ЛАБОРАТОРНАЯ РАБОТА №4. Спектральный анализ

ЦЕЛЬ РАБОТЫ –изучение методов определения основных свойств случайных сигналов, таких, как мощность, спектральная плотность мощности (СПМ) и автокорреляционная функция (АКФ),получение практических навыков анализа случайных сигналов в среде Matlab.

4.1. Теоретические сведения

Случайные сигналы

Наряду с детерминированными (аналитическими*) сигналами существуют стохастические, или *случайные сигналы*. Отличительная черта случайного сигнала состоит в том, что его мгновенные значения заранее непредсказуемы. Характеристики таких сигналов довольно точно можно описать в вероятностном (статистическом) смысле. Важнейшими характеристиками случайного сигнала является его *автокорреляционная функция* и *спектральная плотность мощности*, которые тесно связаны между собой.

Анализ случайных сигналов выполняется в предположении, что рассматриваемый сигнал является *стационарным*. Свойство *стационарности* означает, что статистические характеристики сигнала, такие как среднее значение μи дисперсия (разброс) σ^2,не зависят от времени. Так, периодический сигнал является стационарным, а кратковременный переходной**– нестационарным.

Пример транзиентного сигнала показан на рис. 4.1,а. Чтобы убедиться в его нестационарности, вычислим *кратковременное* среднее значение. Для этого выберем значения сигнала для моментов времени с 0 по 99 и посчитаем среднее, которое отложим на графике в момент времени 0, затем окно наблюдения сместим на один отсчет влево и посчитаем среднее значение сигнала в промежутке от 1 до 100, а результат отложим на графике в момент времени 1 и т.д. Ниже приведен Matlab-код, выполняющий описанное дейстсвие:

```
x   = wavread('letter_T_female_low_8k');
N   = length(x);
Npt = 100;
y   = zeros(1, N);
for n = 1:N-Npt
    y(n) = sum(x(n:n+Npt-1))/Npt;
end
plot(1:N, y)
```

* Аналитические сигналы имеют однозначное математическое описание, например $\sin x$илиx^2.

** Часто употребляют название «транзиентный» сигнал.

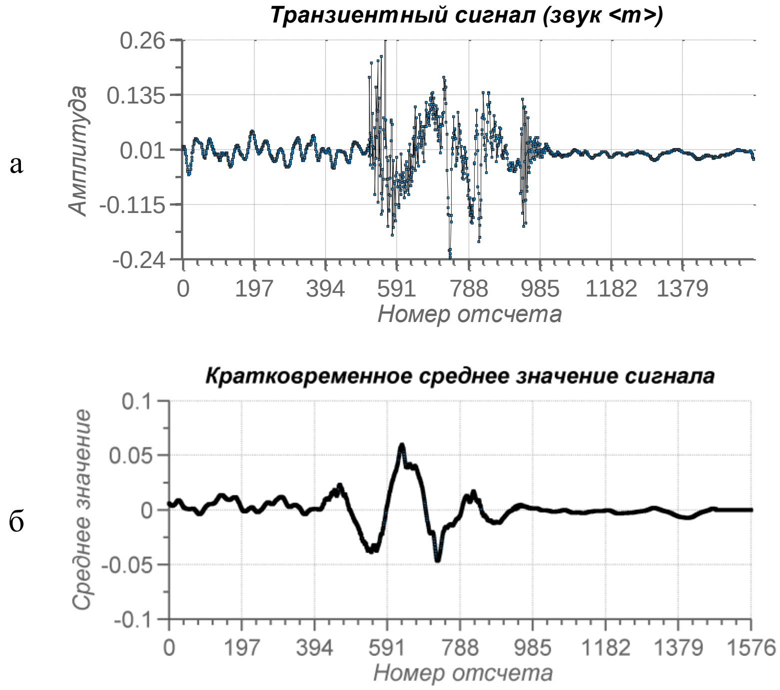


Рис. 4.1.Пример нестационарного сигнала (а) и график изменения его среднего значения во времени (б)

На рис. 4.2 показан пример стационарного сигнала и график изменения его среднего значения. Можно сделать вывод, что среднее значение сигнала практически не меняется со временем, это позволяет утверждать, что представленный сигнал отвечает условию стационарности.

