Обработка сигналов. Лекция 1.

Алексей Орлов

25 ноября 2020 г.

Содержание

1 Введение

Современность отличает мощный процесс цифровизации почти всех сфер деятельности человека. Информационные технологии проникают повсеместно. Аналоговая обработка сигналов повсеместно вытесняется цифровой, практика перевода сигнала в цифровую форму и дальнейшие манипуляции с ним уже «в цифре» является сейчас доминирующей при работе с сигналами самой различной природы. Сегодня цифровая обработка сигналов (digital signal processing - DSP, ЦОС) является ядром множества видов новейших цифровых разработок и различных приложений в информационном обществе (например, цифровая мобильная связь, цифровые видеокамеры, телевидение и системы звукозаписи). Именно поэтому курс нацелен на изучение именно цифровой обработки сигналов, а методы, рассмотренные в курсе, являются универсальными. В ходе курса происходит знакомство не только с теоретической базой, но также осваиваются практические навыки анализа и трансформации цифровых сигналов при помощи системы компьютерной математики MATLAB.

2 Понятие сигнала

Сигнал - это физическая величина, которая содержит в себе определённую информацию. Сигналы такого рода, как акустические или электромагнитные колебания, изменения температуры тела, сейсмические волны, наблюдаемы и могут быть зарегистрированы и преобразованы соответствующим образом в электрические. Но существуют и такие сигналы, обработка которых затруднительна (например, сигналы запаха и вкуса). Хотя и стоит отметить, что разработка подобных приборов, такого как, к примеру, электронный нос, шагнула в последние годы существенно вперед.

По большей части сигналы передают информацию о состоянии или поведении физической системы и часто синтезируются в целях обмена информацией между людьми, а также между людьми и механизмами. Хотя сигналы могут быть представлены многими способами, во всех случаях информацию несут некие изменения. На математическом языке сигналы — это функции одной или более независимых переменных. По общему соглашению независимой переменной в математическом представлении сигнала выступает время, хотя в отдельных примерах независимая переменная в сигнале фактически временем не является.

Если обработку сигнала сравнить с очисткой загрязнённой воды, то основная цель будет заключаться в том, чтобы разделить элементы, содержащиеся в объекте, на нужные и ненужные. Если заранее знать, какими свойствами обладает объект, от которого идёт сигнал, и какие элементы он в себя включает, то извлечение необходимой информации и её обработка не представляют трудности. Однако если мы не знаем заранее свойств объекта, придётся сначала исследовать соответствие между особенностями сигнала и физическими свойствами объекта. Одним словом, становится необходимым анализ сигнала или, иначе говоря, выяснение его происхождения. В этом случае не обойтись без знания теории обработки сигналов. В результате обработки сигнала могут выявиться особенности объекта, не замеченные ранее. Также обработка сигналов крайне важна и для их синтеза, что используется сейчас повсеместно для синтеза речи.

2.1 Виды сигналов

На Рис. 1 мы представили разные виды сигналов. Здесь изображены сигналы, отображающие следующие физические состояния или процессы: а — звук, б — температура, в — сейсмические колебания, г — рельеф поверхности металла. У сигналов, подобных показанным на рисунках а, б и в, независимой переменной является время, а сигналы, подобные показанному на рисунке г, имеют в качестве переменной положение на поверхности материала. Кроме того, возможен разный масштаб по оси абсцисс. В процессе обработки сигналов важно не упустить из виду, с какой физической величиной и с каким масштабом переменной мы имеем дело.

Таким образом, сигналы это: различные физические величины, различные единицы измерений, различные масштабы переменных.

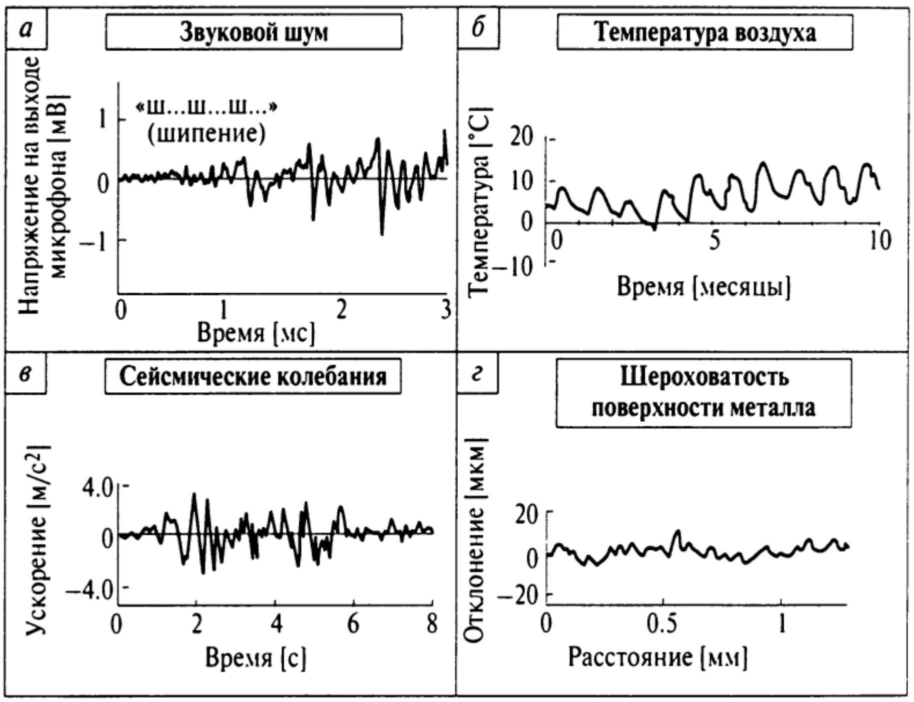


Рис. 1: Различные виды сигналов.

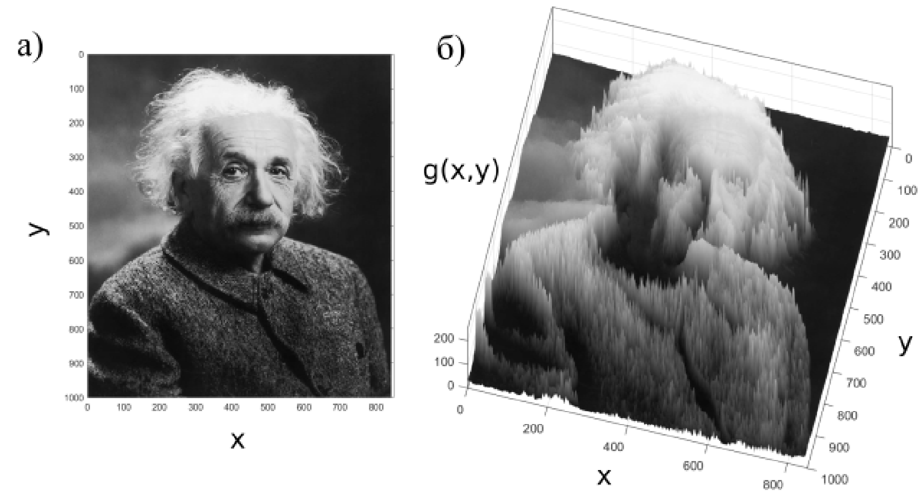


Рис. 2: Представление сигнала изображения. $g(x, y)$ - функция яркости точек изображения.

