

1 ЛАБОРАТОРНАЯ РАБОТА №6. ДИСКРЕТНОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

ЦЕЛЬ РАБОТЫ – Изучение теории и применения дискретного преобразования Фурье с использованием пакета MATLAB.

1.1 Теоретические сведения

1.1.1 Сигнал как вектор

Существует несколько подходов к получению формул дискретного преобразования Фурье (ДПФ). В этом разделе мы рассмотрим способ получения формул ДПФ с позиции линейной алгебры. Подход основан на том, что мы будем отождествлять конечные последовательности, с которыми мы сталкиваемся в ЦОС, с векторами, которые изучаются в линейной алгебре. Иначе говоря, дискретному сигналу $x(n)$ мы будем ставить в соответствие вектор

$$\mathbf{x} = [x(0) \ x(1) \ \dots \ x(N - 1)]^T. \quad (1.1)$$

Таким образом, мы ассоциируем сигнал с точкой в N -мерном пространстве \mathbb{C}^N . В этих обозначениях отдельные значения сигнала $x(n)$ представляют собой *координаты* вектора. Координаты в свою очередь выражаются в отношении к «стандартному» ортонормированному базису. Применительно к представлению (1.1), стандартный базис состоит из векторов вида $\mathbf{e}_0 = [1, 0, 0, \dots, 0]^T$, $\mathbf{e}_1 = [0, 1, 0, \dots, 0]^T$, ..., $\mathbf{e}_{N-1} = [0, 0, 0, \dots, 1]^T$. Можно заметить, что по сути базис $\{\mathbf{e}_k\}_{k=0,1,\dots,N-1}$ совпадает с набором сдвинутых единичных импульсов $\delta(n - k)$.

Для иллюстрации введенных понятий рассмотрим представление дискретного сигнала $x(n) = 2n + 1$ для $n = 0, 1, 2, 3$ в виде вектора.

$$\mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 3 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + 5 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + 7 \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = 1\mathbf{e}_0 + 3\mathbf{e}_1 + 5\mathbf{e}_2 + 7\mathbf{e}_3.$$

Легко видеть, что это представление идентично записи сигнала, как последовательности масштабированных и сдвинутых единичных импульсов:

$$x(n) = 1\delta(n) + 3\delta(n - 1) + 5\delta(n - 2) + 7\delta(n - 3).$$

Используя понятие скалярного произведения можно ввести понятие **нормы** вектора

$$\|\mathbf{x}\| = \sqrt{\sum_{n=0}^{N-1} x^2(n)} = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}. \quad (1.2)$$

Геометрический смысл нормы – это длина вектора. Другими словами, это расстояние между началом координат и точкой пространства, определяемой вектором.

Концепция нормы позволяет также ввести понятие **евклидова расстояния** между двумя векторами \mathbf{x} и \mathbf{y} :

$$d_2(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_{n=0}^{N-1} (x(n) - y(n))^2}. \quad (1.3)$$

Введенные понятия евклидова расстояния и нормы широко используются при конструировании алгоритмов цифровой обработки сигнала. Далее мы продолжим рассматривать сигналы, как векторы и убедимся, что дискретное преобразование Фурье возникает естественным образом, как матрица перехода от одного базиса векторного пространства к другому.

1.1.2 Система базисных векторов

Рассмотрим векторное пространство \mathbb{C}^N . В этом пространстве можно определить систему базисных векторов, которые мы отождествляем с конечными последовательностями – $\{\phi_k(n)\}$, $n, k = 0, 1, \dots, N - 1$. Для простоты будем использовать следующую запись

$$\boldsymbol{\phi}_k = [\phi_k(0) \ \phi_k(1) \ \dots \ \phi_k(N - 1)]^T. \quad (1.4)$$

Любой дискретный сигнал $x(n)$ может быть представлен в следующем виде

$$x(n) = \sum_{k=0}^{N-1} c(k)\phi_k(n), \quad (1.5)$$

где $c(k)$ – коэффициенты разложения сигнала по базису $\{\phi_k\}$, называемые **спектром** сигнала. Выбор $\{\phi_k\}$ определяется соображениями математического и практического удобства.

Чтобы лучше понять выражение (1.5) его лучше представить в векторной форме:

$$\begin{bmatrix} x(0) \\ \vdots \\ x(N-1) \end{bmatrix} = c(0) \begin{bmatrix} | \\ \phi_0 \\ | \end{bmatrix} + c(1) \begin{bmatrix} | \\ \phi_1 \\ | \end{bmatrix} + \dots + c(N-1) \begin{bmatrix} | \\ \phi_{N-1} \\ | \end{bmatrix}. \quad (1.6)$$

Из этого выражения следует, что мы ищем возможность представить дискретный сигнал $x(n)$ в виде линейной комбинации базисных векторов $\{\phi_k\}$, а $c(k)$ выступают в роли весовых коэффициентов.

Выражение (1.6) можно также переписать в матричном виде:

$$\begin{bmatrix} x(0) \\ \vdots \\ x(N-1) \end{bmatrix} = \begin{bmatrix} | & | & \dots & | \\ \phi_0 & \phi_1 & \dots & \phi_{N-1} \\ | & | & \dots & | \end{bmatrix} \cdot \begin{bmatrix} c(0) \\ \vdots \\ c(N-1) \end{bmatrix} \Rightarrow \mathbf{x} = \mathbf{\Phi} \mathbf{c}, \quad (1.7)$$

где $\mathbf{c} = [c(0) \ c(1) \ \dots \ c(N-1)]^T$, а $\mathbf{\Phi}$ – матрица, столбцами которой являются базисные вектора.

Как правило стремятся, чтобы базисная система векторов обладала к следующим свойствам:

1) Вектора системы должны быть линейно независимы, т.е.

$$a_0\phi_0 + a_1\phi_1 + \dots + a_{N-1}\phi_{N-1} \neq \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

при любых значениях коэффициентов a_k , кроме случая, когда все коэффициенты равны нулю.

2) Система векторов должна быть **ортogonalной**

$$\langle \phi_k, \phi_l \rangle = \sum_{n=0}^{N-1} \phi_k(n) \phi_l^*(n) = \begin{cases} E, & k = l \\ 0, & k \neq l \end{cases} \quad (1.8)$$

где $\langle a, b \rangle$ – вычисление скалярного произведения векторов a и b , выражение z^* – означает комплексное число сопряженное с z . В случае, если $E = 1$, то система функций называется *ортонормированной*.

3) Система векторов должна быть *полной*. Это значит, что к ней нельзя добавить ни одного вектора, который был бы ортогонален остальным векторам системы.

1.1.3 Представление сигнала в новом базисе

Теперь можно найти выражение для вычисления коэффициентов $c(k)$ из разложения (1.5):

$$\mathbf{x} = c(0)\phi_0 + c(1)\phi_1 + \dots + c(N-1)\phi_{N-1}$$

Для нахождения $c(k)$ вычислим скалярное произведение на ϕ_k слева и справа

$$\begin{aligned} \langle \mathbf{x}, \phi_k \rangle &= c(0)\langle \phi_0, \phi_k \rangle + c(1)\langle \phi_1, \phi_k \rangle + \dots + c(N-1)\langle \phi_{N-1}, \phi_k \rangle = \\ &= c(k)\langle \phi_k, \phi_k \rangle = c(k)E \end{aligned}$$

или

$$c(k) = \frac{1}{E} \langle \mathbf{x}, \phi_k \rangle = \frac{1}{E} \sum_{n=0}^{N-1} x(n) \phi_k^*(n). \quad (1.9)$$

Результат (1.9) имеет большое практическое значение. Оказывается, если у нас есть система базисных ортогональных векторов, то для того, чтобы выяснить какой вклад в сигнал $x(n)$ вносит вектор ϕ_k достаточно вычислить скалярное произведение сигнала на этот вектор. Другими словами, $c(k)$ является проекцией анализируемого сигнала $x(n)$ на базисный вектор ϕ_k . Если вычислить проекции сигнала $x(n)$ на все N базисных векторов, т.е. если найти все $c(k)$, $k = 0, 1, \dots, N-1$, то мы получим новое *представление* сигнала $x(n)$. В случае использования базиса Фурье, о котором речь пойдет далее, получаемое представление $c(k)$ называется *частотным представлением* сигнала.

1.1.4 Базис ДПФ

Для того, чтобы из общей конструкции (1.9) получить ДПФ необходимо на базисные вектора наложить ограничения. Вспомним, что ДПФ описывает строго N -периодичные сигналы, т.е. $x(n) = x(n + N)$, значит базисные функции должны также удовлетворять этому свойству:

$$\phi_k(n) = \phi_k(n + N). \quad (1.10)$$

Рассмотрим следующее семейство конечных комплексных экспонент:

$$w_k(n) = e^{j\omega_k n}, \quad n = 0, 1, \dots, N - 1 \quad (1.11)$$

где ω_k – различные частоты, которые удовлетворяет нашим требованиям. Для определения этих частот, заметим, что для того чтобы в последовательность $w_k(n)$ длины N укладывалось целое число периодов, она должна удовлетворять условию

$$w_k(0) = w_k(N) = 1,$$

которое учитывая (1.11) принимает следующий вид

$$(e^{j\omega_k})^N = 1.$$

Данное уравнение имеет N различных решений, которые представляют собой N корней из единицы $e^{j2\pi k/N}$, $k = 0, \dots, N - 1$. Если мы определим комплексное число

$$W_N = e^{-j\frac{2\pi}{N}},$$

то (1.11) может быть записано как

$$w_k(n) = W_N^{-nk}, \quad n, k = 0, 1, \dots, N - 1. \quad (1.12)$$

Таким образом:

$$\mathbf{w}_k = \left[1 \ W_N^{-k} \ W_N^{-2k} \ \dots \ W_N^{-(N-1)k} \right]^T. \quad (1.13)$$

На рисунках 1.1-1.4 приведены некоторые примеры векторов Фурье базиса для случая $N = 32$.