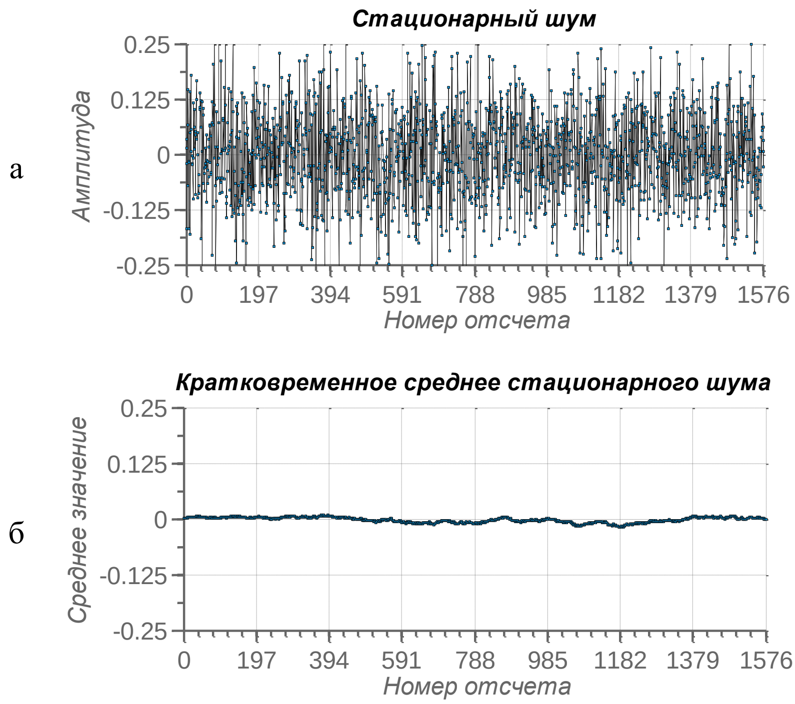


Рис. 4.2.Пример стационарного сигнала (а) и график изменения его среднего значения во времени (б)

Мощность и спектральная плотность мощности

Мгновенная мощность сигнала $x(t)$равна квадрату его амплитуды $|x(t)|^2$. Это утверждение верно, даже когда $x(t)$ принимает комплексные значения. Часто более полезным является понятие *средней мощности*. Пусть дискретный сигнал $x(n)$состоит из отсчетов непрерывного сигнала $x(t)$, тогдасредняя мощность такого сигнала вычисляется как

$$P_a = \frac{1}{N}\sum_{n=0}^{N-1} x^2(n). \tag{4.1}$$

Также мы можем найти среднюю мощность сигнала, если известно его ДПФ:

$$P_a = \frac{1}{N^2}\sum_{k=0}^{N-1} |X(k)|^2. \tag{4.2}$$

Из выражений (4.1) и (4.2) можно сделать вывод, что

$$\sum_{n=0}^{N-1} x^2(n) = \frac{1}{N}\sum_{k=0}^{N-1} |X(k)|^2.$$

Этот результат известен как теорема *Парсеваля*, которая показывает, что энергия сигнала может быть найдена как из временных отсчетов сигнала, так и из компонент преобразования Фурье. Теорема Парсеваля позволяет ввести понятие *периодограммы*:

Периодограмма: $$P_{xx}(k) = \frac{1}{N}|X(k)|^2, \qquad k = 0,1,\dots,N-1. \tag{4.3}$$

Таким образом, периодограмма – это функция, принимающая действительные значения и обладающая свойством симметрии $P_{xx}(k) = P_{xx}(N-k)$вследствие симметричности $X(k)^*$.Зная периодограмму можно записать выражение для средней мощности:

Средняя мощность: $$P_a = \frac{1}{N}\sum_{n=0}^{N-1} x^2(n) = \frac{1}{N}\sum_{k=0}^{N-1} P_{xx}(k). \tag{4.4}$$

---

* ДПФ $X(k)$ действительного сигнала $x(n)$обладает свойствомсимметрии $X(k) = X^*(N-k)$, где «*» знак комплексного сопряжения.

Выражение(4.4) дает нам понять, что также как $x^2(n)$показывает мощность сигнала в отдельной момент времени n, так $P_{xx}(k)$ показывает мощность отдельной k-й частотной компоненты сигнала. Говорят, что периодограмма является *оценкой спектральной плотности мощности* сигнала $x(n)$.

Зачем знать СПМ сигнала?