Рассмотренные выше сигналы имеют не более одной независимой переменной, будь то переменная времени или переменная положения.

Однако существуют сигналы, имеющие более одной независимой переменной, например сигнал изображения. На Рис. 2а представлена фотография, запечатлевшая Альберта Эйнштейна в 1939 году. Если на поверхности изображения провести оси координат (x, y), а яркость его точек выразить функцией $g(x, y)$, то эту функцию вполне можно считать одним из видов сигнала. И действительно, если сигнал изображения $g(x, y)$, представленный на Рис. 2а, выразить объемно, мы получим Рис. 2б. При сканировании данного изображения мы переведем его в цифровую форму, где представление черно-белого изображения по модели Grayscale даст нам матрицу яркостей пикселей. Значения элементов матрицы будут находиться в интервале от 0 до 255, где 255 отвечает белому цвету, а 0 – абсолютно чёрному, отводя таким образом 1 байт на хранение величины яркости каждого пикселя. Сигнал с одной переменной называется одномерным. Сигнал, имеющий две переменные, подобно сигналу изображения, называется *двумерным*.

Кстати, рассмотренные выше сигналы характеризуются тем, что даже если в определенный момент времени (или в определенном месте) мы знаем значение измеренной величины, то последующие ее изменения точно предсказать невозможно. Например, невозможно точно знать температуру окружающей среды в будущем году даже с учетом прогноза (если мы имеем информацию только о температуре прошлого года). Подобные сигналы называют *случайными*.

В то же время существуют сигналы, величину которых можно предсказать в любой момент времени (в любой точке). Например, звук камертона.

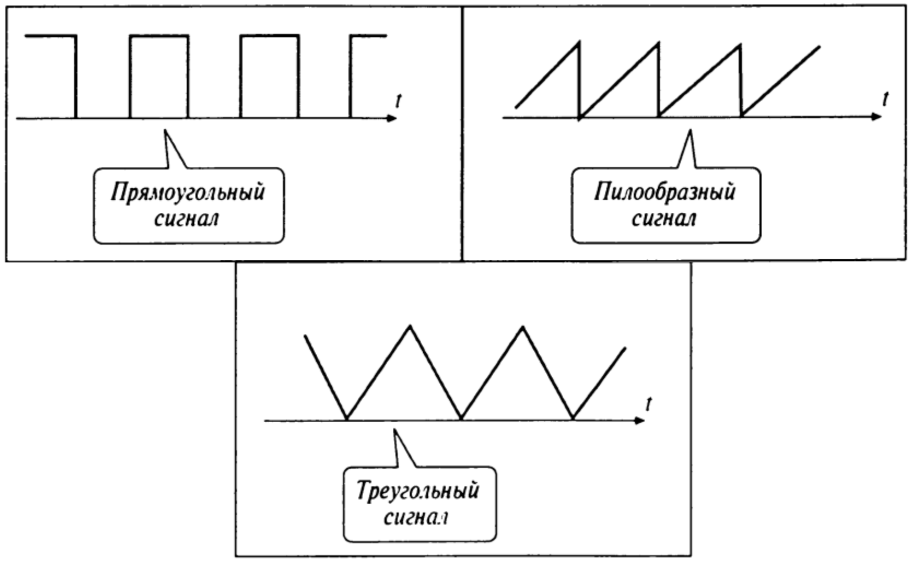


Рис. 3: Основные виды периодических сигналов.

Сколько бы ни было колебаний камертона, возникает чистая звуковая волна одной частоты. Эту волну можно выразить тригонометрической функцией. Поэтому, измерив ее значения в нескольких точках, силу звука можно выразить как функцию времени. Подобные сигналы называются *детерминированными*.

2.2 Детерминированные сигналы

Несомненно, представителем детерминированного сигнала является гармоническое колебание, графически описываемое синусоидой. Синусоида является функцией времени t и записывается в виде:

$$f(t) = A\sin(\omega t + \theta),$$

где величину сигнала определяют коэффициент A, называемый амплитудой, ω — угловая частота, θ — начальная фаза. Можно заметить, что через время T, $2T$ или же $3T$ форма сигнала повторяется. Сигналы, повторяющие свою форму через определенный интервал времени, подобно синусоиде, называют *периодическими*. Если записать выражение периодического сигнала относительного целого числа периодов n $(n = 0, \pm 1, \pm 2, \dots)$ в общем виде, получим:

$$f(t + nT) = f(t)$$

Кстати говоря, функция синуса с периодом $T = 2\pi$ в то же время имеет период, равный 4π, 6π и т.д. Наименьший положительный период называется *основным* периодом. Кроме синусоиды, к часто встречающимся

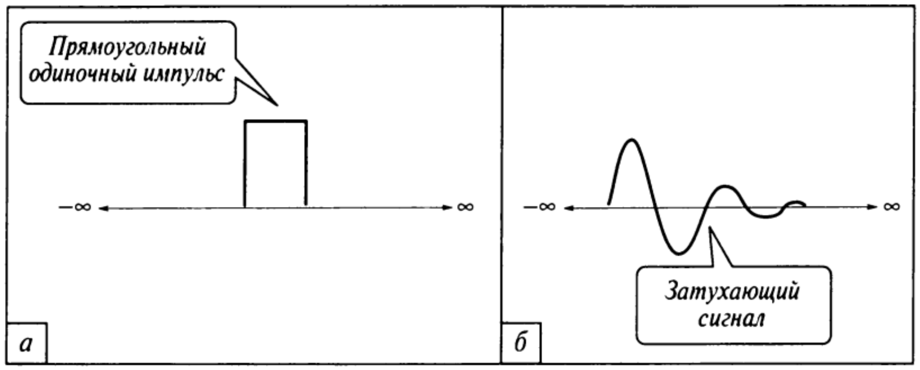


Рис. 4: Виды сигналов.

периодическим сигналам относятся прямоугольный сигнал, пилообразный сигнал, треугольный сигнал (см. Рис. 3).

Сигнал, концентрирующий энергию в коротком интервале времени, подобно единичному сигналу, изображенному на Рис.‘ 4а, называется импульсным сигналом. Сигнал, исчезающий в течение достаточно долгого промежутка времени при ограниченной энергии источника, называется затухающим (Рис. 4б). Периодический сигнал сохраняет энергию бесконечно долго и поэтому, конечно, не является затухающим.

3 Аналоговые и цифровые сигналы

Физические параметры объектов, которые мы исследуем при обработке сигналов, обычно непрерывно изменяются. Например, рассмотрим изменение температуры атмосферы во времени. Поскольку температура меняется непрерывно, теоретически возможно производить измерение через бесконечно малые промежутки времени. Однако, принимая в расчет объем памяти, необходимый для хранения данных измерения, и время на их обработку, невольно задумаешься, насколько подробные измерения нам необходимы. Значение температуры не может внезапно измениться в течение одной секунды или минуты. Следовательно, допустимы измерения через более длительные интервалы времени, что в конечном итоге сокращает объём данных. Чем меньше объем данных, тем меньше времени затрачивается на их обработку в компьютере при меньшем объеме памяти. То же самое можно сказать и о степени точности измерения. Пусть температура атмосферы в данный момент равна 25.27854°C. Нет смысла в проведении измерения с такой высокой точностью. Вполне достаточно определить степень точности до одной десятой градуса, т.е. 25.3°C. В настоящее время в метеорологическом центре данные о температуре атмосферы собираются каждый час и измерения производятся с точностью до одной десятой градуса. Этого вполне достаточно.