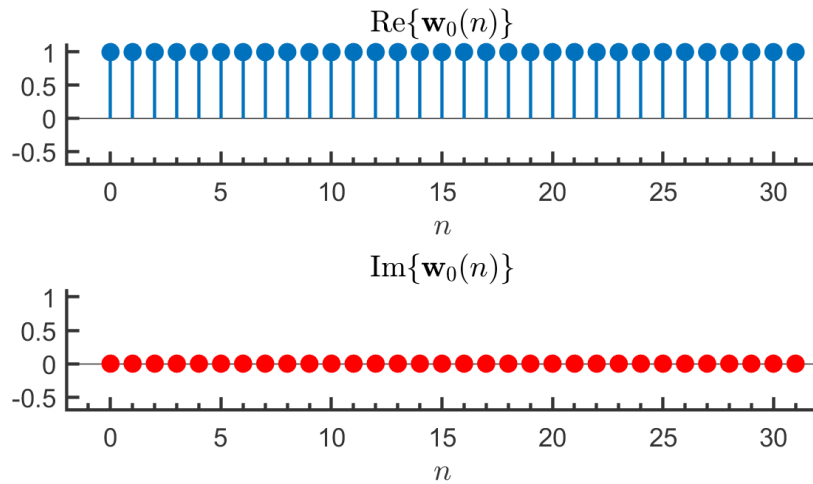


Рисунок 1.1 – Базисный вектор $\mathbf{w}_0 \in \mathbb{C}^{32}$

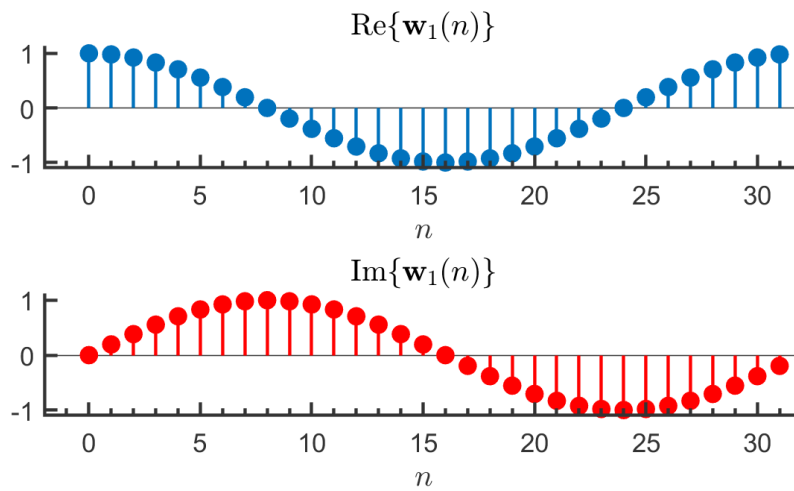


Рисунок 1.2 – Базисный вектор $\mathbf{w}_1 \in \mathbb{C}^{32}$

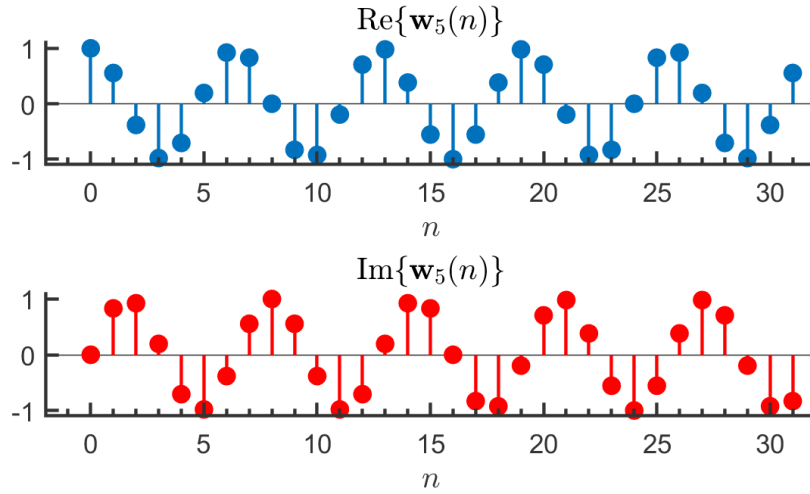


Рисунок 1.3 – Базисный вектор $\mathbf{w}_5 \in \mathbb{C}^{32}$

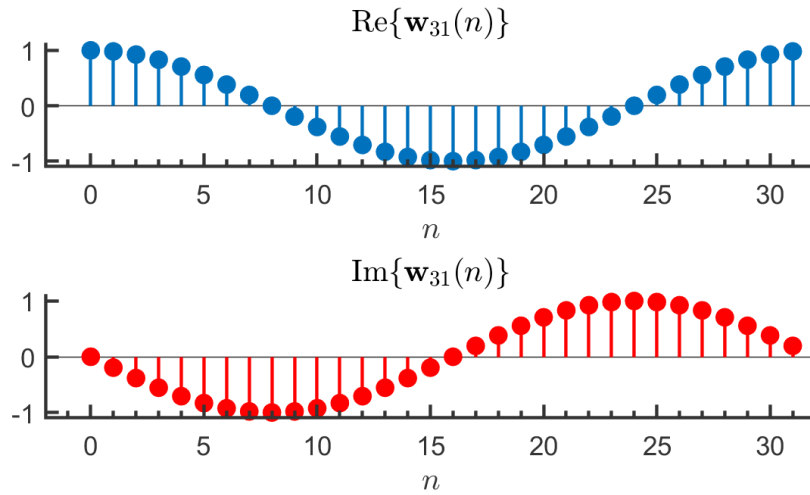


Рисунок 1.4 – Базисный вектор $\mathbf{w}_{31} \in \mathbb{C}^{32}$

1.1.5 Ортогональность базиса ДПФ

Проверим, является ли $\{\mathbf{w}_k\}$ множеством ортогональных векторов (в начале заметим, что $(W_N^{-k})^* = W_N^k$):

$$\langle \mathbf{w}_m, \mathbf{w}_l \rangle = \sum_{i=0}^{N-1} W_N^{(m-l)i} = \begin{cases} N, & m = l \\ \frac{1 - W_N^{(m-l)N}}{1 - W_N^{(m-l)}} = 0, & m \neq l \end{cases} \quad (1.14)$$

поскольку $W_N^{iN} = 1$ для любого $i \in \mathbb{Z}$. В более компактной форме мы можем записать последнее выражение следующим образом:

$$\langle \mathbf{w}_m, \mathbf{w}_l \rangle = N\delta(m - l). \quad (1.15)$$

Таким образом, система векторов $\{\mathbf{w}_k\}_{k=0,1,\dots,N-1}$ является базисом Фурье для векторного пространства \mathbb{C}^N или, что тоже – для пространства сигналов длины N . Также мы можем видеть, что этот базис не ортонормирован, поскольку $\langle \mathbf{w}_k, \mathbf{w}_k \rangle = \|\mathbf{w}_k\|^2 = N$. Однако он легко может быть превращен в ортонормированный, если отмасштабировать исходный на коэффициент $1/\sqrt{N}$.

1.1.6 Явная форма записи ДПФ

Чаще всего в литературе по ЦОС приводят формулы ДПФ, записанные в явной форме, которые имеют вид (1.9), только с конкретизацией базиса ДПФ. В частности, мы имеем две формулы *анализа*:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-\frac{j2\pi nk}{N}} = \sum_{n=0}^{N-1} x(n)W_N^{nk}, \quad (1.16)$$

и *синтеза*

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)e^{\frac{j2\pi nk}{N}} = \frac{1}{N} \sum_{k=0}^{N-1} X(k)W_N^{-nk}, \quad (1.17)$$

где $n, k = 0, 1, \dots, N - 1$.

Выражение (1.16) также называют *прямым ДПФ*, а формулу (1.17) *обратным ДПФ*.

Часто вместо выражений (1.16) и (1.17) используют операторный способ записи ДПФ и ОДПФ:

$$\begin{aligned} X(k) &= \text{DFT}\{x(n)\}, \\ x(n) &= \text{IDFT}\{X(k)\}. \end{aligned}$$

Коэффициенты ДПФ $X(k)$ имеют комплексные значения и чаще всего представляются в полярной форме:

$$X(k) = |X(k)|e^{j \arg X(k)}.$$

Когда на практике выполняют анализ сигнала $x(n)$ с использованием ДПФ (1.16), то на графике отображают только модуль $|X(k)|$, поскольку он несет основную информацию о спектральном составе сигнала. Фазовый спектр $\arg X(k)$ отображают только если он представляет какой-то интерес.

Давайте посмотрим, что же представляет из себя ДПФ некоторых простых сигналов. На рисунке 1.5 показано ДПФ сигнала $x(n) = \cos \frac{2\pi 4n}{32}$.

Частота сигнала $x(n)$ соответствует двум базисным векторам ДПФ – $w_4(n)$ и $w_{28}(n)$ и в результате ортогональности базиса только коэффициенты $X(4)$ и $X(28)$ оказываются ненулевыми. Почему же все-таки так происходит? Рассмотрим этот вопрос подробнее. Ключом к пониманию результата, показанного на рис. 1.5 является представление косинуса в виде суммы двух комплексных экспонент:

$$x(n) = \cos\left(\frac{2\pi 4n}{32}\right) = \frac{1}{2} e^{j\frac{2\pi 4n}{32}} + \frac{1}{2} e^{-j\frac{2\pi 4n}{32}}. \quad (1.18)$$

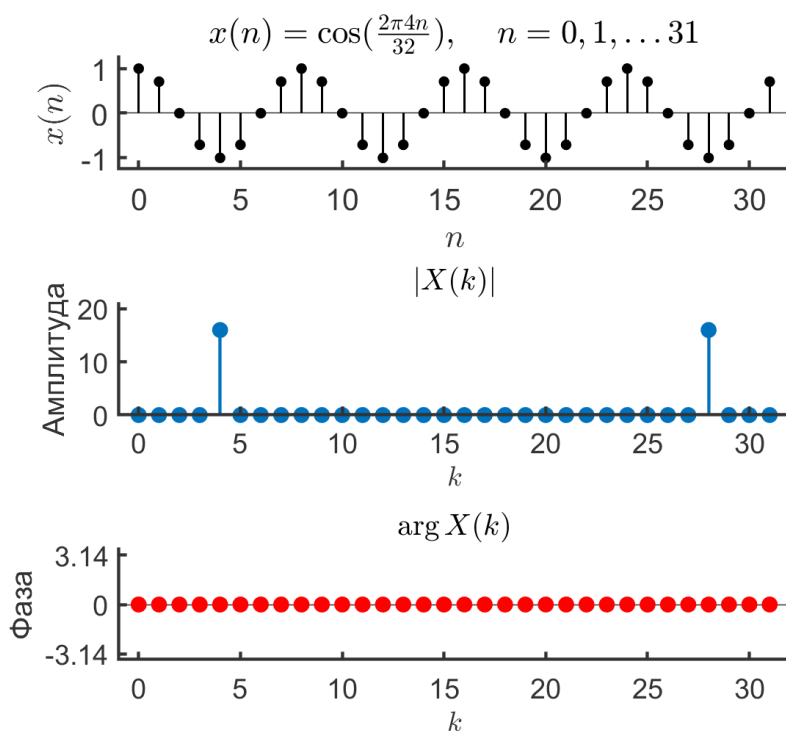


Рис. 1.5 – Синусоидальный сигнал и его ДПФ

Легко заметить, что первая экспонента $e^{j\frac{2\pi 4n}{32}}$ является ничем иным, как базисным вектором ДПФ $w_4(n)$. На первый взгляд вторая экспонента в выражении

(1.18) не совпадает ни с каким вектором из базиса ДПФ, однако это не так. Воспользуемся периодичностью дискретной комплексной экспоненты и прибавим к показателю степени 2π :

$$e^{j\left(\frac{-2\pi 4n}{32} + 2\pi\right)} = e^{j\frac{2\pi(32-4)n}{32}} = e^{j\frac{2\pi 28n}{32}} = w_{28}(n). \quad (1.19)$$

Таким образом, выражение (1.18) можно переписать в виде

$$x(n) = \cos\left(\frac{2\pi 4n}{32}\right) = \frac{1}{2}w_4(n) + \frac{1}{2}w_{28}(n). \quad (1.20)$$

Воспользуемся представлением (1.20) для того, чтобы вычислить ДПФ. Напомним, что мы представляем себе ДПФ, как операцию выполнения проекции исходного вектора $x(n)$ на базис ДПФ (см. выражение (1.9)), т.е.

$$X(k) = \langle x(n), \mathbf{w}_k(n) \rangle \quad (1.21)$$

Теперь перепишем (1.21), учитывая (1.20):

$$X(k) = \left\langle \frac{1}{2}\mathbf{w}_4(n) + \frac{1}{2}\mathbf{w}_{28}(n), \mathbf{w}_k(n) \right\rangle. \quad (1.22)$$

Осталось вспомнить, что вектора, составляющие базис ДПФ являются ортогональными, поэтому, если мы примем во внимание (1.15), то сможем записать (1.22) в окончательном виде

$$X(k) = \frac{N}{2}\delta(k-4) + \frac{N}{2}\delta(k-28). \quad (1.23)$$

Если учесть, что в рассматриваемом примере (1.18) $N = 32$, то тогда нам станет понятно, откуда в графике модуля ДПФ $|X(k)|$ появились две частотные составляющие с амплитудой 16 на индексах с номерами 4 и 28.

