.1.11)$$

Статистическая независимость случайных процессов определяет отсутствие связи между значениями двух случайных величин X и Y . Это означает, что плотность вероятности одной случайной величины не зависит от того, какие значения принимает вторая случайная величина. Двумерная плотность вероятностей при этом должна представлять собой произведение одномерных плотностей вероятностей этих двух величин:

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

Это условие является обязательным условием статистической независимости случайных величин. В противном случае между случайными величинами может существовать определенная статистическая связь. Как линейная, так и нелинейная. Мерой линейной статистической связи является коэффициент корреляции:

$$r_{xy} = [M\{X \cdot Y\} - M\{X\} \cdot M\{Y\}] / \sqrt{D\{X\} D\{Y\}}.$$

Значений r_{xy} могут изменяться в пределах от -1 до +1. В частном случае, если случайные величины связаны линейным соотношением $x=ay+b$, коэффициент корреляции равен ± 1 в зависимости от знака константы a . Случайные величины некоррелированы при $r_{xy}=0$, при этом из выражения для r_{xy} следует:

$$M\{X \cdot Y\} = M\{X\} \cdot M\{Y\}.$$

Из статистической независимости величин следует их некоррелированность. Обратное не очевидно. Так, например, случайные величины $x = \cos \varphi$ и $y = \sin \varphi$, где φ - случайная величина с равномерным распределением в интервале $0 \dots 2\pi$, имеют нулевой коэффициент корреляции, и вместе с тем их зависимость очевидна.

Классификация случайных процессов. Случайные процессы различают по степени однородности их протекания во времени (по аргументу).

Нестационарные процессы. В общем случае значения функций математического ожидания, дисперсии и корреляции могут быть зависимыми от момента времени t , т.е. изменяться во времени. Такие процессы составляют класс нестационарных процессов.

Стационарные процессы. Процесс называют стационарным, если плотность вероятностей процесса не зависит от начала отсчета времени и если на интервале его существования выполняются условия постоянства математического ожидания и дисперсии, а корреляционная функция является функцией только разности аргументов $\tau = t_2 - t_1$, т.е.:

$$\begin{aligned} m_X(t_1) = m_X(t_2) = m_X = \text{const}, \\ D_X(t_1) = D_X(t_2) = D_X = \text{const}, \end{aligned} \quad (17.1.12)$$

$$\begin{aligned} R_X(t_1, t_1 + \tau) \equiv R_X(t_2 - \tau, t_2) = R_X(\tau) \equiv R_X(-\tau), \\ r_X(\tau) = R_X(\tau) / D_X, \quad r_X(0) = 1, \quad |r_X(\tau)| \leq 1, \quad r_X(-\tau) = r_X(\tau). \end{aligned}$$

Последние выражения свидетельствуют о четности корреляционной (а равно и ковариационной) функции и функции корреляционных коэффициентов. Из него вытекает также еще одно свойство смешанных моментов стационарных процессов:

$$|R_X(\tau)| \leq R_X(0), \quad |K_X(\tau)| \leq K_X(0) \equiv D_X.$$

Чем медленнее по мере увеличения значений τ убывают функции $R_X(\tau)$ и $r_X(\tau)$, тем больше эффективный интервал корреляции случайного процесса, и тем медленнее изменяются во времени его реализации.

Среди стационарных процессов выделяют *строго стационарные процессы*, для которых постоянны во времени не только математическое ожидание, дисперсия и корреляция, но и все остальные моменты высших порядков (в частности, асимметрия и эксцесс).

Стационарные случайные процессы наиболее часто встречаются при решении физических и технических задач. Теория стационарных случайных функций разработана наиболее полно и для ее использования обычно достаточно определения стационарности *в широком смысле*: случайная функция считается стационарной, если ее математическое ожидание постоянно, а корреляционная функция зависит только от одного аргумента. Случайные процессы, удовлетворяющие условиям стационарности на ограниченных, интересующих нас интервалах, также обычно относят к числу стационарных в широком смысле и называют квазистационарными.

Эргодические процессы. Строго корректно характеристики случайных процессов оцениваются путем усреднения по ансамблю реализаций в определенные моменты времени (по сечениям процессов). Но большинство стационарных случайных процессов обладает эргодическим свойством. Сущность его заключается в том, что по одной достаточно длинной реализации процесса можно судить о всех его статистических свойствах так же, как по любому количеству реализаций. Другими словами, закон распределения случайных величин в таком процессе может быть одним и тем же как по сечению для ансамбля реализаций, так и по координате развития. Такие процессы получили название эргодических (ergodic). Для эргодических процессов имеет место:

$$m_X(t) = M\{x(t)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt, \quad (17.1.13)$$

$$D_X(t) = M\{x(t) - m_X(t)\}^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (x(t) - m_X(t))^2 dt, \quad (17.1.14)$$

$$R_X(\tau) = M\{x(t)x(t+\tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)x(t+\tau) dt. \quad (17.1.15)$$

Эргодичность является очень важным свойством случайных стационарных, и только стационарных процессов. Математическое ожидание эргодического случайного процесса равно постоянной составляющей любой его реализации, а дисперсия является мощностью его флюктуационной составляющей. Так как определение функций производится по ограниченным статистическим данным одной реализации и является только определенным приближением к соответствующим фактическим функциям процессов, целесообразно называть эти функции статистическими. Заметим, что, как это следует из (17.1.15), вычисление корреляционной функции подобно свертке (с делением на интервал реализации) и может записываться символически:

$$R_X(\tau) = (1/T) x(t) * x(t+\tau).$$

Свойства эргодичности могут проявляться только по отношению к двум первым моментам случайного процесса, что вполне достаточно для использования соответствующих методик исследования процессов. Практическая проверка эргодичности процесса обычно производится проверкой выполнения условия Слуцкого:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T K(\tau) d\tau = 0. \quad (17.1.16)$$

Если ковариационная функция процесса стремится к нулю при возрастании значения аргумента (τ), то процесс относится к числу эргодических, по крайней мере относительно моментов первого и второго порядков.

Пример. Случайная функция задана выражением $Z(t)=X(t)+Y$, где $X(t)$ - стационарная эргодичная функция, Y - случайная величина, некоррелированная с $X(t)$. Эргодична ли функция $Z(t)$?

$$m_Z(t) = m_Z(x)+m_Y, \quad K_Z(\tau) = K_X(\tau)+D_Y.$$

Функция $Z(t)$ стационарна, но не эргодична, так как при $\tau \Rightarrow \infty$ имеет место $K_Z(\tau) \Rightarrow D_Y$.

17.2. ФУНКЦИИ СПЕКТРАЛЬНОЙ ПЛОТНОСТИ [2,25,26].

Каноническое разложение случайных функций. Введем понятие простейшей случайной функции, которая определяется выражением:

$$X(t) = X \cdot \varphi(t), \quad (17.2.1)$$

где X - обычная случайная величина, $\varphi(t)$ - произвольная неслучайная функция. Математическое ожидание простейшей случайной функции:

$$m_X(t) = M\{X\varphi(t)\} = \varphi(t) \cdot M\{X\} = \varphi(t) \cdot m_X, \quad (17.2.2)$$

где m_X - математическое ожидание случайной величины X . При $m_X = 0$ математическое ожидание $m_X(t)$ также равно нулю для всех t и функция (17.2.1) в этом случае называется элементарной случайной функцией. Ковариационная функция элементарной случайной функции определится выражением:

$$K_X(t_1, t_2) = M\{X(t_1)X(t_2)\} = \varphi(t_1)\varphi(t_2) \cdot M\{X^2\} = \varphi(t_1)\varphi(t_2) \cdot D_X. \quad (17.2.3)$$

где D_X - дисперсия случайной величины X .

Центрированную случайную функцию ${}^0X(t)$ можно представить суммой взаимно некоррелированных элементарных случайных функций:

$${}^0X(t) = \sum_{i=1}^M X_i \cdot \varphi_i(t), \quad (17.2.4)$$

Из взаимной некоррелированности элементарных случайных функций следует взаимная некоррелированность величин X_i . Математическое ожидание и ковариационная функция случайной функции ${}^0X(t)$:

$$M\{{}^0X(t)\} = M\left\{\sum_{i=1}^M X_i \cdot \varphi_i(t)\right\} = 0.$$

$$K_x(t_1, t_2) = M\{{}^0X(t_1) {}^0X(t_2)\} = M\left\{\sum_{i,j} X_i \cdot \varphi_i(t_1) X_j \cdot \varphi_j(t_2)\right\} = \sum_{i,j} \varphi_i(t_1) \varphi_j(t_2) M\{X_i X_j\}.$$

В силу взаимной некоррелированности парных значений $X_i X_j$ имеет место $M\{X_i X_j\} = 0$ при $i \neq j$, и все члены суммы в последнем выражении равны нулю, за исключением значений при $i = j$, для которых $M\{X_i X_j\} = M\{X_i^2\} = D_i$. Отсюда:

$$K_x(t_1, t_2) = \sum_{i=1}^M \varphi_i(t_1) \varphi_i(t_2) D_i. \quad (17.2.5)$$

Произвольная нецентрированная случайная функция соответственно может быть представлена в виде

$$X(t) = m_x(t) + {}^0X(t) = m_x(t) + \sum_{i=1}^M X_i \cdot \varphi_i(t), \quad (17.2.6)$$

с математическим ожиданием $m_x(t)$ и с той же самой ковариационной функцией (17.2.5) в силу свойств ковариационных функций, где ${}^0X(t)$ - флюктуационная составляющая случайной функции $X(t)$. Выражение (17.2.6) и является каноническим разложением функции $X(t)$. Случайные величины X_i называются коэффициентами разложения, функции φ_i - координатными функциями разложения. При $t_1 = t_2$ из (17.2.5) получаем функцию дисперсии случайной функции $X(t)$:

$$D_x(t) = \sum_{i=1}^M [\varphi_i(t)]^2 \cdot D_i. \quad (17.2.7)$$

Таким образом, зная каноническое разложение (17.2.6) функции $X(t)$, можно сразу определить каноническое разложение (17.2.5) ее ковариационной функции, и наоборот. Канонические разложения удобны для выполнения различных операций над случайными функциями. Это объясняется тем, что в разложении зависимость функции от аргумента t выражается через неслучайные функции $\varphi_i(t)$, а соответственно операции над функцией $X(t)$ сводятся к соответствующим операциям математического анализа над координатными функциями $\varphi_i(t)$.

В качестве координатных функций разложения, как и при анализе детерминированных сигналов, обычно используются гармонические синус-косинусные функции, а в общем случае комплексные экспоненциальные функции $\exp(j\omega t)$. С учетом последнего предварительно рассмотрим особенности представления случайных функций в комплексной форме.

Комплексные случайные функции. В общем случае случайный процесс может описываться комплексной случайной функцией:

$$Z(t) = X(t) + jY(t), \quad (17.2.8)$$

где $X(t)$ и $Y(t)$ - действительные случайные функции. Соответственно, математическое ожидание комплексной функции:

$$m_z(t) = m_x(t) + j \cdot m_y(t). \quad (17.2.9)$$

Заметим, что комплексное представление случайных функций не более чем удобная для анализа математическая форма их отображения, которая, с использованием выражений Эйлера, всегда может быть переведена в форму вещественных функций. Функции дисперсии, корреляции и ковариации должны представлять собой однозначные и неслучайные вещественные характеристики случайных процессов и функций, независимо от формы их математического представления. Это условие будет выполняться при использовании в выражениях моментов второго порядка операций умножения комплексных функций с комплексно сопряженными функциями. Так, выражение для вычисления корреляционной функции имеет следующий вид:

$$\begin{aligned} R_z(t_1, t_2) &= M\{Z(t_1) \cdot Z^*(t_2)\} = M\{[X(t_1) + jY(t_1)][X(t_2) - jY(t_2)]\} = \\ &= M\{X(t_1)X(t_2) + Y(t_1)Y(t_2) + j[Y(t_1)X(t_2) - X(t_1)Y(t_2)]\} = \\ &= R_x(t_1, t_2) + R_y(t_1, t_2) + j[R_{yx}(t_1, t_2) - R_{xy}(t_1, t_2)]. \end{aligned} \quad (17.2.10)$$

Если действительные и мнимые части комплексной функции некоррелированы, то

$R_{yx} = R_{xy} = 0$ и последний член выражения (17.2.10) также равен нулю.

Аналогичное выражение имеет место и для ковариационной функции. При $t_1 = t_2 = t$ для функции дисперсии комплексной случайной величины имеем:

$$D_z(t) = M\{|Z(t) - m_z(t)|^2\} = D_x(t) + D_y(t), \quad (17.2.11)$$

Все приведенные выражения в общем случае могут использоваться для любых комплексных случайных функций с любым физическим смыслом переменной t .

Финитное преобразование Фурье случайных функций. По аналогии с функциями детерминированных сигналов, отдельно взятая на интервале $0-T$ реализация $x_k(t)$ стационарного случайного процесса ${}^0X(t)$ может быть представлена в виде ряда Фурье:

$$x_k(t) = \sum_{i=-\infty}^{\infty} V_{x,k}(\omega_i) \exp(j\omega_i t), \quad (17.2.12)$$

$$V_{x,k}(\omega_i) = (1/T) \int_0^T x_k(t) \exp(-j\omega_i t) dt, \quad (17.2.13)$$

или, в односторонней тригонометрической форме:

$$x_k(t) = A_{x,k}(0) + 2 \sum_{i=1}^{\infty} (A_{x,k}(\omega_i) \cos(\omega_i t) + B_{x,k}(\omega_i) \sin(\omega_i t)), \quad (17.2.12')$$

$$A_{x,k}(\omega_i) = (1/T) \int_0^T x_k(t) \cos(\omega_i t) dt, \quad (17.2.13')$$

$$B_{x,k}(\omega_i) = (1/T) \int_0^T x_k(t) \sin(\omega_i t) dt. \quad (17.2.13'')$$

где $\omega_i = i \cdot \Delta\omega$ - частоты спектра, $\Delta\omega = 2\pi/T$ - шаг по частоте. Выражения (17.2.13) обычно называют *спектральными характеристиками реализаций*. Из сравнения выражений (17.2.4) и (17.2.12) нетрудно сделать заключение, что выражения (17.2.12) относятся к числу канонических разложений случайных функций, при этом спектральная характеристика $V_{x,k}(\omega)$, а равно и ее составляющие $A_{x,k}(\omega)$ и $B_{x,k}(\omega)$, также являются случайными функциями частоты - единичными реализациями случайных функций $V_x(\omega)$, $A_x(\omega)$ и $B_x(\omega)$. Соответственно, и частотное распределение амплитуд и фаз составляющих гармонических колебаний случайного процесса ${}^0X(t)$ представляет собой случайные функции с соответствующими неслучайными функциями дисперсий.

Если функция ${}^0X(t)$ является дискретной последовательностью случайных величин ${}^0X(n \cdot \Delta t)$ в интервале по n от 0 до N , то, как это и положено для дискретных преобразований Фурье, расчет спектральных характеристик выполняется в Главном частотном диапазоне (до частоты Найквиста $\omega_N = \pi/\Delta t$), с заменой в выражениях (17.2.13) интегрирования на суммирование по n и с соответствующим изменением пределов суммирования в выражениях (17.2.12). Данное пояснение сохраняется и на все дальнейшие выкладки.

Спектральные характеристики единичных реализаций случайных процессов интереса, как правило, не представляют и на практике используются довольно редко. Спектральная характеристика случайной функции ${}^0X(t)$, как ансамбля реализаций, может быть определена осреднением функций (17.2.12-13) по реализациям, в результате которого мы получим те же самые функции (17.2.12-13), только без индексов k . При этом, в силу центрированности стационарной случайной функции ${}^0X(t)$, мы должны иметь:

$$M\{X(t)\} = \sum_{i=-\infty}^{\infty} M\{V_x(\omega_i)\} \exp(j\omega_i t) = 0, \quad (17.2.14)$$

Последнее будет выполняться при условии $M\{V_x(\omega_i)\} = 0$, т.е. математическое ожидание значений спектральной характеристики центрированного стационарного случайного процесса должно быть равно нулю на всех частотах. Другими словами, спектральной характеристики центрированного стационарного случайного процесса не существует. Существуют только спектральные характеристики его отдельных реализаций, которые и используются, например, для моделирования этих реализаций.

