ГЛАВА 5

ЛИНЕЙНЫЕ СИСТЕМЫ

Большинство методов ЦОС основано на стратегии «разделяй и властвуй», называемой *суперпозиция*. При обработке сигнал разделяется на простые компоненты, каждая из которых обрабатывается индивидуально, а результат воссоединяется. Такой подход обладает огромной мощностью, деля одну сложную проблему на несколько более легких. Метод суперпозиции может применяться только к *линейным системам*. Под этим термином понимаются определенные математические правила. К счастью, большинство научных и инженерных приложений попадают в эту категорию. Эта глава описывает основные принципы: что необходимо, чтобы система была линейной, различные способы разделения сигналов на простейшие компоненты, и как суперпозиция применяется в методах цифровой обработки.

Сигналы и системы

Сигнал является описанием, как один параметра изменяется вместе с другим. Например, в электронных схемах напряжение изменятся во времени, или яркость изменяется с расстоянием до изображения. *Система* является процессом, который производит *выходной сигнал* в ответ на *входной сигнал*. Это проиллюстрировано на Рис.5-1. Непрерывные системы вводят и выводят непрерывные сигналы, как это происходит в аналоговой электронике. Дискретные системы вводят и выводят дискретные сигналы, как компьютерные программы, которые манипулируют со значениями, хранящимися в массивах.

Существует несколько правил присваивания имен сигналам. Им не всегда следуют в ЦОС, но они широко распространены и должны быть вами запомнены. Математика будет очень сложной без четкой системы обозначений. Во-первых, для обозначения *непрерывных* сигналов используют круглые скобки, например, x(t) и y(t), в то время как для дискретных сигналов используют квадратные скобки: x[n] и y[n]. Во-вторых, для обозначения сигналов используют буквы в нижнем регистре. Буквы в верхнем регистре зарезервированы для обозначения величин в частотной области. В-третьих, имя, присваиваемое сигналу, обычно описывает параметры, которые он представляет. Например, *напряжение*, зависящее от *времени*, записывают как v(t) и т.п.

Сигналы и системы часто обсуждаются без указания конкретных параметров, которые они представляют. Это похоже на использование x и y в алгебре, без спецификации физического значения переменных. Четвертое правило гласит: если описательное имя не доступно, входной сигнал дискретной системы обозначают через x[n], а выходной через y[n]. Для сигналов непрерывных систем используют обозначения: x(t) и y(t).

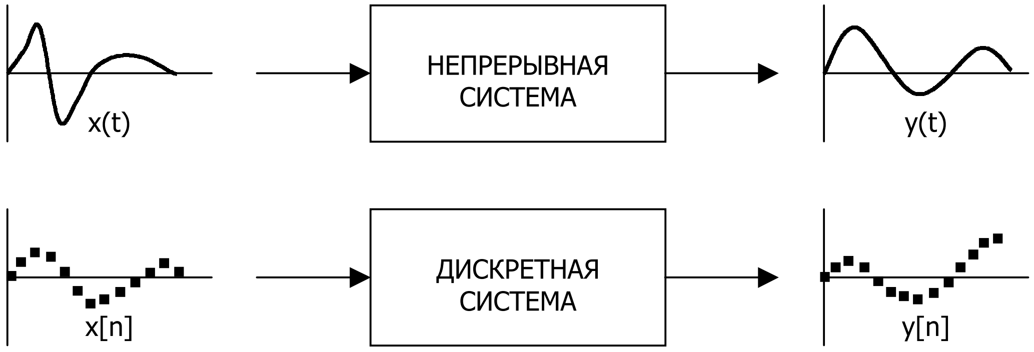


Рис.5-1. Терминология для сигналов и систем.

Существует несколько причин для понимания *системы*. Например, вы хотите *создать* систему для устранения шума в электрокардиограмме, уточнения нечеткого изображения или устранения эхосигналов при звукозаписи. В других случаях система может иметь искажения или помехи, которые необходимо охарактеризовать или измерить. Например, когда вы говорите по телефону, вы встречаетесь с тем, что другой человек слышит что-то похожее на ваш голос. К сожалению, входной сигнал при прохождении через линию передачи редко похож на выходной. Если вы знаете как линия передачи (или система) изменяет сигнал, существует возможность компенсировать этот эффект.

Кроме того, система может представлять собой некоторый физический процесс, который вы хотите изучить и проанализировать. Хорошим примером являются системы радио- и гидролокации. Эти методы основаны на сравнении переданных и отраженных сигналов для нахождения характеристик удаленного объекта. В терминах теории систем, проблема состоит в том, чтобы найти как система меняет переданный сигнал в принятый.

На первый взгляд, понять все существующие в мире системы кажется неразрешимой задачей. К счастью, большинство используемых систем попадают в категорию, называемую *линейные системы*. Это очень важно. *Без* концепции линейной системы, пришлось бы изучать индивидуальные характеристики несвязанных систем. При использовании этого подхода можно сконцентрировать внимание на особенностях линейных систем как единого целого.

Нашей первой задачей является определение свойств, которые делают систему линейной, и как они представляются в электронике, программном обеспечении и других системах обработки сигналов.

Необходимые условия линейности

Система называется *линейной*, если она обладает двумя математическими свойствами: *однородностью* (homogeneity) и *аддитивностью* (additivity). Если вы можете показать, что система отвечает этим двум условиям, можно с уверенностью говорить, что система линейна. С другой стороны, если не доказано, что система обладает одним или обоими свойствами, говорят, что система нелинейная. Третье свойство, *инвариантность относительно смещения* (shift invariance), это нестрогое условие линейной системы, но для большинства методов ЦОС оно является обязательным. Если вы встречаете термин *линейная система* в ЦОС, вы должны понимать, что она включает в себя свойство инвариантности относительно смещения. Давайте рассмотрим эти три свойства с точки зрения математики.

Как показано на Рис.5-2, однородность подразумевает, что изменение в амплитуде входного сигнала отражается соответствующим изменением в амплитуде выходного сигнала. В математических терминах, сигнал x[n] отражается в выходном сигнале y[n], входной сигнал kx[n] отражается в выходном сигнале ky[n], для любого входного сигнала и постоянной k.

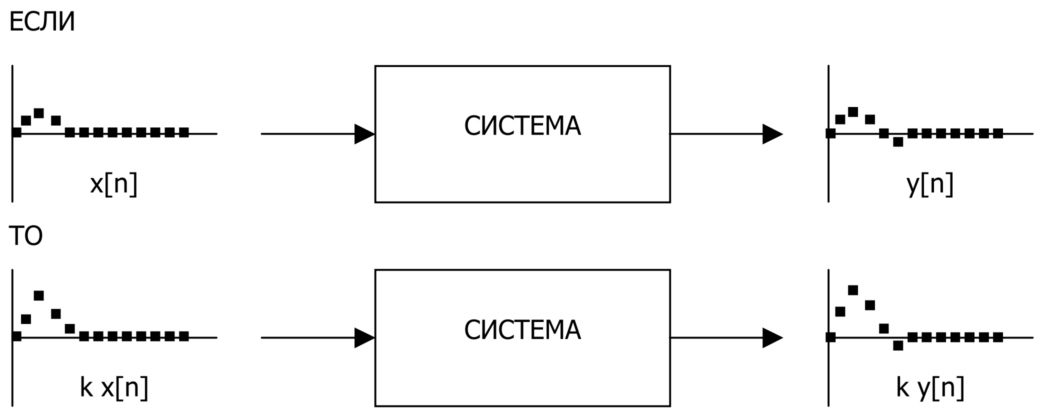


Рис.5-2. Определение однородности.