На рис. 1.6 показан этот же сигнал, но имеющий ненулевой фазовый сдвиг. Можно заметить, что амплитудный спектр сигнала не изменился, однако фазовый сдвиг отразился в фазовом спектре.

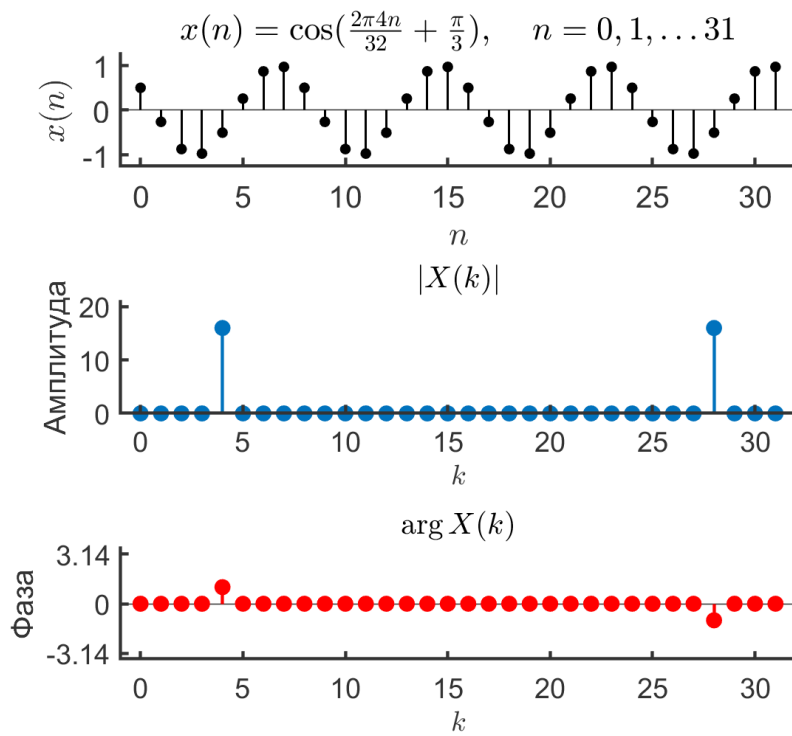


Рис. 1.6 – Синусоидальный сигнал с фазовым сдвигом и его ДПФ

На рис. 1.7 показан пример синусоидального сигнала с частотой, которая не совпадает с частотой ($\omega = \frac{2\pi 4,6}{32}$), которая не совпадает ни с одной из частот базисных векторов ДПФ.

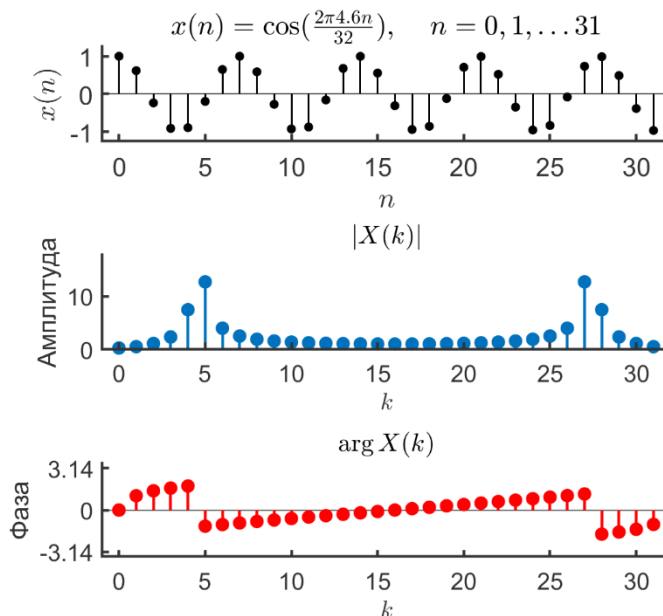


Рис. 1.7 – Временное и частотное представление сигнала с частотой не кратной частоте базисных функций ДПФ

В результате мы видим, что для восстановления такого сигнала нужны *все* базисные функции. Очевидно, что амплитуды компонент больше для тех частот, которые близки к исходной частоте сигнала (в данном случае это частота сигнала находится между $4 \frac{2\pi}{32}$ и $5 \frac{2\pi}{32}$).

1.1.7 Свойства ДПФ

Симметрия ДПФ

Если сигнал $x(n)$ принимает вещественные значения, то его ДПФ $X(k)$ симметрично относительно отсчета $k = N/2$, т.е. удовлетворяет следующим условиям:

$$\operatorname{Re}\{X(k)\} = \operatorname{Re}\{X(N - k)\}, \quad (1.24)$$

$$\operatorname{Im}\{X(k)\} = -\operatorname{Im}\{X(N - k)\}, \quad (1.25)$$

$$|X(k)| = |X(N - k)|, \quad (1.26)$$

$$\arg X(k) = -\arg X(N - k). \quad (1.27)$$

Примеры приведенных типов симметрии мы уже видели на рис. 1.5–1.7.

Выражения (1.24)–(1.27) показывают, что действительная часть $X(k)$ имеет четную симметрию (или просто симметрию), а мнимая – нечетную симметрию (или антисимметрию). Модуль $X(k)$ имеет четную симметрию, а фаза – нечетную симметрию.

Докажем, что

$$\operatorname{Re}[X(k)] = \operatorname{Re}[X(N - k)]. \quad (1.28)$$

Для этого вспомним, что

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi nk}{N}} = \sum_{n=0}^{N-1} x(n) \left[\cos\left(\frac{2\pi nk}{N}\right) - j \sin\left(\frac{2\pi nk}{N}\right) \right],$$

Откуда $\operatorname{Re}\{X(k)\}$ и $\operatorname{Re}\{X(N - k)\}$ можно записать как

$$\begin{aligned}\operatorname{Re}\{X(k)\} &= \sum_{n=0}^{N-1} x(n) \cos\left(\frac{2\pi nk}{N}\right), \\ \operatorname{Re}\{X(N-k)\} &= \sum_{n=0}^{N-1} x(n) \cos\left(\frac{2\pi n(N-k)}{N}\right).\end{aligned}$$

Мы также знаем, что

$$\cos\left(2\pi nk - \frac{2\pi nk}{N}\right) = \cos\left(-\frac{2\pi nk}{N}\right) = \cos\left(\frac{2\pi nk}{N}\right)$$

Таким образом мы доказали справедливость (1.28).

Нечетная симметрия мнимой части доказывается аналогично

$$\operatorname{Im}\{X(k)\} = -\operatorname{Im}\{X(N-k)\}. \quad (1.29)$$

По определению имеем:

$$\begin{aligned}\operatorname{Im}\{X(k)\} &= -\sum_{n=0}^{N-1} x(n) \sin\left(\frac{2\pi nk}{N}\right), \\ \operatorname{Im}\{X(N-k)\} &= \sum_{n=0}^{N-1} x(n) \sin\left(\frac{2\pi n(N-k)}{N}\right).\end{aligned}$$

При этом

$$\sin\left(2\pi nk - \frac{2\pi nk}{N}\right) = \sin\left(-\frac{2\pi nk}{N}\right) = -\sin\left(\frac{2\pi nk}{N}\right),$$

следовательно (1.29) справедливо.

Опираясь на данные свойства симметрии легко показать, что модуль $X(k)$ обладает четной симметрией

$$|X(k)| = |X(N-k)|. \quad (1.30)$$

Вспомним, что

$$\begin{aligned}
|X(k)| &= \sqrt{\operatorname{Re}^2\{X(k)\} + \operatorname{Im}^2\{X(k)\}} = \\
&= \sqrt{\operatorname{Re}^2\{X(N-k)\} + \operatorname{Im}^2\{X(N-k)\}} = |X(N-k)|.
\end{aligned}
\tag{1.31}$$

Свойство (1.30) можно доказать и непосредственно. Для этого распишем левую часть по $|X(k)|$:

$$|X(k)| = \left| \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi nk}{N}} \right|,$$

и правую часть (1.30):

$$|X(N-k)| = \left| \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi n(N-k)}{N}} \right| = \left| \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi nk}{N}} \right|,$$

поскольку $e^{-j\frac{2\pi nN}{N}} = 1$.

Нечетная симметрия фазы $X(k)$:

$$\angle X(k) = -\angle X(N-k),$$

также следует из свойств симметрии действительной и мнимой части $X(k)$, если вспомнить, что

$$\angle X(k) = \operatorname{arctg} \left(\frac{\operatorname{Im}\{X(k)\}}{\operatorname{Re}\{X(k)\}} \right),$$

а

$$\begin{aligned}
\angle X(N-k) &= \operatorname{arctg} \left(\frac{\operatorname{Im}\{X(N-k)\}}{\operatorname{Re}\{X(N-k)\}} \right) = \operatorname{arctg} \left(\frac{-\operatorname{Im}\{X(k)\}}{\operatorname{Re}\{X(k)\}} \right) = \\
&= -\operatorname{arctg} \left(\frac{\operatorname{Im}\{X(k)\}}{\operatorname{Re}\{X(k)\}} \right) = -\angle X(k).
\end{aligned}$$

ДПФ четных и нечетных сигналов

Если $x(n)$ – четный сигнал $x_q(n)$, т.е. $x_q(n) = x_q(-n)$, в случае N -периодических последовательностей имеем $x_q(n) = x_q(N - n)$, то

$$\text{DFT}\{x_q(n)\} = X_q(k) = \sum_{n=0}^{N-1} x_q(n) \cos\left(\frac{2\pi nk}{N}\right). \quad (1.32)$$

Это означает, что четный сигнал раскладывается только по косинусным функциям, которые в свою очередь представляют из себя четные сигналы.

Если $x(n)$ – нечетный сигнал $x_h(n)$, т.е. $x_h(n) = -x_h(-n)$, в случае N -периодических последовательностей имеем $x_h(n) = -x_h(N - n)$, то

$$\text{DFT}\{x_h(n)\} = X_h(k) = -j \sum_{n=0}^{N-1} x_h(n) \sin\left(\frac{2\pi nk}{N}\right). \quad (1.33)$$

Выражение (1.33) показывает, что в разложении $x_h(n)$ участвуют только синусоидальные последовательности, которые сами по себе также представляют собой нечетные сигналы. Другими словами, сигнал $x_h(n)$ имеет нулевую проекцию на функции вида $\cos\left(\frac{2\pi nk}{N}\right)$, поэтому $\text{Re}\{X_h(k)\} = 0$.

Теорема Парсеваля

Нормированная энергия сигнала равна

$$E_x = \sum_{n=0}^{N-1} x^2(n) = \frac{1}{N} \sum_{n=0}^{N-1} |X(k)|^2. \quad (1.34)$$

Правая часть уравнения (1.34) – это среднеквадратическая амплитуда, а левая часть – сумма квадратов амплитуд временного ряда.

