

1. ПОНЯТИЕ О ЦИФРОВОЙ ОБРАБОТКЕ СИГНАЛОВ

1.1. Задачи, решаемые при помощи цифровой обработки сигналов

Прежде всего, необходимо определиться с тем, что же такое сигнал. В зависимости от контекста этому понятию могут придаваться различные значения. Вообще говоря, сигнал – это изменение некоторой физической величины, например, напряжения. В зависимости от области определения говорят о временной, частотной, пространственной форме представления сигнала. Если область определения непрерывна, то говорят о непрерывном, или аналоговом сигнале. В противном случае сигнал – дискретный. Название «аналоговый» дано непрерывным сигналам, потому что они «аналогичны» физическим процессам, встречающимся в действительности. Примером дискретного сигнала может служить, например, совокупность значений напряжения, измеряемых с секундным интервалом. В этом случае сигнал определен лишь в дискретные, то есть отдельные моменты времени.

Область значений сигналов также может быть непрерывной или дискретной. Стрелочный вольтметр – пример прибора, измеряющего непрерывную величину. Записывая его показания с определенной точностью, мы, тем самым, дискретизируем сигнал напряжения по уровню. Процесс дискретизации сигнала по уровню называется квантованием. Сигнал, у которого область определения и область значений дискретны, называется цифровым. Именно с такими сигналами обычно имеют дело в цифровой обработке сигналов (ЦОС) – науке, изучающей методы и алгоритмы обработки цифровых сигналов.

Цифровой сигнал можно представить в виде массива чисел – обычно одномерного или двумерного. Примером одномерного сигнала является речь, примером двумерного – изображение. Вы можете возразить, что и речь и изображение непрерывны. Как же можно представлять их счетной последовательностью чисел? Согласны с Вами, но добавим, что в большинстве глав этой книги (как и обычно в ЦОС) мы будем рассматривать сигналы на выходе аналого-цифрового преобразователя (АЦП), который выполняет операции дискретизации по времени, квантования по уровню и кодирования. Выполнение этих трех операций и связанные с ними проблемы будут рассмотрены в Главе 2. Будем считать, что при их выполнении исходный сигнал не искажается (хотя это и не так). Примером АЦП может являться звуковая карта компьютера.

Сигнал на входе АЦП является непрерывным по уровню и по времени. После дискретизации он становится дискретным по времени, а после квантования – дискретным по уровню. Можно было бы выполнить операцию квантования в первую очередь. Тогда сигнал был бы непрерывным по времени, но дискретным по уровню. Подробнее сигналы описаны в п.1.3.

Элементы массива сигнала обычно называются отсчетами, а если речь идет об изображении, то отсчеты чаще именуют пикселями. Значения отсчетов сигнала дискретны. Их максимальная величина зависит от разрядности АЦП. Для кодирования речи обычно применяется 12-разрядный АЦП, рассматриваемые в книге изображения являются 8-разрядными (256 градаций серого).

Множество задач, решаемых цифровой обработкой сигналов, можно разделить на две большие группы: анализ сигналов и их синтез. Задачей анализа сигналов является выделение небольшой группы значимых параметров, максимально полно описывающих сигнал. Обычно полученные параметры подвергаются обработке. Если анализ сигналов проводился с целью, например, их классификации, то на этом все заканчивается. Принадлежность полученных параметров к некоторому классу позволяет с большой долей уверенности говорить о принадлежности к этому же классу и анализируемого сигнала. Вот несколько примеров анализа сигналов, реально встречающихся в жизни. Все эти процедуры должны выполняться автоматически.