Хорошим примером однородных и неоднородных систем является простой резистор. Если входом системы является напряжение, проходящее через резистор, v(t), а выходом ток на резисторе, i(t), система линейна. Это гарантирует закон Ома; если напряжение уменьшается или увеличивается, произойдет соответствующее изменение тока. Теперь представьте другую систему, в которой входным сигналом является напряжение на резисторе, а выходным мощность, потребляемая на нем, p(t). Поскольку мощность пропорциональна квадрату напряжения, то входной сигнал, увеличенный в *два* раза, вызовет изменение выходного сигнала в *четыре* раза. Такая система неоднородна, и поэтому нелинейная.

Свойство аддитивности показано на Рис.5-3. Предположим, что имеется система, где входной сигнал $x_1[n]$ производит выходной сигнал $y_1[n]$. Далее, предположим, что имеется другой входной сигнал $x_2[n]$, который производит выходной сигнал $y_2[n]$. Система называется аддитивной, если входной сигнал $x_1[n]+x_2[n]$ производит выходной сигнал $y_1[n]+y_2[n]$. Другими словами, сигналы, сложенные на входе, производят сигналы, сложенные на выходе.

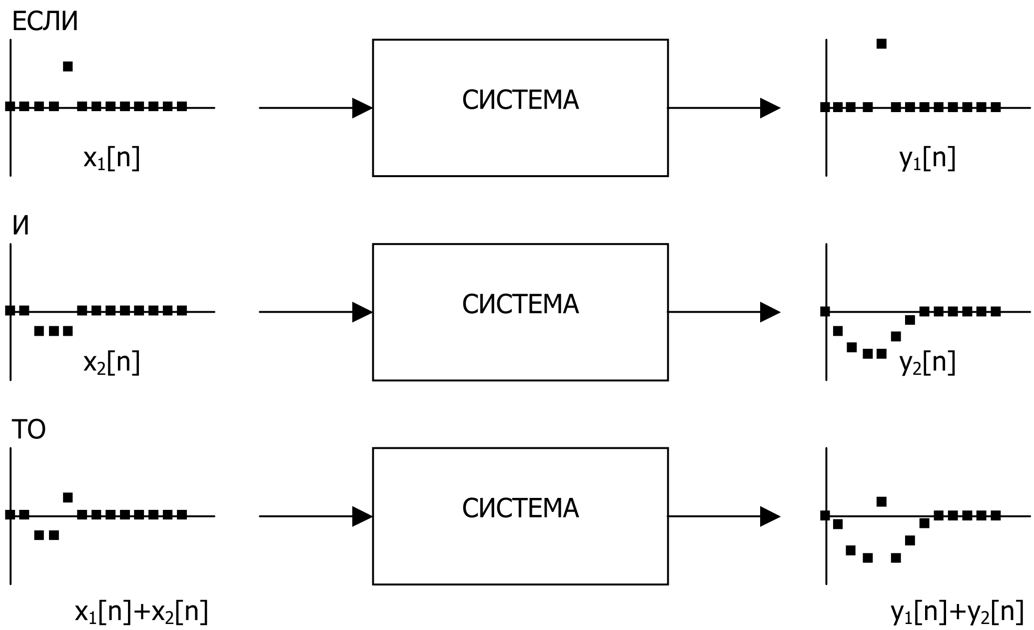


Рис.5-3. Определение аддитивности.

Важной деталью, является то, что сложенные сигналы проходят через систему без взаимодействия. В качестве примера рассмотрим следующий бытовой пример: телефонный разговор с вашими тетей Аней и дядей Борей. Тетя Аня начинает длинную историю о том, как ей плохо живется в этом году. На заднем фоне, дядя Боря кричит на собаку, которая испортила его любимое кресло. Два речевых сигнала суммируются и передаются по телефонной линии. Поскольку система аддитивна, звук, который вы услышите, является суммой двух голосов, и они будут звучать индивидуально. В результате вы слышите тетю Аню и дядю Борю, а не некоторое создание Аняборя :).

Хорошим примером неаддитивной схемы является стадия микшера в радиопередатчике. Присутствуют два сигнала: аудиосигнал, содержащий голос или музыку, и несущая, которая может распространяться в пространстве для приема антенной. Два сигнала складываются и приобретают нелинейность, например, через плоскостной диод. Это выражается в соединении сигналов в некоторый третий сигнал, модулированную радиоволну, способную передавать информацию на большие расстояния.

Как показано на Рис.5-4, инвариантность относительно смещения подразумевает, что сдвиг входного сигнала отражается в идентичном сдвиге выходного сигнала. В формальных терминах, если входной сигнал x[n] отражается в выходном сигнале y[n], то входной сигнал x[n+s] отразится в выходном сигнале y[n+s]. Обратите внимание на то, как записывается этот сдвиг, такая форма будет использоваться в следующих главах. Добавление константы s к независимой переменной x может вызвать опережение или запаздывание сигнала. Например, сигнал сдвигается на две выборки *влево*, если s=2, и на две выборки *вправо*, если s=-2.

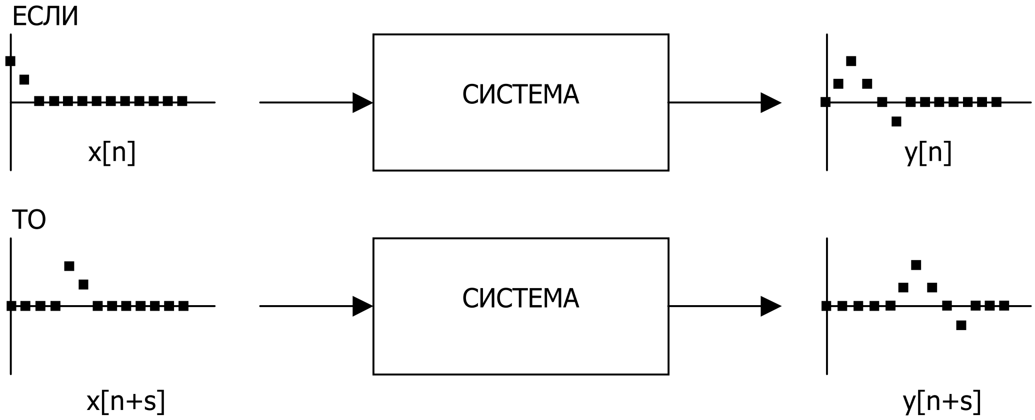


Рис.5-4. Определение инвариантности относительно смещения.

Инвариантность относительно смещения важна тем, что она означает, что свойства системы не изменяются во времени. Большинство встречаемых систем являются инвариантными относительно смещения. Это удачное обстоятельство, поскольку гораздо сложнее иметь дело с системами, характеристики которых меняются во время работы. Например, представьте, что вы создали цифровой фильтр, делающий речевой сигнал более чистым и компенсирующий эффекты ухудшения телефонной линии. Но как только наступила зима, вы узнаете, что характеристики телефонной линии изменились из-за температуры. Ваш компенсирующий фильтр бесполезен, поскольку не работает в этих условиях. Такая ситуация требует более сложных алгоритмов, которые могут *адаптироваться* к изменяющимся условиям.

Почему однородность и аддитивность являются критичными для линейных систем, в то время как инвариантность относительно смещения это что-то дополнительное. Ответ: поскольку линейность является более обширной концепцией, включающей в себя нечто большее, чем сигналы и системы. Например, представьте, что фермер продает апельсины по 2$ за ящик и яблоки по 5$ за ящик. Если он продает только апельсины, он получит 20$ за 10 ящиков, 40$ за 20 ящиков, это означает, что бартер *однородный*. Если он продаст 20 ящиков апельсинов и 10 ящиков яблок он получит: 20x2$ + 10x5$ = 90$, то же количество денег, если бы он продал их отдельно. Это означает, что сделка *аддитивна*. Однако, поскольку в этом примере нет сигналов и систем, инвариантность относительно смещения не имеет значения. Инвариантность относительно смещения является дополнительным аспектом в том случае, когда вопрос касается сигналов и систем.