Знание спектральной плотности мощности сигнала позволяет ответить на вопрос, какие частотные составляющие сигнала имеют наибольшую энергию. Часто записанный сигнал представляет собой смесь нескольких источников. В этом случае если каждый источник имеет свою «собственную» частоту, то по СПМ можно судить о количестве имеющихся источников: оно будет равно числу максимумов (пиков) функции СПМ. Ниже приведен Matlab-код, моделирующий описанную ситуацию. Имеется два источника сигнала: первый – частотой 500 Гц, второй – 1125 Гц. Записанный сигнал представляет собой смесь сигналов, поступающих от двух источников на фоне шума*.

```
Npt = 256;
Fs  = 8000; % частота дискретизации 8 кГц
n   = 0:Npt-1;
f1  = 500;  % частота первого источника (в Гц)
f2  = 1125; % частота первого источника (в Гц)
x       =     0.4*cos(2*pi*f1/Fs*n)   +   0.2*cos(2*pi*f2/Fs*n)   +
0.2*randn(1,Npt);
plot(n, x)
```

По временной форме сигнала (рис. 4.3) трудно судить о количестве источников. Однако это возможно сделать, если построить периодограмму сигнала (рис. 4.4) и по оси абсцисс отложить аналоговые частоты. Периодограмма показывает, что исходный сигнал состоял из двух синусоидальных компонент с разными амплитудами и частотами (в нашем случае каждая из компонент генерировалась своим источником). Также периодограмма позволяет оценить частоту синусоидальных компонент. Наличие в периодограмме компонент с малыми амплитудами объясняется присутствием шума во входном сигнале.

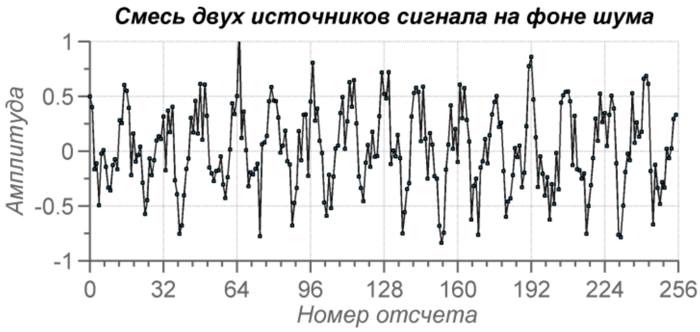


Рис. 4.3. Временная форма смеси сигналов

* Гауссовский шум с нулевым средним значением и СКО, равным единице, генерируется в Matlabкомандой `randn(1,N)`.

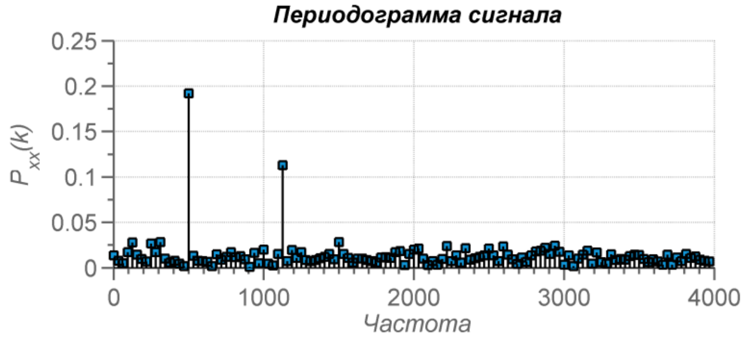


Рис. 4.4. Периодограмма смеси сигналов

Автокорреляционная функция сигнала

Объясним понятие автокорреляционной функции сигнала $x(n)$, $n = 0,1,\dots,N-1$. Предположим, что сигнал $x(n)$имеет за пределами своего определения периодическое продолжение. Это предположение весьма обоснованно, особенно для стационарных сигналов. АКФ сигнала $x(n)$определяется следующим образом:

$$\varphi_{xx}(m) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)x(n+m), \qquad m = 0,1,\dots,N-1. \tag{4.5}$$

Можно заметить, что мощность сигнала сконцентрирована в первом отсчете АКФ:

$$P_a = \frac{1}{N}\sum_{n=0}^{N-1} x^2(n) = \varphi_{xx}(0).$$

Однако более важным является другой результат: *дискретное преобразование Фурье, взятое от автокорреляционной функции сигнала, равно периодограмме этого сигнала*:

$$\text{ДПФ}\{\varphi_{xx}(m)\} = \frac{1}{N}|X(k)|^2 = P_{xx}(k), \qquad k = 0,1,\dots,N-1.$$