- Определение типа объекта по некоторой совокупности сигналов. Например, таких сигналов, как шум двигателя, визуальные очертания объекта, изображение на экране радара. Для решения этой задачи необходимо обрабатывать как одномерные, так и двумерные сигналы.
- Анализ речевого сигнала с целью идентификации говорящего.
- Выделение речевого сигнала, наблюдаемого в условиях преднамеренных помех. Такая задача возникает при попытке внешнего прослушивания разговора в помещении, которое «защищено», например, генератором шума.
- Определение параметров перехваченного радиосигнала: частоты излучения, типа модуляции, начала и конца посылок. Здесь задача осложняется тем, что сигнал наблюдается в условиях сильных шумов, так как радиолиния не настроена: ведь сигнал не предназначался для нас.
- Анализ сейсмических сигналов помогает предсказать землетрясения.
- Изучение отклика геодезического зонда позволяет обнаружить полезные ископаемые и определить характеристики залежей.
- Анализ электрокардиограммы способствует диагностированию болезней, в том числе на ранних стадиях их возникновения.
- Анализ свойств материалов (рентгенография, спектроскопия...).
- Анализ трафика в локальных и глобальных сетях передачи данных. (Да- да, и здесь применяются методы цифровой обработки сигналов).
- Анализ временных рядов (колебания численности населения, курсов акций...).

Как видите, задач анализа сигналов, реально требующих решения - множество. Здесь перечислены лишь немногие из них. Можно сформулировать общие требования к средствам анализа сигналов:

- средства анализа не должны требовать вмешательства в процесс человека;
- анализ должен осуществляться быстро. В каждом конкретном случае под словом «быстро» можно понимать разные вещи. Например, анализ курса акций допустимо проводить в течении нескольких часов, тогда как анализ радиоизлучений в эфире необходимо осуществлять в реальном масштабе времени (и по несколько излучений одновременно);
- извлекаемые из сигнала параметры должны быть физически значимы. Поэтому, наибольшее распространение получил частотный (спектральный) анализ, а также современные методы анализа – частотно-временной, масштабно-временной и их комбинации. (К сожалению, вынуждены вас предупредить: эти методы анализа в этом учебнике рассмотрены не будут).

Часто имеет место процедура, обратная анализу - синтез сигналов по совокупности описывающих их параметров. В этом случае целью обработки сигналов может быть, например:

- более экономное представление сигнала (его сжатие);
- удаление из сигналов нежелательных шумов;
- улучшение качества звучания речи, музыки, модификация ее спектра ;
- изменение свойств изображения (контраста, цветовой насыщенности, резкости).

В результате синтеза получается реконструированный (восстановленный) сигнал. Если реконструированный сигнал является точной копией исходного (возможно, лишь задержанной на некоторое время и умноженной на константу), то схема анализа-синтеза называется схемой с полным восстановлением. Схемы с полным восстановлением хорошо изучены в теории и почти не встречаются на практике. Причина этого в том, что в большинстве случаев значения коэффициентов преобразования изменяются непрерывно. Числа же в компьютере имеют ограниченный диапазон представления. Поэтому, даже и в отсутствии обработки коэффициентов реконструированный в процессе синтеза сигнал может не являться точной копией исходного сигнала. На практике обычно можно говорить о «почти полном» восстановлении. При этом на разность между восстановленным и исходным сигналом (ошибку) накладываются какие-либо ограничения. Например, при кодировании аудио эти ограничения могут заключаться в том, чтобы сигнал ошибки был неслышим. При этом восстановленный сигнал может существенно отличаться от исходного в математическом смысле, но быть неразличим с ним человеческими органами слуха.

Анализ и синтез сигналов может проводиться различными способами. Одним из наиболее важных инструментов являются линейные преобразования сигналов, заключающиеся в умножении сигналов на некоторую матрицу. В этом случае процедуры анализа и синтеза могут быть симметричными (то есть заключаться в умножении на матрицу, может быть даже на одну и ту же). Процесс анализа называется прямым преобразованием, синтеза - обратным преобразованием. Числа, полученные в результате умножения на матрицу прямого преобразования, называются коэффициентами преобразования или коэффициентами трансформанты. В главе 6 будет рассмотрен важный частный случай линейного преобразования – дискретное преобразование Фурье.

Важнейшим инструментом анализа, да и синтеза сигналов является фильтр – линейная времянезависимая система. Цифровой фильтр может рассматриваться как последовательность чисел – коэффициентов фильтра. Нахождение этих коэффициентов, исследование их чувствительности к округлениям и т.д. составляет едва ли не большую часть теории ЦОС. Фильтры будут изучаться в 4 и 5 главах. Понятно, что процесс фильтрации сигнала также может быть описан в терминах умножения сигнала на матрицу.