Статическая линейность и синусоидальная точность

Однородность, аддитивность и инвариантность относительно смещения являются важными, поскольку они обеспечивают математическую основу для определения линейных систем. К сожалению, эти свойства не могут сказать, *что такое* линейная система. В этом могут помочь свойства *статической линейности* (static linearity) и *синусоидальной точности* (sinusoidal fidelity). Они не особенно важны с математической точки зрения, а относятся к тому, как человек воспринимает и понимает линейные системы. Обратите особенное внимание на этот раздел.

Статическая линейность определяет как действует линейная система, когда сигналы не изменятся, т.е. когда они потенциального типа или *статичны*. Статический отклик линейной системы очень прост: *выходной сигнал равен входному сигналу, умноженному на константу*. То есть, график возможных входных величин, построенный напротив соответствующих выходных, является прямой линией, проходящей через начало координат. На Рис.5-5 показаны две распространенные линейные системы. Закон Ома для резисторов и закон Гука для упругости. Для сравнения с ними, на Рис.5-6. показана статичная зависимость для двух нелинейных систем: p-n плоскостного диода и магнитных свойств железа.

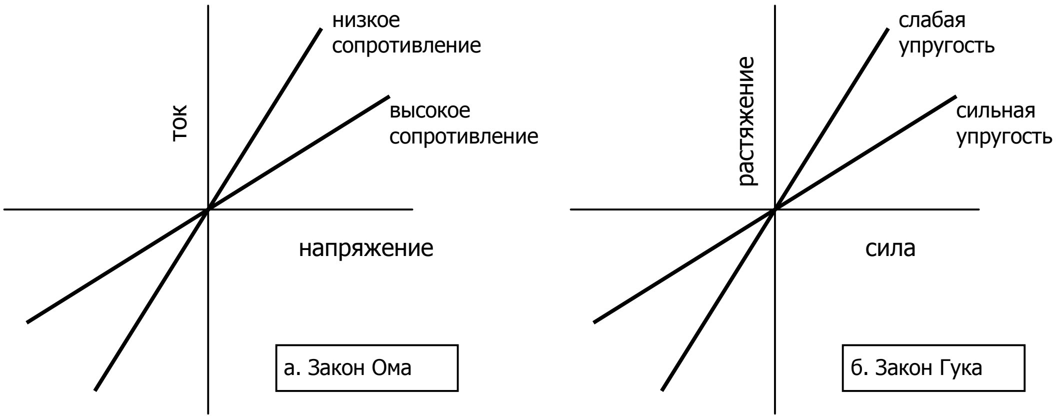


Рис.5-5. Два примера статичной линейности.

Все линейные системы обладают свойством статичной линейности. Обратное обычно тоже верно, но не всегда. Существуют системы, обладающие статичной линейностью, но не являющиеся линейными по отношению к изменяющимся сигналам. Однако широко распространенный класс систем может быть полностью изучен на основе только статичной линейности. В таких системах неважно, является ли входной сигнал статичным или изменяющимся. Они называются системами с *отсутствием последействия* (memoryless systems), поскольку в них выход зависит от текущего состояния входа, а не его истории. Например, мгновенный ток резистора зависит только от мгновенного напряжения, протекающего через него, а не от того, как сигнал пришел к такому значению. Если система обладает статичной линейностью и отсутствием последействия, то она должна быть линейной. Это позволяет понять (и доказать) линейность таких простых систем.

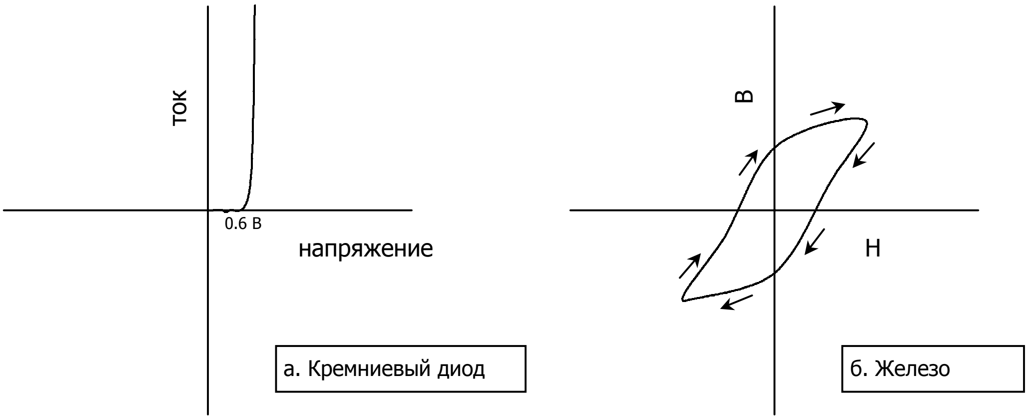


Рис.5-6. Два примера нелинейных систем.

Важной характеристикой линейных систем является их поведение при работе с синусоидальными сигналами, свойство, которое мы будем называть *синусоидальной точностью*. *Если на входе линейной системы синусоидальный сигнал, то на выходе будет также синусоидальная волна такой же частоты, как и входная.* Только синусоидальные сигналы обладают таким свойством. Например, не причин предполагать, что прямоугольный сигнал, поданный на вход линейной системы, будет воспроизводить на ее выходе также прямоугольный сигнал. Хотя синусоида на входе гарантирует синусоиду на выходе, она может отличиться по *амплитуде* и *фазе*. Это должно быть известно вам из знаний по электронике: цепь может быть описана ее *частотным откликом*, зависимостью коэффициента усиления и фазы от частоты.

Теперь рассмотрим обратный вопрос: если система всегда производит синусоидальный сигнал в ответ на синусоидальный сигнал на входе, является ли это гарантией того, что система линейна? Ответ отрицательный, но исключения редки и обычно очевидны. Например, представьте, что внутри системы скрыт демон, задача которого ввести вас в заблуждение. Демон использует осциллограф для исследования входного сигнала и синусоидальный генератор, производящий выходной сигнал. Когда на вход системы подается синусоидальный сигнал, демон быстро измеряет его частоту и настраивает свой генератор для воспроизведения соответствующего выходного сигнала. Конечно, система нелинейна, поскольку не аддитивна. Чтобы показать это, подайте на вход системы сумму двух синусоид. Демон может ответить только одночастотной синусоидой. Этот пример не такой запутанный, как вы можете подумать, *петля фазовой автоподстройки* действует по тому же принципу.

Для получения лучшего представления о свойстве линейности, представьте техника, который пытается определить, обладает ли электронный прибор линейностью. При синусоидальной волне на входе он будет ожидать соответствующую синусоиду на выходе. Например, выход не может быть срезанным сверху или снизу, верхняя половина не должна отличаться от нижней, не должно быть искажений в местах пересечения сигнала нулевого уровня и т.д. Следующим шагом техник будет изменять амплитуду сигнала и исследовать воздействие на выходной сигнал. Если система линейна, то амплитуда выходного сигнала должна следовать амплитуде на входе. Наконец, техник будет менять частоту входного сигнала и убеждаться в том, что соответственно изменяется и частота выходного. С изменением частоты соответствующим образом будут выглядеть изменения амплитуды и фазы выходного сигнала. На некоторых частотах выходной сигнал может быть даже нулевым, например, для сигнала с нулевой частотой. Если техник увидит все эти вещи, он придет к заключению, что система линейна. В то время как это заключение математически не строго, уровень достоверности достаточно высок.

Примеры линейных и нелинейных систем

В Таблице 5-1 приведены примеры наиболее распространенных линейных и нелинейных систем. При просмотре этого списка помните о математическом взгляде на свойство линейности (*однородность*, *аддитивность* и *инвариантность относительно смещения*). В то же время, большинство ученых и инженеров рассматривают линейные системы в терминах *статической линейности* и *синусоидальной точности*.