Для произвольных нецентрированных случайных процессов $X(t)$, при записи последних в форме $X(t) = m_x(t) + {}^0X(t)$, будем соответственно иметь преобразование Фурье:

$$m_x(t) + {}^0X(t) \Leftrightarrow m_x(\omega) + V_x(\omega) = m_x(\omega),$$

т.е., по существу, функцию спектра (или спектральной плотности) неслучайной функции математического ожидания случайного процесса, естественно, в пределах той точности, которую может обеспечить выборочный ансамбль реализаций. Это лишний раз подтверждает отсутствие в спектрах случайных процессов какой-либо информации о флюктуационной составляющей процессов и говорит о том, что фазы спектральных составляющих в реализациях процесса являются случайными и независимыми.

С учетом вышеизложенного, под спектрами случайных процессов (или спектральной плотностью при интегральном преобразовании Фурье) повсеместно понимается не преобразования Фурье собственно случайных функций, а преобразования Фурье функций мощности случайных процессов, поскольку функции мощности не зависят от соотношения фаз спектральных составляющих процессов.

Спектры мощности случайных функций определяются аналогично спектрам мощности детерминированных сигналов. Средняя мощность случайного процесса $X(t)$, зарегистрированного в процессе одной реализации на интервале $0-T$, с использованием равенства Парсеваля может быть вычислена по формуле:

$$P_T = \int_0^T [x^2(t)/T] dt = \int_{-\infty}^{\infty} [|X_T(f)|^2/T] df,$$

где $X(f)$ – спектральная плотность единичной реализации $x(t)$. При увеличении интервала T энергия процесса на интервале неограниченно нарастает, а средняя мощность стремится к определенному пределу:

$$P = \int_{-\infty}^{\infty} \left[\lim_{T \rightarrow \infty} \frac{1}{T} |X_T(f)|^2 \right] df,$$

где подынтегральная функция представляет собой спектральную плотность средней мощности данной реализации случайного процесса:

$$W(f) = \lim_{T \rightarrow \infty} \frac{1}{T} |X_T(f)|^2.$$

Очень часто это выражение называют просто спектром мощности. Плотность мощности является вещественной, неотрицательной и четной функцией частоты. В общем случае, плотность мощности необходимо усреднять по множеству реализаций, но для эргодических процессов допустимо усреднение по одной достаточно длительной реализации. Средняя мощность любой реализации центрированного процесса равна его дисперсии:

$$D_x = \int_{-\infty}^{\infty} W(f) df.$$

Спектр функций случайных процессов. При представлении ковариационной функции на интервале $0-T$, шаг по спектру функции с учетом четности ковариационной функции устанавливается равным $\Delta\omega = \pi/T$, $\omega_i = i \cdot \Delta\omega$, а спектр определяется обычно непосредственно по косинусам в односторонней форме:

$$K_x(\tau) = D_x(0)/2 + \sum_{i=1}^{\infty} D_x(\omega_i) \cos(\omega_i\tau), \quad (17.2.15')$$

$$D_x(\omega_i) = (2/T) \int_0^T K_x(\tau) \cos(\omega_i\tau) d\tau, \quad (17.2.16')$$

где $D_x(\omega_i)$ в соответствии с (17.2.5) - дисперсии случайных величин $V_x(\omega_i)$, а равно и $A_x(\omega_i)$ и $B_x(\omega_i)$, в разложениях (17.2.12). В комплексной форме, как обычно:

$$K_x(\tau) = \sum_{i=-\infty}^{\infty} D_x(\omega_i) \exp(j\omega_i\tau), \quad (17.2.15)$$

$$D_x(\omega_i) = (2/T) \int_0^T K_x(\tau) \exp(-j\omega_i\tau) d\tau, \quad (17.2.16)$$



Рис. 17.2.1. Спектры случайных функций.

Спектры случайных функций всегда ограничены ($D(\omega) \neq \infty$) и неотрицательны ($D(\omega) \geq 0$), при двустороннем представлении всегда четные ($D(-\omega) = D(\omega)$). Пример спектров в одно- и двустороннем представлении приведен на рис. 17.2.1.

Дисперсия стационарного случайного процесса $X(t)$ может определяться по формуле (17.2.15) при $\tau = 0$:

$$D_x = \sum_{i=-\infty}^{\infty} D_x(\omega_i), \quad (17.2.17)$$

т.е. дисперсия стационарного случайного процесса равна сумме дисперсий всех случайных гармоник ее спектрального разложения.

Обобщенной характеристикой спектра случайного процесса служит эффективная ширина спектра, определяемая по формуле:

$$B_k = (\Delta\omega/D_{\max}) \sum_{i=0}^{\infty} D_x(\omega_i) = \Delta\omega \cdot D_x/D_{\max}, \quad (17.2.18)$$

где D_{\max} — максимальное значение функции $D_x(\omega_i)$. Отметим, что ширина спектра является практической характеристикой случайного процесса и вычисляется, как правило, для реальных частот по одностороннему спектру процесса. При вычислении по двустороннему спектру, где значение D_{\max} соответственно в два раза меньше, чем в одностороннем спектре, величина B_k завышается в два раза, если суммирование осуществлять по всему спектру. Поэтому пределы суммирования в выражении (17.2.18) не изменяются вне зависимости от того, какой вид спектра используется.

При использовании предельного перехода $T \Rightarrow \infty$ и соответственно интегралов Фурье в выражениях (17.2.15), двусторонние функции дисперсий $D(\omega_i)$ заменяются функциями $S(\omega)$, а односторонние — функциями $G(\omega)$, которые называют соответственно дву- и односторонними функциями *спектральной плотности* случайных процессов. Такое же индексирование в научно-технической литературе применяют и для спектров корреляционных функций, а зачастую и для дискретных преобразований ковариационных функций вместо $D(\omega_i)$, хотя последнее применительно к ковариационным функциям более точно отражает физическую сущность величин. Но оно может считаться вполне приемлемым для сохранения общности математических описаний.

Эффективная ширина спектра для функций спектральной плотности случайных процессов:

$$B_k = \int_0^{\infty} G_x(f) df / G_x(f)_{\max} = \int_0^{\infty} S_x(f) df / S_x(f)_{\max} = K_x(0) / S_x(f)_{\max}. \quad (17.2.18')$$

Соотношение неопределенности связывает эффективную ширину спектра B_k с эффективным интервалом ковариации T_k . Для его определения найдем произведение $B_k T_k$ случайного процесса с использованием формул (17.1.7) и (17.2.18'):

$$B_k T_k = 2 \int_0^{\infty} |K_x(\tau)| d\tau / S_x(f)_{\max}. \quad (17.2.19)$$

Оценка этого произведения и приводит к соотношению неопределенности:

$$B_k T_k \geq 1/2. \quad (17.2.20)$$

Следовательно, с уменьшением эффективной ширины спектра увеличивается эффективный интервал ковариации случайного процесса, и наоборот.

Взаимные спектральные функции. Статистическая связь двух случайных процессов $X(t)$ и $Y(t)$ оценивается по функциям взаимной ковариации $K_{xy}(\tau)$ или $K_{yx}(\tau)$. Функции взаимной ковариации в общем случае являются произвольными и соответственно функции

взаимного спектра представляют собой комплексные выражения:

$$S_{xy}(\omega_i) = (1/T) \int_0^T K_{xy}(\tau) \exp(-j\omega_i\tau) d\tau, \quad (17.2.21)$$

при этом:

$$S_{xy}(-\omega) = S_{xy}^*(\omega) = S_{yx}(\omega).$$

Квадратурным аналогом нормированной взаимной ковариационной функции или функции коэффициентов ковариации двух процессов (17.1.11) в спектральной области является *функция когерентности*, которая определяется выражением:

$$\gamma_{xy}^2(\omega) = |S_{xy}(\omega)|^2 / (S_x(\omega)S_y(\omega)), \quad (17.2.22)$$

и для любых ω удовлетворяет неравенствам

$$0 \leq \gamma_{xy}^2(\omega) \leq 1. \quad (17.2.23)$$

Функция когерентности обычно используется при анализе линейных систем преобразования входной функции $X(t)$ в выходную функцию $Y(t)$ (рассмотрено ниже).

Теорема Винера-Хинчина. Рассмотрим сигнал $q(t)$, представляющий собой одну реализацию случайного стационарного эргодического процесса длительностью T . Для сигнала $q(t)$ может быть определен спектр $Q(\omega)$. Если сдвинуть на τ реализацию процесса, то получим спектр $Q(\omega)\exp(j\omega\tau)$. Для вещественных сигналов $Q(\omega) = Q^*(\omega)$ равенство Парсеваля по энергии взаимодействия двух сигналов

$$\int_{-\infty}^{\infty} x(t) y^*(t) dt = \int_{-\infty}^{\infty} X(f) Y^*(f) df.$$

может быть записано в следующей форме:

$$\int_{-\infty}^{\infty} q(t)q(t+\tau) dt = (1/2\pi) \int_{-\infty}^{\infty} Q(\omega)Q^*(\omega) \exp(j\omega\tau) d\omega.$$

Поделим обе части данного равенства на T и перейдем к пределу при $T \Rightarrow \infty$, при этом в его левой части мы увидим выражение для функции корреляции, а в правой части - преобразование Фурье спектра мощности сигнала:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} q(t)q(t+\tau) dt = \lim_{T \rightarrow \infty} \frac{1}{2\pi T} \int_{-\infty}^{\infty} |Q(\omega)|^2 \exp(j\omega\tau) d\omega,$$

$$R(\tau) = (1/2\pi) \int_{-\infty}^{\infty} W(\omega) \exp(j\omega\tau) d\omega. \quad (17.2.24)$$

Отсюда следует, что корреляционная функция случайного стационарного эргодического процесса представляет собой обратное преобразование Фурье его спектра мощности, и наоборот:

$$W(\omega) = \int_{-\infty}^{\infty} R(\tau) \exp(-j\omega\tau) d\tau. \quad (17.2.25)$$

В этом состоит суть теоремы Винера-Хинчина. Функции $W(\omega)$ и $R(\tau)$ являются четными, а соответственно в тригонометрической форме:

$$R(\tau) = 2 \int_0^{\infty} W(f) \cos(2\pi f\tau) df, \quad W(f) = 2 \int_0^{\infty} R(\tau) \cos(2\pi f\tau) d\tau.$$

Так как ковариационные функции стационарных процессов являются частным случаем корреляционных функций, то эти выражения действительны и для ФАК, а следовательно спектральные функции случайных процессов, рассмотренные выше как преобразования Фурье ковариационных функций, являются спектрами мощности флюктуирующей составляющей процессов. С этих позиций дисперсия случайных процессов представляет собой среднюю мощность его флюктуаций

$$K(\tau=0) = \sigma^2 = (1/2\pi) \int_{-\infty}^{\infty} W(\omega) d\omega,$$

т.е., равна суммарной мощности всех его частотных составляющих процессов.

В заключение данного раздела отметим, что спектральные плотности процессов и спектры плотности мощности, это одно и то же понятие. Оба термина используются достаточно широко в научно-технической литературе. Учитывая то обстоятельство, что понятие мощности по своему смыслу больше связано с энергетическими понятиями, а понятие спектральной плотности - с анализом сигналов и систем, при дальнейшем рассмотрении случайных сигналов и процессов будем использовать, в основном, понятие спектральной плотности или (для дискретных величин) спектров случайных сигналов и процессов.

17.3. ПРЕОБРАЗОВАНИЯ СЛУЧАЙНЫХ ФУНКЦИЙ [1, 26, 27].

Системы преобразования случайных функций. Пусть имеется система преобразования с одним входом, на который поступает (подается) входная случайная функция $X(t)$ - функция *воздействия* или *возбуждения*, и с одним выходом, с которого снимается выходная функция $Z(t)$ - *отклик* или *выходная реакция* системы. Система осуществляет преобразование $X(t) \Rightarrow Z(t)$ и описывается определенным *системным оператором* трансформации T - функцией, алгоритмом, набором правил преобразования входного сигнала в выходной. Символическое обозначение операции преобразования:

$$Z(t) = T[X(t)].$$

При выполнении преобразования конкретных реализаций случайного процесса может использоваться также более информативное символическое отображение операции преобразования:

$$z(t) = h(\tau) * x(t-\tau),$$

где $h(\tau)$ - математическая функция импульсного отклика системы на единичное входное воздействие. Последнее выражение, по существу, представляет собой краткую запись операции свертки, которой реализуется линейное преобразование. В интегральной форме:

$$z(t) = \int_0^{\infty} h(\tau) \cdot x(t-\tau) d\tau.$$

Импульсный отклик определяет соответствующую частотную передаточную характеристику системы: $h(\tau) \Leftrightarrow H(\omega)$.

Для неслучайных (детерминированных) входных сигналов соотношение между выходными и входными сигналами всегда однозначно задается системным оператором. В случае реализации на входе системы случайного входного процесса (случайного сигнала) тоже существует однозначное соответствие процессов на выходе и входе системы, однако при этом одновременно происходит изменение статистических характеристик выходного сигнала (математического ожидания, дисперсии, ковариационной функции и пр.).

Линейные и нелинейные системы составляют два основных класса систем обработки сигналов. Термин линейности означает, что система преобразования сигналов должна иметь произвольную, но в обязательном порядке линейную связь между входным сигналом (возбуждением) и выходным сигналом (откликом). В нелинейных системах связь между входным и выходным сигналом определяется произвольным нелинейным законом.

Основные системные операции линейных систем, из которых могут быть сформированы любые линейные операторы преобразования, это операции скалярного умножения, сдвига и сложения сигналов:

$$s(t) = c \times a(t), \quad s(t) = a(t-\Delta t), \quad s(t) = a(t)+b(t).$$

Для нелинейных систем выделим важный тип безинерционных операций нелинейной трансформации сигнала, результаты которой зависят только от его входных значений. К ним относятся, например, операции квадратирования и логарифмирования сигнала:

$$y(t) = [s(t)]^2, \quad y(t) = \log[s(t)].$$

Система считается линейной, если ее реакция на входные сигналы аддитивна (вы-

полняется принцип суперпозиции сигналов) и однородна (выполняется принцип пропорционального подобия).

Принцип *аддитивности* требует, чтобы реакция на сумму двух входных сигналов была равна сумме реакций на каждый сигнал в отдельности:

$$T[a(t)+b(t)] = T[a(t)]+T[b(t)].$$

Принцип *однородности* или пропорционального подобия требует сохранения однозначности масштаба преобразования при любой амплитуде входного сигнала:

$$T[c \times a(t)] = c \times T[a(t)].$$

Другими словами, отклик линейной системы на взвешенную сумму входных сигналов должен быть равен взвешенной сумме откликов на отдельные входные сигналы независимо от их количества и для любых весовых коэффициентов, в том числе комплексных.

Примеры линейных операций преобразования:

1. Умножение на заданную функцию: $Z(t) = f(t) \cdot Y(t)$.

2. Дифференцирование: $Z(t) = dX(t)/dt$.

3. Интегрирование: $Z(t) = \int_0^t X(v) dv$.

Линейные системы могут быть неоднородными, если они осуществляют какое-либо линейное однородное преобразование с прибавлением (вычитанием) заданной функции, т.е. операцию вида $Z(t) = T[X(t)] = T_o[X(t)] + f(t)$.

Двухвходовая система описывается системным оператором T , который связывает два входных воздействия, соответственно $X(t)$ и $Y(t)$, с выходной реакцией $Z(t)$. Система считается линейной, если принципы аддитивности и однородности выполняются для обоих входов, т.е.:

$$Z(t) = T[c \cdot X(t), c \cdot Y(t)] = c \cdot T[X(t), Y(t)],$$

$$Z(t) = T[X_1(t)+X_2(t), Y_1(t)+Y_2(t)] = T[X_1(t), Y_1(t)]+T[X_2(t), Y_2(t)].$$

Двухвходовая система может применяться, например, для суммирования двух случайных процессов с разными коэффициентами усиления их значений.

При выполнении линейного преобразования $Z(t) = T[X(t)]$ обычно ставится задача определения характеристик распределения $Z(t)$ по известным характеристикам $X(t)$.