Из сказанного выше можно сделать вывод о том, что наука ЦОС тесно связана с линейной алгеброй. Для успешного освоения теории и алгоритмов обработки сигналов вы должны хорошо ориентироваться в методах работы с векторами и матрицами. Другой основой ЦОС является радиотехника – наука, предметом которой является обработка аналоговых сигналов. Многие способы обработки аналоговых и цифровых сигналов схожи. Иногда бывает достаточно лишь заменить знак интеграла суммированием, а дифференциальные уравнения - разностными. Перед изучением ЦОС важно освежить свои знания и в теории связи. Ведь и само понятие сигнала было введено связистами.

Особенностью науки ЦОС является то, что она гораздо «ближе к железу», чем многие другие. От специалиста ЦОС часто требуется знание внутреннего устройства (архитектуры) реализующей алгоритм микросхемы. Многие методы разрабатываются с учетом конкретных схемотехнических решений. Поэтому в учебнике описаны особенности работы с цифровыми процессорами обработки сигналов (ЦПОС) и программируемыми логическими интегральными схемами (ПЛИС).

Еще три-четыре года назад реализовать свой алгоритм на ЦПОС вы могли, лишь досконально изучив язык ассемблера конкретного процессора. Однако, развитие программных средств в последние годы привело к тому, что в настоящее время достаточно сложные задачи можно программировать на языках высокого уровня – на Си и даже на С++. Современные компиляторы и компоновщики генерируют код для ЦПОС, достаточно эффективный в большинстве случаев. Использование языков программирования высокого уровня позволяет инженеру сосредоточиться на содержательной части алгоритма, существенно уменьшает время создания приложений. В главе 8 будет рассмотрена самая современная среда разработки для процессоров фирмы Analog Devices – Visual DSP++2.0.

И еще одно замечание. Предположим, Вы разработали алгоритм «на бумаге» и хотите проверить его эффективность. Что делать? Можно, конечно, написать программу на C/C++. Однако, гораздо более простой и быстрый путь заключается в использовании для моделирования системы Matlab. Во всем мире Matlab является стандартом де-факто при экспериментах с алгоритмами ЦОС. Она включает в себя богатейшие библиотеки готовых функций и процедур, в том числе и для обработки сигналов, изображений. В последней версии этой системы (на момент написания книги это V.6.1) имеется даже библиотека написания приложений для ЦПОС фирмы Texas Inst.. Входной язык программирования этой среды во многом похож на Си. Кроме того, в состав Matlab входит Simulink, в котором реализована концепция графического программирования, позволяющая создавать программы путем рисования блок-схем алгоритмов. Классический путь создания новых устройств такой: Matlab -> С -> ассемблер ЦПОС.

Итак, специалист ЦОС должен быть как хорошим математиком, так и опытным программистом, практиком. Кроме того, ему необходимы знания предметной области. Для того чтобы вы представили себе разнообразие решаемых при помощи ЦОС задач, приведем перечень основных направлений работы 4-й Международной конференции «Цифровая обработка сигналов и ее применение», прошедшей в 2002 г. в Москве:

- теория сигналов, методы кодирования, алгоритмы преобразований, цифровая фильтрация и спектральный анализ;
- цифровая обработка и передача многомерных сигналов;
- цифровая обработка и передача речевых и звуковых сигналов;
- цифровая обработка и передача сигналов в системах телекоммуникаций, радиолокации и гидроакустике;
- цифровая обработка сигналов в системах управления;
- цифровая обработка и передача изображений;
- цифровая обработка и передача измерительной информации;
- цифровое телерадиовещание;
- нейросетевая обработка сигналов;
- защита информации и цифровая обработка сигналов;
- элементная база, аппаратные средства и системы ЦОС;
- проектирование систем ЦОС;
- программно-аппаратные средства ЦОС на базе стандартов MPEG, DAB, DVB, W-CDMA и других;
- IP-телефония, видеоконференцсвязь, мультимедиа, Интернет;
- ЦОС в учебном процессе.