<u>Примеры *линейных систем*</u>

Распространение волн, например, звуковые и электромагнитные волны

Электрические схемы, состоящие из резисторов, конденсаторов и емкостей

Электронные схемы, например, усилители и фильтры

Механические перемещения от взаимодействия масс, пружин и амортизаторов (демпферов)

Системы, описанные дифференциальными уравнениями, например, цепи резистор-конденсатор-емкость

Умножение на постоянный коэффициент, например, усилители или аттенюаторы сигнала

Вариации сигнала, например, эхо, резонанс, размытость изображения

Системы неизменного режима, где выход всегда эквивалентен входу

Нуль-системы, где выход всегда равен нулю, несмотря на тип входного сигнала

Дифференцирующие и интегрирующие и аналогичные операции первой разности и текущей сумме для дискретных сигналов

Системы малых возмущений в противоположность нелинейным системам, например, небольшой сигнал усиленный надлежаще смещенным транзистором

Свертка, математическая операция при которой каждое значение выходного сигнала рассчитывается как сумма значений на входе, умноженная на набор весовых коэффициентов

Рекурсия, метод похожий на свертку, за исключением того, что рассчитываемые величины выходного сигнала используют сложение со значениями входного.

Примеры *нелинейных систем*

Системы, которые не обладают статической линейностью, например, напряжение и мощность на резисторе: $P=V^2R$, радиационное излучение энергии горячего объекта зависящее от его температуры: $R=KT^4$, интенсивность светового излучения сквозь тонкий полупрозрачный материал: $I=e^{-\alpha T}$, и т.д.

Системы, которые не обладают синусоидальной точностью, такие как электронные цепи для пикового детектора, возведение в квадрат, преобразование синусоидальной волны в меандр, удвоение частоты и т.д.

Электронные искажения, такие как ограничение, ступенчатые искажения поворотные искажения.

Умножение одного сигнала на другой, например, амплитудная модуляция и автоматическая регулировка усиления.

Явление **гистерезиса**, например, плотность магнитного потока в сравнении с напряженностью магнитного поля в железе, или механическое напряжение в сравнении с напряжением в вулканизированной резине.

Системы с порогом, например, цифровые логические элементы или сейсмические вибрации, сила которых достаточна для измельчения горных пород.

Таблица 5-1

Специальные свойства линейности

Линейность *коммутативна*, это свойство относится к двум или более системам. Рис.5-7. показывает основную идею коммутативности. Представьте себе две системы соединенные *последовательно*, т.е. выход одной системы является входом следующей. Если каждая система линейна, то их соединение также будет линейно. Свойство коммутативности устанавливает, что порядок систем может быть реорганизован без воздействия на характеристики всей комбинации. Этот принцип часто используется в электронных схемах. Например, представьте схему, состоящую их двух каскадов, один из которых отвечает за усиление сигнала, а другой за фильтрацию. Что лучше, усилить сигнал, а потом его отфильтровать, или сначала отфильтровать, а потом усилить? Оба каскада линейны, их порядок не оказывает никакого влияния на конечный результат. При этом нужно помнить, что реальная электроника обладает *нелинейными эффектами*, которые могут оказать влияние на выбор порядка каскадов в комбинации, в их числе: перекрестные помехи, смещение постоянной составляющей, внутренний шум, искажение скорости нарастания и др.

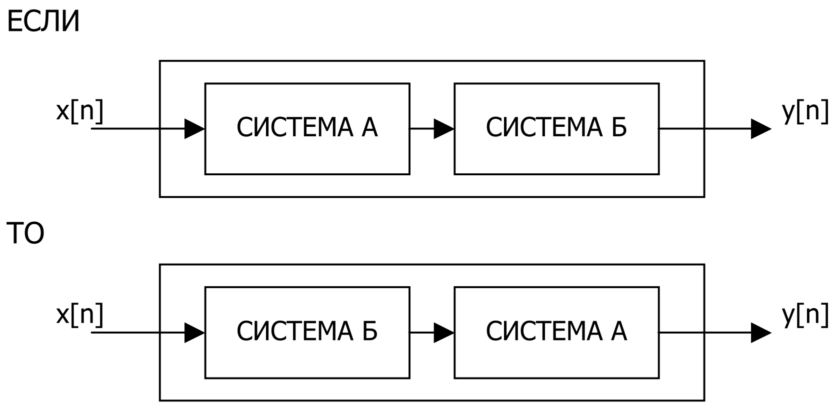


Рис.5-7. Свойство коммутативности для линейных систем.

Рис.5-8. показывает следующее свойство линейных систем: кратность входов и выходов. Система с кратным числом входов и/или выходов будет линейной, если состоит из *линейных подсистем* и *сумм сигналов*. Сложность не существенна, главным условием является отсутствие *чего-либо нелинейного* внутри системы.

Для того чтобы понять, что означает линейность для систем с кратными входами и выходами, рассмотрим следующий эксперимент. Начнем с размещения сигнала на одном из входов, в то время как на другие входы будут удерживаться на нуле. Это вызовет на кратных выходах некоторую комбинацию сигналов. Далее повторим процедуру, размещая другой сигнал на другом входе. Также как и раньше, удержим все другие входы на нуле. Этот второй входной сигнал вызовет на кратных выходах появление другой комбинации сигналов. Чтобы закончить эксперимент разместим на оба сигнала на соответствующих входах. Сигналы, появляющиеся на выходах будут просто суммой (суперпозицией) выходных сигналов, когда входные были приложены отдельно.

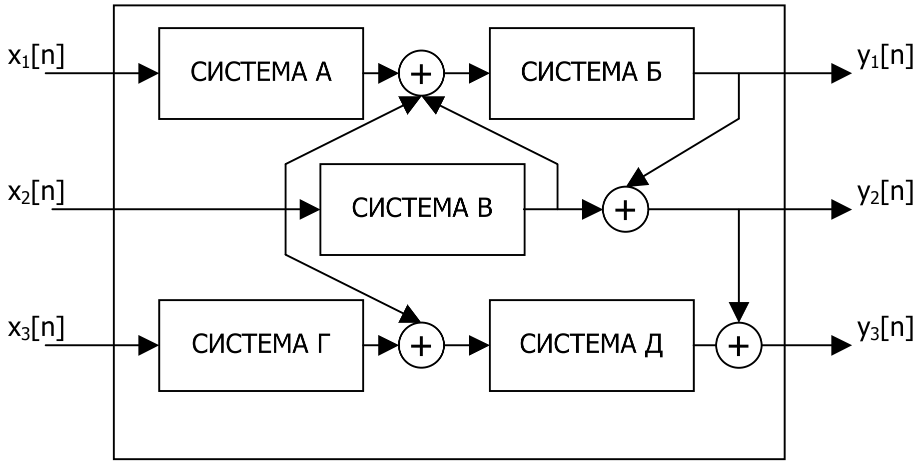
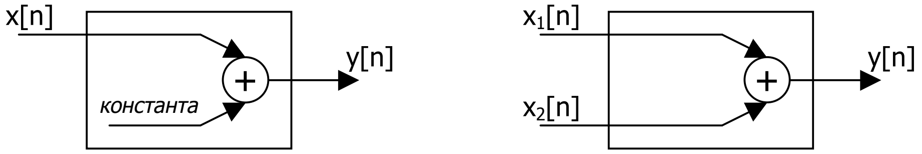


Рис.5-8. Свойство линейности для систем с кратными входами и выходами.

Использование умножения в линейных системах часто затруднительно. Это вызвано тем, что умножение может быть как линейным, так и нелинейным, в зависимости от того, на что умножается сигнал. Рис.5-9. показывает оба случая. Система, которая умножает входной сигнал на *константу* линейна. Такой системой является усилитель или аттенюатор, в зависимости от того, больше единицы константа или меньше, соответственно. Наоборот, умножение сигнала на *другой сигнал* нелинейно. Представьте себе синусоиду, умноженную на синусоиду с другой частотой. Полученная форма волны будет не синусоидальной.