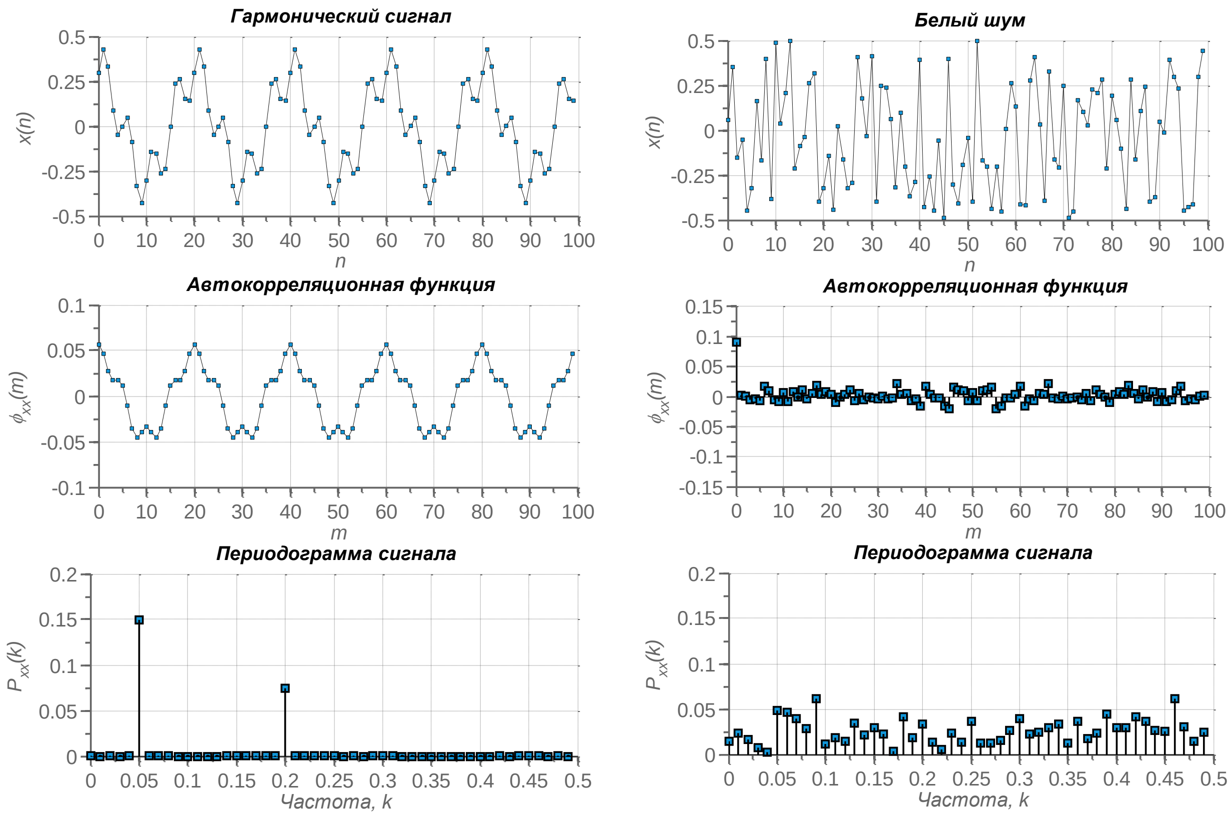
Повторим, что последнее утверждение справедливо *только* в том случае, когда сигнал $x(n)$имеет периодическое продолжение,т.е. когда $x(N) = x(0)$, $x(N+1) = x(1)$и т. д.* Автокорреляционная функция φ_{xx} показывает, насколько сильна связь между отсчетами сигнала. Например $\varphi_{xx}(1)$показывает, насколько сильно зависимы соседние отсчеты в сигнале, $\varphi_{xx}(2)$обнаруживает связь между отсчетами $x(n)$и $x(n \pm 2)$, т. е. между отсчетами, отстоящими друг

* В Matlabможно легко получить периодическое продолжение сигнала, используя операцию конкатенации. Например, если вектор-строка `sig`содержит отсчеты сигнала, то периодическое продолжение сигнала можно получить командой `periodic_sig = [sigsig];`

от друга на два интервала дискретизации. Таким образом, с помощью АКФ можно выявить скрытую в сигнале периодичность. Допустим, что сигнал содержит гармоническое колебание с периодом в 50 отсчетов. Можно уверенно сказать, что автокорреляционная функция $\varphi_{xx}(m)$такого сигнала будет иметь локальный максимум в точке $m = 50$.

Отметим, что ни $\varphi_{xx}(m)$,ни $P_{xx}(k)$не содержат никакой информации о фазе сигнала $x(n)$. По этой причине сдвиг $x(n)$во временной области не влияет на автокорреляционную функцию и периодограмму.