Большинство сигналов, существующих в природе, являются по своей форме *аналоговыми*, что часто означает непрерывное изменение во времени, и описывающими изменение физических величин (например, звуковые волны). Сигналы, применяемые в ЦОС, обычно получаются из аналоговых сигналов, дискретизованных через равные интервалы времени и преобразованных в цифровой вид.

Итак, сигнал, выражающий непрерывно изменяющуюся величину, называется аналоговым сигналом, а ступенчатое представление сигнала — *дискретизацией*. Дискретизация может производиться как по времени, так и по значению величины сигнала (Рис. 5). В первом случае сигнал называется *дискретным*, а операцию его получения часто называют *выборкой*, во втором — *квантованием*. Дискретные сигналы определены в дискретные моменты времени и, как следствие, имеют независимую переменную с дискретными значениями. Тем самым они отождествляются с числовыми последовательностями. Такие сигналы, как речь или изображение, могут иметь как непрерывное, так и дискретное представление, и при выполнении определенных условий, эти представления полностью эквивалентны. Если сигнал, подвергнутый дискретизации по времени и по значению, затем представляется в цифровом виде, то такое преобразование аналогового сигнала в цифровой называется *аналого-цифровым преобразованием*.

Аналоговый сигнал, полученный от датчика, посредством *аналого-цифрового преобразователя* (АЦП) преобразуется в числовые значения в двоичной системе счисления, т.е. предстает в виде нулей и единиц. Например, раньше при записи на компакт-диск преобразованный звуковой сигнал под воздействием лазерного луча записывался в виде цифрового сигнала. Частота выборки звукового сигнала составляла 44.1 кГц, а число бит в записываемом числе было равно 16. Цифры на диске записывались в виде наличия или отсутствия углубления, называемого питом. Вы, наверное, знаете, что по сравнению с записью обычного аналогового сигнала на кассете или пластинке цифровая запись характеризуется высоким отношением сигнал-шум и широким динамическим диапазоном (отношение минимального сигнала к максимальному неискаженному сигналу) и обеспечивает высокое качество воспроизведения звука. Но чтобы его воспроизвести, цифровой сигнал необходимо снова преобразовать в аналоговый. Этот процесс называется *цифро-аналоговым преобразованием*.

Компьютеры обладают высокой скоростью вычисления и обработки информации. Поэтому сейчас они повсеместно применяются для целей обработки цифровых сигналов, почти полностью вытеснив традиционные методы обработки аналоговых сигналов посредством электронной аппаратуры. Что касается аналого-цифрового преобразования, важно предусмотреть оптимальное количество уровней квантования, а также установить необходимую частоту выборки.

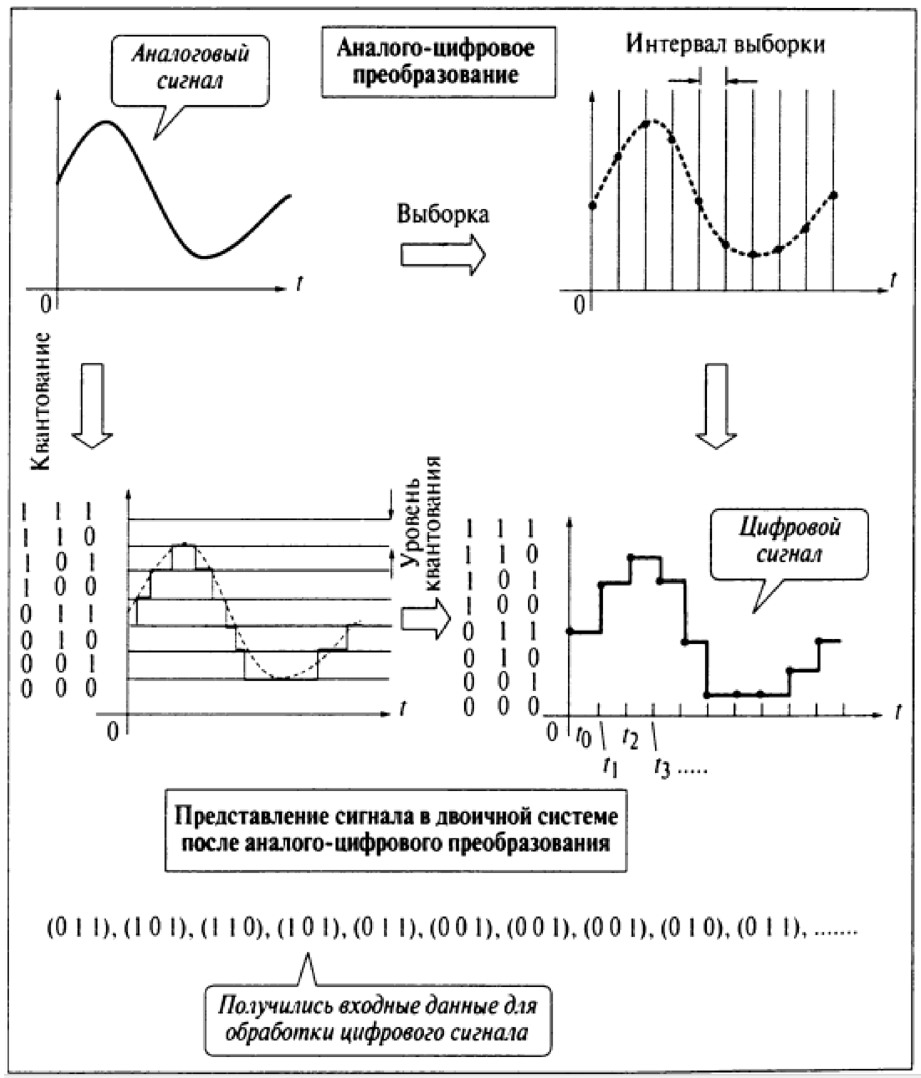


Рис. 5: Преобразование аналогового сигнала в цифровой.

3.1 Преимущества цифровой обработки сигналов

Обработка цифрового сигнала обычно нужна для устранения интерференции или шума, получения спектра данных или преобразования сигнала в более удобную форму. Помимо того, что в настоящее время ЦОС используется во многих областях, где раньше применялись аналоговые методы, появились совершенно новые области применения, где было сложно или невозможно пользоваться аналоговыми устройствами. Привлекательность ЦОС обусловлена такими основными преимуществами.

- *Гарантированная точность.* Точность определяется только числом задействованных битов.

- *Совершенная воспроизводимость.* Можно идентично воспроизвести каждый элемент, поскольку отсутствуют отклонения, обусловленные устойчивостью отдельных составляющих. Например, используя методы ЦОС, цифровые записи можно копировать или воспроизводить многократно без ухудшения качества сигнала.

- Отсутствует искажение характеристик из-за температуры или старости.

- Полупроводниковые технологии позволяют повысить надежность, уменьшить размеры, снизить стоимость, понизить энергопотребление и увеличить скорость работы.