Математическое ожидание выходного сигнала:

$$m_z(t) = M\{Z(t)\} = M\{T[X(t)]\}.$$

Из теории линейных систем: *Линейный оператор можно выносить за знак математического ожидания.* Отсюда следует:

$$m_z(t) = T[M\{X(t)\}] = T[m_x(t)], \quad (17.3.1)$$

т.е. для определения функции математического ожидания выходного сигнала $Z(t)$ достаточно выполнить преобразование тем же системным оператором функции математического ожидания входного сигнала $X(t)$:

$$m_z(t) = h(\tau) * m_x(t-\tau). \quad (17.3.2)$$

Корреляционная функция выходного сигнала:

$$R_z(t_1, t_2) = M\{Z(t_1)Z(t_2)\} = M\{T_1[X(t_1)]T_2[X(t_2)]\},$$

где T_1 и T_2 - один и тот же оператор T по переменным соответственно t_1 и t_2 , что позволяет вынести его за знак математического ожидания, сохраняя переменные:

$$R_z(t_1, t_2) = T_1 T_2 [M\{X(t_1)X(t_2)\}] = T_1 T_2 [R_x(t_1, t_2)], \quad (17.3.3)$$

т.е. при известной функции корреляции входного сигнала функция корреляции выходного сигнала находится двойным преобразованием тем же оператором по двум аргументам.

При определении функции $R_z(\tau)$ следует учесть порядок преобразования. Для произведения выходных сигналов $z(t)$ и $z(t+\tau)$ линейной системы можно записать:

$$z(t) \cdot z(t+\tau) = \int_0^\infty \int_0^\infty h(\alpha)h(\beta) x(t-\alpha) x(t+\tau-\beta) d\alpha d\beta.$$

Если взять математические ожидания от обеих частей этого равенства, то, с учетом соотношения в подинтегральном выражении

$$M\{x(t-\alpha) x(t+\tau-\beta)\} = -R_x(t-\alpha-t+\tau+\beta) = R_x(\tau+\alpha-\beta),$$

получим:

$$R_z(\tau) = \int_0^{\infty} \int_0^{\infty} h(\alpha)h(\beta) R_x(\tau+\alpha-\beta) d\alpha d\beta \equiv R_x(\tau) * h(\tau+\alpha) * h(\tau-\beta). \quad (9.3.4)$$

Таким образом, функция корреляции выходного сигнала равна функции корреляции входного сигнала, свернутой дважды, в прямом и обратном направлении, с импульсным откликом системы, что сохраняет четность корреляционной функции выходного сигнала. Аналогичное заключение действительно и для ковариационных функций.

Заметим, что для свертки импульсных откликов, производя замену $\tau-\beta = t$, мы имеем равенство:

$$h(\tau+\alpha) * h(\tau-\beta) = h(t+\alpha+\beta) * h(t) = h(t) * h(t+\gamma) = R_h(t),$$

где $R_h(t)$ - функция корреляции импульсного отклика системы. Отсюда:

$$R_z(\tau) = R_x(\tau) * R_h(\tau). \quad (17.3.5)$$

т.е. функция корреляции выходного сигнала равна свертке функции корреляции входного сигнала с функцией корреляции импульсного отклика системы. Это означает появление в случайном сигнале на выходе системы определенной ковариационной зависимости, вызванной инерционностью системы, причем радиус ковариации выходного сигнала обратно пропорционален верхней частоте, пропускаемой системой.

Функции взаимной корреляции входного и выходного сигналов определяются аналогично:

$$R_{zx}(t_1, t_2) = T_1[R_x(t_1, t_2)], \quad R_{xz}(t_1, t_2) = T_2[R_x(t_1, t_2)]. \quad (17.3.6)$$

Для функции R_{xz} входного и выходного сигналов имеем:

$$\int_0^{\infty} x(t) \cdot z(t+\tau) dt = \int_0^{\infty} \int_0^{\infty} h(\alpha) x(t) x(t+\tau-\alpha) d\alpha dt.$$

$$R_{xz}(\tau) = \int_0^{\infty} h(\alpha) R_x(\tau-\alpha) d\alpha \equiv R_x(\tau) * h(\tau-\alpha). \quad (9.3.7)$$

т.е. функция взаимной корреляции входного и выходного сигналов равна свертке функции корреляции входного сигнала с функцией импульсного отклика системы.

Другая взаимно корреляционная функция R_{yx} может быть получена из соотношения:

$$R_{zx}(\tau) = R_{xz}(-\tau) \equiv R_x(\tau) * h(\tau+\alpha). \quad (17.3.8)$$

Отметим, что для статистически независимых случайных величин при одностороннем импульсном отклике $h(\tau) = 0$ при $\tau < 0$ функция $R_{xz}(\tau)$ также является односторонней и равна 0 при $\tau < 0$, а функция R_{zx} соответственно равна 0 при $\tau > 0$.

Спектральные соотношения, которые характеризуют систему в целом по отношению к преобразованию случайных сигналов, это соотношения спектральных плотностей случайных сигналов (спектров мощности) на входе и выходе.

Применяя преобразование Фурье к выражениям (17.3.5), для спектра мощности выходного сигнала получаем:

$$S_z(f) = S_x(f) |H(f)|^2. \quad (17.3.9)$$

Спектр мощности случайного сигнала на выходе системы равен спектру мощности входного сигнала, умноженному на квадрат модуля частотной характеристики фильтра. С учетом четности ковариационных функций спектр мощности выходного сигнала также является четной действительной функцией и содержит только амплитудную характеристику системы.

Аналогично, для взаимного спектра мощности сигналов на основе выражений (17.3.7-8) имеем:

$$S_{xz}(f) = S_x(f) H(f). \quad (17.3.10)$$

$$S_{zx}(f) = S_x(f) H(-f). \quad (17.3.10')$$

Взаимный спектр сигналов при одностороннем импульсном отклике является комплексным и содержит как амплитудную, так и фазовую характеристику системы.

Отметим, что с использованием выражения (17.3.10) можно производить определение частотной характеристики и импульсного отклика системы:

$$H(f) = S_{xz}/S_x \Leftrightarrow h(t).$$

Дисперсия выходного сигнала может быть определена с использованием формул (17.3.4,9) по функциям ковариации:

$$\sigma_z^2 = K_z(0) = \int_{-\infty}^{\infty} S_x(f) |H(f)|^2 df \equiv K_x(0) \int_0^{\infty} h^2(t) dt = \sigma_x^2 \int_0^{\infty} h^2(t) dt, \quad (17.3.11)$$

Если сигнал нецентрированный и значение дисперсии входного сигнала неизвестно, то по аналогичным формулам вычисляется сначала *средний квадрат* выходного сигнала или так называемая *средняя мощность сигнала*:

$$\bar{z}^2 = \overline{z^2(t)} = R_z(0) \equiv \bar{x}^2 \int_0^{\infty} h^2(t) dt \equiv \int_0^{\infty} S_x(f) |H(f)|^2 df. \quad (17.3.12)$$

Средняя мощность выходного сигнала равна средней мощности входного сигнала, умноженной на квадрат площади импульсной реакции системы (для цифровых систем - сумму квадратов коэффициентов импульсного отклика). Для центрированных случайных сигналов средняя мощность равна дисперсии сигналов. Для нецентрированных выходных сигналов:

$$\sigma_z^2 = \bar{z}^2 - \bar{z}^2 \equiv (\bar{x}^2 - \bar{x}^2) \int_0^{\infty} h^2(t) dt. \quad (17.3.13)$$

Функция когерентности дает оценку точности принятой линейной модели системы. Когерентность входного и выходного сигналов системы оценивается по формуле:

$$\gamma_{xz}^2(f) = |S_{xz}(f)|^2 / [S_x(f) \cdot S_z(f)]. \quad (17.3.14)$$

Если функции $S_x(f)$ и $S_z(f)$ отличны от нуля и не содержат дельта-функций, то для всех f значения функции когерентности заключены в интервале:

$$0 \leq \gamma_{xz}^2(f) \leq 1.$$

Для исключения дельта-функций на нулевой частоте определение функции когерентности производится по центрированным сигналам. Для линейных систем с постоянными параметрами функция когерентности равна 1, в чем нетрудно убедиться, если в формулу (17.3.14) подставить выражения S_{xz} и S_z , определенные через S_x в формулах (17.3.9-10). Для совершенно не связанных сигналов функция когерентности равна нулю. Промежуточные между 0 и 1 значения могут соответствовать трем ситуациям:

1. Система осуществляет преобразование $x(t) \Rightarrow z(t)$, но в измерениях этих сигналов или одного из них присутствует внешний шум. Так, например в сигналах, зарегистрированных с ограничением по разрядности, появляется шум квантования (округления значений).

2. Система не является строго линейной. Это может наблюдаться, например, при определенном ограничении по разрядности вычислений в цифровых системах, при накоплении ошибки в рекурсивных системах и т.п.

3. Выходной сигнал $z(t)$ помимо $x(t)$ зависит еще от каких-то входных или внутренних системных процессов.

Величина $1 - \gamma_{xz}^2(f)$ задает долю среднего квадрата сигнала $z(t)$ на частоте f , не связанную с сигналом $x(t)$.

Аналогично можно вычислить функцию когерентности двух реализаций $x(t)$ и $y(t)$. Значения функции будут указывать на степень линейной зависимости одной реализации от другой, хотя это и не означает обязательности наличия какой-либо причинно-следственной

связи между реализациями. Функция когерентности γ_{xy} сохраняется при точных однотипных линейных преобразованиях функций $x(t)$ и $y(t)$, что позволяет производить ее определение не измеряя самих величин $x(t)$ и $y(t)$.

Использование функций когерентности в практических методах анализа случайных данных подробно рассмотрено в работе /л4/.

Преобразования случайных функций.

Сложение случайных функций. При сложении случайных функций, в общем случае, с произвольными постоянными коэффициентами a и b , и образовании случайной функции суммы

$$Z(t) = a \cdot X(t) + b \cdot Y(t)$$

функция математического ожидания процесса $Z(t)$:

$$m_z(t) = M\{Z(t)\} = M\{aX(t) + bY(t)\} = a \cdot M\{X(t)\} + b \cdot M\{Y(t)\} = a \cdot m_x(t) + b \cdot m_y(t). \quad (17.3.15)$$

Корреляционная функция суммы вычисляется аналогично и равна:

$$\begin{aligned} R_z(t_1, t_2) &= M\{Z(t_1) \cdot Z(t_2)\} = M\{[aX(t_1) + bY(t_1)][aX(t_2) + bY(t_2)]\} = \\ &= M\{a^2 X(t_1)X(t_2) + b^2 Y(t_1)Y(t_2) + ab[X(t_1)Y(t_2) + Y(t_1)X(t_2)]\} = \\ &= a^2 R_x(t_1, t_2) + b^2 R_y(t_1, t_2) + ab \cdot [R_{xy}(t_1, t_2) + R_{yx}(t_1, t_2)]. \end{aligned} \quad (17.2.16)$$

Для некоррелированных функций $X(t)$ и $Y(t)$ функции взаимной корреляции R_{xy} и R_{yx} обнуляются. Аналогичную форму записи имеют и ковариационные функции (как частный случай корреляционных функций при центрировании случайных процессов). Выражения легко обобщаются на сумму любого числа случайных функций. В частности, для корреляционной функции стационарной случайной функции $Z(t) = \sum_i a_i X_i(t)$ при $t_2 - t_1 = \tau$ имеем:

$$R_z(\tau) = \sum_i a_i^2 R_{x_i}(\tau) + \sum_i \sum_{j, j \neq i} a_i a_j R_{x_i x_j}(\tau). \quad (17.3.16')$$

При сложении случайной функции $X(t)$ с неслучайной функцией $y(t)$ математическое ожидание и корреляционная функция суммы $Z(t) = X(t) + y(t)$ равны:

$$m_z(t) = m_x(t) + y(t), \quad R_z(t_1, t_2) = R_x(t_1, t_2). \quad (17.3.17)$$

При сложении случайной функции $X(t)$ с некоррелированной случайной величиной Y математическое ожидание и корреляционная функция суммы $Z(t) = X(t) + Y$:

$$m_z(t) = m_x(t) + m_y, \quad R_z(t_1, t_2) = R_x(t_1, t_2) + D_y. \quad (17.3.18)$$

Произведение случайной и неслучайной функций $X(t)$ и $f(t)$. Математическое ожидание и корреляционная функция выходного сигнала:

$$m_z(t) = M\{Z(t)\} = M\{f(t) \cdot X(t)\} = f(t) \cdot M\{X(t)\} = f(t) \cdot m_x(t). \quad (17.3.19)$$

$$\begin{aligned} R_z(t_1, t_2) &= M\{f(t_1)X(t_1) f(t_2)X(t_2)\} = f(t_1)f(t_2)M\{X(t_1)X(t_2)\} = \\ &= f(t_1)f(t_2) \cdot R_x(t_1, t_2). \end{aligned} \quad (17.3.20)$$

Если $f(t) = \text{const} = C$ и $Z(t) = C \cdot X(t)$, то соответственно имеем:

$$m_z(t) = C \cdot m_x(t), \quad R_z(t_1, t_2) = C^2 \cdot R_x(t_1, t_2). \quad (17.3.21)$$

Производная от случайной функции $Z(t) = dX(t)/dt$. Если функция $X(t)$ является непрерывной и дифференцируемой, то математическое ожидание производной:

$$m_z(t) = M\{Z(t)\} = M\{dX(t)/dt\} = d(M\{X(t)\})/dt = dm_x(t)/dt, \quad (17.3.22)$$

т.е. математическое ожидание производной от случайной функции равно производной от ее математического ожидания. Для корреляционной функции имеем:

$$R_z(t_1, t_2) = M\{(dX(t_1)/dt_1)(dX(t_2)/dt_2)\} = \frac{\partial^2}{\partial t_1 \partial t_2} M\{X(t_1)X(t_2)\} = \frac{\partial^2}{\partial t_1 \partial t_2} R_x(t_1, t_2), \quad (17.3.23)$$

т.е. корреляционная функция производной случайной функции равна второй смешанной частной производной от корреляционной функции исходной случайной функции.

Интеграл от случайной функции $Z(t) = \int_0^t X(v)dv$.

$$m_z(t) = M\{Z(t)\} = M\left\{\int_0^t X(v)dv\right\} = \int_0^t M\{X(v)\}dv = \int_0^t m_x(v)dv, \quad (17.3.24)$$

т.е. математическое ожидание интеграла от случайной функции равно интегралу от ее математического ожидания. Для корреляционной функции имеем:

$$\begin{aligned} R_z(t_1, t_2) &= M\left\{\int_0^{t_1} X(t_1)dt_1 \int_0^{t_2} X(t_2)dt_2\right\} = M\left\{\int_0^{t_1} \int_0^{t_2} X(t_1)X(t_2)dt_1 dt_2\right\} = \\ &= \int_0^{t_1} \int_0^{t_2} M\{X(t_1)X(t_2)\}dt_1 dt_2 = \int_0^{t_1} \int_0^{t_2} R_x(t_1, t_2)dt_1 dt_2, \end{aligned} \quad (17.3.25)$$

т.е. корреляционная функция интеграла от случайной функции равна двойному интегралу от корреляционной функции исходной случайной функции.

Преобразования стационарных случайных функций выполняются по вышеприведенным формулам и приводят к следующим результатам (вместо корреляционных функций приводятся ковариационные функции, которые обычно используются на практике).

Математическое ожидание выходного сигнала $Z(t)$ входной стационарной случайной функции $X(t)$ по (17.3.2):

$$m_z = h(\tau) * m_x = m_x \int_{-\infty}^{\infty} h(\tau) d\tau, \quad (17.3.26)$$

Отсюда следует, что математическое ожидание выходных сигналов системы равно математическому ожиданию входных сигналов, умноженному на площадь (или сумму коэффициентов) импульсного отклика системы, т.е. на коэффициент усиления системой постоянной составляющей. Если система не пропускает постоянную составляющую сигналов (площадь или сумма коэффициентов импульсного отклика системы равна нулю), то случайный выходной сигнал всегда будет иметь нулевое математическое ожидание.