В качестве характерных примеров рассмотрим АКФ и периодограмму гармонического сигнала и белого шума (рис. 4.5).



а б

Рис. 4.5.Временная форма, АКФ и периодограмма гармонического сигнала $x(n) = 0{,}3\cos(2\pi \cdot 0{,}05n) + 0{,}15\sin(2\pi \cdot 0{,}2n)$, $n = 0, 1, \ldots, 99$(а) и белого шума (б)

В обоих случаях частота дискретизации сигнала равна 1 Гц. Периодический сигнал содержит две синусоидальные компоненты: одна с периодом $T_1 = 20$ отсчетов (частота $f_1 = 1/T_1 = 0{,}05$), другая с периодом $T_2 = 5$ отсчетов (частота $f_2 = 1/T_2 = 0{,}20$). Таким образом, в сигнале имеется частота основного тона f_1и ее 4-я гармоника, поскольку $f_2 = 4f_1$. С помощью АКФ можно обнаружить период частоты основного тона, как правило, он соответствует первому максимуму$\varphi_{xx}(m)$для $m > 0$. В нашем случае мы наблюдаем максимум при $m = 20$, что в точности соответствует периоду основного колебания. Периодограмма гармонического сигнала имеет два спектральных пика, показывая тем самым, что вся энергия сигнала сосредоточена в двух гармонических компонентах. *Белый шум* (рис. 4.5, б) представляет собой случайный сигнал. Анализ его периодограммы показывает, что мощность белого шума равномерно распределена во всем частотном диапазоне. Каждый последующий отсчет белого шума не зависит от предыдущего. Именно этот факт отражает АКФ$\varphi_{xx}(m)$ белого шума, в которой имеется пиктолько для $m = 0$. При любом другом сдвиге mАКФ близка к нулю, что говорит о статистической независимости отсчетов белого шума.

Оценка СПМ сигнала

Представим, что дана запись длительного стационарного сигнала и необходимо определить его СПМ. Прямой метод определения СПМ основан на вычислении квадрата модуля преобразования Фурье этого сигнала (см. формулу (4.3)). Полученная таким образом периодограмма, которую иногда называют *выборочным спектром*, представляет собой *оценку* СПМ. На практике такой метод практически не применяется, поскольку полученная СПМ является *статистически несостоятельной*. Это означает, что наличие в СПМ пика на какой-либо частоте не гарантирует нам наличие гармонической компоненты с такой частотой в сигнале. Например, на рис. 4.5 периодограмма белого шума содержит максимум на частоте 0,09 Гц, однако это не значит, что в сигнале присутствует стационарный гармонический сигнал с такой частой и амплитудой.

Для получения *устойчивой* (достоверной) оценки СПМ применяют метод усреднения периодограмм. Метод заключается в выполнении следующих действий:

1) разбить исходный сигнал $x(n)$ на K перекрывающихся или неперекрывающихся сегментов (рис. 4.6);

2) для каждого сегмента $x_i(n)$выполнить расчет периодограммы $P_{xx}^{(i)}(k)$;

3) усреднить полученные периодограммы $P_{xx}(k) = \frac{1}{K}\sum_{i=1}^{K} P_{xx}^{(i)}(k)$.

Данный метод имеет регулируемые параметры $NSAMP$ – число отсчетов на сегмент – и $NSHIFT$ – число отсчетов, на которое необходимо сдвинуть начало следующего сегмента. Ниже приводится каркас Matlab-кода, который необходим для реализации метода усреднения периодограмм.

```
Npt   = length(x);
Nsamp = 50;
Nshift= 10;
K     = floor((Npt-Nsamp)/Nshift)+1;
for i=1:K
    x_i = x(1+(i-1)*Nshift:Nsamp+(i-1)*Nshift);
% Вычисления с сегментом x_i
    % ...
end
```