- *Большая гибкость.* Системы ЦОС можно запрограммировать и перепрограммировать на выполнение различных функций без изменения оборудования. Это, пожалуй, одна из самых важных особенностей ЦОС.

- *Превосходная производительность.* ЦОС можно использовать для выполнения функций, которые невозможны при аналоговой обработке сигналов. Например, можно получить линейную фазовую характеристику и реализовать сложные алгоритмы адаптивной фильтрации.

- В некоторых случаях информация уже может быть записана в цифровом виде, и обрабатывать ее можно только методами ЦОС.

4 Представление дискретных сигналов

Математически дискретные сигналы представляются последовательностями чисел. Числовая последовательность x, n-й член в которой обозначают через $x[n]$, формально записывается как

$$x = \{x[n]\}, \qquad -\infty < n < \infty, \tag{1}$$

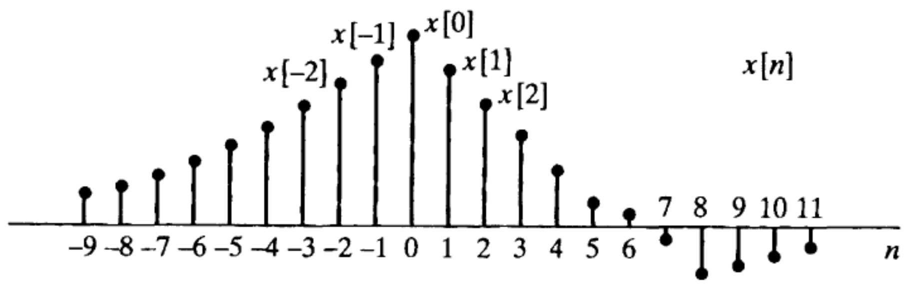


Рис. 6: Графическое представление дискретного сигнала.

где n - целое число. Заметим, что мы используем квадратные скобки для обозначения аргумента последовательности, а аргумент непрерывной функции заключается в круглые скобки.

При преобразовании аналогового сигнала в дискретную форму численное значение n-го члена последовательности равно величине аналогового сигнала $x_a(t)$ в момент времени nT, т.е.

$$x[n] = x_a(nT), \qquad -\infty < n < \infty. \tag{2}$$

Число T называют шагом дискретизации, а обратное к нему — частотой дискретизации. Хотя последовательность может возникать не только при преобразовании сигналов, ее член $x[n]$ удобно называть n-м отсчетом. Кроме того, обозначение последовательности 1 довольно громоздко, и мы будем говорить о «последовательности $x[n]$», подобно выражению «аналоговый сигнал $x_a(t)$», хотя запись $x[n]$, строго говоря, относится к отдельному члену последовательности. Это видится довольно удобным — использовать круглые скобки для аналогового непрерывного сигнала и квадратные скобки для дискретного.

Дискретные сигналы (т. е. последовательности) обычно изображают, как показано на Рис. 6. Абсцисса графика показана в виде непрерывной прямой, но важно отдавать себе отчет, что величина $x[n]$ определена только при целых значениях аргумента n. Неверно думать, что $x[n]$ равно нулю, если n не является целым числом, просто при таких аргументах $x[n]$ не определено.

В качестве примера на Рис. 7а показан фрагмент речевого сигнала, соответствующего вариациям акустического давления, в виде функции от времени, а на рис.7б приведена последовательность отсчетов этого сигнала. В то время как исходный речевой сигнал определен в любой момент времени t, последовательность несет информацию об этом сигнале только в определенные моменты времени. Тут естественным образом возникает вопрос, с какой частотой необходимо производить выборку, чтобы можно было восстановить исходный сигнал с желаемой степенью точности.

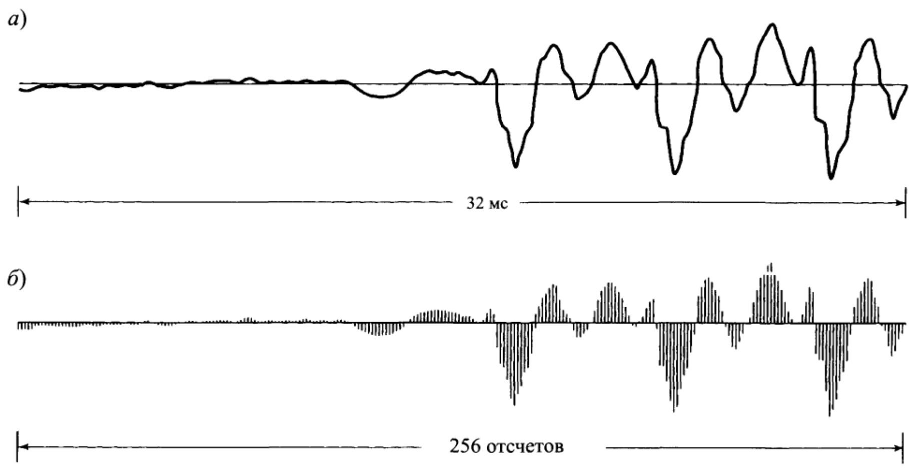


Рис. 7: а) фрагмент непрерывного речевого сигнала; б) последовательность отсчетов, полученная из фрагмента а), при шаге $T = 125$ мкс.

4.1 Проблема выборки

Проблема выбора интервала дискретизации - одна из самых основных проблем, с которой мы с самого начала сталкиваемся при обработке сигналов. Рассмотрим эту проблему более подробно.

Разберемся, какой интервал дискретизации выборки сигнала нам нужен. Первым делом рассмотрим выборку синусоиды (Рис. 8). Если соединить выделенные черным цветом точки выборки на Рис. 8а, то форма синусоиды четко просматривается, и видно, что интервал дискретизации достаточно мал. А что произойдет, если его расширить? На Рис. 8б представлен случай, когда период сигнала и интервал дискретизации выборки совпадают. Видно, что значения выборки не отображают форму сигнала. Следовательно, выбранный интервал выборки слишком велик.

Попробуем его уменьшить до половины периода сигнала, как показано на Рис. 8в. В этом случае есть вероятность, что преобразуются только нулевые значения сигнала, и полученная информация будет неполной. А если еще сузить интервал дискретизации выборки, как показано на Рис. 8г, тогда будет более или менее возможно определить первоначальную форму синусоиды.