Сумма двух стационарных случайных функций $X(t)$ и $Y(t)$ дает стационарную случайную функцию $Z(t)$, при этом:

$$m_z = m_x + m_y, \quad D_z = D_x + D_y + 2K_{xy}(0). \quad (17.3.27)$$

$$K_z(t_1, t_2) = K_z(\tau) = K_x(\tau) + K_y(\tau) + K_{xy}(\tau) + K_{yx}(\tau). \quad (17.3.28)$$

Сумма стационарной случайной и неслучайной функций $X(t)$ и $y(t)$ нестационарна по математическому ожиданию:

$$m_z(t) = m_x + y(t), \quad K_z(\tau) = K_x(\tau). \quad (17.3.29)$$

Произведение стационарной случайной и неслучайной функций $X(t)$ и $y(t)$ - нестационарная случайная функция, так как:

$$m_z(t) = y(t) \cdot m_x, \quad D_z(t) = y^2(t) \cdot D_x. \quad (17.3.30)$$

$$K_z(t, \tau) = y(t)y(t+\tau)K_x(\tau). \quad (17.3.31)$$

Производная от стационарной случайной функции - стационарная случайная функция с математическим ожиданием $m_z = 0$ и ковариационными функциями:

$$K_z(t_1, t_2) = \frac{\partial^2}{\partial t_1 \partial t_2} K_x(t_1 - t_2) = -\frac{d^2}{d\tau^2} K_x(\tau) = K_z(\tau). \quad (17.3.32)$$

$$K_{zx}(\tau) = d(K_x(\tau))/d\tau, \quad K_{xz}(\tau) = -d(K_x(\tau))/d\tau. \quad (9.3.33)$$

Из выражения (17.3.32) следует также, что для дифференцируемости $X(t)$ необходимо, чтобы ее ковариационная функция была дважды дифференцируемой по τ .

Интеграл от стационарной случайной функции - нестационарная случайная функция с математическим ожиданием $m_z(t) = \int_0^t m_x(t)dt$ и функцией ковариации:

$$K_z(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} K_x(u_1 - u_2) du_1 du_2. \quad (17.3.34)$$

17.4. МОДЕЛИ СЛУЧАЙНЫХ СИГНАЛОВ И ПОМЕХ [2, 28].

Наиболее распространенными моделями случайных сигналов и помех являются теле-

графный сигнал, белый шум, гауссовский случайный процесс, гауссовский шум.

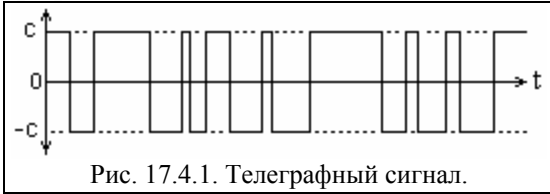


Рис. 17.4.1. Телеграфный сигнал.

туд c и $-c$, причем перемены знака внутри любого интервала $(t, t+\tau)$ происходят с интенсивностью α в случайные моменты времени и не зависят от процессов в смежных временных интервалах. Если считать случайной величиной телеграфного сигнала значение n - количество перемен знака внутри интервала τ , то распределение вероятностей значений n будет описываться законом Пуассона:

$$P(n) = (\alpha|\tau|)^n \exp(-\alpha|\tau|)/n! \quad (17.4.1)$$

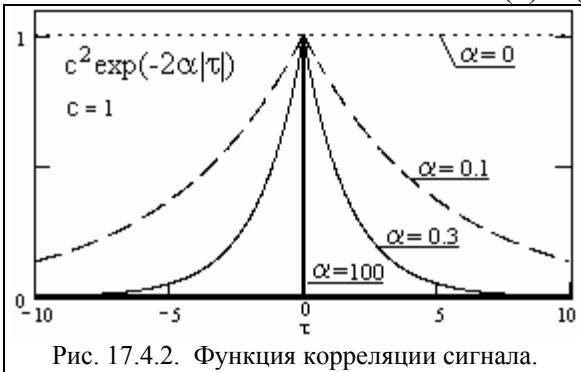


Рис. 17.4.2. Функция корреляции сигнала.

При вычислении корреляционной функции телеграфного сигнала каждое отдельное произведение $x_k(t)x_k(t+\tau)$ равно либо c^2 , либо $-c^2$ в зависимости от совпадения или несовпадения знаков $x_k(t)$ и $x_k(t+\tau)$, причем вероятность c^2 равна сумме вероятностей $P(0)+P(2)+P(4)+\dots$, а вероятность $-c^2$ определяется соответственно суммой вероятностей $P(1)+P(3)+P(5)+\dots$.

Следовательно:

$$R_x(\tau) = M\{x_k(t)x_k(t+\tau)\} = c^2 \sum_{n=0}^{\infty} (-1)^n P(n) =$$

$$= c^2 \exp(-\alpha|\tau|) \sum_{n=0}^{\infty} (-1)^n (\alpha|\tau|)^n / n! = c^2 \exp(-2\alpha|\tau|). \quad (17.4.2)$$

Параметр α полностью определяет ковариационные и спектральные свойства телеграфного сигнала. При $\alpha \Rightarrow 0$ характеристики сигнала приближаются к характеристикам постоянной составляющей, при $\alpha \Rightarrow \infty$ - к характеристикам белого шума.

Интервал ковариации сигнала:

$$T_k = 2 \int_0^{\infty} (R_x(\tau)/c^2) d\tau = 2/\alpha. \quad (17.4.3)$$

Отсюда следует, что чем больше α , тем меньше время ковариации процесса. При $\alpha \Rightarrow 0$ $T_k \Rightarrow \infty$ и процесс вырождается в детерминированный (стремится к постоянной составляющей). При $\alpha \Rightarrow \infty$ $T_k \Rightarrow 0$ и процесс вырождается в белый шум с некоррелированными отсчетами даже на соседних временных точках.

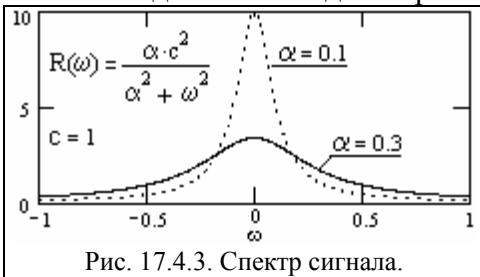


Рис. 17.4.3. Спектр сигнала.

Двусторонняя спектральная плотность сигнала:

$$S_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau) \exp(-j\omega\tau) d\tau = \alpha c^2 / (\alpha^2 + \omega^2). \quad (17.4.4)$$

Односторонняя спектральная плотность:

$$G_x(\omega) = 2 \int_0^{\infty} R_x(\tau) \exp(-j\omega\tau) d\tau = 2\alpha c^2 / (\alpha^2 + \omega^2). \quad (17.4.5)$$

Ширина спектра телеграфного сигнала:

$$B_k = \int_0^{\infty} G_x(\omega) d\omega / G_x(0) \equiv \int_0^{\infty} S_x(\omega) d\omega / S_x(0) = \alpha\pi. \quad (17.4.6)$$

Отсюда следует, что спектр случайного процесса тем шире, чем меньше интервал ковариации процесса.

Белый шум является стационарным случайным процессом $x(t)$ с постоянной спек-

тральной плотностью $G_x(f) = \sigma^2$, равной дисперсии значений $x(t)$. Другими словами, все спектральные составляющие белого шума имеют одинаковую энергию (как белый цвет содержит все цвета видимого спектра).

По своему физическому смыслу спектральная плотность - это мощность процесса, которая приходится на 1 Гц полосы частот. Но тогда идеального белого шума на практике не может существовать, так как для него должно было бы выполняться условие:

$$R_x(0) = \int_0^{\infty} G_x(f) df = (\sigma^2/2) \cdot \delta(0) = \infty, \quad (17.4.7)$$

т.е. мощность белого шума и его дисперсия равны бесконечности, а значения шума не коррелированы для любых $|\tau| \neq 0$, так как корреляционная функция представляет собой идеальный дельта-импульс. Тем не менее многие помехи в радиотехнике, в технике связи и в других отраслях рассматривают как белый шум, если выполняется следующее соотношение между шириной спектров полезных сигналов и шумов

$$B_{\text{к.сигнал}}/B_{\text{к.шум}} \ll 1,$$

и спектральная плотность шумов слабо изменяется в интервале спектра сигнала.

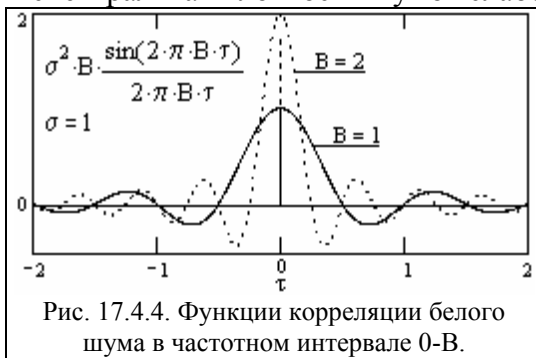


Рис. 17.4.4. Функции корреляции белого шума в частотном интервале 0-B.

Если частотный диапазон спектра, на котором рассматриваются сигналы и помехи, равен 0-B, то спектральная плотность шума задается в виде: $G_x(f) = \sigma^2$, $0 \leq f \leq B$; $G_x(f) = 0$, $f > B$, (17.4.8) при этом корреляционная функция шума определяется выражением:

$$R_x(\tau) = \sigma^2 B \cdot \sin(2\pi B\tau) / 2\pi B\tau. \quad (17.4.9)$$

Эффективная шумовая ширина спектра:

$$B_k = R_x(0)/G_x(f)_{\text{max}} = B. \quad (17.4.10)$$

Эффективное шумовое время ковариации:

$$T_k = 2 \int_0^{\infty} |R_x(\tau)| d\tau / R_x(0). \quad (17.4.11)$$

Реальное шумовое время ковариации целесообразно определить по ширине главного максимума функции $R_x(\tau)$, в котором сосредоточена основная часть энергии шумов, при этом $T_k = 1/B$ и $B_k T_k = 1$, т.е. соотношение неопределенности выполняется.

Как следует из всех этих выражений и наглядно видно на рис. 17.4.4, при ограничении частотного диапазона в шумах появляется определенная ковариация между значениями и чем меньше частотный диапазон шумов, тем больше их радиус ковариации. По существу, ограничение частотного диапазона шумов определенным диапазоном эквивалентно фильтрации белого шума частотным фильтром с соответствующей шириной полосы пропускания, при этом, в полном соответствии с выражением (17.3.7), корреляционная функция импульсного отклика фильтра переносится на шум.

Гауссовский шум возникает при суммировании статистически независимых белых шумов и имеет следующую функцию корреляции:

$$R_x(\tau) = a \exp(-2\pi\sigma^2\tau^2). \quad (17.4.12)$$

Спектральная плотность шумов:

$$S_x(f) = (a/\sigma\sqrt{2\pi}) \exp(-f^2/2\sigma^2), \quad -\infty < f < \infty. \quad (17.4.13)$$

Эффективные шумовые ширина спектра и время ковариации:

$$B_k = \sigma\sqrt{2\pi}/2 = 1.25\sigma, \quad T_k = 1/\sigma\sqrt{2\pi} = 0.4/\sigma. \quad (17.4.14)$$

Соотношение неопределенности превращается в равенство: $B_k T_k = 1/2$.

Гауссовские случайные процессы преобладают в практических задачах. Случайный процесс $x(t)$ называется гауссовским, если для любого набора фиксированных моментов времени t_n случайные величины $x(t_n)$ подчиняются многомерному нормальному распределе-

нию. Плотность вероятностей мгновенных значений $x(t)$ эргодического гауссовского процесса определяется выражением:

$$p(x) = (\sigma_x \sqrt{2\pi})^{-1} \exp(-(x-m_x)^2/2\sigma^2). \quad (17.4.15)$$

Среднее значение и его оценка по достаточно большому интервалу T :

$$m_x = \int_{-\infty}^{\infty} xp(x) dx, \quad m_x \approx (1/T) \int_0^T x(t) dt.$$

При нулевом среднем (или при центрировании функции $x(t)$ для упрощения расчетов) дисперсия не зависит от t и равна:

$$\sigma_x^2 = \int_{-\infty}^{\infty} x^2 p(x) dx.$$

Оценка дисперсии при больших T :

$$\sigma_x^2 \approx (1/T) \int_0^T x^2(t) dt = \int_{-\infty}^{\infty} S_x(f) df = 2 \int_0^{\infty} S_x(f) df = \int_0^{\infty} G_x(f) df. \quad (17.4.16)$$

Следовательно, плотность вероятностей гауссовского процесса полностью характеризуется спектральной плотностью, по которой можно определить значение дисперсии процесса. На вид спектральных плотностей и соответствующих им ковариационных функций никаких ограничений не накладывается.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988.- 448 с.
2. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. - М.: Мир, 1989. - 540 с.
25. Сергиенко А.Б. Цифровая обработка сигналов. - СПб.: Питер, 2003. - 608 с.
26. Вероятностные методы в вычислительной технике: Учебное пособие для вузов./ А.В.Крайников и др. - М.: Высшая школа, 1986. - 312 с.
26. Вероятностные методы в вычислительной технике: Учебное пособие для вузов./ А.В.Крайников и др. - М.: Высшая школа, 1986. - 312 с.
27. Гурский Е.И. Теория вероятностей с элементами математической статистики: Учебное пособие для вузов. - М.: Высшая школа, 1971.- 328 с.
28. Игнатов В.А. Теория информации и передачи сигналов. - М.: Советское радио, 1979.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

Тема 18: МНОГОМЕРНЫЕ СИГНАЛЫ И СИСТЕМЫ.

Когда ты смотришь в бездну, бездна смотрит в тебя.

Фридрих Ницше. Немецкий философ-моралист, XIX в.

Человек и бездна – две бесконечномерные системы в разных функциональных пространствах с одной точкой пересечения. И лучше держаться от этой точки подальше.

Эрик Трубов. Русский геофизик-оптимист, XX в.

Содержание: Введение. 18.1. Двумерные и многомерные сигналы. 18.2. Двумерные системы. Импульсный отклик систем. Двумерная свертка. 18.3. Частотные характеристики сигналов и систем. Частотный отклик системы. Импульсный отклик системы. Свойства двумерного преобразования Фурье. 18.4. Дискретизация двумерных сигналов. Прямоугольный растр дискретизации. Дискретные преобразования Фурье. Интерполяционный ряд восстановления двумерного сигнала. Произвольный растр дискретизации. Интерполяция дискретных сигналов. Прямоугольный и гексагональный растры дискретизации. 18.5. Частотный анализ многомерных сигналов. Периодические последовательности. Конечные последовательности. Многомерные последовательности. Литература.

ВВЕДЕНИЕ.

Обработка многомерных сигналов, используя в частных случаях методы обработки одномерных сигналов, имеет и существенные особенности. Это объясняется тремя факторами. Во-первых, математические методы описания многомерных систем далеки от совершенства и завершенности. Во-вторых, при решении многомерных задач используется значительно больший объем данных. И в третьих, многомерные системы обладают большим числом степеней свободы и, соответственно, значительно большей гибкостью. Так, например, при дискретизации информации в одномерном случае устанавливается только частота отсчетов, а в многомерном не только частота, но и форма раstra дискретизации. С другой стороны, многомерные полиномы разлагаются на множители только в частном случае, а, следовательно, многие одномерные методы не обобщаются на случай многомерных задач.

Ниже будут рассматриваться сигналы и системы с размерностью два и более, при этом основное внимание будет уделяться двумерным задачам, имеющим широкое распространение в геофизической практике. Повышение размерности выше двух не приводит к качественным отличиям от двумерных случаев, кроме повышения сложности вычислений.

Многомерная информация в своем абсолютном большинстве, это дискретная информация в цифровой форме – многомерные массивы данных. Многомерные непрерывные функции используются только в чисто теоретических исследованиях. Даже двумерных данных, непрерывных (аналоговых) по обоим аргументам практически не существует. С учетом этого ниже рассматриваются, в основном, многомерные сигналы в дискретной форме.

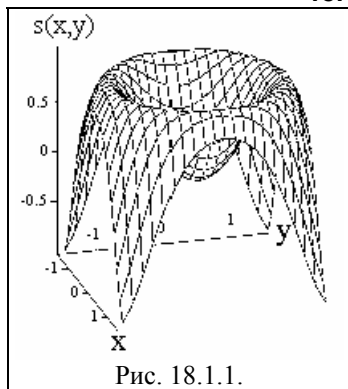
18.1. ДВУМЕРНЫЕ И МНОГОМЕРНЫЕ СИГНАЛЫ [9].