Другой известной затруднительной ситуация относится к паразитным сигналам, появляющимся в электронике, например, смещение постоянной составляющей или тепловой шум. Линейно или нелинейно добавление этих посторонних сигналов? Ответ зависит от того, где были порождены загрязняющие сигналы. Если они появились *внутри* системы, то процесс будет нелинейным, так как синусоидальный сигнал не воспроизводит на выходе чистую синусоиду. С другой стороны, посторонний сигнал может быть рассмотрен как отдельный вход в системе с кратными входами, вошедший в систему *извне*. Такой процесс будет линейным, поскольку требуется только добавление сигнала внутри системы.



а. линейная система

б. нелинейная система

Рис.5-9. Линейность при умножении.

Суперпозиция: основа ЦОС

Когда мы имеем дело с линейными системами, известен лишь один способ для объединения сигналов – это *масштабирование* (умножение сигнала на константу) с последующим *сложением*. Например, сигнал не может быть умножен на другой сигнал. На Рис.5-10. показан пример: три сигнала $x_0[n]$, $x_1[n]$, $x_2[n]$ складываются для образования четвертого сигнала. Этот процесс объединения сигналов через масштабирование и сложение называется *синтез.*

Декомпозиция (разложение) является операцией обратной синтезу, когда один сигнал разбивается на две или более слагаемые компоненты. Этот процесс более сложен, чем синтез, поскольку существует бесконечное множество вариантов разложения сигнала. Например, числа 15 и 25 могут синтезировать (образовать) только число 40, в то время как само число 40 может быть разложено на 1+39, 2+38, -30.5+60.0+10.5 и т.д.

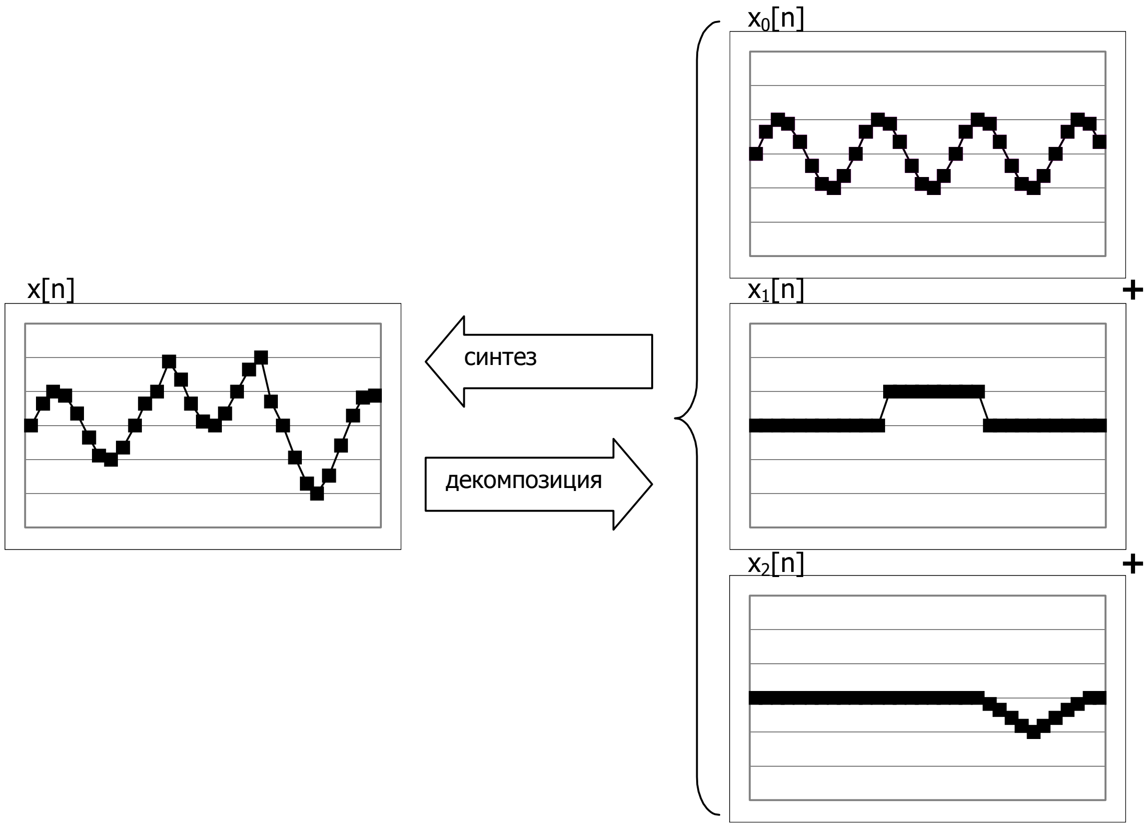


Рис.5-10. Иллюстрация синтеза и декомпозиции сигнала.

Теперь мы подошли к сердцу ЦОС: принципу *суперпозиции*, общей стратегии для понимания как могут анализироваться сигналы и системы. Предположим, имеется входной сигнал x[n], который пропускается через линейную систему, порождающую выходной сигнал y[n]. Как показано на Рис.5-11., входной сигнал может быть разложен на несколько простейших сигналов: $x_0[n]$, $x_1[n]$ и $x_2[n]$ и т.д. Будем называть их *компонентами входного сигнала*. На следующем шаге мы пропускаем каждую компоненту входного сигнала через линейную систему отдельно, порождая набор *компонент выходного сигнала*: $y_0[n]$, $y_1[n]$, $y_2[n]$ и т.д. Эти компоненты используются для синтеза выходного сигнала y[n].