Но возникает подозрение, что существуют какие-то другие синусоиды, проходящие через все точки выборки. Если это так, то возможно ли достоверно воспроизвести первоначальную синусоиду? Однако далее в курсе будет обосновано, что по ряду значений выборки можно достоверно воспроизвести только одну синусоиду (конечно, при условии, что ее период больше чем в два раза превышает интервал выборки). Исходя из вышеприведенного результата стало ясно, что интервал дискретизации выборки должен быть меньше половины периода. Если установить связь выборки с частотой, то

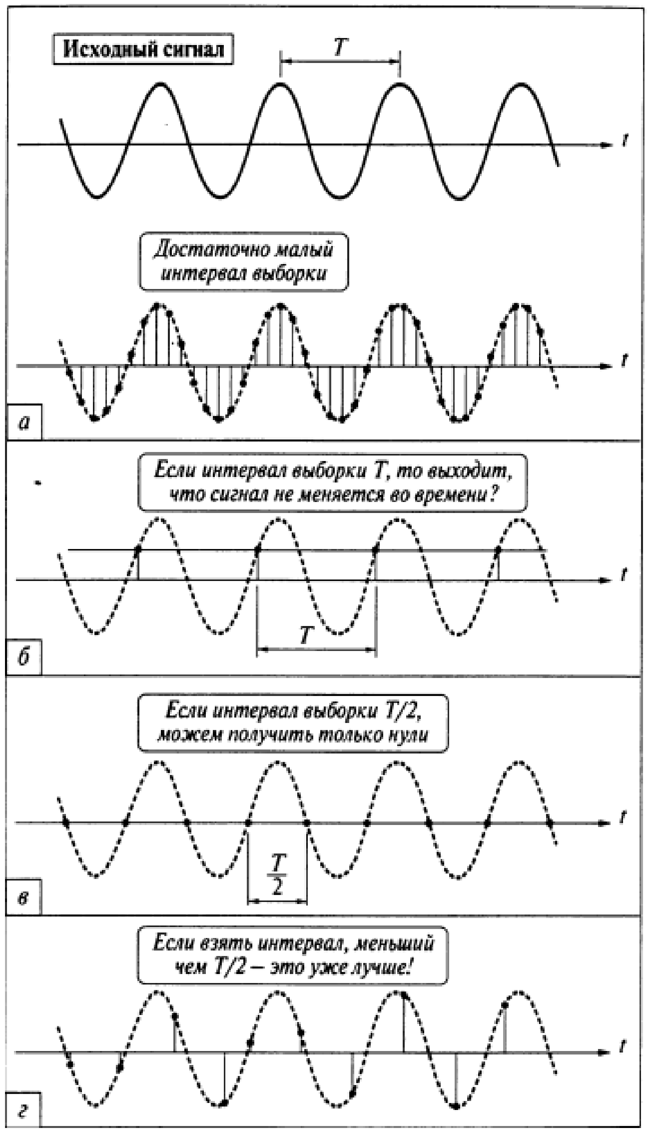


Рис. 8: Выборка синусоиды.

получится, что по отношению к синусоиде с частотой f_c необходимо следование значений выборки с частотой, большей чем $2f_c$. Эта частота выборки $2f_c$ называется *частотой Найквиста*. В последующих лекциях будет подробно показано, что любой сигнал можно выразить суммой синусоид различной частоты.

Для сигнала с наивысшей частотой f_c необходима большая частота следования значений выборки, чем частота Найквиста $2f_c$.

В отечественной литературе частота f_c определяется теоремой Котельникова, которая лежит в основе всей импульсной связи. Она показывает, при каких условиях передача непрерывной функции может быть сведена к передаче отдельных импульсов или кодовых комбинаций.

5 Стандартные сигналы и операции над ними

При анализе дискретных сигналов и в системах обработки сигналов над последовательностями совершается ряд основных преобразований (тут нужно вспомнить, что понятие последовательности в нашем случае синонимично понятию дискретного сигнала). Произведение и сумма двух последовательностей $x[n]$ и $y[n]$ определяются почленно, т.е. $z[n] = x[n] \cdot y[n]$ – произведение, а $w[n] = x[n] + y[n]$ – сумма этих последовательностей. Произведением последовательности $x[n]$ на число a считается последовательность, получающаяся из $x[n]$ в результате умножения каждого его члена на a. Последовательность $y[n]$ называют задержанной, или сдвинутой версией последовательности $x[n]$, если

$$y[n] = x[n - n_0], \tag{3}$$

где n_0 - целое число.

Некоторые из последовательностей особо важны при обсуждении теории дискретных сигналов и систем. Они представлены на Рис. 9 и описываются ниже.

Единичный импульс (Рис. 9а) определяется как

$$\delta[n] = \begin{cases} 0, & n \neq 0, \\ 1, & n = 0. \end{cases} \tag{4}$$

Как мы убедимся, единичный импульс играет ту же роль в теории дискретных сигналов, что и дельта-функция Дирака в теории непрерывных сигналов. Последовательность с единственным ненулевым отсчетом удобно называть дискретным импульсом, или просто импульсом. Стоит отметить, что дискретный импульс — понятие существенно более простое и понятное, чем дельта-функция.

К одному из важных свойств единичного импульса относится тот факт, что любая последовательность может быть выражена в виде линейной комбинации сдвинутых импульсов. Например, последовательность $p[n]$ из Рис. 10

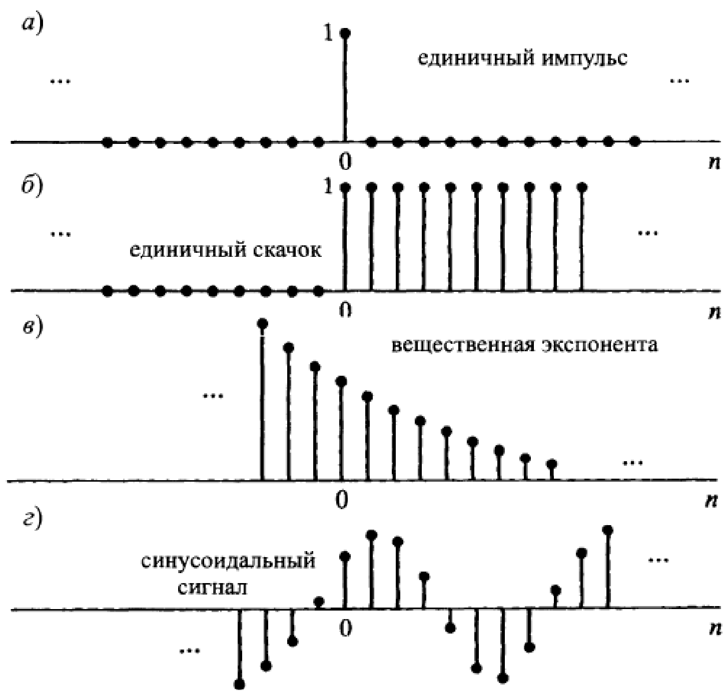


Рис. 9: Некоторые стандартные последовательности.

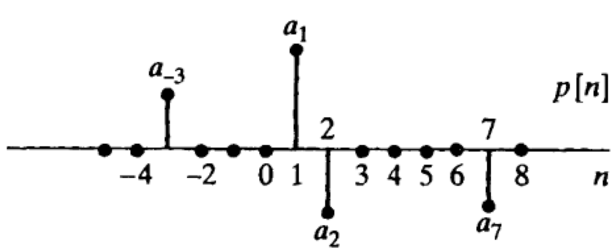


Рис. 10: Пример последовательности, представляющейся в виде конечной линейной комбинации сдвинутых импульсов.

представляется в виде:

$$p[n] = a_{-3}\delta[n+3] + a_1\delta[n-1] + a_2\delta[n-2] + a_7\delta[n-7] \tag{5}$$

а для произвольной последовательности справедливо соотношение:

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k]. \tag{6}$$

Данная формула будет использована для представления дискретных линейных систем.