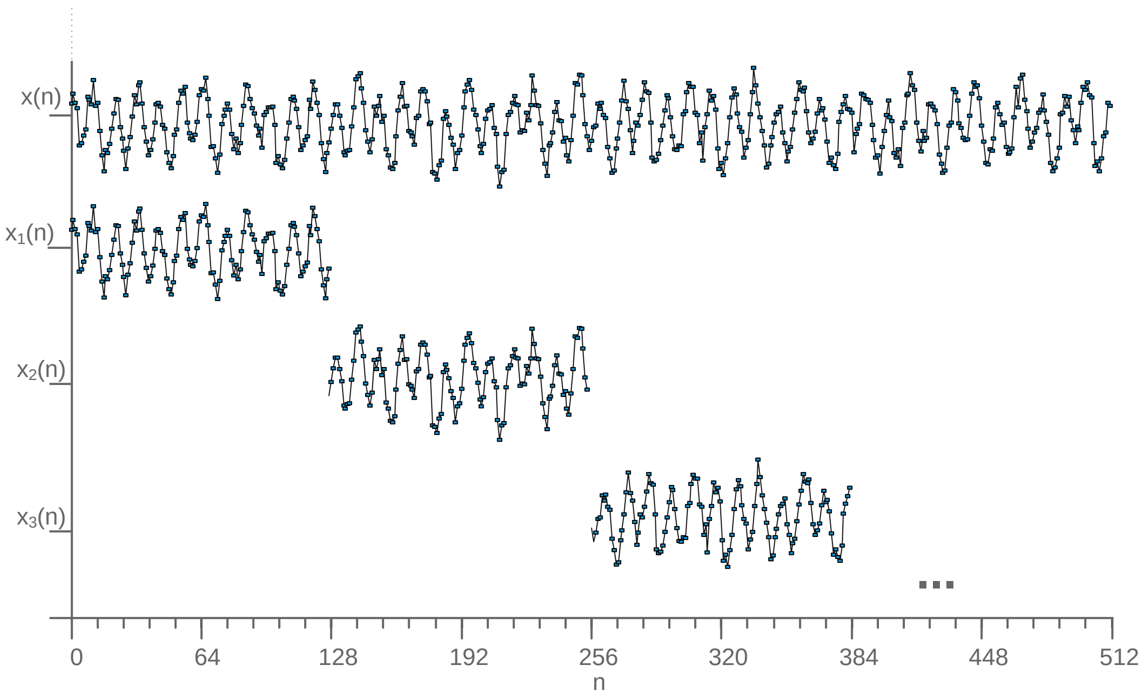


Рис. 4.6. Сегментация сигнала (без перекрытия)

Количество сегментов выбирается в зависимости от необходимой точности (надежности) спектральной оценки. При малом значении параметра $NSAMP$ получается больше сегментов, по которым производится усреднение, и, следовательно, ошибка уменьшается, но падает частотное разрешение. Увеличение параметра $NSAMP$ повышает частотное разрешение (спектр становится более детализированным), но, естественно, за счет увеличения ошибки, возникающей из-за уменьшения числа усредняемых сегментов.

Кратковременный анализ речевых сигналов

На рис.4.7показанавременная форма речевого сигнала. Из графика видно, что свойства речевого сигнала, например, характер возбуждения на вокализованных и невокализованных участках, пиковая амплитуда и др., изменяются во времени. Вокализованными называются те сегменты речи, где задействованы голосовые связки диктора. Такие звуки как 'а','о','ж',являются вокализованными, в то время как звуки 'с','ш','щ' являются невокализованными.

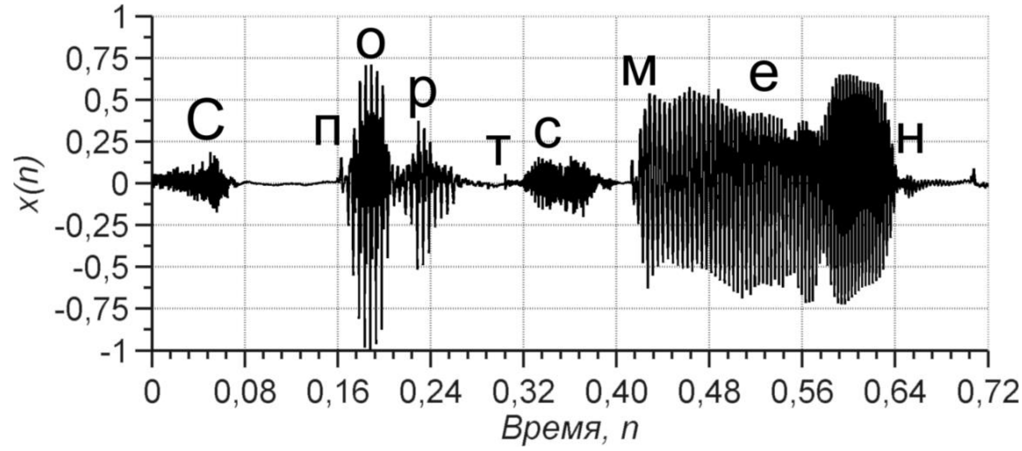


Рис. 4.7. Речевой сигнал (слово «спортсмен»)