1.2. Функциональная схема ЦОС. Сравнение между аналоговыми и дискретными методами обработки сигналов.

При всем разнообразии решаемых ЦОС задач последовательность их решения обычно одна и та же. Это позволяет говорить о некоторой унифицированной схеме обработки сигналов, которая изображена на Рис.1.1.

Как видно из Рис.1.1, основным элементом схемы ЦОС является вычислитель. В качестве него чаще всего применяется процессор, специально спроектированный для решения задач цифровой обработки – ЦПОС. Иногда более выгодным оказывается использование программируемой логики - ПЛИС. Вычислитель интенсивно использует для работы память, в которой хранятся необходимые константы, промежуточные и окончательные результаты вычислений (память данных), а также сами исполняемые команды (память программ). В зависимости от архитектуры вычислителя память может быть единой либо делиться по функциональному предназначению, внутренней (встроенной в вычислитель) либо внешней.

Так как реальные сигналы обычно аналоговые, то для обработки их в вычислителе необходимо предварительное выполнение операции аналогово-цифрового преобразования, выполняемой АЦП. АЦП обычно представляет собой отдельную микросхему, хотя иногда и встраивается в ЦПОС (примером является ADSP-2151msp).

Результат обработки сигнала, выдаваемый пользователю, также обычно аналоговый. Например, это может быть очищенная от шумов речь или музыка после цифрового эквалайзера. Задачу цифроаналогового преобразования выполняет ЦАП, как и АЦП выполняемый обычно в виде отдельной микросхемы.

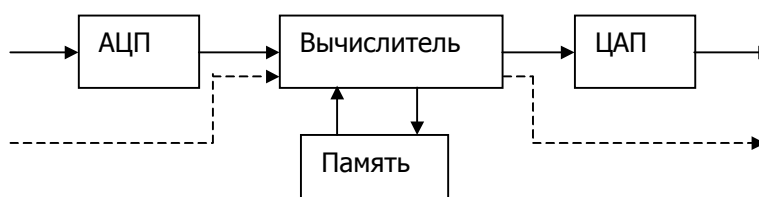


Рис.1.1. Функциональная схема цифровой обработки сигналов.

В некоторых случаях выходной сигнал схемы ЦОС остается в дискретном виде (например, цифровой индикатор). Также и на входе схемы сигнал может быть цифровым. В этих случаях устройства АЦП/ЦАП не нужны.

Надо отметить, что в современных устройствах цифровые методы обработки сигналов соседствуют с аналоговыми. Причина этого вовсе не в какой-то там «отсталости» или консерватизме разработчиков, а в том, что аналоговые методы обработки могут быть лучше цифровых.

Приведем пример, показывающий преимущество аналоговых методов перед цифровыми. Пусть нам необходимо обмениваться речевыми сигналами по стандартному двухпроводному каналу тональной частоты. Аналоговая конструкция оконечного устройства совершенно проста – обычный телефонный аппарат. При этом требования к качеству канала связи минимальны. Теперь посмотрим, что мы получим в случае цифрового построения прибора. Пусть отсчеты речи берутся с частотой 8кГц, каждый отсчет кодируется 12 битами. Получается цифровой поток 96кбит/с. Ясно, что по каналу с полосой пропускания 3,1кГц такой поток передать нельзя. Приходится применять сложные методы сжатия речи, модуляции/демодуляции, помехоустойчивого кодирования. Получившийся сигнал оказывается чувствительным к ошибкам в канале связи. Однако, если нам нужны какие-либо дополнительные сервисы, то альтернативы цифровым методам передачи не будет. Например, для засекречивания информации с гарантированной стойкостью, передачи помимо речи других видов информации, архивации сообщений с возможностью быстрого доступа разработчик будет вынужден прибегнуть к цифровой конструкции оконечного аппарата.

Другой причиной использования наряду с цифровыми аналоговых методов является то, что, как мы увидим в Главе 2, перед АЦП сигнал должен быть ограничен по спектру фильтром. Возникает вопрос об оптимальном соотношении цифровой и аналоговой фильтрации. Оптимальном с точки зрения минимизации потребляемой мощности или занимаемой микросхемой площади. Этот вопрос был детально исследован японскими учеными, и результаты их исследований приведены на графике (см. Рис.1.2.)