Единичный скачок (Рис. 9б) определяется формулой:

$$u[n] = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0. \end{cases} \tag{7}$$

Представление единичного скачка в терминах импульса получается в результате суммирования сдвинутых импульсов, как в уравнении 6. Все ненулевые члены единичного скачка равны 1, поэтому

$$u[n] = \delta[n] + \delta[n-1] + \delta[n-2] + \cdots = \sum_{k=-\infty}^{0} \delta[n-k] \tag{8}$$

С другой стороны, единичный импульс может быть выражен через единичный скачок как разность:

$$\delta[n] = u[n] - u[n-1] \tag{9}$$

Экспоненциальная последовательность наиболее важна при представлении и анализе линейных стационарных дискретных систем. В общем виде такие последовательности записываются как

$$x[n] = A\alpha^n. \tag{10}$$

Если A и α – вещественные числа, то соответствующая последовательность тоже называется вещественной. Если $0 < \alpha < 1$ и A положительно, то значения последовательности положительны и убывают при росте n, как на Рис. 9в. Когда $-1 < \alpha < 0$, знаки членов последовательности чередуются, но их абсолютные значения всё равно убывают. Наконец, при $|\alpha| > 1$ последовательность возрастает по абсолютной величине с ростом n.

Довольно часто стандартные последовательности комбинируются для получения новых. Экспоненциальную последовательность, члены которой равны нулю при $n < 0$, можно определить как

$$x[n] = \begin{cases} A\alpha^n, & n \geq 0, \\ 0, & n < 0. \end{cases} \tag{11}$$

что имеет довольно неуклюжий вид. Более просто такая последовательность задается как выражение $x[n] = A\alpha^n u[n]$.

Синусоидальная последовательность тоже играет не последнюю роль. Мы уже говорили о ней ранее, проблему выборки. В общей форме она имеет вид

$$x[n] = A\cos(\omega_0 n + \phi), \qquad \forall n \in \mathbb{Z}, \tag{12}$$

где A и ϕ — вещественные константы Рис. 9г.

Экспоненциальная последовательность $A\alpha^n$ с комплексным α имеет вещественную и мнимую части, являющиеся взвешенными синусоидами. Более точно, если $\alpha = |\alpha|e^{j\omega_0}$ и $A = |A|e^{j\phi}$ (где j – мнимая единица), то последовательность $A\alpha^n$ может быть записана одним из следующих способов:

$$\begin{gathered} x[n] = A\alpha^n = |A|e^{j\phi}|\alpha|^n e^{j\omega_0 n} = |A||\alpha|^n e^{j\omega_0 n+\phi} = \\ |A||\alpha|^n \cos(\omega_0 n + \phi) + j|A||\alpha|^n \sin(\omega_0 n + \phi) \end{gathered}. \tag{13}$$

Эта последовательность осциллирует с экспоненциально растущей огибающей, если $|\alpha| > 1$, или с экспоненциально уменьшающейся огибающей при $|\alpha| < 1$.

Когда $\alpha = 1$, последовательность называется комплексной *экспоненциальной последовательностью*:

$$x[n] = Ae^{j\omega_0 n+\phi} = A\cos(\omega_0 n + \phi) + j|A|\sin(\omega_0 n + \phi), \tag{14}$$

т.е. как вещественная, так и мнимая часть последовательности меняется синусоидально в зависимости от n. По аналогии со случаем непрерывного времени величину ω_0 называют (круговой) частотой комплексной синусоиды или комплексной экспоненты, а ϕ — ее фазой. Заметим, однако, что n — безразмерное целое число. Поэтому ω_0 должна измеряться в радианах. Если мы хотим прослеживать тесную аналогию со случаем непрерывного времени, нам необходимо уточнить единицу частоты как радианы на отсчет, а n измерять в отсчетах.

Тот факт, что переменная n в формуле (14)) всегда принимает только целые значения, подводит нас к некоторым важным отличиям в свойствах дискретных и непрерывных комплексных экспоненциальных и синусоидальных последовательностей. Разница между непрерывной и дискретной комплексной экспонентой особенно заметна при частоте $\omega_0 + 2\pi$. В этом случае

$$x[n] = Ae^{j(\omega_0+2\pi)n} = Ae^{j\omega_0 n}e^{j2\pi n} = Ae^{j\omega_0 n}. \tag{15}$$

Более общий факт: можно легко убедиться, что комплексные экспоненциальные последовательности с частотами $(\omega_0 + 2\pi r)$ при $r \in \mathbb{Z}$ неотличимы одна от другой. Аналогичное утверждение справедливо для синусоидальных последовательностей:

$$x[n] = A\cos\left((\omega_0 + 2\pi r)n + \phi\right) = A\cos(\omega_0 n + \phi). \tag{16}$$

Отметим, что при рассмотрении комплексных экспоненциальных сигналов вида $x[n] = Ae^{j\omega_0 n}$ или вещественных синусоидальных сигналов типа $x[n] =$

$A\cos(\omega_0 n+\phi)$ мы должны ограничиться частотами, лежащими в интервале длины 2π.

Следующее важное отличие дискретных комплексных экспонент и синусоид от непрерывных касается их периодичности. В непрерывном случае как синусоидальный, так и комплексный экспоненциальный сигнал является периодической функцией, период которой равен 2π, деленному на частоту. В дискретном случае последовательность считают периодичной, если

$$x[n] = x[n+N], \quad \forall n, \tag{17}$$

где период N — обязательно целое число. Проверяя это условие для дискретных синусоид, получим

$$A\cos(\omega_0 n + \phi) = A\cos(\omega_0 n + \omega_0 N + \phi), \tag{18}$$

откуда

$$\omega_0 N = 2\pi k, \quad k \in \mathbb{Z} \tag{19}$$

Аналогичное утверждение имеет место и для комплексной экспоненциальной последовательности $Ae^{j\omega_0 n}$. Она будет N-периодичной, только если

$$e^{j\omega_0(n+N)} = e^{j\omega_0 n}. \tag{20}$$

Это равенство верно тогда и только тогда, когда $\omega_0 N = 2\pi k$. Следовательно, комплексная экспоненциальная и синусоидальная последовательности не обязательно меняются периодично в зависимости от n с периодом $2\pi/\omega_0$. Свойство их периодичности зависит от значения частоты ω_0.

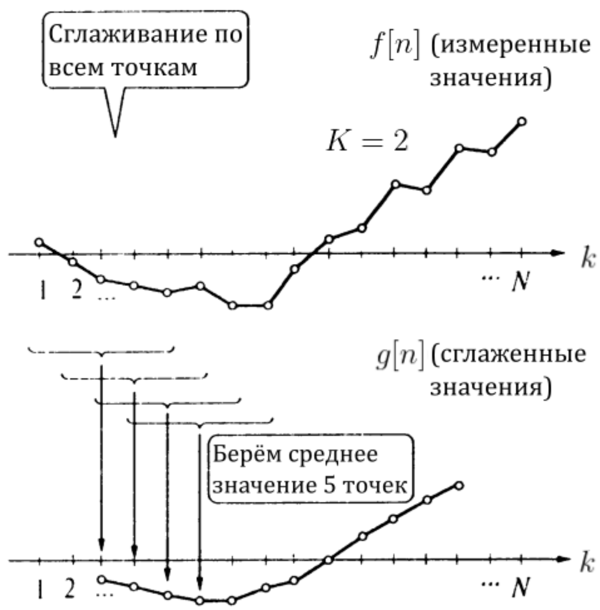
5.1 Операция скользящего усреднения

Если мы хотим увидеть приблизительную динамику изменения сигнала, необходимо сделать его «гладким», удалив незначительный шум, входящий в сигнал, и устранив мелкие колебания сигнала. Этот вид обработки называют *сглаживанием сигнала.* Простым способом сглаживание можно осуществить, воспользовавшись так называемой операцией *скользящего усреднения.* Эта операция берёт некоторую область до и после рассматриваемой точки и, учитывая численные значения измерений, входящих в эту область, вычисляет среднее значение.