В основе большинства методов обработки речи лежит предположение о том, что свойства речевого сигнала с течением времени медленно изменяются. Это предположение приводит к методам *кратковременного анализа*, в которых сегменты речевого сигнала выделяются и обрабатываются так, как если бы они были короткими участками отдельных звуков с отличающимися свойствами. *Сегменты*, которые иногда называют *кадрами анализа*, или *фреймами*, обычно пересекаются. Результатом обработки на каждом интервале является число или совокупность чисел. Следовательно, подобная обработка приводит к новой, зависящей от времени последовательности, которая может служить характеристикой речевого сигнала.

В общем виде большинство методов кратковременного анализа могут быть описаны выражением

$$Q(n) = \sum_{m=-\infty}^{\infty} T[x(m)]w(n-m). \tag{4.6}$$

Смысл этого выражения в следующем: сигнал $x(m)$подвергается преобразованию $T[\cdot]$, затемрезультирующая последовательность умножается на последовательность значений временного окна (весовая функция) с центром на отсчете n. Результаты суммируются по всем ненулевым значениям.

Примером такого анализа может служить измерение кратковременной мощности сигнала, которая определяется как

$$P(n) = \frac{1}{N+1} \sum_{m=n-N/2}^{n+N/2} x^2(m). \tag{4.7}$$

Таким образом, кратковременная мощность в момент времени nесть просто средняя сумма квадратов отсчетов в диапазоне от $n - N/2$до $n + N/2$. Из (4.6) видно, что в (4.7)$T[\cdot]$представляет собой возведение в квадрати деление на $(N + 1)$, а

$$w(n) = \begin{cases} 1, & -N/2 \le n \le N/2, \\ 0 & \text{в противном случае.} \end{cases} \qquad (4.8)$$

Вычисление кратковременной мощности иллюстрирует рис. 4.8. Отметим, что окно «скользит» вдоль последовательности квадратов значений сигнала, ограничивая длительность интервала, используемого в вычислениях.

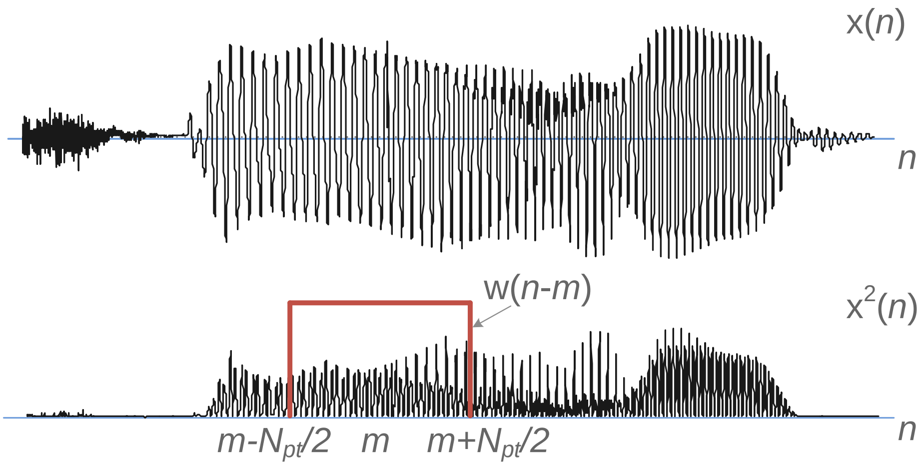


Рис. 4.8. Иллюстрация вычисления функции кратковременной мощности

Основное значение функции $P(n)$в применении к обработке речи состоит в том, что эта величина позволяет отличить вокализованные речевые сегменты от невокализованных.

Часто в методах кратковременного анализа вместо прямоугольного окна (см. формулу (4.8) и рис. 4.8) используют окна, имеющие колоколообразный вид. Эти окна имеют значения, близкие к единице в центре сегмента и близкие к нулю по краям. Применением этих окон стараются ослабить краевые эффекты, которые возникают, когда на границах сегмента сигнал имеет большую амплитуду.