Например, рассмотрим ДПФ для сигнала длины 2:

$$\begin{aligned} X(0) &= x(0) + x(1), \\ X(1) &= x(0) - x(1). \end{aligned}$$

Легко убедиться, что для данного примера теорема Парсеваля выполняется:

$$X^2(0) + X^2(1) = (x(0) + x(1))^2 - (x(0) - x(1))^2 = 2(x^2(0) + x^2(1)).$$

Линейность ДПФ

Если последовательности $x_1(n)$ и $x_2(n)$ имеют N -точечные ДПФ $X_1(k)$ и $X_2(k)$, то

$$\text{DFT}\{\alpha x_1(n) + \beta x_2(n)\} = \alpha X_1(k) + \beta X_2(k).$$

Используя это свойство важно убедиться, что исходные последовательности имеют одинаковую длину. Если $x_1(n)$ и $x_2(n)$ различной длины, то более короткая последовательность должны быть дополнена нулями. Например, если $x_1(n)$ имеет длину N_1 , а $x_2(n)$ имеет длину N_2 , причем $N_2 > N_1$, то $x_1(n)$ можно рассматривать как последовательность длины N_2 , которая имеет на конце $N_2 - N_1$ нулей, таким образом от обеих последовательностей можно вычислить ДПФ длины N_2 .

Свойство сдвига во временной и частотной области

Справедливо следующее соотношение:

$$\text{DFT}\{x(n - m)\} = X(k)e^{-j\frac{2\pi mk}{N}}. \quad (1.35)$$

Выражение (1.35) показывает, что сдвиг временного сигнала на m отсчетов соответствует умножению его ДПФ-образа на комплексную экспоненту $e^{-j\frac{2\pi mk}{N}}$. Однако, поскольку представление сигнала посредством ДПФ подразумевает периодическое продолжение сигнала, то в результате сдвига произойдет циклический сдвиг. Поэтому выражение (1.35) можно переписать, как

$$\text{DFT}\{x(\langle n - m \rangle_N)\} = X(k)e^{-j\frac{2\pi mk}{N}}, \quad (1.36)$$

где $\langle a \rangle_N$ обозначает операцию нахождения остатка от деления числа a на N .

Аналогичным образом можно показать, что

$$\text{DFT}\left\{x(n)e^{j\frac{2\pi mk}{N}}\right\} = X(\langle k - m \rangle_N). \quad (1.37)$$

Свойство частотного сдвига проиллюстрировано на рис. 1.8.

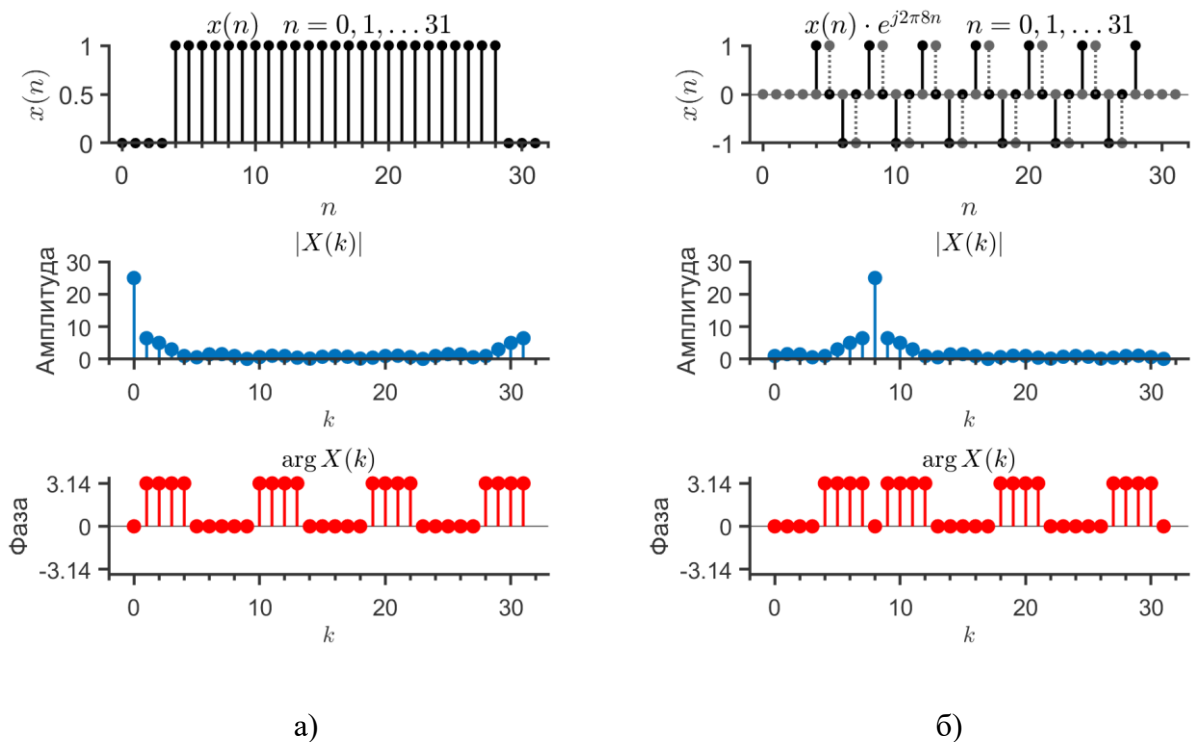


Рис. 1.8 – Свойство частотного сдвига ДПФ: а) исходный сигнал и его спектр; б) сигнал после умножения на экспоненту и его спектр

Свойство обращения времени

Если $X(k)$ есть ДПФ вещественного сигнала $x(n)$, то обращенный во времени сигнал $x(\langle -n \rangle_N)$ имеет следующее ДПФ:

$$\text{DFT}[x(\langle -n \rangle_N)] = X^*(k).$$

Заметим, что $|X(k)| = |X^*(k)|$. Таким образом амплитудный спектр сигнала не меняется, при обращении времени.

Круговая (периодическая) свертка

ДПФ можно использовать для вычисления круговой свертки:

$$\text{DFT}\{x_1(n) \circledast x_2(n)\} = X_1(k)X_2(k), \quad (1.38)$$

где $x_1(n), x_2(n)$ – конечные периодические последовательности одинаковой длины.

Выражение (1.38) можно записать как:

$$x_3(n) = x_1(n) \circledast x_2(n) = \sum_{m=0}^{N-1} x_1(m)x_2(\langle n - m \rangle_N) = \text{IDFT}\{X_1(k)X_2(k)\}.$$

Приведем небольшой пример. Так линейная (нециклическая) свертка последовательности x_0, x_1, x_2, x_3 с другой последовательностью y_0, y_1, y_2, y_3 дает

$$\begin{aligned} x * y &= x_0y_0, \\ & x_0y_1 + x_1y_0, \\ & x_0y_2 + x_1y_1 + x_2y_0, \\ & x_0y_3 + x_1y_2 + x_2y_1 + x_3y_0, \\ & x_1y_3 + x_2y_2 + x_3y_1, \\ & x_2y_3 + x_3y_2, \\ & x_3y_3. \end{aligned}$$

Круговая (циклическая) свертка этих же последовательностей имеет вид

$$\begin{aligned} x \circledast y &= x_0y_0 + x_1y_3 + x_2y_2 + x_3y_1, \\ & x_0y_1 + x_1y_0 + x_2y_3 + x_3y_2, \\ & x_0y_2 + x_1y_1 + x_2y_0 + x_3y_3, \\ & x_0y_3 + x_1y_2 + x_2y_1 + x_3y_0, \\ & x_0y_0 + x_1y_3 + x_2y_2 + x_3y_1, \\ & x_0y_1 + x_1y_0 + x_2y_3 + x_3y_2, \\ & x_0y_2 + x_1y_1 + x_2y_0 + x_3y_3. \end{aligned}$$

Если исходные последовательности дополнить нулями, то с помощью ДПФ можно вычислить линейную свертку.

Теорема о свертке в частотной области

Следующее уравнение – это формулировка теоремы о свертке в частотной области:

$$\text{DFT}[x_1(n)x_2(n)] = \frac{1}{N} X_1(k) \circledast X_2(k). \quad (1.39)$$

Выражение (1.38) дает основания утверждать, что свертка во временной области эквивалентна умножению в частотной, тогда как (1.39) приводит к

наблюдению о том, что свертка в частотной области эквивалентна умножению во временной.

1.1.8 Перестановка коэффициентов ДПФ

Необходимо всегда помнить, что ДПФ является N -периодической последовательностью, поэтому для него справедливо равенство:

$$X(k + mN) = X(k), \quad \forall k \in \mathbb{Z}. \quad (1.40)$$

Рассмотрим 8-точечный сигнал

$$x(n) = \cos(2\pi fn), \quad f = 1/8, \quad n = 0, 1, \dots, 7.$$

Его ДПФ будет иметь вид:

$$X = [0 \ 4 \ 0 \ 0 \ 0 \ 0 \ 0 \ 4],$$

т.е. коэффициенты ДПФ расположены следующим образом

$$X(0), X(1), X(2), X(3), X(4), X(5), X(6), X(7).$$

Формально все правильно, но интуитивно воспринимается нехорошо. Мы видим, что в конце массива имеется ненулевой компонент, что может наводить на мысль, что в сигнале есть какая-то компонента с высокой частотой. На самом деле более предпочтительно представление, в котором имеются отрицательные частоты. По аналогии с тем, как мы привыкли расписывать косинус по формуле Эйлера:

$$\cos \omega t = \frac{1}{2} (e^{j\omega t} + e^{-j\omega t}),$$

исходя из которой, мы видим, что косинус состоит из двух экспонент с положительной и отрицательной частотой. В этом же смысле и для ДПФ предпочтительнее следующий порядок выходных коэффициентов:

$$X(-4), X(-3), X(-2), X(-1), X(0), X(1), X(2), X(3).$$

Как найти $X(-k)$? Нужно использовать свойство периодичности (1.40). Тогда получим:

$$X(5) = X(-3), X(6) = X(-2), X(7) = X(-1), X(4) = X(-4).$$

В Matlab есть функция выполняющая данный сдвиг – `fftshift`.

```
>> X = fft(cos(2*pi*1/8*(0:7)))
X =
     0     4     0     0     0     0     0     4
>> Xs = fftshift(X)
Xs =
     0     0     0     4     0     4     0     0
k= -4    -3    -2    -1     0     1     2     3
```

На рис. 1.9 приведен пример вычисления ДПФ от дискретного прямоугольного импульса, показан его амплитудный спектр, а также выполнена операция перестановки коэффициентов ДПФ. Полученное представление более наглядно, поскольку лучше отражает тот факт, что основная энергия данного сигнала сконцентрирована в низкочастотном диапазоне (т.е. около частоты с индексом 0).