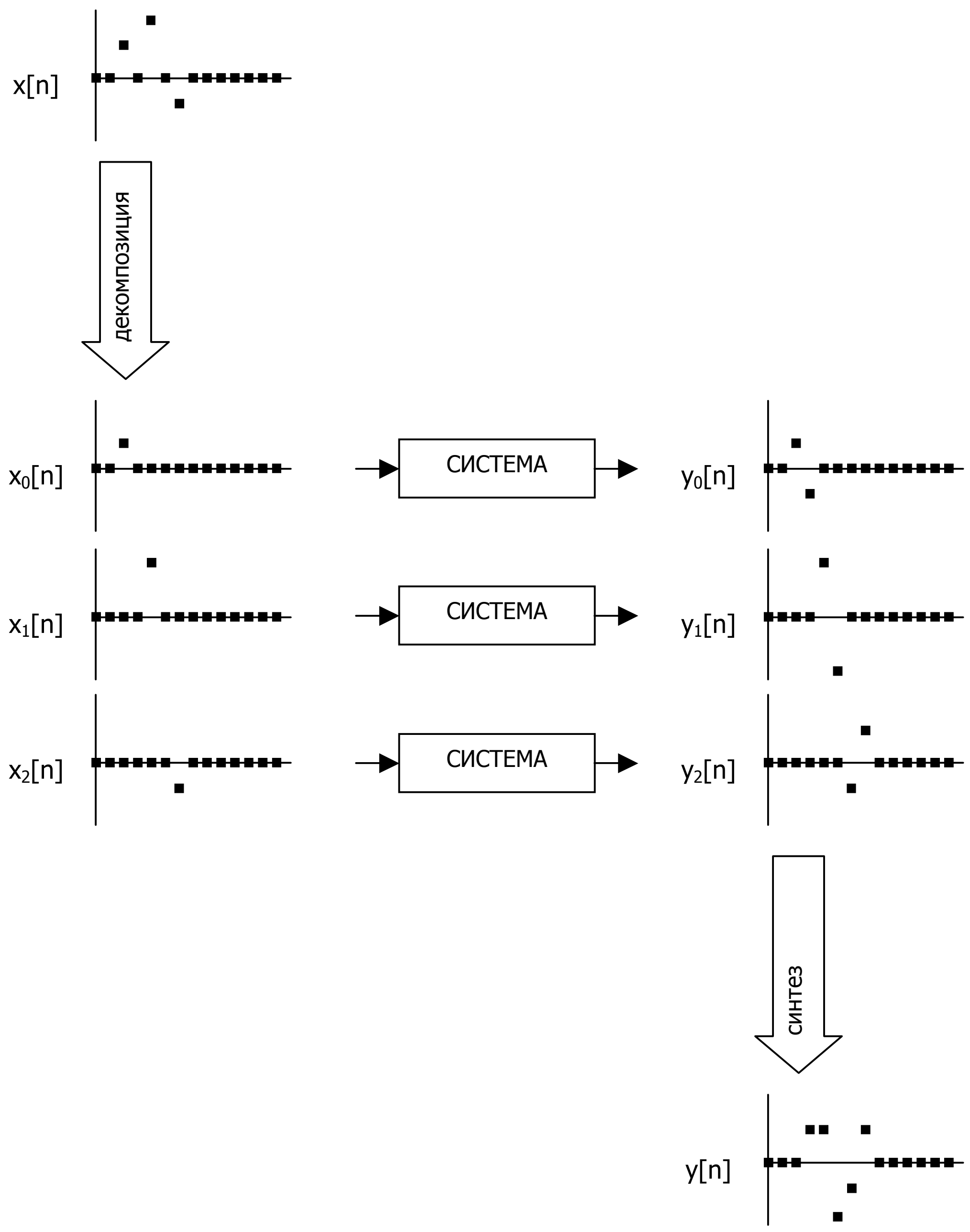


Рис.5-11. Фундаментальная концепция в ЦОС – суперпозиция.

Важное замечание: выходной сигнал, полученный таким методом идентичен сигналу, полученному прямым пропусканием входного сигнала через систему. Это очень важно! Вместо того чтобы пытаться понять, как изменяются в системе сложные сигналы, необходимо просто знать, как изменяются простые. На жаргоне цифровой обработки сигналов, входные и выходные сигналы выглядят как суперпозиция (сумма) простейших сигналов. Это являются фундаментом всей цифровой обработки сигналов.

В качестве простого примера этой идеи, рассмотрим следующий пример: попытайтесь умножить в уме число 2041 на 4. Как вы будете это делать? Можно представить число 2041 в виде отдельных палочек и каждую из них учетверить. Более простым способом является использование суперпозиции. Число 2041 может быть разложено на сумму чисел 2000+40+1. Каждая из этих компонент может быть отдельно умножена на 4, а затем собрана в конечный результат, т.е. ответ равен 8000+160+4 = 8164.

Наиболее распространенные разложения

Помните о том, что цель данного метода состоит в замещении сложной задачи несколькими простыми. Если разложение не упрощает ситуацию, то ничего не достигнуто. В цифровой обработке сигналов существует два основных способа разложения: *импульсное разложение* и *разложение Фурье*. Существуют также другие разложения, которые используются редко. Ниже приведены два основных и три редких разложения сигнала.

Импульсное разложение

Как показано на Рис.5-12, импульсное разложение делит сигнал из N выборок на N составляющих, каждая из которых содержит N выборок. Каждая составляющая сигнала содержит одну точку исходного сигнала, а остальные выборки приравнены к нулю. Единственная ненулевая точка называется *импульсом*. Важность импульсного разложения состоит в том, что оно позволяет рассматривать сигнал по одной точке, аналогично тому, как системы характеризуются ответом на импульсы. Зная, как система отвечает на импульс, ее выход может быть рассчитан для любого входного сигнала. Этот метод называется *сверткой*, и ему посвящены следующие две главы.

Ступенчатое разложение

Ступенчатое разложение, показанное на Рис.5-13, также разбивает сигнал из N выборок на N составляющих, каждая из которых содержит N выборок. Каждая составляющая сигнала это *ступень*, первые выборки которой нулевые, а остальные содержат некоторое значение. Рассмотрим разложение сигнала из N выборок, $x[n]$, на составляющие: $x_0[n]$, $x_1[n]$, $x_2[n]$, …, $x_N[n]$. K-ая составляющая сигнала, $x_K[n]$, состоит из первых k-1 нулевых выборок, в то время как оставшиеся выборки содержат значение $x[k] - x[k-1]$. Например, пятая составляющая состоит из нулевых точек до четвертой выборки включительно, а оставшиеся выборки содержат значение x[5] – x[4] (разницу между значениями выборок 4 и 5 исходного сигнала). В специальном случае для $x_0[n]$ все выборки содержат значение равно x[0]. В то время как импульсное разложение рассматривает сигналы по одной точке, ступенчатое разложение характеризует сигналы *разностью* между смежными выборками. Аналогично, системы характеризуются ответом на *изменение* во входном сигнале.

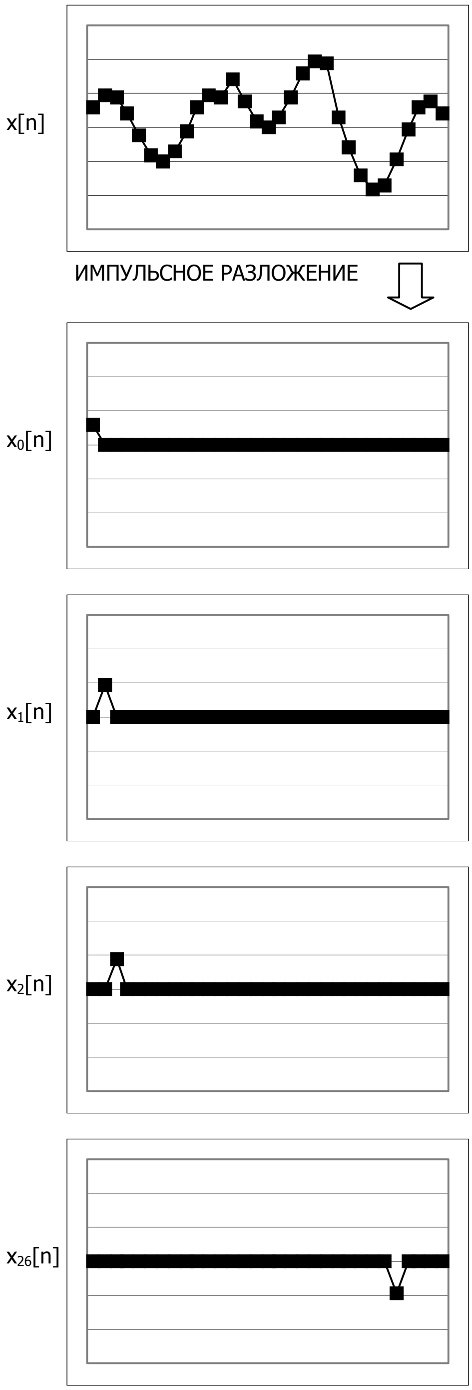


Рис.5-12. Импульсное разложение.

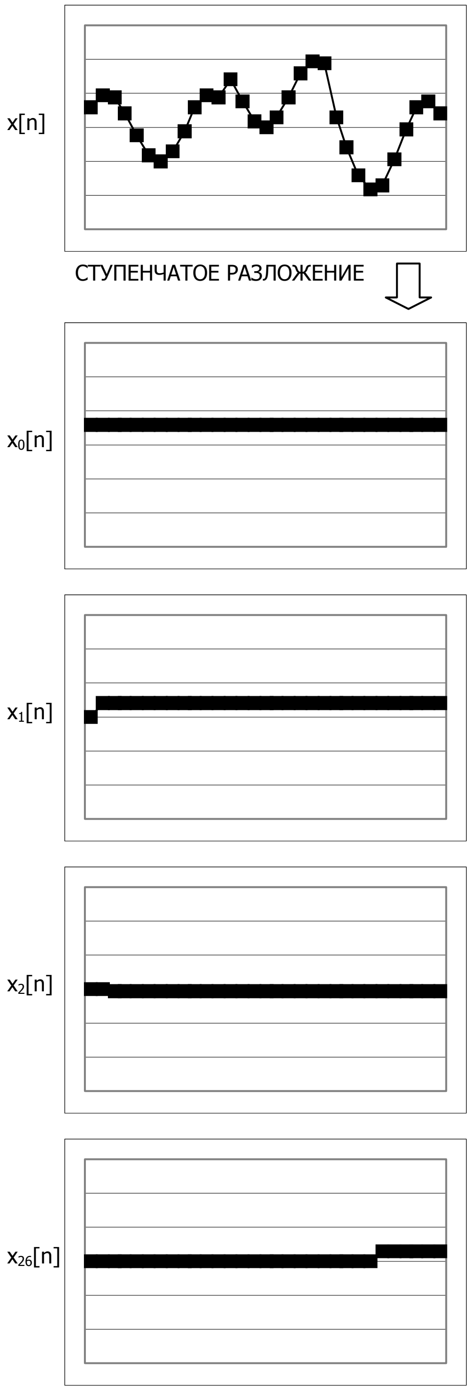


Рис.5-13. Ступенчатое разложение.