Рис. 18.1.1.

Понятие многомерного сигнала. Многомерные сигналы представляют собой функции P независимых переменных при $P > 1$. В общем случае, сигнал может быть непрерывным, дискретным или смешанным. Понятия непрерывности и дискретности аналогичны одномерным сигналам. Что касается смешанного сигнала, то это многомерный сигнал, который описывается функцией некоторого количества непрерывных и некоторого количества дискретных переменных. Пример смешанного двумерного сигнала: ансамбль непрерывных сигналов, изменяющихся во времени (t - вторая переменная), снимаемых с набора сейсмических приемников сейсмотрассы (номера датчиков - первая переменная).

В общем случае, двумерный непрерывный сигнал представляет собой функцию, значения которой зависят от двух независимых переменных (аргументов, координат):

$$s(x,y) = \sin(x^2 + y^2), \quad -\infty < x, y < \infty \quad (18.1.1)$$

График функции (в пределах одного периода) приведен на рис. 18.1.1.

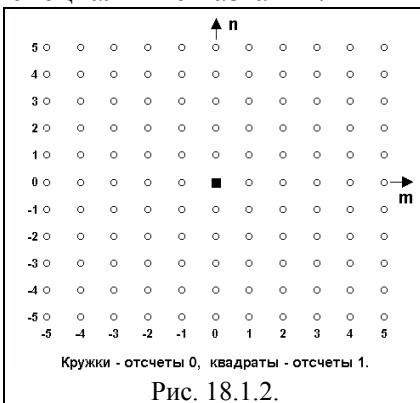
Двумерный дискретный сигнал (цифровой массив) - это функция, определенная на совокупности пар числовых значений координат с определенным шагом дискретизации Δx и Δy . В общем случае, при различной физической размерности аргументов x и y , значения Δx и Δy не равны друг другу:

$$s_{n,m} = s(n\Delta x, m\Delta y), \quad -\infty < n, m < \infty. \quad (18.1.1')$$

Элемент последовательности $s_{n,m}$ представляет собой отсчет двумерной функции s в координатной точке $(x=n\Delta x, y=m\Delta y)$, где значения x и y – независимые переменные (аргументы) функции. Для числовых массивов значения шага дискретизации по аргументам также могут приниматься равными 1 (независимо от размерности) и использоваться аргументация $s(n,m) \equiv s_{n,m}$. Результаты геофизических съемок какого-либо одного геофизического параметра по поверхности земли относятся к двумерным функциям: дискретным - если это отсчеты в отдельных точках по определенной координатной сети (x,y) , или смешанным - если это непрерывная регистрация данных по профилям (например - мощности экспозиционной дозы гамма излучения горных пород при аэросъемке). Но в настоящее время геофизические съемки относятся даже не к двумерным, а к многомерным функциям, так как регистрируется, как правило, сразу несколько физических параметров геологических сред. Так, например, при спектрометрической съемке естественной радиоактивности горных пород регистрируется содержание в горных породах урана, тория и калия, в гравиразведке - трехкоординатный вектор силы тяжести, и т.п. Если на какой-либо площади проведена съемка нескольких видов геофизики, то их результаты также могут рассматриваться в совокупности, как многомерная функция физических параметров данной геологической среды.

По определениям (18.1.1) двумерные функции и сигналы, равно как и многомерные, имеют бесконечную протяженность по координатам. На практике мы всегда имеем дело с конечными координатами наших данных. Учитывая это, будем считать, что значения наших сигналов за пределами определенных координат равны нулю.

Отметим некоторые двумерные последовательности (функции, сигналы), имеющие специальные названия.



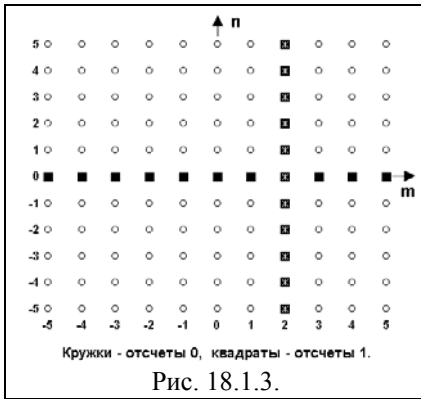
Двумерный единичный импульс $\delta(n\Delta x, m\Delta y) = \delta_{n,m}$ или единичный отсчет:

$$\begin{aligned} \delta_{n,m} &= 1, \quad \text{при } n = m = 0. \\ &= 0, \quad \text{при } n \neq 0, m \neq 0. \end{aligned}$$

$$\delta_{n,m} = \delta_n \delta_m,$$

где δ_n, δ_m - одномерные единичные импульсы (импульсы Кронекера) по координатам n и m . Стилизованное графическое представление двумерного единичного импульса приведено на рис. 18.1.2.

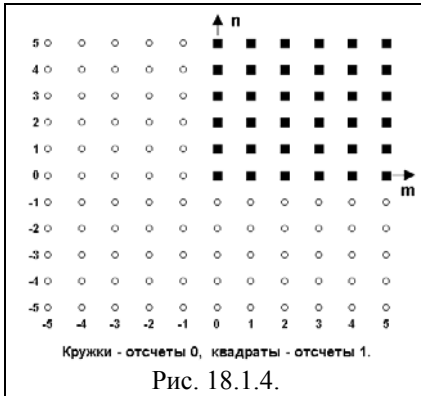
Произвольное расположение двумерного единичного импульса по координатам n_1, m_1 соответственно записывается в виде: $\delta((n-n_1)\Delta x, (m-m_1)\Delta y) = \delta_{n-n_1, m-m_1}$. Попутно напомним, что математическая запись импульса Кронекера обозначает не единичный отсчет, а функцию, определяющую место положения единичного отсчета и нулевые значения по остальным координатам (аргументам).



Двумерный линейный импульс представляет собой последовательность единичных отсчетов по одной координате: $s(n,m) = \delta(n)$ или $s(n,m) = \delta(m)$.

На рис. 18.1.3 приведены два двумерных линейных импульса, первый - по координате $m = 0$: $s(n,m) = \delta(m)$, и второй импульс по координате $n = 2$: $s(n,m) = \delta(n-2)$.

Очевидно, что для P -мерных случаев точно таким же образом могут быть определены P -мерные единичные импульсы, P -мерные линейные импульсы, P -мерные площадные импульсы и т.д., хотя понятие импульса, заимствованное из теории одномерных сигналов, здесь несколько не к месту.



Двумерная единичная ступенька $u(n,m)$, представленная на рис. 18.1.4, определяется выражением:

$$u(n,m) = \begin{cases} 1, & \text{при } n \geq 0 \text{ и } m \geq 0, \\ 0, & \text{в остальных случаях.} \end{cases}$$

$$u(n,m) = u(n) u(m),$$

где $u(j)$ представляют собой единичные ступеньки соответственно по координатам n и m : $u(j)=1$ при $j \geq 0$, $u(j)=0$ при $j < 0$. Двумерная единичная ступенька отлична от нуля в одном квадранте (n,m) - плоскости.

Экспоненциальная последовательность: $s(n,m) = a^n b^m$, $-\infty < n, m < \infty$, где a и b в общем случае комплексные

числа. При $a = \exp(j\omega_1)$, $b = \exp(j\omega_2)$, $|a|=1$, $|b|=1$:

$$s(n,m) = \exp(jn\omega_1 + jm\omega_2) = \cos(n\omega_1 + m\omega_2) + j\sin(n\omega_1 + m\omega_2).$$

Экспоненциальные последовательности, как и в одномерном случае, являются собственными функциями двумерных линейных систем, инвариантных к сдвигу.

Разделимые последовательности. Разделимой называют последовательность, которую можно представить в виде произведения одномерных последовательностей. Так, для двумерной разделимой последовательности:

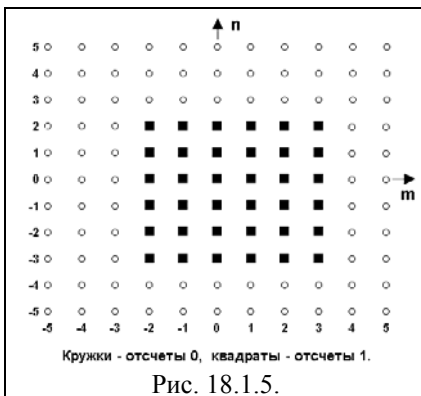
$$s(n,m) = s(n) s(m).$$

Разделение возможно для немногих практических сигналов. Однако любое двумерное множество с конечным числом ненулевых отсчетов разлагается на конечную сумму разделимых последовательностей:

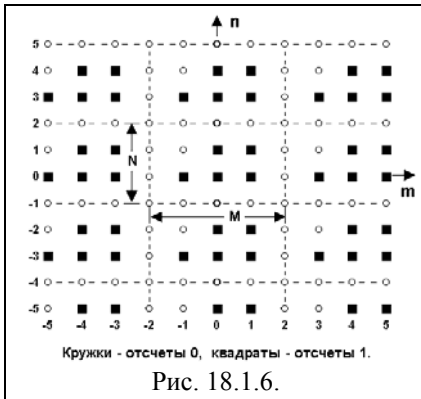
$$s(n,m) = \sum_{i=1}^N s_{i-n}(n) s_{i-m}(m),$$

где N - число ненулевых строк или столбцов массива. В крайнем случае, для этого достаточно выразить $s(n,m)$ в виде суммы отдельных строк:

$$s(n,m) = \sum_{i=1}^N s(n,i) \delta(m-i). \quad (18.1.2)$$



Конечные последовательности. Важным классом сигналов являются последовательности конечной протяженности, для которых сигнал равен нулю вне определенной области, называемой *опорной областью* сигнала. На рис.18.1.5 условно представлена двумерная последовательность конечной протяженности, значения которой отличны от нулевых только внутри ограниченной прямоугольной области $-3 \leq n \leq 2$, $-2 \leq m \leq 2$. Опорная область сигнала может быть произвольной формы и выходить за пределы сигнала, частично включая нулевые отсчеты. Отсчеты за пределами опорной области считаются равными нулю.



Периодические последовательности. Двумерные последовательности могут быть периодическими, регулярно повторяющимися в пространстве. Последовательность, удовлетворяющая условиям:

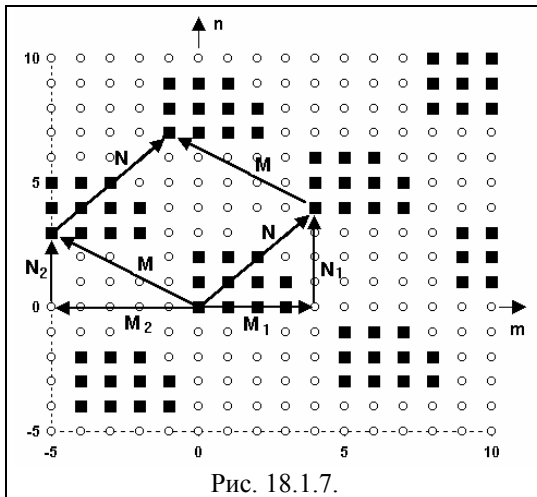
$$\begin{aligned} s(n, m+M) &= s(n, m), \\ s(n+N, m) &= s(n, m), \end{aligned} \quad (18.1.3)$$

обладает периодичностью в двух направлениях, по n и по m . Значения M и N называют интервалами периодичности сигнала соответственно по координатам m и n (*горизонтальными* и *вертикальными* интервалами периодичности). Прямоугольная форма области периода (пример на рис.18.1.6) наиболее

удобна при обработке данных, но не является единственно возможной.

Для двумерных последовательностей условия (18.1.3) могут рассматриваться как частный случай общих условий периодичности:

$$\begin{aligned} s(n+N_1, m+M_1) &= s(n, m), \\ s(n+N_2, m+M_2) &= s(n, m), \\ D &= N_1M_2 - N_2M_1 \neq 0. \end{aligned} \quad (18.1.4)$$



Упорядоченные пары (N_1, M_1) и (N_2, M_2) представляют собой смещения от отсчетов одного периода к соответствующим отсчетам других периодов и могут рассматриваться как векторы \mathbf{N} и \mathbf{M} , которые образуют области периодов в форме параллелограмма. Линейная независимость векторов обеспечивается при ненулевом определителе D , а количество отсчетов в пределах периода равно $|D|$. Пример периодической последовательности с векторами $(4,4)$ и $(3,-5)$ приведен на рис. 18.1.7.

Понятие периодичности можно обобщить на многомерные сигналы. P -мерный сигнал $s(\bar{n})$ будет представлять собой P -мерную периодическую последовательность, если существует P линейно неза-

висимых P -мерных целочисленных \mathbf{N} -векторов периодичности, с которыми выполняется условие:

$$s(\bar{n}) = s(\bar{n} + \bar{N}_i), \quad i = 1, 2, 3, \dots, P. \quad \bar{P}$$

Столбцы векторов \mathbf{N}_i образуют матрицу периодичности \mathbf{N} размером $P \times P$. Векторы периодичности матрицы линейно независимы при наличии у матрицы ненулевого определителя. Абсолютное значение определителя равно числу отсчетов в периоде. Последовательность $s(\bar{n})$ прямоугольно периодична для случаев диагональной матрицы \mathbf{N} . Если функция $s(\bar{n})$ периодична с матрицей периодичности \mathbf{N} , то для любого целочисленного вектора \mathbf{P} имеет место $s(\bar{n} + \mathbf{P} \mathbf{N}) = s(\bar{n})$, и матрица $\mathbf{P} \mathbf{N}$ также будет матрицей периодичности для $s(\bar{n})$. Отсюда следует, что любая многомерная периодическая последовательность имеет не единственную матрицу периодичности.

18.2. ДВУМЕРНЫЕ СИСТЕМЫ.

Системы осуществляют преобразование сигналов. Формализованная система - это оператор (операция) отображения входного сигнала на выходной: $z(x, y) = T[s(x, y)]$.

Базовыми операциями в системах, комбинациями которых осуществляются преобразования, являются операции скалярного умножения, сдвига и сложения:

$$\begin{aligned} z(n, m) &= c s(n, m), \\ z(n, m) &= s(n-N, m-M), \\ z(n, m) &= s(n, m) + u(n, m). \end{aligned}$$

Используя базовые операции, любую двумерную последовательность можно разло-

жить на сумму взвешенных двумерных единичных импульсов:

$$s(n,m) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} s(i,j) \delta(n-i,m-j). \quad (18.2.1)$$

Обобщением скалярного умножения является пространственное маскирование:

$$z(n,m) = c_{n,m} s(n,m). \quad (18.2.2)$$

Правая часть равенства (18.2.2) представляет собой поэлементное произведение входного сигнала на совокупность чисел $c_{n,m}$.

Кроме линейных операций в системах используются также безынерционные нелинейные преобразования с независимым нелинейным воздействием на значения отсчетов входной последовательности. Пример операции - возведение в квадрат:

$$z_{n,m} = (s_{n,m})^2.$$

Линейные системы. Система считается линейной при выполнении двух условий:

1. Пропорциональное изменение входного сигнала вызывает пропорциональное изменение выходного сигнала.
2. Суммарный сигнал двух входных последовательностей дает суммарный сигнал двух соответствующих выходных последовательностей.

Другими словами, если оператор $T[s(x,y)]$ описывает линейную систему и имеет место $z(x,y) = T[s(x,y)]$, $q(x,y) = T[u(x,y)]$, то $T[as(x,y)+bu(x,y)] = az(x,y)+bq(x,y)$. Линейные системы подчиняются принципу суперпозиции сигналов.

В выражении (18.2.1) значения $s(i,j)$ можно рассматривать как скалярные множители для соответствующих единичных импульсов. Применяя оператор преобразования $T[.]$ к левой и правой части (18.2.1), получаем:

$$\begin{aligned} T[s(n,m)] = y(n,m) &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} s(i,j) T[\delta(n-i,m-j)], \\ z(n,m) &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} s(i,j) h_{ij}(n,m), \end{aligned} \quad (18.2.3)$$

где $h_{ij}(n,m)$ - отклик системы в точке (n,m) на единичный импульс в точке (i,j) . Если импульсный отклик $h_{ij}(n,m)$ определен для всех точек (i,j) , то отклик системы на произвольный многомерный сигнал, как и для одномерных систем, находится с помощью суперпозиции.