4.2. Порядок выполнения работы

4.2.1. Разработайте в Matlabфункцию `[P, P_dB]=calc_power(x)`для вычисления мощности сигнала по формуле (4.1). Функция также должна выдавать значение мощности в децибелах $P_{dB} = 10\log_{10} P_a$. Посчитайте мощность следующих сигналов при$n = 0, 1 \dots ,127$:

$$x_1(n) = 0{,}5 \sin\left(2\pi\frac{3}{128}n\right),$$
$$x_2(n) = 0{,}5 \sin\left(2\pi\frac{30}{128}n\right),$$
$$x_3(n) = 0{,}3\cos\left(2\pi\frac{6}{128}n\right) + 0{,}4\cos\left(2\pi\frac{10}{128}n\right).$$

4.2.2. Разработайте в Matlabфункцию длявычисление периодограммы сигнала (см. формулу(4.3)). Постройте периодограммы сигналов из задания 4.2.1. Вычислите мощность сигналов из полученных периодограмм по формуле(4.2).

4.2.3. Предположим, что зарегистрированный сигнал представляет собой сумму полезного сигнала и белого шума($n = 0, 1, \ldots, 511$):

$$x(n) = 0{,}4 \cos\left(2\pi \frac{9}{128} n\right) + A \cdot r(n).$$

Определите максимальное отношение сигнал/шум (ОСШ), при котором по периодограмме $x(n)$можно обнаружить в нем наличие полезного сигнала. Отношение сигнал/шум (англ. *signal-to-noise ratio*, *SNR*) – безразмерная величина, равная отношению мощности полезного сигнала к мощности шума:

$$\mathrm{SNR(dB)} = 10 \log_{10} \frac{P_{signal}}{P_{noise}}.$$

Для определения требуемого ОСШ необходимо выполнить несколько экспериментов, плавно изменяя параметр A и наблюдая изменение периодограммы сигнала. Будем считать, что сигнал можно обнаружить, если уровень полезного сигнала на 10 дБ больше уровня компонент спектра, образуемых шумовым сигналом. Белый шум $r(n)$можно получить в Matlabпри помощи функции

```
rand(1, 512)
```

4.2.4. Оцените СПМ сигнала на выходе фильтра при помощи метода усреднения периодограмм. Параметры $NSAMP$ и $NSHIFT$ подберите самостоятельно. На вход фильтра поступает 4096 отсчетов белого шума. Коэффициенты фильтра:

```
b = [0.00482434 0 -0.0192973 0 0.02894606 0 -0.0192973 0
0.00482434];
a = [1 -2.06958023 3.99771255 -4.3894077 4.45285533 -2.9060422 1.75168470
- 0.5862147 0.18737949];
```

4.2.5. Разработайте в Matlabфункцию вычисления АКФ сигнала. Примените функцию для определения периода основного тона речевого сигнала (wav-файл для анализа хранится в папке "\\1-311-kafevs\Docs\ТиПЦОС\Материалы для ЛР№4\007.wav"). Wav-файл можно загрузить командой`[x, fs] =wavread('путь_к_файлу')`. Входной сигнал разбивается на перекрывающиеся секции по 320 отсчетов. Перекрытие секций составляет 160 отсчетов. Для каждой секции сигнала вычисляется АКФ и находится номер отсчета, содержащий первый максимум функции. Номер этого отсчета и является *оценкой* периода основного тона. Полученные значения отобразите на графике:

по оси абсцисс – номер сегмента, по оси ординат – значение периода основного тона.

4.2.6. Разработайте в Matlabфункцию для отделения вокализованных участков речи от невокализованных:

```
[x_v, x_u] = vu_separate(x, N, threshold)
```

Сепарация (разделение) происходит на основе анализа кратковременной мощности сигнала(см.(4.7)). На вход функции поступает речевой сигнал `x`, длина окна анализа на котором считается мощность `N` и значение порога `threshold`. Если мощность сигнала в момент времени n больше порога, то считается, что отсчет $x(n)$ является вокализованным, в противном случае – невокализованным. Сепарация выполняется следующим образом:

а) для входного сигнала $x(n)$ рассчитывается кратковременная мощность $P(n)$ по выражению (4.7);

б) рассчитывается признак вокализованности:

$$v(n) = \begin{cases} 1, & P(n) > \text{thershold}, \\ 0, & P(n) \le \text{thershold}; \end{cases}$$

в) вычисляется сепарация сигнала:

$$x_v(n) = x(n) \cdot v(n),$$
$$x_u(n) = x(n) \cdot \big(1 - v(n)\big).$$

Удобно мощность сигнала посчитать в децибелах:

$$P_{dB}(n) = 10 \log_{10} P(n)$$

и порог для сепарации также подбирать в децибелах.