Например, в GSM полоса пропускания канала составляет 100кГц, а требуемое число бит АЦП – 14. Значит, предпочтительнее применять цифровые методы фильтрации. В DECT полоса пропускания 700кГц, требуемое разрешение – 12 бит. Здесь выгоднее применять смешанные методы фильтрации.

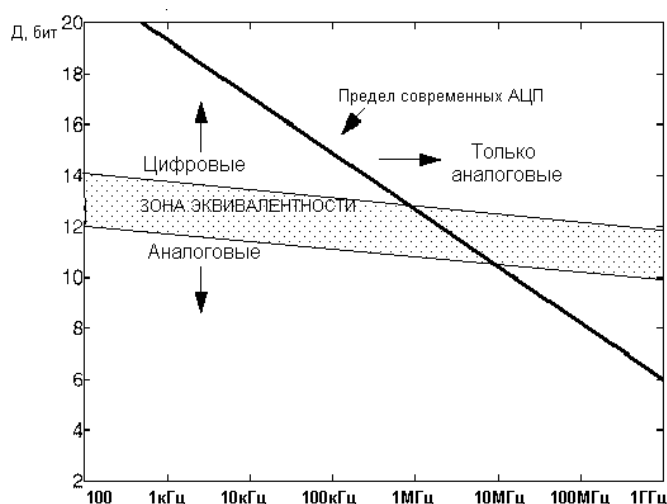


Рис.1.2. Предпочтительные области применения аналоговых и цифровых методов.

1.3. Аналоговые и дискретные сигналы

Сигналы описываются математическими функциями. Аналоговые сигналы чаще всего описываются непрерывной функцией $x(t)$, где t - время. Для обозначения сигналов используются строчные буквы. Если функция $x(t)$ принимает любые значения, то это – сигнал непрерывный по времени, но дискретный по уровню. Дискретные по времени сигналы описываются функцией $x[nT]$, где T - период между отсчетами сигнала, n - номер отсчета (может принимать и отрицательные значения). Отметим, что аргумент дискретных сигналов записывается в квадратных скобках, а аналоговых - в круглых.

Часто значение T несущественно и опускается при описании сигнала: $x[n]$. Если функция $x[n]$ может принимать любые значения, то это - сигнал дискретный по времени, но непрерывный по уровню. Дискретный по времени и уровню сигнал называется цифровым. На Рис.1.3 представлены все четыре возможных типа сигнала.

Далее будем рассматривать только цифровые сигналы, если это не оговорено особо. Цифровой сигнал называют еще последовательностью. По аналогии с функциями существуют различные способы задания последовательностей: аналитический (формульный), графический, табличный и комбинация этих способов.

Рассмотрим некоторые важные примеры сигналов.

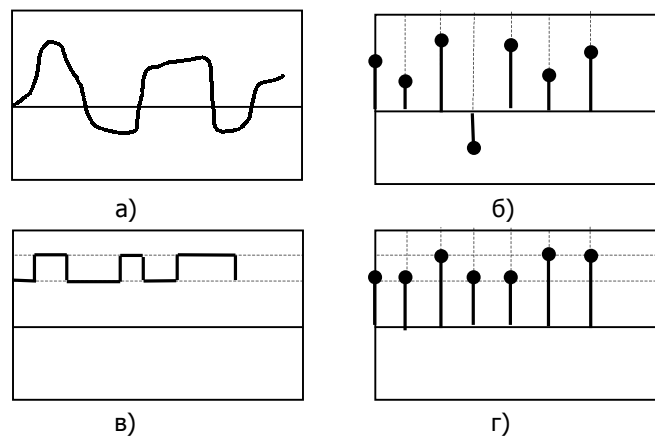


Рис.1.3. Виды сигналов:

**а) аналоговый; б) непрерывный по уровню и дискретный по времени;
в) непрерывный по времени и дискретный по уровню; г) цифровой.**

1. Единичный импульс (аналог дельта-функции Дирака) приведен на Рис.1.4(а). Он задается выражением

$$\delta[n] = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases} \quad (1.1)$$

Значение единичного импульса состоит в том, что любой сигнал может быть представлен, как суперпозиция единичных импульсов, как мы увидим в п.1.4. Кроме того, импульсная характеристика цифрового фильтра определяется как реакция фильтра на воздействие единичного импульса (см.п.4.2).