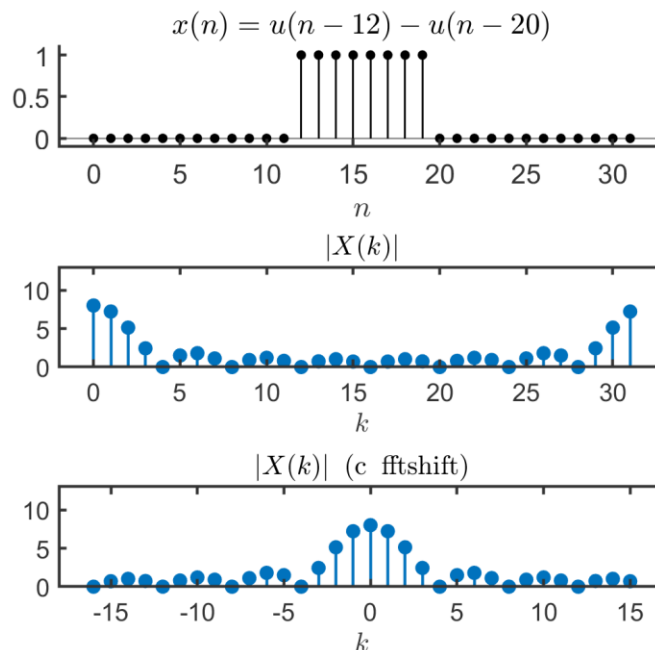


Рис. 1.9 – Пример использования перестановки `fftshift`

1.1.9 Какое значение имеет индекс k в обозначении ДПФ $X(k)$?

После вычисления ДПФ от последовательности $x(n)$, которая обычно получается путем дискретизации сигнала $x(t)$ нас может интересовать значение частоты компоненты с максимальной амплитудой $|X(k)|$. Выводя график $|X(k)|$ обычно по горизонтальной оси мы откладываем частотный индекс k . Однако как этот индекс соотносится с реальной физической частотой, которая содержалась в исходном сигнале $x(t)$?

Предположим, что в результате дискретизации $x(t)$ мы получили дискретный сигнал $x(n)$ длины N такой, что в нем содержится ровно один период функции косинус. Это можно записать следующим образом

$$x(n) = \cos\left(\frac{2\pi \cdot 1}{N} n\right), \quad n = 0, 1, \dots, N - 1. \quad (1.41)$$

Мы специально ввели умножение на единицу, чтобы показать, что сигнал содержит ровно один период функции косинус. Сигнал (1.41) можно также записать, используя понятие нормированной частоты (см. выражение **Ошибка! Источник ссылки не найден.**):

$$x(n) = \cos(2\pi f n), \quad (1.42)$$

где $f = 1/N$.

Из предыдущих разделов мы уже знаем, что ДПФ от сигнала (1.41) будет равно

$$X(k) = \frac{N}{2} \delta(k - 1) + \frac{N}{2} \delta(k - (N - 1)), \quad k = 0, 1, \dots, N - 1. \quad (1.43)$$

Т.е. основная информация попадет на частотный отсчет с номером $k = 1$, второй частотный отсчет $k = N - 1$, на котором также содержится информация является лишь следствием симметрии ДПФ (1.28). Теперь вернемся к дискретному сигналу (1.41) и попробуем ответить на вопрос: какова была реальная аналоговая частота F этого сигнала до дискретизации? Естественно, чтобы ответить на этот вопрос мы должны знать частоту дискретизации f_s . Поскольку мы считаем ее известной, то можем легко вычислить, умножая нормированную частоту на частоту дискретизации:

$$F = f \cdot f_s = \frac{f_s}{N}. \quad (1.44)$$

Зафиксируем этот результат. Дискретный сигнал (1.41) при частоте дискретизации f_s соответствует аналоговому сигналу с частотой $F = f_s/N$. Получается, что значение ДПФ $X(1)$ для соответствует частоте $F = f_s/N$. Если бы исходный сигнал имел нормированную частоту $f = 2/N$ (т.е. в N отсчетов укладывалось бы 2 периода), то ДПФ $X(k)$ содержало бы информацию на отсчете с номером $k = 2$, а аналоговая частота сигнала равнялась $F = 2 \cdot f_s/N$. В общем случае, для сигнала с нормированной частотой $f = m/N$ имеем ненулевой отсчет ДПФ $X(m)$ и частотой равной $F = m \cdot f_s/N$.

Таким образом мы выяснили, что отсчету ДПФ $X(k)$ с целым индексом k соответствует аналоговая частота

$$F_k = k \frac{f_s}{N}. \quad (1.45)$$

Таким образом, зная частоту дискретизации сигнала мы всегда можем отмасштабировать частотную ось на графике, чтобы вместо частотного индекса k на нем отображались аналоговые частоты F_k .

Расстояние между двумя последовательными значениями k часто называют **частотным разрешением ДПФ**

$$\Delta f = f_s/N. \quad (1.46)$$

Допустим, что максимальное значение $|X(k)|$ соответствует индексу k_m , тогда значение аналоговой частоты равно

$$\frac{k_m f_s}{N} \text{ Гц.}$$

1.1.10 Применение MATLAB для расчета ДПФ сигнала

Ниже приведен пример как можно при помощи MATLAB вычислить ДПФ сигнала

$$x(n) = \sin^3\left(2\pi \frac{2}{N} n\right), \quad n = 0, 1, \dots, N - 1, \quad (1.47)$$

для $N = 40$. В примере также показано, как построить график сигнала во временной области и отобразить действительную и мнимую часть ДПФ.

```
f = @(n,N)(sin(2*pi*(2/N*(0:N-1))).^3);

N = 40;
n = 0:N-1;      % Временные отсчеты
x = f(n,N);     % Вычисление сигнала
X = fft(x);     % Вычисление ДПФ
k = 0:N-1;     % Частотный индекс

figure('Position',[100 100 800 450]);
subplot(2,2,[1 2]);
stem(n,x,'LineWidth',1.5);

xlabel('$n$', 'interp', 'latex');
title('$x(n)$', 'interp', 'latex');

subplot(2,2,3);
stem(k,real(X),'LineWidth',1.5);
xlabel('$k$', 'interp', 'latex');
title('$\mathrm{Re}\{X(k)\}$', 'interp', 'latex');

subplot(2,2,4);
stem(k,imag(X),'r','LineWidth',1.5);
xlabel('$k$', 'interp', 'latex');
title('$\mathrm{Im}\{X(k)\}$', 'interp', 'latex');

FileName = 'Task_1.png';
print('-dpng', '-opengl', '-r300',FileName);
```

Результат работы программы представлен на рис. 1.10.

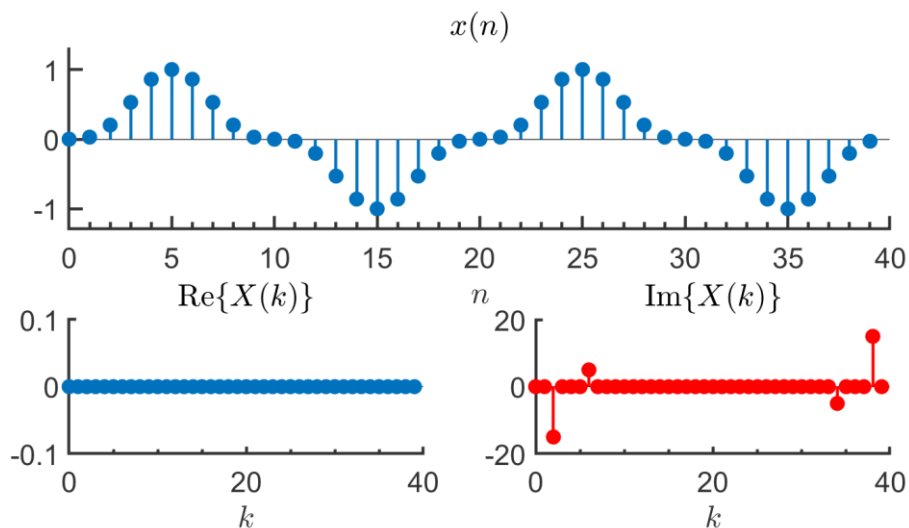


Рис. 1.10 – ДПФ сигнала (1.19)

Поскольку исходный сигнал $x(n)$ является нечетным его ДПФ $X(k)$ является мнимой функцией (см. выражение (1.33)). Полученный график это подтверждает, поскольку $\text{Re}\{X(k)\} = 0$. Также на основании того, что исходный сигнал представляет собой вещественную функцию его мнимая часть его ДПФ обладает свойством антисимметричности (1.25), что также видно из приведенного графика.

1.1.11 Гармоническая форма ДПФ

В некоторых задачах удобно представлять формулу обратного ДПФ не с использованием комплексных экспонент (см. формулу (1.17)), а в гармонической форме:

$$x(n) = \sum_{k=0}^{N/2} M_k \cos\left(\frac{2\pi nk}{N} + \varphi_k\right), \quad (1.48)$$

Значения амплитуд M_k и фаз φ_k можно непосредственно определить из коэффициентов ДПФ $X(k)$. По сути, M_k представляет из себя амплитудный спектр, а φ_k – фазовый спектр. Расчетные формулы имеют следующий вид:

$$M_k = \sqrt{\left(\text{Re} \hat{X}(k)\right)^2 + \left(\text{Im} \hat{X}(k)\right)^2}, \quad \varphi_k = \arctg \frac{\text{Im}\{\hat{X}(k)\}}{\text{Re}\{\hat{X}(k)\}}, \quad (1.49)$$

где

$$\text{Re} \hat{X}(k) = \frac{\text{Re} X(k)}{N/2}, \quad \text{Im} \hat{X}(k) = \frac{\text{Im} X(k)}{N/2}, \quad (1.50)$$

за исключением двух специальных случаев:

$$\text{Re} \hat{X}(0) = \frac{\text{Re} X(0)}{N}, \quad \text{Re} \hat{X}(N/2) = \frac{\text{Re} X(N/2)}{N}. \quad (1.51)$$

Чтобы понять откуда берется масштабирование (1.50)-(1.51) рассмотрим вспомним пример разложения (1.18). В этом примере рассматривался сигнал

$$x(n) = \cos\left(\frac{2\pi 4n}{32}\right)$$

и было показано, что его ДПФ имеет вид:

$$X(k) = \frac{N}{2} \delta(k - 4) + \frac{N}{2} \delta(k - 28).$$

Теперь покажем, что, используя $X(k)$, и гармоническую форму ДПФ можно выполнить восстановление сигнала $x(n)$. Очевидно, что из всей суммы

(1.48) ненулевым будет только значением при $k = 4$. В соответствии с (1.50) найдем, что

$$\operatorname{Re}\{\hat{X}(4)\} = \frac{\operatorname{Re}\{X(4)\}}{N/2} = 1, \quad \operatorname{Im}\{\hat{X}(4)\} = \frac{\operatorname{Im}\{X(4)\}}{N/2} = 0.$$

Далее вычислим амплитуду 4-й гармоники в соответствии с (1.49):

$$M_4 = \sqrt{1^2 + 0^2} = \mathbf{1}, \quad \varphi_4 = \operatorname{arctg} \frac{0}{1} = \mathbf{0}.$$

В итоге, подставляя найденные значения в выражение (1.48) получим:

$$x[n] = \sum_{k=0}^{N/2} M_k \cos\left(\frac{2\pi nk}{N} + \varphi_k\right) = \mathbf{1} \cdot \cos\left(\frac{2\pi 4n}{N} + \mathbf{0}\right).$$

Удобство гармонической формы ДПФ (1.48) состоит в том, что сигнал представляется с помощью действительных, а не комплексных функций. Причем параметры M_k и φ_k имеют простую и понятную интерпретацию. Мы ещё раз убедились, что помощью ДПФ можно представить сигнал в виде суммы гармонических колебаний (с амплитудами M_k и фазами φ_k). Для правильного вычисления фазы в MATLAB предусмотрена функция `atan2(imX, reX)`, которая возвращает значение угла в диапазоне $[-\pi, \pi]$.