Четно-нечетное разложение

Четно-нечетное разложение, показанное на Рис.5-14, разбивает сигнал на две составляющих сигнала, одна из которых имеет **четную симметрию**, а другая **нечетную симметрию**. Говорят, что сигнал из N выборок имеет четную симметрию, если он зеркально отражается от точки N/2. Таким образом, значение в точке x[N/2+1] равно значению в точке x[N/2-1], значение x[N/2+2] равно x[N/2-2] и т.д. Подобным образом, нечетная симметрия достигается, если значение в точке x[N/2+1] равно значению в точке -x[N/2-1], значение x[N/2+2] равно -x[N/2-2] и т.д. Такое определение предполагает, что сигнал составлен из четного числа выборок, и индекс изменяется от 0 до N-1. Разложение рассчитывается из отношений:

$$x_E[n] = \frac{x[n] + x[N-n]}{2}$$
$$x_O[n] = \frac{x[n] - x[N-n]}{2}$$

Уравнение 5-1. Выражения для расчета четно-нечетного разложения.

Определение симметрии может показаться странным, поскольку настоящий центр симметрии сигнала не N/2, а N/2-1/2 (между двумя выборками). Это требует дополнительного объяснения.

Такое разложение является важной частью концепции ЦОС, называемой *круговой симметрией*. Она базируется на том, что *конец* рассматриваемого сигнала соединен с его *началом*. Как точка x[5] следует за точкой x[4], так за точкой x[N-1] следует точка x[0]. Это похоже на картину змеи, кусающей свой хвост. Когда четные и нечетные сигналы рассматриваются в данном ключе, реально существуют две точки симметрии: x[N/2] и x[0]. Например, в четном сигнале симметрия вокруг x[0] означает, что значение в точке x[1] равно значению в точке x[N-1], точка x[2] равна точке x[N-2] и т.д. В нечетном сигнале точка 0 и N/2 равны соответствующим точкам в исходном сигнале.

Какая мотивация в рассмотрении последней выборки сигнала как следующей за первой выборкой? Ведь при обычном сборе данных нет ничего для поддержки такого кругового представления? Известно, что первая и последняя точка в сигнале имеют наиболее меньшую связь, чем любые другие две точки в сигнале. Отсутствующей деталью в этой головоломке является метод ЦОС называемый *Фурье-анализом*. Математика Фурье-анализа (гармонического анализа) по сути рассматривает сигнал как бесконечный и периодический. Более детально это будет рассмотрено в десятой главе. Сейчас важно понять, что выражение 5-1 дает корректное разложение, поскольку четные и нечетные составляющие могут быть сложены для восстановления исходного сигнала.

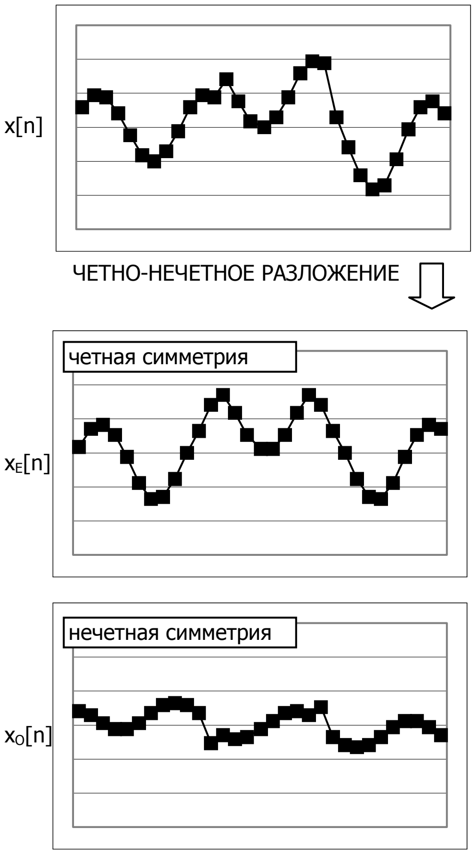


Рис.5-14. Четно-нечетное разложение.

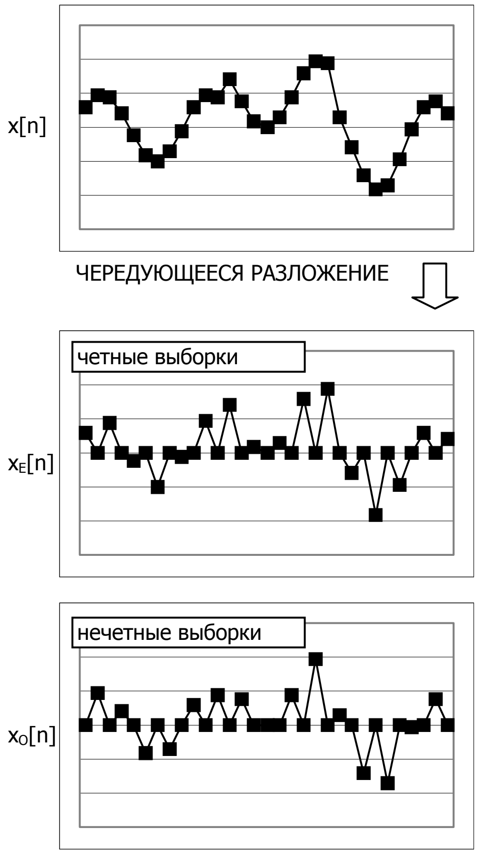


Рис.5-15. Чередующееся разложение.

Чередующееся разложение

Как показано на Рис.5-15, чередующееся разложение разделяет сигнал на две составляющих, *четные выборки* сигнала и *нечетные выборки* сигнала (не путайте с четной и нечетной симметрией). Для нахождения сигнала четных выборок, возьмите начальный сигнал и установите значения всех нечетных выборок в ноль. Чтобы найти сигнал нечетных выборок, возьмите исходный сигнал и установите в ноль значения всех четных выборок. Это достаточно просто.

На первый взгляд это разложение выглядит тривиальным и неинтересным. Однако чередующееся разложение является основой очень важного алгоритма ЦОС – быстрого преобразования Фурье (БПФ). Процедура расчета разложения Фурье была известна еще несколько сотен лет назад. К сожалению, она очень медленна, и для ее расчета требуется от нескольких минут до нескольких часов. БПФ представляет собой семейство алгоритмов, разработанных в 60-х годах для уменьшения времени расчета этого разложения. Эта стратегия является изящным примером ЦОС: разбить сигнал на элементарные составляющие повторяя применение чередующегося разложения; вычислить разложение Фурье для каждой составляющей; объединить результаты в итоговый ответ. Такой метод уменьшает время вычисления в сотни и тысячи раз!

Разложение Фурье

Разложение Фурье является слишком «математическим», и поэтому не таким очевидным и понятным. Рис.5-16 показывает пример использования этого метода.