Инвариантность к сдвигу. Система инвариантна к сдвигу, если сдвиг входной последовательности приводит к такому же сдвигу выходной последовательности:

$$T[s(n-N,m-M)] = z(n-N,m-M).$$

Линейность и инвариантность к сдвигу являются независимыми свойствами системы. Так, пространственное маскирование линейно, но не инвариантно к сдвигу, а безынерционные операторы нелинейны, но инвариантны к сдвигу.

В дальнейшем ограничимся рассмотрением только систем, широко распространенных при решении практических задач - линейных и инвариантных к сдвигу (ЛИС-системы).

Импульсный отклик на произвольно расположенный входной импульс, как следует из выражения (18.2.3), описывается выражением:

$$h_{ij}(n,m) = T[\delta(n-n_i,m-m_j)].$$

Для частного случая $i = j = 0$ имеем:

$$h_0(n,m) = T[\delta(n,m)].$$

Используя принцип инвариантности к сдвигу, получим:

$$h_{ij}(n,m) = h_0(n-i,m-j) = h(n-i,m-j), \quad (18.2.4)$$

т.е. импульсный отклик на произвольно расположенный входной импульс равен сдвинутому импульсному отклику на входной импульс, расположенный в начале координат.

Двумерная свертка. Подставляя (18.2.4) в выражение (18.2.3), получаем:

$$z(n,m) = \sum_i \sum_j s(i,j) h(n-i,m-j). \quad (18.2.5)$$

Двумерная дискретная свертка (18.2.5), является аналогом одномерной дискретной свертки. При замене переменных $n-i = k$, $m-j = l$, получим:

$$z(n,m) = \sum_k \sum_l h(k,l) s(n-k,m-l), \quad (18.2.5')$$

т.е. двумерная свертка коммутативна, как и одномерная. В такой же мере она обладает свойством ассоциативности по отношению к последовательности операций свертки нескольких функций (результат не зависит от порядка свертки) и свойством дистрибутивности по отношению к операции свертки с суммой функций (результат аналогичен сумме сверток с каждой функцией). Эти свойства определяют и основное свойство двумерных (и многомерных) линейных систем при их параллельном и/или последовательном соединении – результирующая система также является линейной.

Для упрощения символического аппарата двумерную свертку обозначают индексом (**):

$$z(n,m) = h(k,l) ** s(n-k,m-l).$$

При обобщении этого выражения на многомерные системы, в векторной форме:

$$z(\bar{n}) = h(\bar{k}) ** s(\bar{n} - \bar{k}).$$

Разделимые системы. Если импульсный отклик системы может быть разделен:

$$h(k,l) = h(k) h(l), \quad (18.2.6)$$

то выражение (18.2.5') принимает вид:

$$z(n,m) = \sum_k h(k) \sum_l h(l) s(n-k,m-l), \quad (18.2.7)$$

или: $z(n,m) = \sum_k h(k) g(n-k,m), \quad g(n-k,m) = \sum_l h(l) s(n-k,m-l).$

Массив $g(n,m)$ вычисляется одномерной сверткой столбцов массива $s(n,m)$ при $n = \text{const}$ (сечения массива по координатам n) с откликом $h(l)$, с последующим вычислением выходного массива $z(n,m)$ одномерной сверткой строк $g(n,m)$ при $m = \text{const}$ с откликом $h(k)$. Результат не изменится, если сначала выполнять свертку по строкам, а затем по столбцам. Система с откликом вида (18.2.6) называется разделимой. Отметим, что в разделимой системе входной и выходной сигнал не обязаны быть разделимыми.

Аналогичные разделимые системы могут существовать и в многомерном варианте.

Устойчивость системы. Интерес для практики представляют только устойчивые системы, обеспечивающие определенный конечный результат системной операции на конечные входные сигналы. Необходимым и достаточным условием устойчивости системы является абсолютная суммируемость ее импульсного отклика: $\sum_k \sum_l |h(k,l)| < \infty$.

Специальные двумерные системы. На практике используются также системы с несколькими входами и/или выходами.

Допустим, система имеет i -входы и j -выходы, линейна и инвариантна к сдвигу по переменной t . Если на i -вход системы поступает одномерный единичный импульс $\delta_i(t)$ при нулевых сигналах на остальных входах, то j -выходные сигналы будут импульсным откликом системы $h_{ij}(t)$. При известном полном ансамбле значений h_{ij} для всех i -входов, для произвольной комбинации входных сигналов $s_i(t)$ сигнал на j -выходе будет определяться выражением:

$$z_j(t) = \sum_i \sum_k h_{ij}(k) s_i(t-k). \quad (18.2.8)$$

18.3. ЧАСТОТНЫЕ ХАРАКТЕРИСТИКИ СИГНАЛОВ И СИСТЕМ.

Частотный отклик системы. Допустим, что двумерная ЛИС-система имеет импульсный отклик $h(k\Delta x, l\Delta y)$. Подадим на вход системы сигнал вида комплексной синусоиды:

$$s(n,m) = \exp(jn\Delta x \cdot \omega_x + jm\Delta y \cdot \omega_y),$$

где ω_x и ω_y – значения частоты сигнала соответственно по координатам x и y . Принимая $\Delta x = 1$, $\Delta y = 1$ и выполняя двумерную свертку (18.2.5), получаем:

$$\begin{aligned} z(n,m) &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(k,l) \exp[j\omega_x \cdot (n-k) + j\omega_y \cdot (m-l)] = \\ &= \exp(jn \cdot \omega_x + jm \cdot \omega_y) \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(k,l) \exp(-jk \cdot \omega_x - jl \cdot \omega_y) = H(\omega_x, \omega_y) \exp(jn \cdot \omega_x + jm \cdot \omega_y). \\ H(\omega_x, \omega_y) &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(k,l) \exp(-jk \cdot \omega_x - jl \cdot \omega_y). \end{aligned} \quad (18.3.1)$$

Таким образом, выходной сигнал представляет собой комплексную синусоиду с теми же значениями частоты, что и у входного сигнала, с изменением амплитуды и фазы за счет

комплексного множителя $H(\omega_x, \omega_y)$, который носит название частотного отклика (частотной характеристики) системы. Для дискретных сигналов частотный отклик периодичен с периодом 2π по обоим частотным переменным:

$$H(\omega_x + 2\pi k, \omega_y + 2\pi l) = H(\omega_x, \omega_y).$$

Пример расчета частотного отклика системы.

Определить частотную характеристику системы с импульсным откликом:

$$h(0,0) = 0.25, \quad h(0, \pm 1) = 0.125, \quad h(\pm 1, 0) = 0.125, \quad h(\pm 1, \pm 1) = 0.0625.$$

Частотный отклик:

$$H(\omega_x, \omega_y) = \sum_{n=-1}^1 \sum_{m=-1}^1 h(n,m) \exp(-jn\omega_x - jm\omega_y) = 0.25 + 0.125[\exp(-j\omega_x) + \exp(j\omega_x) + \exp(-j\omega_y) + \exp(j\omega_y)] + \\ + 0.0625[\exp(-j\omega_x - j\omega_y) + \exp(-j\omega_x + j\omega_y) + \exp(j\omega_x - j\omega_y) + \exp(j\omega_x + j\omega_y)] = 0.25(1 + \cos \omega_x)(1 + \cos \omega_y).$$

Система является примером двумерного фильтра нижних частот. Частотный отклик системы на плоскости (ω_x, ω_y) , приведенный на рис. 18.3.1, имеет осевую симметрию с коэффициентом передачи 1 в центре $(\omega_x=0, \omega_y=0)$ со спадом до нуля при $\omega_x = \pm \pi$ и $\omega_y = \pm \pi$.

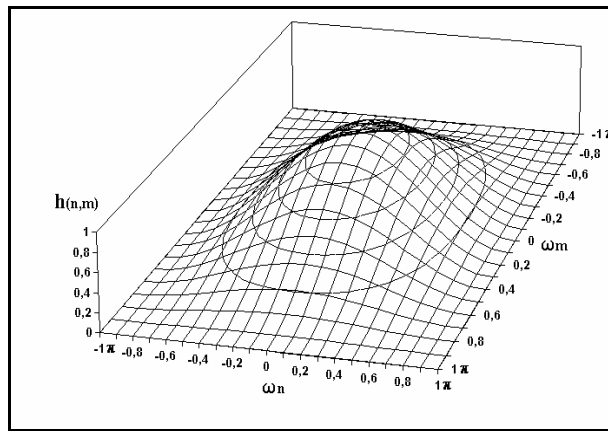


Рис. 18.3.1. Частотная характеристика ФНЧ.

При разделимости импульсного отклика частотный отклик многомерных систем также является разделимой функцией:

$$h(k,l) = q(k)g(l) \Leftrightarrow Q(\omega_x)G(\omega_y) = H(\omega_x, \omega_y)$$

$$Q(\omega_x) = \sum_k q(k) \exp(-jk \cdot \omega_x), \quad G(\omega_y) = \sum_l g(l) \exp(-jl \cdot \omega_y).$$

Импульсный отклик системы. Выражение (18.3.1), по существу, описывает разложение функции $H(\omega_x, \omega_y)$ в двумерный ряд Фурье с коэффициентами разложения в виде отсчетов импульсного отклика $h(k,l)$, т.е. прямое преобразование Фурье. Очевидно, что обратным преобразованием Фурье с интегрированием в пределах одного периода из частотного отклика $H(\omega_x, \omega_y)$ можно получить импульсный отклик системы:

$$h(k,l) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} H(\omega_x, \omega_y) \exp(jk\omega_x + jl\omega_y) d\omega_x d\omega_y. \quad (18.3.2)$$

Пример расчета импульсного отклика фильтра.

Определить импульсный отклик идеального фильтра низких частот с прямоугольной частотной характеристикой вида: $H(\omega_x, \omega_y) = 1$ при $|\omega_x| \leq a < \pi$, $|\omega_y| \leq b < \pi$; $H(\omega_x, \omega_y) = 0$ в остальных случаях.

Импульсный отклик:
$$h(k,l) = \frac{1}{4\pi^2} \int_{-a}^a \int_{-b}^b \exp(jk\omega_x + jl\omega_y) d\omega_x d\omega_y.$$

Система разделима:
$$h(k,l) = \frac{1}{2\pi} \int_{-a}^a \exp(jk\omega_x) d\omega_x \cdot \frac{1}{2\pi} \int_{-b}^b \exp(jl\omega_y) d\omega_y = \frac{\sin ak}{\pi k} \cdot \frac{\sin bl}{\pi l}.$$

Пример расчета неразделимого импульсного отклика.

Определить импульсный отклик идеального кругового фильтра нижних частот:

$$H(\omega_x, \omega_y) = 1 \text{ при } \omega_x^2 + \omega_y^2 < R^2 < \pi^2; \quad H(\omega_x, \omega_y) = 0 \text{ в остальных случаях.}$$

Вычисления по круговой области целесообразно выполнять в полярных координатах: $\omega = \sqrt{\omega_x^2 + \omega_y^2}$,

$\varphi = \arctg(\omega_y/\omega_x)$, $\phi = \arctg(m/n)$, при этом выражение 18.3.2 переписывается в следующем виде:

$$\begin{aligned} h(n,m) &= \frac{1}{4\pi^2} \int_0^R \int_0^{2\pi} \omega \exp[j\omega \cos(\phi-\varphi)] d\varphi d\omega = \\ &= \frac{1}{2\pi} \int_0^R \omega J_0(\omega \sqrt{n^2 + m^2}) d\omega = \frac{R}{2\pi} J_1(R \sqrt{n^2 + m^2}) / \sqrt{n^2 + m^2}, \end{aligned}$$

где $J_0(\dots)$, $J_1(\dots)$ - функции Бесселя 1-го рода 0-го и 1-го порядков соответственно.

На рис. 18.3.2 приведена пространственная форма импульсного отклика фильтра, расчет которой проведен при $R = 1$ с ограничением по $N = 10$ и $M = 10$, и сечения отклика по координате m .

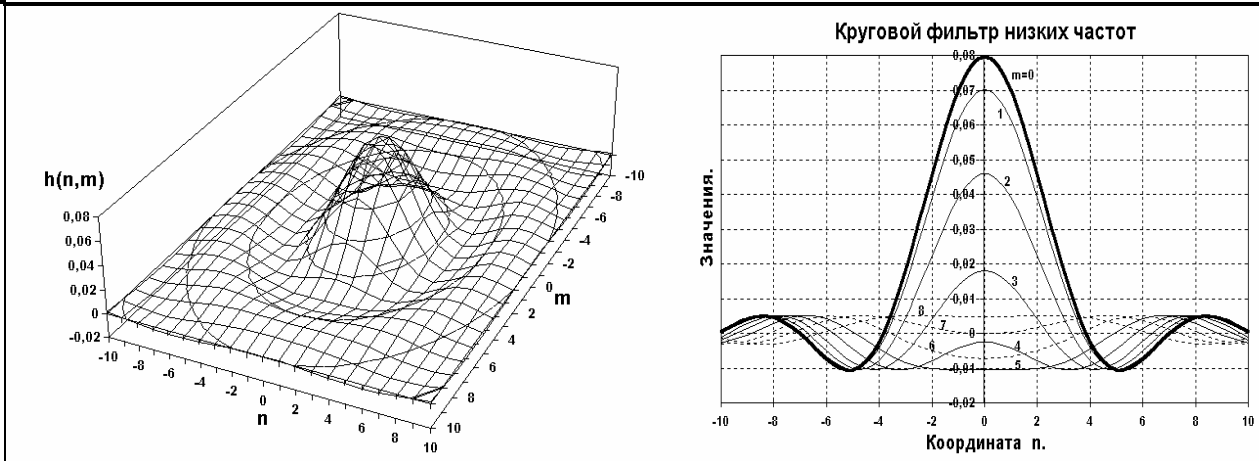


Рис. 18.3.2. Круговой низкочастотный фильтр (справа - сечения по координате m).

Свойства двумерного преобразования Фурье. Вышеприведенные преобразования импульсного отклика в частотный отклик и наоборот представляют собой двумерные дискретные преобразования Фурье с прямоугольным растром дискретизации информации, эквивалентные одномерным преобразованиям. На двумерные преобразования с прямоугольным растром переносятся и другие свойства одномерных систем. В частности:

1. Фурье-преобразования сигналов.

$$S(\omega_x, \omega_y) = \sum_n \sum_m s(n,m) \exp(-jn\omega_x - jm\omega_y). \quad (18.3.3)$$

$$s(n,m) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} S(\omega_x, \omega_y) \exp(jn\omega_x + jm\omega_y) d\omega_x d\omega_y. \quad (18.3.4)$$

2. Теорема о свертке.

$$z(n,m) = h(n,m) ** s(n,m) \Leftrightarrow H(\omega_x, \omega_y) S(\omega_x, \omega_y) = Z(\omega_x, \omega_y).$$

$$z(n,m) = c(n,m) s(n,m) \Leftrightarrow C(\omega_x, \omega_y) ** S(\omega_x, \omega_y) = Z(\omega_x, \omega_y).$$

3. Основные свойства Фурье-преобразования.

1) Линейность (в том числе для любых комплексных чисел a и b):

$$a \cdot s(n,m) + b \cdot z(n,m) \Leftrightarrow aS(\omega_x, \omega_y) + bZ(\omega_x, \omega_y).$$

2) Пространственный сдвиг:

$$s(n-N, m-M) \Leftrightarrow S(\omega_x, \omega_y) \exp(-jN\omega_x - jM\omega_y).$$

3) Дифференцирование:

$$dS(\omega_x, \omega_y)/d\omega_x \Leftrightarrow -jn s(n,m),$$

$$dS(\omega_x, \omega_y)/d\omega_y \Leftrightarrow -jm s(n,m),$$

$$d^2S(\omega_x, \omega_y)/(d\omega_x d\omega_y) \Leftrightarrow -nm s(n,m).$$

4) Комплексное сопряжение:

$$x^*(n,m) \Leftrightarrow S^*(-\omega_x, -\omega_y).$$

Вещественная и мнимая части Фурье-образов последовательностей $s(n,m)$:

$$\begin{aligned} S(\omega_x, \omega_y) &= S^*(-\omega_x, -\omega_y). \\ \operatorname{Re} [S(\omega_x, \omega_y)] &= \operatorname{Re} [S(-\omega_x, -\omega_y)]. \\ \operatorname{Im} [S(\omega_x, \omega_y)] &= -\operatorname{Im} [S(-\omega_x, -\omega_y)]. \end{aligned}$$

5) Теорема Парсеваля:

$$\sum_n \sum_m s(n, m) s^*(n, m) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} S(\omega_x, \omega_y) S^*(\omega_x, \omega_y) d\omega_x d\omega_y.$$

В частности, при $s(n, m) = s(n, m)$:

$$\sum_n \sum_m |s(n, m)|^2 = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |S(\omega_x, \omega_y)|^2 d\omega_x d\omega_y,$$

где левая часть уравнения представляет собой полную энергию дискретного сигнала $s(n, m)$, а функция $|S(\omega_x, \omega_y)|^2$ - спектральную плотность энергии сигнала.