Общая система скользящего среднего имеет вид:

$$\begin{aligned} y[n] = &\frac{1}{M_1+M_2+1}\sum_{k=-M_1}^{M_2} x[n-k] = \frac{1}{M_1+M_2+1}\times \\ &\times \left(x[n+M_1] + x[n+M_1-1] + \cdots + x[n] + \cdots + x[n-M_2]\right] \end{aligned} \tag{21}$$

Она вычисляет n-й отсчет входной последовательности как среднее арифметическое (M_1+M_2+1) отсчетов входной последовательности, расположенных вокруг n-го.



$$g[n] = \frac{1}{5}(f[n-2] + f[n-1] + f[n] + f[n+1] + f[n+2])$$

Рис. 11: Методика нахождения скользящего среднего.

Допустим, мы имеем последовательность из N точек измерений цифрового сигнала: $f[1], f[2], \ldots f[N]$ (Рис. 11):

$$f[n] = \sum_{k=1}^{N} f[k]\delta[n-k] \tag{22}$$

Для нахождения скользящего среднего в окрестности рассматриваемой точки n берем среднее арифметическое от K предыдущих и последующих точек, включая точку n, полагая таким образом $M_1 = M_2 = K$. Новые значения $g[n]$ для каких-либо точек n, представленные как средние значения этих $2K+1$ точек, включающих точки n, определяем как значения сглаживания:

$$\begin{aligned} g[n] =& \frac{1}{2K+1}\sum_{k=-K}^{K} f[n-k] = \frac{1}{2K+1}\times \\ &\times (f[n+K] + f[n+K-1] + \cdots + f[n] + \cdots + f[n-K]) \end{aligned} \tag{23}$$

Кстати, обратим внимание на то, что на первых и на последних точках n оси абсцисс невозможно вычислить значение сглаживания. Область, где это возможно сделать при выборе $M_1 = M_2 = K$, определяется следующим образом:

$$n = 1+K, 2+K, \ldots, N-K \tag{24}$$

На Рис. 12 представлен график средней температуры, атмосферы, измеряемой каждый день в течение одного года. Попробуем сделать обработку методом скользящего среднего. Из графиков видно, что если число точек K, которые мы учитываем, достаточно большое, то форма сигнала сглаживается. Если число K слишком мало, то эффект сглаживания слабый, но если K слишком велико, то форма сигнала становится невыразительной. Как видно из этого факта, взять скользящее среднее – это все равно что исключить из сигнала быстроколеблющуюся или, иначе говоря, высокочастотную его составляющую. Таким образом, область исключаемой частоты в зависимости от значения K меняется. Однако если мы возьмем область учитываемых точек только перед рассматриваемой точкой, а не до и после, положив $M_1 = 0$ то получим такой же эффект сглаживания.

Это определяется соотношением

$$\begin{aligned} g[n] =& \frac{1}{M_2+1}\sum_{k=0}^{M_2} f[n-k] = \frac{1}{M_2+1}\times \\ &(f[n-M_2] + f[n-M_2+1] + \cdots + f[n]), \end{aligned} \tag{25}$$

при этом $n = 1+M_2, 2+M_2, \ldots, N-M_2$. Все точки из этого интервала, используемые при вычислении скользящего среднего, мы рассматриваем с одинаковой степенью значимости. Однако во многих случаях необходимо учитывать, что, чем ближе точка к рассматриваемой точке n, тем выше её значимость, и, соответственно, с отдалением эта значимость уменьшается.

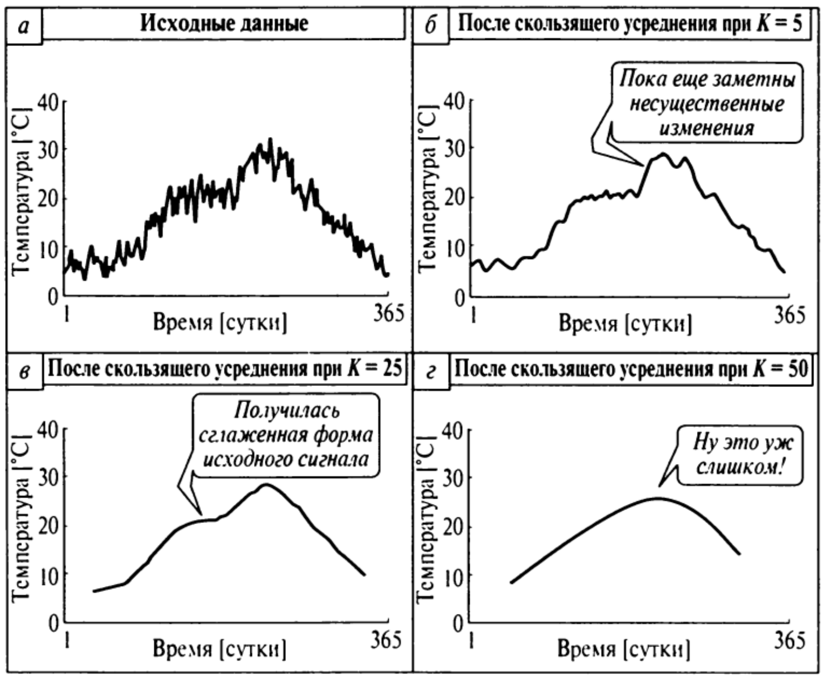


Рис. 12: Сглаживание графиков температуры атмосферы методом скользящего среднего.

Когда мы производим сглаживание, нужно определить вес каждой точки в соответствии с ее значимостью. Это можно записать следующим выражением:

$$g[n] = \sum_{k=-K}^{K} w[k] f[n-k] \tag{26}$$

Чтобы не исказить величину усредняемой функции, примем следующее условие:

$$\sum_{k=-K}^{K} w[k] = 1, \tag{27}$$

где $w[n]$ - весовая функция точек. В качестве весовой функции обычно используется *функция распределения Гаусса.*

5.2 Подавление шумов

Очень часто сигнал искажается шумами. Чтобы его восстановить, нужно каким-то образом уменьшить шумовую составляющую. Как это сделать? Очевидно, что начинать надо не с обработки сигналов. Сначала нужно определить причины возникновения шумов и постараться устранить их. Например, возникли шумы во время передачи электрического сигнала по длинному кабелю. В первую очередь необходимо проверить напряжение, осмотреть состояние кабеля и попытаться устранить причину возникновения шумов. Если же это сделать не удалось, может быть, придется заменить кабель на световод. Возможно, несмотря на все наши старания, шумы снова не удается устранить, тогда придется применить способы обработки сигналов. В данном случае будет рассмотрена обработка аналогового сигнала, недискретизованного и неоцифрованного.