1.1.12 Частотный анализ сигнала при помощи ДПФ

Для анализа амплитудного спектра сигнала удобно отобразить на графике вместо частотных индексов реальные частоты (см. выражение (1.45)). Рассмотрим следующий пример. Пусть задан сигнал, состоящий из трех компонент

$$x(n\Delta T) = A_1 \sin(2\pi f_1 n\Delta T) + A_2 \sin(2\pi f_2 n\Delta T) + A_3 \sin(2\pi f_3 n\Delta T) \quad (1.52)$$

где A_1, A_2, A_3 – амплитуды, а f_1, f_2, f_3 – частоты синусоидальных составляющих, $n\Delta T$ – дискретное время.

Необходимо построить амплитудный спектр сигнала $x(n\Delta T)$ и отобразить на графике реальные частоты. Ниже приведен MATLAB-код, выполняющий данную задачу.

```
fs = 1000;           % Частота дискретизации
N = 100;            % Длина сигнала
t = (0:N-1)*(1/fs); % Дискретное время

A = [0.1 0.3 0.05]; % Вектор амплитуд
f = [10 100 400];   % Вектор частот
```

```

x = A(1)*sin(2*pi*f(1).*t)+...
    A(2)*sin(2*pi*f(2).*t)+...
    A(3)*sin(2*pi*f(3).*t);

k = N/2 + 1;           % Частотные индексы
freq = (0:k-1)*(fs/N); % Частотная сетка
X = fft(x);           % Вычисление ДПФ
abs_X = abs(X(1:N/2+1)); % Половина амплитудного спектра

subplot(311);
stem(t,x,'filled','k','LineWidth',1.5,'MarkerSize',4); grid on;
xlabel('\it Время, с');
ylabel('\it Амплитуда');
title('$x(t) = A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t) + A_3 \sin(2\pi f_3 t)$','interp','latex');

subplot(312);
stem([0:k-1],abs_X,"filled",'LineWidth',1.5); grid on;
xlabel('\it Частотный индекс, k');
ylabel('\it Амплитуда');
title('$|X(k)|$','interp','latex');

subplot(313);
stem(freq,abs_X,"filled",'LineWidth',1.5);grid on;
xlabel('\it Частота, Гц');
ylabel('\it Амплитуда');
title('$|X(k)|$','interp','latex');

print('-dpng', '-opengl', '-r200', 'fft_real_freq.png');

```

Результат работы программы представлен на рис. 1.11.

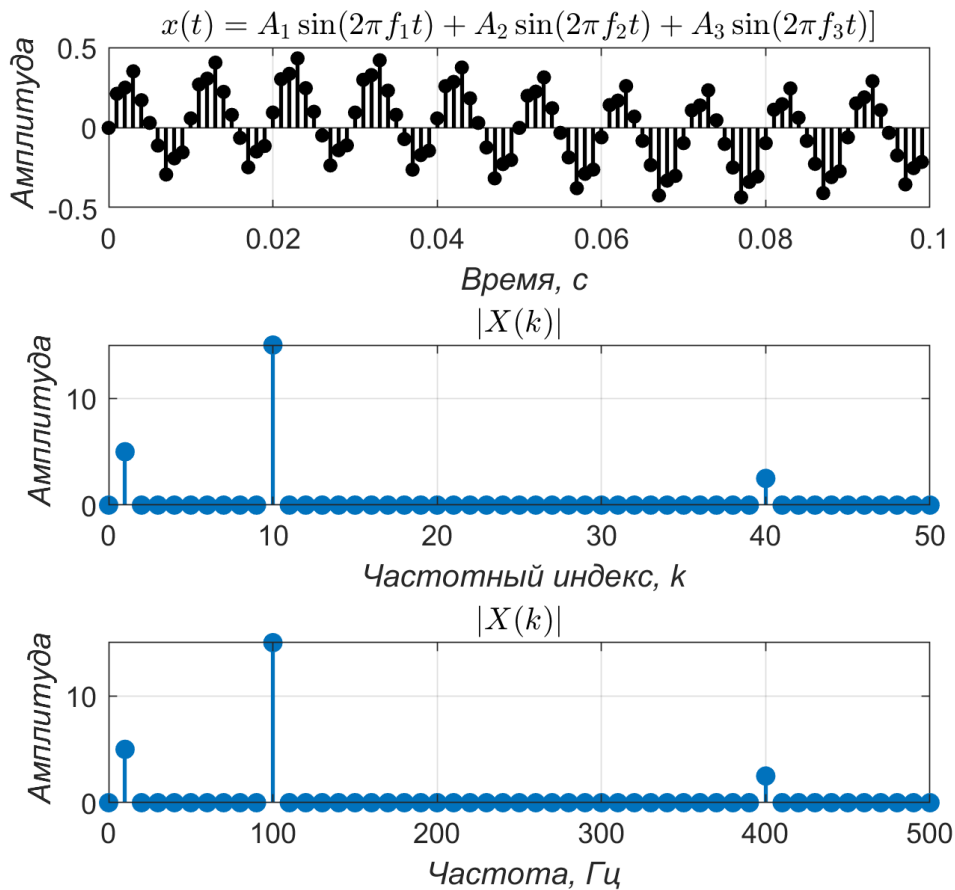


Рис. 1.11 – ДПФ сигнала (1.52)

В спектре $|X(k)|$ отчетливо видны все три компоненты сигнала (1.52). Нижний график также позволяет выявить реальные частоты компонент, содержащихся в сигнале. Остается лишь один нюанс – мы не видим в спектре реальных амплитуд частотных компонент. Так, например, значение $|X(1)|$, которое соответствует компоненте с частотой 10 Гц, равно 5, хотя из приведенного кода видно, что реальная амплитуда компоненты с частотой 10 Гц равняется $A_1 = 0,1$. Чтобы понять почему так произошло следует обратиться к выражению (1.43), из которого становится понятно, что дискретный синусоидальный сигнал с единичной амплитудой будет в спектре ДПФ давать амплитудное значение $N/2$. Соответственно, чтобы увидеть реальные значения амплитуд синусоидальных компонент сигнала нужно отсчеты ДПФ $X(k)$ разделить на $N/2$. Единственное исключение будет только для компоненты $X(0)$, который необходимо для нормировки делить на N .

Таким образом, нормировка амплитудного спектра описывается выражением:

$$X_{norm}(k) = \begin{cases} X(k)/N, & k = 0, \\ X(k)/(N/2), & k \neq 0. \end{cases} \quad (1.53)$$

После выполнения описанной нормировки амплитудный спектр рассматриваемого сигнала будет выглядеть, как показано на рис. 1.12.

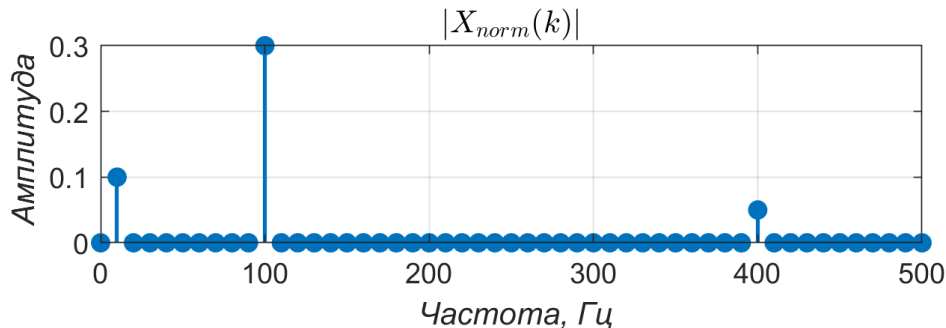


Рис. 1.12 – Амплитудный спектр сигнала (1.52) после нормировки (1.53)

Можно заметить, что компоненты спектра на рис. 1.12 совпадают с амплитудами, которые мы задавали в приведенном выше коде: $|X_{norm}(1)| = A_1 = 0,1$, $|X_{norm}(10)| = A_2 = 0,3$, а $|X_{norm}(40)| = A_3 = 0,05$. Приведенный пример может показаться достаточно тривиальным, поскольку мы заранее знали количество синусоидальных компонент, которые содержатся в сигнале. Однако, если представить, что нам известны лишь отсчеты сигнала во временной области, показанные на верхнем графике рис. 1.11, то окажется, что без такого инструмента как ДПФ далеко не так просто определить какие именно синусоидальные компоненты входят в состав сигнала. При помощи же ДПФ мы не только смогли определить количество компонент, но и нашли значение их частот и амплитуд.

1.1.13 Применение ДПФ: интерполяция сигнала во временной области

Пусть есть рассчитанное ДПФ от N -точечной последовательности $x(n)$:

$$X(k) = \text{DFT}\{x(n)\}.$$

Если мы хотим получить в M раз больше отсчетов во временной области, то мы должны добавить $(N - 1)M$ нулей в середину массива $X(k)$ и таким образом получить последовательность $X_{int}(k)$ длиной NM . Далее необходимо к $X_{int}(k)$ обратное ДПФ, чтобы получить интерполированную версию сигнала во временной области:

$$x_{int}(n) = \text{IDFT}\{X_{int}(k)\}.$$

Чтобы быть уверенными, что интерполированная последовательность будет иметь только вещественные значения нужно обеспечить симметричную комплексную-сопряженность коэффициентов $X_{int}(k)$. Если последовательность $X(k)$ содержит ненулевой отсчет $X(N/2)$, т.е. компоненту соответствующую $f_s/2$, то необходимо выполнить следующие шаги, чтобы гарантировать сохранность комплексно-сопряженной симметрии:

1) вычисляя N -точечное ДПФ последовательности $x(n)$, получите N частотных отсчетов $X(k)$;

2) создайте массив $X_{int}(k)$ длиной NM , заполненный нулями;

3) присвойте $X_{int}(k) = X(k)$ для $0 \leq k \leq \frac{N}{2} - 1$;

4) присвойте $X_{int}(\frac{N}{2})$ и $X_{int}(MN - \frac{N}{2})$ значения равные $X(\frac{N}{2})/2$;

5) присвойте $X_{int}(MN - N + k) = X(k)$ для $\frac{N}{2} + 1 \leq k \leq N - 1$;

6) вычисляя NM -точечное обратное ДПФ получите интерполированную версию последовательности $x_{int}(n)$.

7) в результате описанной процедуры интерполяции и происходит ослабление амплитуды на коэффициент $1/M$; для восстановления исходной амплитуды последовательность необходимо умножить на M .

Ниже приведен пример того, как должно происходить формирования спектра сигнала для выполнения интерполяции в случае, когда $N = 8$, а $M = 2$.