2. Единичный скачок (ступенчатый сигнал) показан на Рис.1.4(б); ему соответствует выражение

$$x[n] = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0. \end{cases} \quad (1.2)$$

3. Сдвиг последовательности $x[n]$ на k отсчетов записывается в виде $y[n] = x[n - k]$. При этом, если $k > 0$, то сдвиг будет вправо (рис.1.4(в)), если $k < 0$, то влево.

4. Периодической последовательностью называют последовательность, удовлетворяющую условию $x[n] = x[n + mN]$, где m, N - натуральные числа, N - период последовательности. Из определения периодической последовательности следует, что ее достаточно задать на одном интервале $[0; N - 1]$. Периодическая последовательность показана на рис.1.4(г).

Одной из основных проблем цифровой обработки сигналов является нахождение компактного (разреженного) представления сигнала, наилучшим образом подходящего для той или иной задачи. Наиболее применимы аддитивные модели сигнала, то есть представление сигнала в виде взвешенной суммы коэффициентов преобразования

$$x[n] = \sum_{i=0}^{N-1} a_i g_i[n], \quad (1.3)$$

где a_i - коэффициенты, $g_i[n]$ - векторы преобразования.

Векторы преобразования – это элементарные сигналы, строительные кирпичики, из которых строятся реальные, сложные сигналы. В этом случае говорят о разложении или декомпозиции сигнала по векторам преобразования. Чем больше эти кирпичики похожи на анализируемый сигнал, тем больше будет малых по величине коэффициентов преобразования (или, коэффициентов трансформанты). Тогда ими можно пренебречь и использовать для синтеза небольшое количество коэффициентов, значения которых превышают некоторый порог. На этом основано большинство алгоритмов сжатия сигналов.

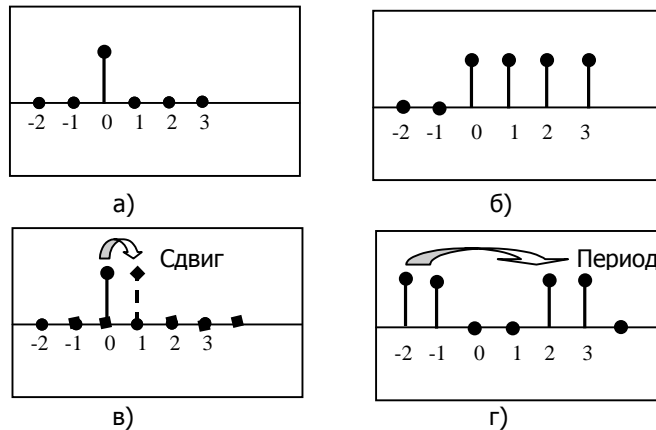


Рис.1.4. Примеры дискретных сигналов:
а) единичный импульс; б) единичный скачок;
в) сдвиг последовательности; г) периодическая последовательность.

Вообще говоря, декомпозиция (1.3) является основой большинства алгоритмов цифровой обработки сигналов. При этом вначале выполняется декомпозиция, затем коэффициенты преобразования подходящим образом обрабатываются, затем выполняется синтез сигнала.

Величина N в выражении (1.3) – это количество отсчетов векторов трансформанты. Оно необязательно равно числу отсчетов в исходном сигнале. Если векторов трансформанты больше, то говорят об избыточном (а иногда - о передискретизированном) представлении. Если векторов трансформанты меньше, то дискретный сигнал в общем случае точно представить нельзя. В случае, когда вектора преобразования образуют базис, их число равно числу отсчетов сигнала. Модели сигналов, задаваемые равенством (1.3) могут быть разделены на две группы: параметрические и непараметрические. В непараметрических моделях векторы преобразования $g[n]$ фиксированы и не зависят от параметров анализируемого сигнала. Примером таких моделей может послужить представление сигналов в любом базисе. Параметрические модели сигналов исключительно успешно применяются, например, при сжатии речи. В отрезке речевого сигнала тем или иным способом выделяются частота основного тона, признак тон/шум и другие параметры, которые кодируются. Декодер способен восстановить переданный сигнал на основе полученных параметров.