18.4. ДИСКРЕТИЗАЦИЯ ДВУМЕРНЫХ СИГНАЛОВ [9].

Прямоугольный растр дискретизации. Из способов обобщения одномерной периодической дискретизации на двумерный случай наиболее простым является периодическая дискретизация в прямоугольных координатах:

$$s(n, m) = s_a(n\Delta x, m\Delta y),$$

где Δx и Δy - горизонтальный и вертикальный интервалы дискретизации двумерного непрерывного сигнала $s_a(x, y)$ с непрерывными координатами x и y . Ниже значения Δx и Δy , как и в одномерном случае, принимаются равными 1.

Дискретизация двумерного, а в общем случае и многомерного сигнала, также приводит к периодизации его спектра и наоборот. Сохраняется также и условие информационной равноценности координатного и частотного представлений дискретного сигнала при равном количестве точек дискретизации в главных диапазонах сигнала. Для прямоугольной дискретизации связь фурье-преобразований непрерывного и дискретного сигналов устанавливается аналогично одномерной дискретизации.

Интегральные преобразования Фурье аналоговых сигналов в непрерывной шкале частот Ω_x и Ω_y :

$$S_a(\Omega_x, \Omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_a(x, y) \exp(-j\Omega_x x - j\Omega_y y) dx dy. \quad (18.4.1)$$

$$s_a(x, y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_a(\Omega_x, \Omega_y) \exp(j\Omega_x x + j\Omega_y y) d\Omega_x d\Omega_y. \quad (18.4.2)$$

Дискретные преобразования Фурье:

$$S(k, l) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} s(n, m) \exp(-jn2\pi k/N - jm2\pi l/M), \quad (18.4.3)$$

$$S(k, l) = \sum_{n=0}^{N-1} \exp(-jn2\pi k/N) \sum_{m=0}^{M-1} s(n, m) \exp(-jm2\pi l/M), \quad (18.4.3')$$

$$s(n, m) = \frac{1}{NM} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} S(k, l) \exp(-jn2\pi k/N - jm2\pi l/M). \quad (18.4.4)$$

$$s(n, m) = \frac{1}{NM} \sum_{k=0}^{N-1} \exp(-jn2\pi k/N) \sum_{l=0}^{M-1} S(k, l) \exp(-jm2\pi l/M). \quad (18.4.4')$$

Выражения (18.4.3') и (18.4.4') показывают, что двумерное ДПФ по прямоугольному растру дискретизации данных может вычисляться с помощью одномерных последовательных ДПФ. Вторые суммы выражений являются одномерными ДПФ сечений функций $s(n, m)$ и $S(k, l)$ по линиям n и k соответственно, а первые - одномерными ДПФ вычисленных функций в сечениях по m и l . Другими словами, исходные матрицы значений $s(n, m)$ и $S(k, l)$ пересчитываются сначала в промежуточные матрицы с ДПФ по строкам (или по столбцам), а проме-

жуточные - в окончательные с ДПФ по столбцам (или соответственно по строкам).

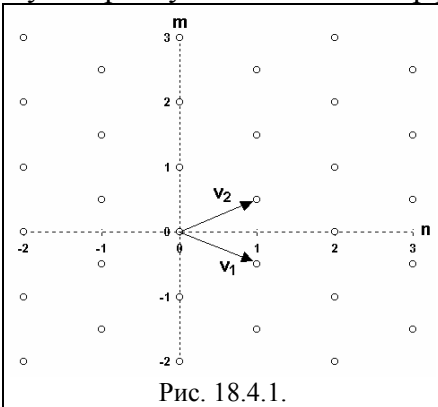
Интерполяционный ряд восстановления двумерного сигнала. Если непрерывный сигнал $s_a(x,y)$ является сигналом с ограниченным спектром, а периоды дискретизации выбраны достаточно малыми и спектры соседних периодов не перекрываются:

$$S_a(\Omega_x, \Omega_y) = 0 \quad \text{при } |\Omega_x| \geq \pi/\Delta x, \quad |\Omega_y| \geq \pi/\Delta y,$$

то, как и в одномерном случае, сигнал $s_a(x,y)$ может быть восстановлен по дискретному сигналу с использованием двумерного аналога ряда Котельникова-Шеннона:

$$s_a(x,y) = \sum_n \sum_m s(n,m) \frac{\sin\pi(x/\Delta x - n)}{\pi(x/\Delta x - n)} \frac{\sin\pi(y/\Delta y - m)}{\pi(y/\Delta y - m)}. \quad (18.4.5)$$

Сигнал с неограниченным спектром также может быть дискретизирован, однако в этом случае имеет место наложение спектров в смежных периодах, при этом высокие частоты, большие частоты Найквиста, будут "маскироваться", как и в одномерном случае, под низкие частоты главного периода. Эффект "отражения" от границ периода дает еще более сложную картину вследствие интерференции частот, отраженных по разным координатам.



Произвольный растр дискретизации.

Понятие прямоугольной дискретизации обобщается на произвольный растр дискретизации с линейно независимыми векторами $\mathbf{v}_1 = (v_{11}, v_{21})^T$ и $\mathbf{v}_2 = (v_{12}, v_{22})^T$, где T - индекс транспонирования (рис. 18.4.1). Координаты двумерного периодического множества отсчетов на плоскости (x,y) :

$$x = v_{11}n + v_{12}m,$$

$$y = v_{21}n + v_{22}m.$$

С использованием векторных обозначений:

$$\bar{\mathbf{z}} = \bar{\mathbf{V}} \bar{\mathbf{n}}$$

где $\bar{\mathbf{z}} = (x,y)^T$, $\bar{\mathbf{n}} = (n,m)^T$, $\bar{\mathbf{V}} = (\mathbf{v}_1 | \mathbf{v}_2)$ - матрица дискретизации.

Определитель матрицы $\bar{\mathbf{V}}$ не равен нулю, если вектора \mathbf{v}_1 и \mathbf{v}_2 линейно независимы. При дискретизации непрерывного сигнала $s_a(x,y)$ матрицей $\bar{\mathbf{V}}$ формируется дискретный сигнал:

$$s(\bar{\mathbf{n}}) \Leftarrow s_a(\bar{\mathbf{n}} \bar{\mathbf{V}}).$$

Двумерное интегральное преобразование Фурье непрерывного сигнала по непрерывному вектору $\bar{\Omega} = (\Omega_1, \Omega_2)^T$:

$$S_a(\bar{\Omega}) = \int_{-\infty}^{\infty} s_a(\bar{\mathbf{z}}) \exp(-j \bar{\Omega}^T \bar{\mathbf{z}}) d\bar{\mathbf{z}}, \quad (18.4.6)$$

$$s_a(\bar{\mathbf{z}}) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} S_a(\bar{\Omega}) \exp(j \bar{\Omega}^T \bar{\mathbf{z}}) d\bar{\Omega}, \quad (18.4.7)$$

Данные интегралы являются двойными, поскольку дифференциалы $d\bar{\mathbf{z}}$ и $d\bar{\Omega}$ являются векторами.

Преобразование Фурье дискретного сигнала:

$$S(\bar{\omega}) = \sum_n s(\bar{\mathbf{n}}) \exp(-j \bar{\omega}^T \bar{\mathbf{n}}), \quad (18.4.8)$$

$$s(\bar{\mathbf{n}}) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} S(\bar{\omega}) \exp(j \bar{\omega}^T \bar{\mathbf{n}}) d\bar{\omega}. \quad (18.4.9)$$

где: $\bar{\omega} = (\omega_x, \omega_y)^T$.

Выражение $s(\bar{\mathbf{n}})$ может быть получено дискретизацией выражения $s_a(\bar{\mathbf{z}})$ (18.4.7):

$$s(\bar{\mathbf{n}}) = s_a(\bar{\mathbf{V}} \bar{\mathbf{n}}) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} S_a(\bar{\Omega}) \exp(j \bar{\Omega}^T \bar{\mathbf{V}} \bar{\mathbf{n}}) d\bar{\Omega}.$$

После подстановки в это выражение значения $\bar{\omega} = \bar{\Omega} \bar{\mathbf{V}}^T$, получаем:

$$s(\bar{\mathbf{n}}) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \frac{1}{|\det \bar{\mathbf{V}}|} S_a(\bar{\omega} / \bar{\mathbf{V}}^T) \exp(j \bar{\omega}^T \bar{\mathbf{n}}) d\bar{\omega}.$$

Или, с учетом периодичности по квадратным областям плоскости $\bar{\omega}$:

$$s(\bar{\mathbf{n}}) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \frac{1}{|\det \bar{\mathbf{V}}|} \sum_{\bar{\mathbf{k}}} S_a((\bar{\omega} - 2\pi \bar{\mathbf{k}}) / \bar{\mathbf{V}}^T) \exp(j \bar{\omega}^T \bar{\mathbf{n}}) d\bar{\omega}, \quad (18.4.10)$$

где $\bar{\mathbf{k}}$ - вектор целочисленных значений периодов дискретизированной функции по осям ω_x и ω_y . Сравнивая последнее выражение с выражением (18.4.9), получаем:

$$S(\bar{\omega}) = \frac{1}{|\det \bar{\mathbf{V}}|} \sum_{\bar{\mathbf{k}}} S_a((\bar{\omega} - 2\pi \bar{\mathbf{k}}) / \bar{\mathbf{V}}^T),$$

$$S(\bar{\Omega} \bar{\mathbf{V}}^T) = \frac{1}{|\det \bar{\mathbf{V}}|} \sum_{\bar{\mathbf{k}}} S_a(\bar{\Omega} - \bar{\mathbf{U}} \bar{\mathbf{k}}), \quad (18.4.11)$$

где $\bar{\mathbf{U}}$ - матрица периодичности:

$$\bar{\mathbf{U}}^T \bar{\mathbf{V}} = 2\pi \bar{\mathbf{I}}, \quad (18.4.12)$$

которой задаются два линейно независимых вектора периодичности спектра, $\bar{\mathbf{I}}$ - единичная матрица 2 x 2. Выражение (18.4.11) определяет связь между преобразованиями Фурье дискретных и аналоговых сигналов.

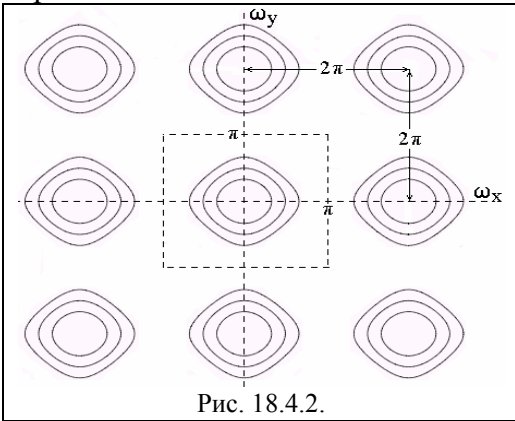


Рис. 18.4.2.

Как и в одномерном случае, интервалы дискретизации Δx и Δy определяют главный период двумерного спектра соответственно по осям ω_x и ω_y и частоты Найквиста: $\omega_{xN} = \pi/\Delta x$ и $\omega_{yN} = \pi/\Delta y$. Спектр дискретного сигнала также является периодическим продолжением спектра аналогового сигнала. Для исключения искажений спектра (наложения спектров боковых периодов на главный период) предельные частоты сигнала должны быть меньше частот Найквиста. На рис. 18.4.2 приведен пример центральной части спектра дискретного сигнала при $\Delta x=1$ и $\Delta y=1$.

В случае прямоугольной дискретизации:

$$\bar{\mathbf{V}} = \begin{bmatrix} \Delta x & 0 \\ 0 & \Delta y \end{bmatrix}, \quad \det \bar{\mathbf{V}} = \Delta x \Delta y, \quad (18.4.13)$$

$$\bar{\mathbf{U}} = \begin{bmatrix} 2\pi/\Delta x & 0 \\ 0 & 2\pi/\Delta y \end{bmatrix}. \quad (18.4.14)$$

Интерполяция дискретных сигналов. Для сигнала с ограниченным спектром изменением матрицы дискретизации $\bar{\mathbf{V}}$ можно подобрать матрицу периодичности $\bar{\mathbf{U}}$ таким образом, чтобы в правой части выражения (18.4.11) не было перекрытия спектров. Тогда для значений по точкам $\bar{\Omega} \bar{\mathbf{V}}^T$ области C главного периода спектра выражение (18.4.11) упрощается:

$$S(\bar{\Omega} \bar{\mathbf{V}}^T) = S_a(\bar{\Omega}) / |\det \bar{\mathbf{V}}|. \quad (18.4.15)$$

$$S_a(\bar{\Omega}) = |\det \bar{\mathbf{V}}| S(\bar{\Omega} \bar{\mathbf{V}}^T) = |\det \bar{\mathbf{V}}| S(\bar{\omega}), \quad \bar{\Omega} \in C. \quad (18.4.16)$$

Из выражения (18.4.16) следует, что при корректной дискретизации непрерывной двумерной функции ее спектр с точностью до нормировочного множителя $|\det \bar{\mathbf{V}}|$ может быть восстановлен по спектру дискретной функции. Соответственно, выполнив обратное преобразование Фурье левой и правой части равенства (18.4.16), получим уравнение восстановления непрерывной функции по ее дискретному варианту (многомерный аналог интерполяционного ряда Котельникова-Шеннона):

$$s_a(\bar{z}) = \frac{|\det \bar{\mathbf{V}}|}{4\pi^2} \sum_{\bar{\mathbf{n}}} s(\bar{\mathbf{n}}) \int_{\mathcal{C}} \exp(j \bar{\Omega}^T (\bar{z} - \bar{\mathbf{V}} \bar{\mathbf{n}})) d\bar{\Omega}.$$

$$s_a(\bar{z}) = \sum_{\bar{\mathbf{n}}} s(\bar{\mathbf{n}}) f(\bar{z} - \bar{\mathbf{V}} \bar{\mathbf{n}}), \quad (18.4.17)$$

где $f(\dots)$ – интерполяционная функция:

$$f(\bar{z} - \bar{\mathbf{V}} \bar{\mathbf{n}}) = \frac{|\det \bar{\mathbf{V}}|}{4\pi^2} \int_{\mathcal{C}} \exp(j \bar{\Omega}^T (\bar{z} - \bar{\mathbf{V}} \bar{\mathbf{n}})) d\bar{\Omega}. \quad (18.4.18)$$

Все приведенные векторные уравнения могут быть обобщены на P -мерные функции с заменой константы $4\pi^2$ там, где она встречается, на $(2\pi)^P$.

Прямоугольный и гексагональный растры дискретизации. В принципе, сигнал с ограниченным спектром можно представить по различным растрам дискретизации. Выбор раstra обычно производят из условия минимальной плотности отсчетов на плоскости, т.е. минимизацией величины $|\det \bar{\mathbf{V}}|$, при котором обеспечивается отсутствие наложений для частот анализируемых сигналов.