Таблица 1.1 – Модификация спектра сигнала, выполняемая для интерполяции

k	$X_{int}(k)$	k	$X_{int}(k)$
0	$X(0)$	8	0
1	$X(1)$	9	0
2	$X(2)$	10	0
3	$X(3)$	11	0
4	$X(4)/2$	12	$X(4)/2$
5	0	13	$X(5)$
6	0	14	$X(6)$
7	0	15	$X(7)$

1.1.14 Кратковременное преобразование Фурье

Кратковременное преобразование Фурье (КВПФ) (англ. *STFT – short-time Fourier transform*) широко используется для анализа, модификации и синтеза звуковых сигналов. КВПФ можно рассматривать как преобразование со скользящим окном, которое имеет вид:

$$X(k, l) = \sum_{n=0}^{N-1} h(n)x(n + lL)e^{-j\omega_k n}, \quad (1.54)$$

где $x(t)$ – входной дискретный сигнал (в данном случае мы используем переменную t для обозначения глобального индекса отсчетов сигнала), $h(n)$ – ограниченная во времени оконная функция, а $\omega_k = \frac{2\pi k}{M}$, $k = 0, 1, \dots, M - 1$ – частотный индекс, L – временной шаг анализа (расстояние между соседними фреймами), l – номер фрейма анализа. Легко заметить, что (1.54) является вычислением дискретного преобразования Фурье (ДПФ) для сигнала $h(n)x(n + lL)$. Таким образом, $X(k, l)$ это последовательность локализованных во времени спектров.

На рис. 1.13 приведен пример как речевой сигнал $x(t)$ разбивается на фреймы умножением на оконную функцию $h(n)$. В данном случае n – это «локальный» индекс времени внутри одного фрейма, который пробегает значения от 0 до $N - 1$. В приведенном на рис. 1.13 примере $N = 1024$, а шаг анализа $L = 512$.

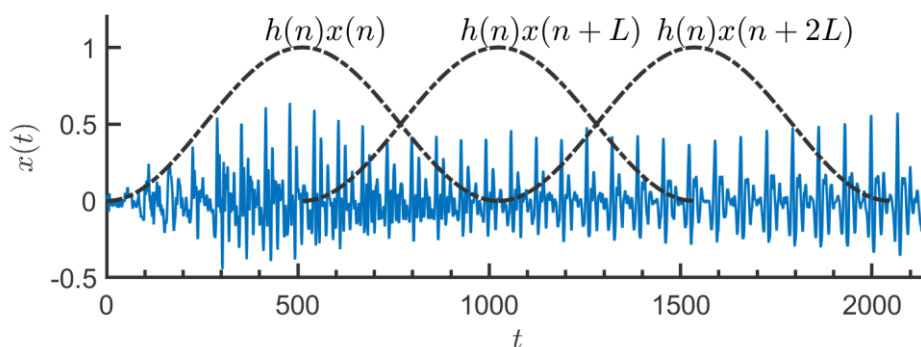


Рис. 1.13 –Разбиение сигнала $x(t)$ на фреймы

Спектрограмма (частотно-временное представление сигнала)

КВПФ (1.54) позволяет получить **частотно-временное представление** сигнала $x(n)$. Вместо одномерного сигнала $x(n)$ у нас появляется функция двух переменных $X(k, l)$, которая отражает частотный состав сигнала (k – индекс частоты) на фрейме с номером l :

$$X(k, l) = \text{DFT}\{h(n)x(n + lL)\}, \quad n = 0, 1, \dots, N - 1. \quad (1.55)$$

Отображая поверхность $X(k, l)$ можно увидеть динамику изменения частотного состава сигнала во времени. Учитывая, что $X(k, l)$ принимает комплексные значения графически отображают модуль $|X(k, l)|$. Двумерное представление $|X(k, l)|$ называют *спектрограммой* сигнала. Пример спектрограммы для речевого сигнала («да») показан на рис. 1.14.

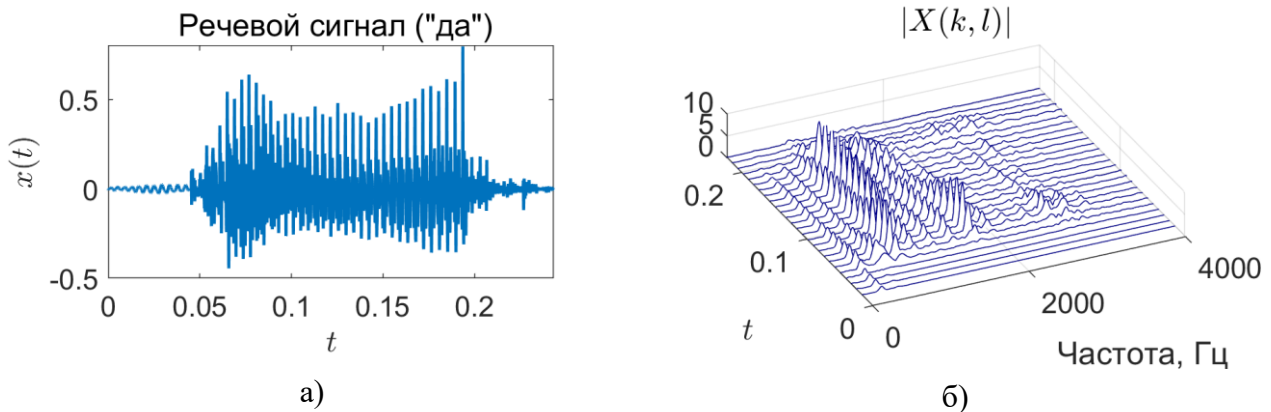


Рис. 1.14 – Речевого сигнал («да»): а) представление во временной области; б) спектрограмма сигнала

Величины $|X(k, l)|$ как правило имеют большой динамический диапазон и изменяются в пределах от 10^{-5} до 1, поэтому для удобства графического отображения часто переходят к логарифмическим единицам – децибелам:

$$\text{Spec}_{dB}(k, l) = 20 \log_{10}|X(k, l)|. \quad (1.56)$$

Величина $\text{Spec}_{dB}(k, l)$ как правило находится в пределах от -100 до 0 дБ. На практике чаще всего спектрограмму отображают в виде тепловой карты. **Тепловая карта** представляет собой способ визуализации функции двух переменных в виде двумерного изображения, при котором значение функции отображается цветом. На рис. 1.15 показан пример отображения спектрограммы речевого сигнала в виде тепловой карты.

Из приведенных на рис. 1.14,б и 1.15 спектрограмм видно, что сигнал звонкого звука /а/ состоит из нескольких частотных компонент, которые разнесены друг от друга по частоте на величину основного тона f_0 . В приведенном примере частота основного тона находится на уровне 250 Гц. Спектрограмма дает

представление о том, что частота основного тона не является постоянной, а изменяется в процессе фонации. Компоненты сигнала, с частотами, которые кратны частоте основного тона, называются *гармониками*. Анализируя приведенные спектрограммы, мы можем заметить, что в процессе произнесения звука /а/ амплитуды гармоник плавно изменяются. Изменение гармоник связано с изменением *огубающей спектра* речевого сигнала, которая в свою очередь отвечает за тембр речи.

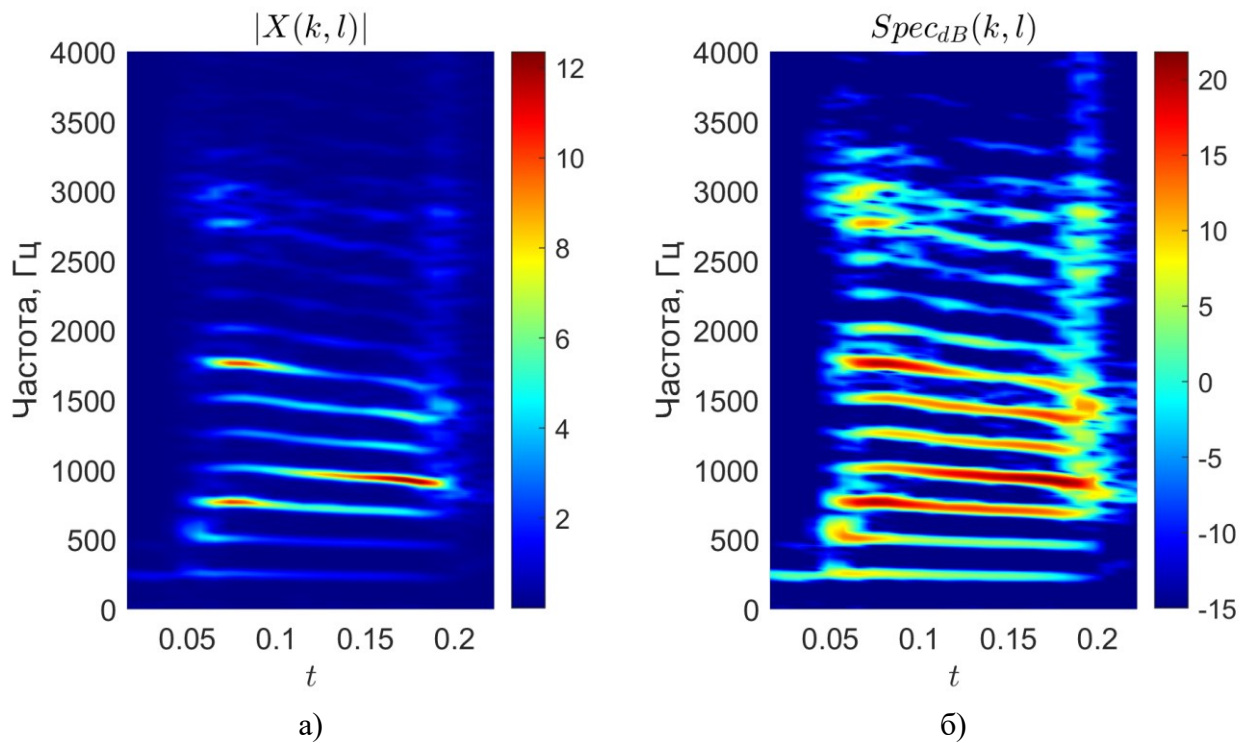


Рис. 1.15 – Отображение спектрограммы речевого сигнала («да») в виде тепловой карты:
 а) линейный масштаб; б) логарифмический масштаб

На практике строят спектрограммы двух видов – широкополосные и узкополосные. Способ получения спектрограммы в обоих случаях одинаков (см. выражение (1.55)), отличие заключается в выборе параметра N – длины анализируемого фрейма, которая в свою очередь определяет частотное и временное разрешение спектрограммы. Ниже приведен MATLAB-код, который позволяет построить спектрограмму речевого сигнала.

```
[x,fs] = audioread('female_speech.wav'); % чтение файла

Ln = length(x);

% Параметры анализа
L=32; % перекрытие между фреймами
N=round(fs/60); % размер фрейма для узкополосной спектрограммы
% N=round(fs/180); % размер фрейма для широкополосной спектрограммы
```



```

N_bins=round(N/2)+1; % число эффективных частотных отсчетов
w = hann(N);          % оконная функция Ханна для анализа

N_frames=floor((Ln-N)/L); % число фреймов

Spec = zeros(N_bins,N_frames);

for l = 1:N_frames
    fprintf('Фрейм %d из %d\n',l,N_frames);

    x_frame = x(1+(l-1)*L:(l-1)*L+N);
    x_frame = x_frame.*w;

    % Получение X(k,m) по формуле (4.50)
    X_kl = fft(x_frame);
    X_kl = X_kl(1:N_bins)/(N/2);

    Spec(:,m) = 20*log10(abs(X_kl));
end

%% Визуализация
figure('Position',[100 100 800 550]);

subplot(3,1,[1 2]);
pcolor(Spec);
shading flat;
colormap(jet);
set(gca,'CLim',[-75 -15]);
ylabel('Частотный индекс \it k');
xlabel('Временной индекс \it l');

subplot(3,1,3);
plot(0:Ln-1,x);
xlabel('Номер отсчета,\it n');
ylabel('Амплитуда');
xlim([0 Ln-1]);

```

При помощи приведённого кода получены широкополосная (рис. 1.16) и узкополосная спектрограмма речевого сигнала (рис. 1.17). Анализируемый сигнал имел частоту дискретизации $f_s = 16000$ Гц. В случае широкополосной спектрограммы размер фрейма $N = 89$ отсчетов и частотное разрешение по формуле (1.46) равнялось $\Delta f = f_s/N \approx 180$ Гц. На широкополосной спектрограмме хорошо видны периоды основного тона речевого сигнала. Но разрешение по частоте недостаточно для представления тонкой структуры спектра речи. Другими

словами, на широкополосной спектрограмме не видны отдельные гармоники основного тона. Обратите внимание на то, что для шипящих звуков /sh/ основная энергия сконцентрирована в области верхних частот. Во временной области эти сигналы имеют шумоподобный вид. Широкополосная спектрограмма по существу представляет собой аппроксимацию огибающей спектра, как функции времени.