Любой сигнал из N точек может быть разложен на сигналы, половина из которых является синусоидальными сигналами, а другая косинусоидальными. Самая нижняя частота косинусоидальной волны (названная на рисунке $x_{c0}[n]$) становится на протяжении N выборок нулевой периодичности, т.е. постоянным сигналом. Последующие косинусоидальные компоненты: $x_{c1}[n]$, $x_{c2}[n]$, $x_{c3}[n]$ делают на протяжении N выборок один, два и три периода, соответственно, и т.д. Это относится и к синусоидальным компонентам. Поскольку частота каждой компоненты фиксирована, единственным изменяющимся параметром является амплитуда каждой синусоидальной и косинусоидальной волны.

Разложение Фурье важно по трем причинам.

Во-первых, широкий диапазон сигналом состоит из накладывающихся синусоид. Примером могут служить звуковые сигналы. Разложение Фурье производит прямой анализ информации, содержащейся в сигналах этого типа.

Во-вторых, линейные системы отвечают на синусоидальный сигнал определенным образом: синусоида на входе порождает синусоиду на выходе. При этом подходе системы характеризуются тем, как они изменяют свою амплитуду и фазу пропускаемых сигналов. Поскольку входной сигнал может быть разложен на синусоиды, знание того, как система будет реагировать на них, позволит найти выходной сигнал.

В-третьих, разложение Фурье является основой широкой области математики, называемой Фурье-анализом, а также более сложных преобразований Лапласа и z-преобразования. Большинство современных алгоритмов ЦОС также базируются на некоторых аспектах этого метода.

Почему можно разложить произвольный сигнал на синусоиды? Как амплитуды этих синусоид определяются для конкретного сигнала? Какие виды систем могут быть созданы с помощью данного метода? Ответы на эти вопросы будут даны в следующих главах. Детали разложения Фурье достаточно сложны, чтобы объяснить их в кратком обзоре. Сейчас важно понять, что когда все компоненты синусоид будут сложены вместе, исходный сигнал действительно восстанавливается. Более подробно этот вопрос будет рассмотрен в Главе 8.

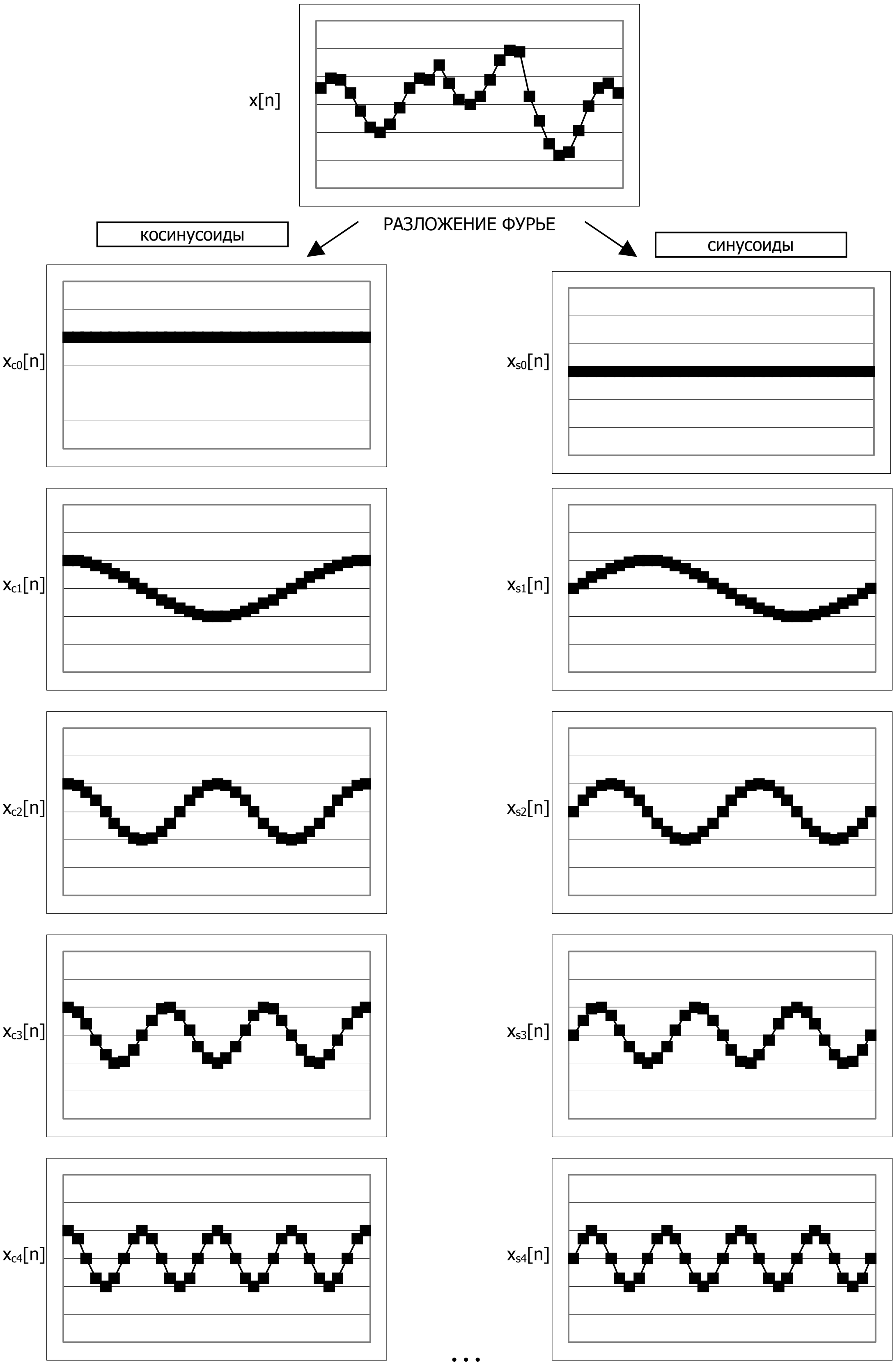


Рис.5-16. Иллюстрация разложения Фурье.

Альтернатива линейности

Чтобы оценить важность линейных систем, предположим, что существует только *одна* главная стратегия для анализа нелинейных систем. Она состоит в том, чтобы *привести* (*сделать похожей*) нелинейную систему к линейной. Существует три основных способа, чтобы сделать делать:

Первый, пренебречь нелинейностью. Если нелинейность достаточно мала, система может быть аппроксимирована к линейной. Ошибки, возникающие в результате этого допущения, сравнимы с шумом или ими просто пренебрегают.

Второй, работайте с небольшими сигналами. Множество нелинейных систем проявляют линейность при работе с сигналами с малой амплитудой. Например, транзисторы нелинейны в полном рабочем диапазоне, но обеспечивают линейное усиление, когда сигналы удерживаются в пределах нескольких милливольт. Операционные усилители реализуют этот подход. При использовании высокого коэффициента усиления разомкнутой цепи вместе с отрицательной обратной связью, входной сигнал в ОУ (т.е. разница между инвертирующим и неинвертирующим входом) удерживается на уровне нескольких милливольт. Этот микромасштабный сигнал обладает прекрасной линейностью в противоположность нелинейной цепи.

Третий, применение линеаризующих преобразований. Например, рассмотрим два сигнала, которые при перемножении образуют третий: a[n] = b[n] x c[n]. Взятие логарифма от этих сигналов преобразует нелинейный процесс умножения в линейный процесс сложения: log(a[n]) = log(b[n]) + log(c[n]). Такой метод имеют причудливое название: *гомоморфная обработка сигналов*. Например, визуальное изображение может быть смоделировано как отражение (двумерный сигнал) умноженное на окружающее освещение (другой двумерный сигнал). Гомоморфные методы делают сигнал освещения более однородным, позволяя таким образом улучшить изображение.

В следующих главах будут рассмотрены два основных метода обработки сигналов: *свертка* и *Фурье-анализ*. Оба метода основаны на стратегии, рассмотренной в этой главе: (1) разложение сигнала на простые слагаемые компоненты, (2) обработка компонентов некоторым способом, (3) синтез компонент в окончательный результат. Это и есть ЦОС!