На практике для двумерных сигналов используют, как правило, только два варианта растров дискретизации - прямоугольный и гексагональный. Прямоугольному варианту соответствуют диагональные матрицы дискретизации и периодичности (18.4.13-14). Для гексагональной дискретизации, пример которой приведен на рис. 18.4.1, в частном случае при $\Delta t = \Delta x \sqrt{3}$ каждый отсчет располагается на равном расстоянии от шести ближайших отсчетов, при этом матрицы дискретизации:

$$\bar{\mathbf{V}} = \begin{bmatrix} 1 & 1 \\ \sqrt{3} & -\sqrt{3} \end{bmatrix}, \quad \bar{\mathbf{U}} = \begin{bmatrix} 2\pi & 2\pi \\ 2\pi/\sqrt{3} & -2\pi/\sqrt{3} \end{bmatrix}.$$

Допустим, имеем сигнал с частотным спектром, ограниченным круговой областью частот ω_r :

$$S_a(\omega_x, \omega_y) = 0 \text{ при } \omega_x^2 + \omega_y^2 > \omega_r^2.$$

Круговая область частот вписывается без перекрытий в квадрат со стороной $2\omega_r$ или в шестиугольник со стороной $2\omega_r/\sqrt{3}$. Матрицы дискретизации:

$$\bar{\mathbf{V}}_{\text{пр}} = \begin{bmatrix} \pi/\omega_r & 0 \\ 0 & \pi/\omega_r \end{bmatrix}, \quad \det \bar{\mathbf{V}} = \pi^2/\omega_r^2,$$

$$\bar{\mathbf{V}}_{\text{гекс}} = \begin{bmatrix} \pi/(\omega_r\sqrt{3}) & \pi/(\omega_r\sqrt{3}) \\ \pi/\omega_r & -\pi/\omega_r \end{bmatrix}, \quad \det \bar{\mathbf{V}} = 2\pi^2/(\omega_r^2\sqrt{3}).$$

Поскольку плотность отсчетов пропорциональна $1/|\det \bar{\mathbf{V}}|$, то отсюда следует, что для представления одного и того же сигнала гексагональный растр дискретизации требует на 13.4% меньше отсчетов по сравнению с прямоугольным. Эффективность "гексагональной" матрицы возрастает при увеличении размерности сигнала. Так, при 4-мерном сигнале для "гексагональной" матрицы требуется в 2 раза меньше отсчетов, чем для "прямоугольной".

18.5. МНОГОМЕРНЫЙ СПЕКТРАЛЬНЫЙ АНАЛИЗ [9].

Периодические последовательности. Двумерная последовательность $s(n,m)$ прямоугольно периодична, если

$$s_{\text{п}}(n,m) = s(n,n+M) = s(n+N,m)$$

для всех (n,m) при целочисленных значениях N и M . Минимальные значения N и M , при которых выполняется данное равенство, называют горизонтальным и вертикальным периодами функции $s_{\text{п}}(\dots)$, которыми ограничивается прямоугольная область $R_{N,M}$ главного периода, содержащая NM независимых отсчетов: $0 \leq n \leq N-1$, $0 \leq m \leq M-1$.

Последовательность $s_{\text{п}}(n,m)$ с периодами N и M можно представить в виде конечной суммы (ряда Фурье) комплексных синусоид с кратными частотами:

$$s_{\Pi}(n,m) = (1/NM) \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} S_{\Pi}(k,l) \exp(j2\pi nk/N + j2\pi ml/M). \quad (18.5.1)$$

$$S_{\Pi}(k,l) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} s_{\Pi}(n,m) \exp(-j2\pi nk/N - j2\pi ml/M). \quad (18.5.2)$$

Пример.

Разложить в ряд Фурье периодический сигнал: $s_{\Pi}(n,m) = \delta(n,m)$, $0 \leq n \leq 4$, $0 \leq m \leq 2$ (единичные импульсы с периодом: $N = 5$, $M = 3$).

$S_{\Pi}(k,l) = \sum_{n=0}^4 \sum_{m=0}^2 \delta(n,m) \exp(-j2\pi nk/5 - j2\pi ml/3) = 1$ для всех k и l , т.е. равномерная частотная характеристика в главном диапазоне. Соответственно сам сигнал может быть записан в виде двумерного ряда Фурье:

$$s_{\Pi}(n,m) = (1/20) \sum_{k=0}^4 \sum_{l=0}^2 \exp(j2\pi nk/5 + j2\pi ml/3).$$

Конечные последовательности. Если $s(n,m)$ представляет собой последовательность конечной протяженности, имеющей опорную область $R_{N,M}$, то периодическую последовательность $s_{\Pi}(n,m)$ с главным периодом $R_{N,M}$ можно сформировать периодическим продолжением $s(n,m)$:

$$s_{\Pi}(n,m) = \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} s(n-aN, m-bM),$$

$$s(n, m) = s_{\Pi}(n, m), \quad (n, m) \in R_{N,M}.$$

$$= 0, \quad \text{в остальных случаях.}$$

Отсюда следует, что любой финитный сигнал может быть полностью определен своим периодическим продолжением и опорной областью.

Аналогично можно записать и для частотной области:

$$S_{\Pi}(k, l) = \sum_a \sum_b S(k-aN, l-bM).$$

$$S(k, l) = S_{\Pi}(k, l), \quad 0 \leq k \leq N, \quad 0 \leq l \leq M$$

$$= 0, \quad \text{в остальных случаях.}$$

Отсюда значения $s(n,m)$ и $S(k,l)$ можно вычислить с использованием выражений (18.5.1-2) путем последовательности операций:

$$s(n,m) \Rightarrow s_{\Pi}(n,m) \Leftrightarrow S_{\Pi}(k,l) \Rightarrow S(k,l).$$

Практически это означает, что для получения ДПФ последовательности конечной протяженности достаточно из выражений (18.5.1-2) для рядов Фурье убрать знак периодичности, при этом следует помнить, что вычисление отсчетов $s(\cdot)$ вне опорной области приведет к вычислению значений не отсчетов $s(\cdot)$, а отсчетов $s_{\Pi}(\cdot)$ периодического продолжения сигнала $s(\cdot)$.

Таким образом:

1. Дискретизация сигнала в пространственной области вызывает периодизацию частотного спектра сигнала.
2. Дискретизация частотного спектра сигнала вызывает периодизацию его пространственного представления.
3. Прямое и обратное ДПФ сигнала ограниченной протяженности автоматически означает периодизацию как его спектра, так и его пространственного представления.
4. Сигналы, ограниченные в пространстве, можно точно отобразить отсчетами их фурье-преобразования.
5. Частотное представление сигнала с ограниченным спектром обратным фурье-преобразованием может быть точно переведено в пространственную область.
6. Ограниченность как пространственного сигнала, так и его спектра является обязательным условием корректного ДПФ, т.к. в противном случае периодизация сигнала может привести к искажению его спектрального и пространственного представления.

Многомерные последовательности. Определение ДПФ для P -мерной последователь-

ности с опорной областью $R_P = \{\bar{\mathbf{n}} : 0 \leq n_i \leq N_i - 1, i=1,2,3, \dots, P\}$ производится введением диагональной матрицы значений N_i :

$$\bar{\mathbf{N}} = \begin{bmatrix} N_1 & 0 & \dots & 0 \\ 0 & N_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & N_p \end{bmatrix},$$

при этом P -мерное ДПФ записывается в виде:

$$S(\bar{\mathbf{k}}) = \sum_{\bar{\mathbf{n}} \in R_P} s(\bar{\mathbf{n}}) \exp(-j \bar{\mathbf{k}}^T 2\pi \bar{\mathbf{n}} / \bar{\mathbf{N}}). \quad (18.5.3)$$

$$s(\bar{\mathbf{n}}) = \frac{1}{|\det \bar{\mathbf{N}}|} \sum_{\bar{\mathbf{k}} \in R_P} S(\bar{\mathbf{k}}) \exp(j \bar{\mathbf{k}}^T 2\pi \bar{\mathbf{n}} / \bar{\mathbf{N}}). \quad (18.5.4)$$

Кратко рассмотрим особенности многомерных ДПФ (на примере двумерных последовательностей).

ДПФ суммы двух последовательностей с опорной областью на $R_{N,M}$ равно сумме их ДПФ:

$$as(n,m) + bz(n,m) \Leftrightarrow aS(k,l) + bZ(k,l),$$

но при этом все ДПФ должны быть одного размера и этот размер должен быть достаточным, чтобы включить всю опорную область суммарной последовательности $as(n,m) + bz(n,m)$. Практически это означает, что $x(\cdot)$ и $z(\cdot)$ должны иметь одну и ту же опорную область. Опорная область каждой последовательности при необходимости дополняется нулями.

Операция свертки двух функций в пространственной области отображается операцией умножения фурье-образов функций в частотной области, однако при этом линейная свертка полных пространственных сигналов при ее вычислении через ДПФ в силу периодического продолжения пространственных функций переходит в циклическую свертку (как и для одномерных сигналов). Результат свертки зависит от периодов N и M .

Допустим, что $s(n,m)$ имеет опорную область R_{P_1, P_2} , а $h(n,m)$ - R_{Q_1, Q_2} . Результат линейной свертки:

$$s(n,m) = \sum_k \sum_l h(k,l) s(n-k, m-l).$$

Опорная область последовательности $s(n,m)$:

$$0 \leq n \leq P_1 + Q_1 - 1, \quad 0 \leq m \leq P_2 + Q_2 - 1.$$

Следовательно, наложения периодов результата свертки не произойдет и циклическая свертка в главном частотном диапазоне будет равна линейной свертке при опорной области ДПФ:

$$N \geq P_1 + Q_1 - 1, \quad M \geq P_2 + Q_2 - 1.$$

Пример.

Заданы последовательности (начало координат в нижнем левом углу): $s(n,m) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$, $h(n,m) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

Вычислить свертку $z(n,m) = s(n,m) ** h(n,m)$.

ДПФ размера 2×2 для $s(n,m)$.

$$S(0,0) = s(0,0) + s(1,0) + s(0,1) + s(1,1) = 2 + 1 + 1 + 0 = 4$$

$$S(1,0) = s(0,0) - s(1,0) + s(0,1) - s(1,1) = 2 - 1 + 1 - 0 = 2$$

$$S(0,1) = s(0,0) + s(1,0) - s(0,1) - s(1,1) = 2 + 1 - 1 - 0 = 2$$

$$S(1,1) = s(0,0) - s(1,0) - s(0,1) + s(1,1) = 2 - 1 - 1 + 0 = 0$$

После аналогичного вычисления $H(k,l)$ и перемножения $S(k,l) = S(k,l) H(k,l)$:

$$S(k,l) = \begin{bmatrix} 2 & 0 \\ 4 & 2 \end{bmatrix}, \quad H(k,l) = \begin{bmatrix} 1 & -1 \\ 3 & 1 \end{bmatrix}, \quad Z(k,l) = \begin{bmatrix} 2 & 0 \\ 12 & 2 \end{bmatrix}.$$

После обратного ДПФ размера 2×2 получим результат циклической свертки: $z(n,m) = \begin{bmatrix} 3 & 2 \\ 4 & 3 \end{bmatrix}$.

Дополним опорные области $s(\cdot)$ и $h(\cdot)$ до размера 4×4 (для исключения искажения спектра, в принципе, достаточен размер 3×3), и повторим вычисления:

$$s(n,m) = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{vmatrix}, \quad h(n,m) = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{vmatrix}$$

$$S(k,l) = \begin{vmatrix} 3+j & 2 & 1+j & 2+2j \\ 2 & 1-j & 0 & 1+j \\ 3-j & 2-2j & 1-j & 2 \\ 4 & 3-j & 2 & 3+j \end{vmatrix}, \quad H(k,l) = \begin{vmatrix} 2+j & 1 & j & 1+2j \\ 1 & -j & -1 & j \\ 2-j & 1-2j & -j & 1 \\ 3 & 2-j & 1 & 2+j \end{vmatrix}$$

$$S(k,l) = \begin{vmatrix} 5+5j & 2 & -1+j & -2+6j \\ 2 & -1-j & 0 & -1+j \\ 5-5j & -2-6j & -1-j & 0 \\ 12 & 5-5j & 2 & 5+5j \end{vmatrix}, \quad s(n,m) = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 2 & 3 & 1 & 0 \end{vmatrix}$$

Сравнение данного результата с ДПФ размером 2 x 2 позволяет наглядно видеть эффект цикличности свертки.

В настоящее время имеются разнообразные и весьма эффективные алгоритмы ДПФ. Для прямого вычисления Р-мерного ДПФ требуется $(N_1 \times N_2 \times \dots \times N_P)^2$ операций умножения и сложения. Для многомерного ДПФ, как и для одномерного, существуют алгоритмы быстрых преобразований Фурье. Простейший из них в двумерном ДПФ - разбиение на строки и столбцы, который мы уже рассматривали. Аналогично, Р-мерное ДПФ может заменяться Р-операциями одномерных ДПФ, при этом общее количество операций умножения и сложения сокращается.

ЛИТЕРАТУРА

9. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. – М.: Мир, 1988. – 488 с.

[Главный сайт автора](#) ~ [Лекции по сигналам](#) ~ [Практикум](#)

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru. Буду благодарен.

Copyright ©2005 Davydov A.V.

ЛИТЕРАТУРА

1. Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов. - М.: Высшая школа, 1988. - 448 с.
2. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. – М.: Мир, 1989. – 540 с.
3. Васильев Д.В. Радиотехнические цепи и сигналы: Учебное пособие для вузов. - М.: Радио и связь, 1982. - 528 с.
4. Гольденберг Л.М. и др. Цифровая обработка сигналов: Справочник. - М.: Радио и связь, 1985. - 312 с.
5. Гольденберг Л.М. и др. Цифровая обработка сигналов: Учебное пособие для вузов. - М.: Радио и связь, 1990. - 256 с.
6. Горбенко Л.А., Месенжник Я.З. Кабели и провода для геофизических работ. - М.: Энергия, 1977. - 192 с.
7. Гроднев И.И., Фролов Н.А. Коаксиальные кабели связи. - М.: Радио и связь, 1983. - 209 с.
8. Гутников В.С. Фильтрация измерительных сигналов. – Л.: Энергоатомиздат, 1990. – 192 с.
9. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. – М.: Мир, 1988. – 488 с.
10. Дмитриев В.И. Прикладная теория информации: Учебник для вузов. - М.: Высшая школа, 1989. - 325 с.
11. Зиновьев А.Л., Филиппов Л.И. Введение в теорию сигналов и цепей: Учебное пособие для вузов. - М.: Высшая школа, 1975. - 264 с.
12. Игнатов В.А. Теория информации и передачи сигналов. - М.: Советское радио, 1979.
13. Канасевич Э.Р. Анализ временных последовательностей в геофизике. - М.: Недра, 1985. - 300 с.
14. Купер Дж., Макгиллем А. Вероятностные методы анализа сигналов и систем. – М.: Мир, 1989. – 376 с.
15. Лосев А.К. Линейные радиотехнические цепи: Учебник для вузов. - М.: Высшая школа, 1971. - 560 с.
16. Макс Ж. Методы и техника обработки сигналов при физических измерениях: В 2-х томах. - М.: Мир, 1983.
17. Никитин А.А. Теоретические основы обработки геофизической информации: Учебник для вузов. - М.: Недра, 1986. - 342 с.
18. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. – М.: Связь, 1979. – 416 с.
19. Отнес Р., Эноксон Л. Прикладной анализ временных рядов. – М.: Мир, 1982. – 428 с.
20. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. – М.: Мир, 1978. – 848 с.
21. Рапопорт М.Б. Вычислительная техника в полевой геофизике: Учебник для вузов. - М.: Недра, 1993. - 350 с.
22. Сиберт У.М. Цепи, сигналы, системы. – М.: Мир, 1988. – 336 с.
23. Стрижевский Н.З. Коаксиальные видеолнии. – М.: Радио и связь, 1988. – 200 с.
24. Хемминг Р.В. Цифровые фильтры. – М.: Недра, 1987. – 221 с.
25. Сергиенко А.Б. Цифровая обработка сигналов. – СПб.: Питер, 2003. – 608 с.
26. Вероятностные методы в вычислительной технике: Учебное пособие для вузов./ А.В.Крайников и др. - М.: Высшая школа, 1986. - 312 с.
27. Гурский Е.И. Теория вероятностей с элементами математической статистики: Учебное пособие для вузов. - М.: Высшая школа, 1971.- 328 с.
28. Игнатов В.А. Теория информации и передачи сигналов. - М.: Советское радио, 1979.
29. Сато Ю. Обработка сигналов. Первое знакомство. – Изд.: ДОДЭКА, 2002.
30. Корн Г., Корн Е. Справочник по математике для научных работников и инженеров. – М.: Наука, 1984.

О замеченных опечатках, ошибках и предложениях по дополнению: davpro@yandex.ru.

Copyright ©2005 Davydov A.V.