Если шум высокочастотный и величина шума незначительна, то, используя один из способов сглаживания, не так уж трудно выделить искомый сигнал, и наоборот, при большой шумовой составляющей и невысокой ее частоте способ сглаживания становится неэффективным. Тем не менее в случае периодического сигнала, даже если он искажен шумами, все-таки существует эффективный способ их подавления, называемый *синхронной фильтрацией.*

Этот способ заключается в суммировании выборок сигнала в одной и той же точке периода, т.е. с одной и той же фазой. Так как шумы имеют случайный характер, то в результате усреднения они подавляются, а сигнал, наоборот, выделяется.

Итак, обозначим рассматриваемый сигнал $f(t)$. Сигнал $f(t)$ содержит полезную периодическую составляющую сигнала $s(t)$ и шумовую составляющую $p(t)$

$$f(t) = s(t) + p(t). \tag{28}$$

Сигнал в k-м периоде обозначим как $f_k(t)$. Шумовая составляющая сигнала в каждом периоде отличается, поэтому записывается как $p_k(t)$. Однако если брать один и тот же момент времени в периоде сигнала, то периодическая составляющая сигнала $s(t)$ всегда одна и та же. Следовательно, принятый сигнал $f_k(t)$ выражается как:

$$f_k(t) = s(t) + p_k(t). \tag{29}$$

Итак, многократно приняв сигнал $f_k(t)$, определим его среднее значение по периодам. Пусть N — число периодов, тогда:

$$\frac{1}{N}\sum_{k=1}^{N} f_k(t) = \frac{1}{N}\sum_{k=1}^{N}\{s(t) + p_k(t)\} = \frac{1}{N}\sum_{k=1}^{N} s(t) + \frac{1}{N}\sum_{k=1}^{N} p_k(t) \tag{30}$$

Рассмотрим это на практическом примере. На Рис. 13 видно, что с увеличением числа периодов N, в течение которых производится усреднение, шумовая составляющая уменьшается, а периодическая составляющая сигнала восстанавливается. Почему так происходит? Обратим внимание на правую часть выражения 30. Первое слагаемое, являясь периодической функцией, и после суммирования N раз остается функцией $s(t)$. А что же происходит со вторым слагаемым?

В этом месте давайте ненадолго отвлечемся и рассмотрим пример с игральными костями. Бросим несколько раз игральные кости и определим среднее значение от числа выпавших очков. Если число бросаний невелико, то среднее значение четко не выявляется, но с увеличением числа бросаний постепенно устанавливается вполне определенное число. Почему? Потому что вероятность выпадания каждой грани игральной кости одинакова. Следовательно, с ростом числа попыток частота появления каждой грани становится примерно одной и той же. Значит, среднее значение, которое мы можем получить, бросая игральную кость, приближается к $(1 + 2 + 3 + 4 + 5 + 6)/6 = 21/6 = 3.5$.

На Рис. 14 в правой части представлен график функции плотности вероятности, которая отображает среднее значение шумов и вероятность появления каждого значения. Одним словом, вероятность появления значений, близких к нулю, высока, а для значений, удаленных от нуля, вероятность появления уменьшается. Это распределение вероятности хорошо известно как *нормальное распределение* (распределение Гаусса). В случае игральных костей вероятность появления той или иной грани одинакова, и поэтому, как вы, наверное, знаете, это распределение вероятности называется *равномерным распределением*. Среднее значение независимых случайных значений с ростом их числа приближается к некоторому определенному числу. В случае шумов это среднее значение обычно равно нулю. Следовательно, сумма большого числа измерений шумовой составляющей приближается к нулю.

Из этого следует, что в ранее рассматриваемом выражении (30) второе слагаемое, являющееся шумовой составляющей, должно приближаться к нулю. Следовательно, даже в случае сильно «зашумленного» сигнала при суммировании с одной и той же фазой в результате усреднения форма сиг-

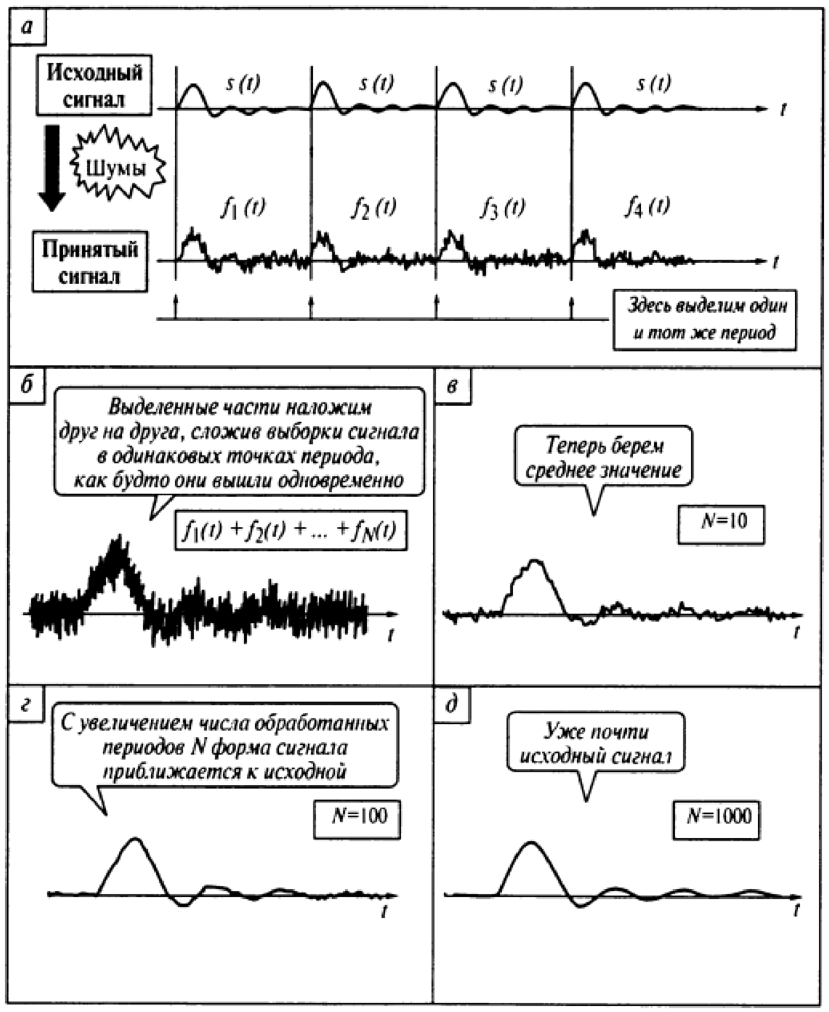


Рис. 13: Подавление шумов.

Рис. 14: Гауссовское свойство шумов.

нала выражается следующим соотношением:

$$\frac{1}{N}\sum_{k=1}^{N} f_k(t) \xrightarrow{N\to\infty} s(t) \tag{31}$$

Шумовая составляющая уменьшается, проявляется полезная составляющая сигнала, и таким образом происходит подавление шумов.

Изложенный выше способ увеличения отношения сигнал/помеха в отечественной теории и практике известен как одна из форм «метода накопления».