Промежуточное положение между параметрическими и непараметрическими представлениями занимают адаптивные преобразования сигналов, когда значения $g[n]$ выбираются в зависимости от сигнала из большой (но фиксированной!) библиотеки векторов.

1.4. Линейные системы в цифровой обработке сигналов

Под системой в ЦОС понимается любое устройство, порождающее отклик в ответ на воздействие. Важным классом систем являются линейные системы, которые обладают двумя свойствами: однородностью и аддитивностью. Методы анализа и синтеза устройств ЦОС во многом базируются на теории линейных систем. Это объясняется их простотой, хорошей изученностью, а также тем, что в рамках линейной модели существует возможность учета и нелинейных эффектов. С линейными системами вы уже познакомились в курсе линейных радиотехнических устройств. Канал связи также является линейной системой.

Свойство однородности означает, что при умножении воздействия на некоторую константу, отклик системы также будет умножен на эту константу, то есть

$$\text{если } y[n]=T(x[n]), \text{ то } ky[n]=T(kx[n]), \quad (1.4)$$

где через T обозначено производимое системой действие. Выражение (1.4) означает, что зависимость выхода от входа в линейной системе описывается уравнением прямой, например, $u = Ri$ - закон Ома.

Другим свойством линейных систем является свойство аддитивности. Пусть реакция системы на воздействие $x_1[n]$ есть $y_1[n]$, а на воздействие $x_2[n]$ - $y_2[n]$. Тогда реакция системы на сумму воздействий есть сумма реакций:

$$T(x_1[n]+x_2[n])=y_1[n]+y_2[n]. \quad (1.5)$$

Свойства однородности и аддитивности могут быть объединены вместе в один, важнейший принцип, соблюдающийся в линейных системах – принцип суперпозиции (наложения):

$$T(ax_1[n]+bx_2[n])=ay_1[n]+by_2[n]. \quad (1.6)$$

Таким образом, если объединение сигналов происходило путем их масштабирования и сложения, они всегда могут быть разделены. Представление сигнала в виде взвешенной суммы элементарных сигналов (1.3) возможно именно благодаря принципу суперпозиции.

При построении систем часто требуется свойство инвариантности к сдвигу (времянезависимости). Это свойство означает, что сдвиг входной последовательности на n отсчетов приведет к соответствующему сдвигу на n отсчетов реакции системы. Линейная система может быть как инвариантной к сдвигу, так и неинвариантной. Примером линейной времянезависимой системы является фильтр. Примером системы, чьи свойства изменяются во времени, могут служить банки фильтров в многоскоростных системах ЦОС.

В качестве элементарных строительных блоков в выражении (1.3) чаще всего выступают тригонометрические функции. Причина этого заключается в том, что они являются собственными функциями линейных систем. Линейная система не изменяет формы этих функций: если на вход системы подать синусоидальный сигнал с частотой f , то и на выходе будет синусоидальный сигнал с той же частотой и, возможно, другой амплитудой. Никакие другие функции не обладают таким свойством.

Линейные системы обладают свойством коммутативности. В случае каскадного соединения нескольких таких систем порядок их следования не имеет значения.

В реальности большинство систем нелинейные, хотя бы из-за эффектов округления вещественных чисел. Однако, обычно этой нелинейностью пренебрегают, аппроксимируя реальную систему линейной. Иногда существует возможность рассматривать нелинейную систему в небольшом диапазоне изменения входного воздействия, в пределах которых систему можно считать линейной. Наконец, нелинейную операцию перемножения сигналов можно заменить линейной, после взятия логарифма произведения. Это – так называемая гомоморфная обработка сигналов, широко применяемая в ЦОС речи и изображений.