В случае узкополосной спектрограммы размер фрейма $N = 267$ отсчетов, а частотное разрешение равнялось $\Delta f = f_s/N \approx 60$ Гц. Если сравнить её с узкополосной спектрограммой, то легко заметить, что на ней отчётливо видны отдельные гармоники спектра. При этом разрешение по времени ухудшилось.

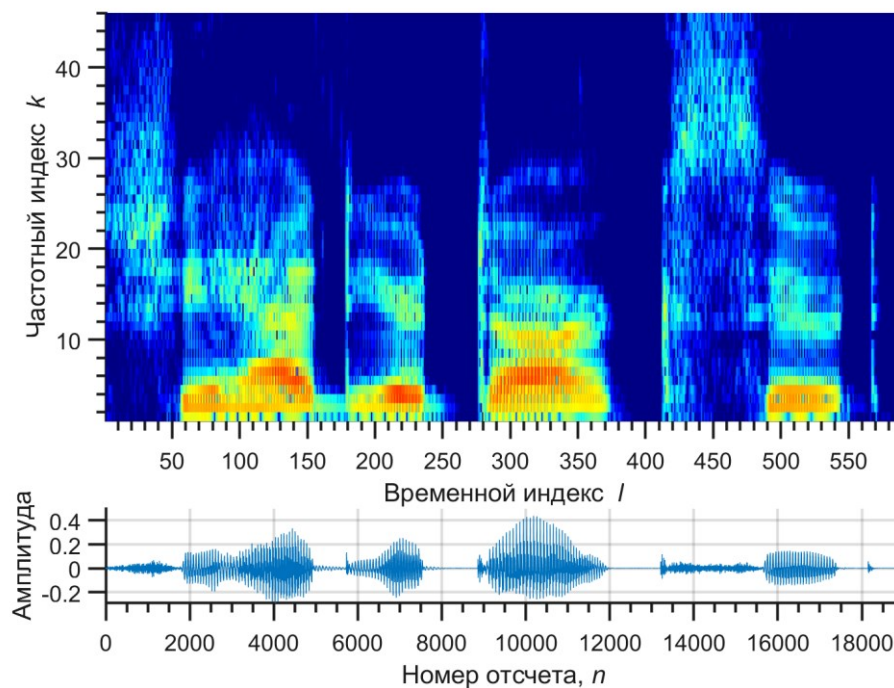


Рис. 1.16 – Широкополосная спектрограмма речевого сигнала («she had your dark suit»)

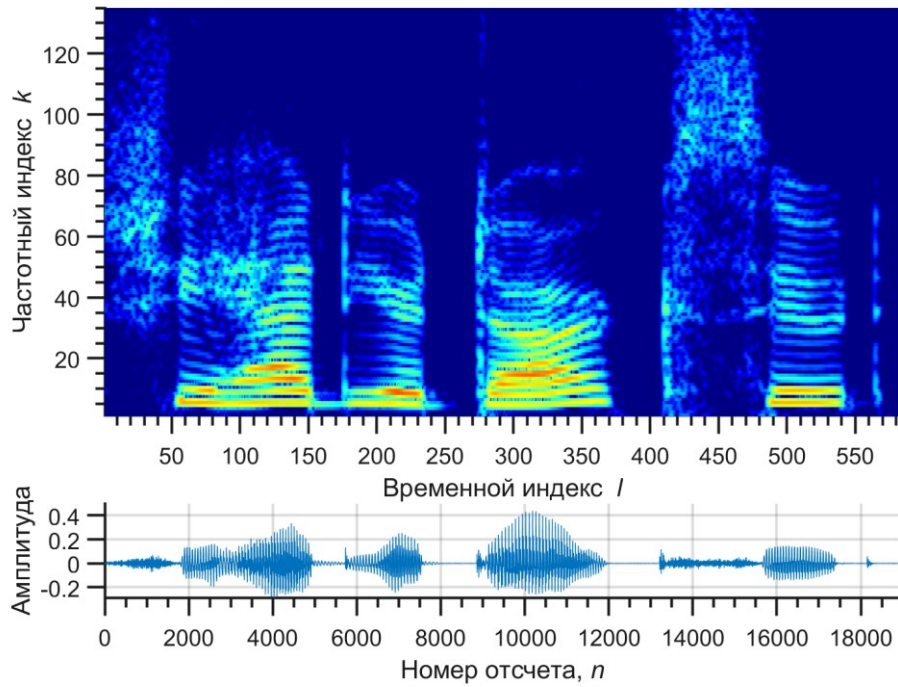


Рис. 1.17 – Узкополосная спектрограмма речевого сигнала («she had your dark suit»)

КВПФ: восстановление сигнала

Выражение (1.54) представляет собой *операцию анализа*. Чтобы выполнять модификацию сигнала на основе КВПФ необходимо иметь выражение и для синтеза сигнала. В идеальном случае синтез должен приводить к полному восстановлению сигнала (при условии, что спектральные составляющие $X(k, l)$ остаются неизменными). Для восстановления нужно использовать ОДПФ:

$$\hat{x}_l(n) = \text{IDFT}\{X(k, l)\} = h(n)x(n + lL), \quad 0 \leq n \leq N - 1.$$

Это выражение записано относительно внутрифреймового индекса n , его можно переписать относительно глобального временного индекса $t = n + lL$

$$\hat{x}_l(t - lL) = h(t - lL)x(t), \quad lL \leq t \leq lL + N - 1. \quad (1.57)$$

Таким образом у нас есть последовательность фреймов $\hat{x}_l(t - lL)$ для всех l . Нужно объединить их, чтобы в результате получился исходный сигнал $x(t)$. Для этого нам необходимо устранить эффект умножения на окно $h(t - lL)$ – см. выражение (1.57). Процесс восстановления сигнала производится методом пере-

крытия с суммированием (англ. *OLA – overlap-add*), с применением синтезирующего окна $g(n)$. Мы умножаем каждый восстановленный фрейм $\hat{x}_l(t - lL)$ на синтезирующее окно и складываем все полученные фреймы:

$$\begin{aligned}\hat{x}_l(t) &= \sum_l g(t - lL) \cdot \text{IDFT}\{X(k, l)\} = \\ &= \sum_l g(t - lL)h(t - lL)x(t).\end{aligned}\tag{1.58}$$

Таким образом, для восстановления всего сигнала, каждый фрейм, восстановленный ОДПФ, взвешивается на синтезирующее окно и добавляется к соседним окнам с перекрытием во времени. В выражении (1.58) $x(t)$ может быть вынесен за скобки:

$$\hat{x}_l(t) = x(t) \left(\sum_l g(t - lL)h(t - lL) \right).$$

Значит полное восстановление сигнала достигается, если $h(n)$ и $g(n)$ удовлетворяют условию:

$$\sum_l g(t - lL)h(t - lL) = 1.\tag{1.59}$$

Рассмотрим один из подходов к получению окон удовлетворяющих (1.59). Пусть есть окно $h_{PR}(n)$ удовлетворяющее условию полного восстановления, т.е.

$$\sum_l h_{PR}(t - lL) = 1.\tag{1.60}$$

Тогда для анализа может быть выбрана произвольная оконная функция $h(n)$, а синтезирующее окно должно вычисляться следующим образом:

$$g(n) = \frac{h_{PR}(n)}{h(n)}.\tag{1.61}$$

Известно, что свойству (1.60) отвечает треугольное окно, окна семейства Блэкмена-Харриса, Хэмминга и Ханна. Рассмотрим пример, когда в качестве окна $h_{PR}(n)$ выбирается окно Ханна. Для того, чтобы избежать нулевых отсчетов в начале и конце сгенерируем окно длины $N + 2$, а затем отбросим первый и последний отсчеты:

$$h_{PR}(n) = 0,5 - 0,5 \cos\left(\frac{2\pi n}{N + 2}\right), \quad n = 1, 2, \dots, N.$$

В качестве окна анализа $h(n)$ возьмем окно Хэмминга:

$$h(n) = 0,54 - 0,5 \cos\left(\frac{2\pi n}{N}\right), \quad n = 0, 1, \dots, N - 1. \quad (1.62)$$

Окно синтеза $g(n)$ рассчитывается по выражению (1.61). На рис. 1.18 показаны полученные окна анализа и синтеза, для значения параметров $L = 256$, $N = 2L$,

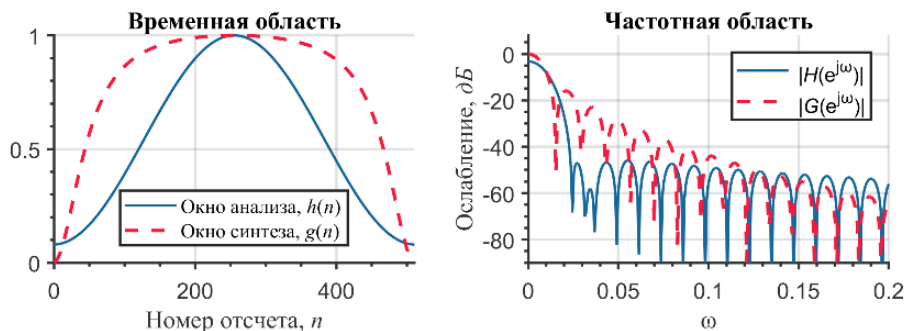


Рис. 1.18 – Окна анализа и синтеза во временной и частотной областях

В следующей таблице приводятся расчетные формулы для нескольких оконных функций, часто применяемых в цифровой обработке сигналов.

Таблица 1.2 – Оконные функции

Название	Расчетная формула
Окно Блэкмана	$w(n) = 0,42 - 0,5 \cos \frac{2\pi n}{N} + 0,08 \cos \frac{4\pi n}{N}$.
Окно Ханна	$w(n) = 0,5 - 0,5 \cos \frac{2\pi n}{N}$.
Окно Блэкмана-Харриса	$w(n) = 0,35875 - 0,48829 \cos \frac{2\pi n}{N} +$

$$+0,14128 \cos \frac{4\pi n}{N} - 0,01168 \cos \frac{6\pi n}{N}.$$

1.1.15 Спектральный анализ сигнала с использованием оконной функции

Модель сигнала для ДПФ предполагает, что вне заданного интервала сигнал имеет периодическое продолжение. Это означает, что если мы найдем ДПФ сигнала $X(k) = \text{DFT}\{x(n)\}$, а затем воспользуемся формулой (1.17) для его восстановления, но будем подставлять в нее значения $n > N$, то в результате получим периодическое продолжение сигнала. Этот процесс показан на рис. 1.19.

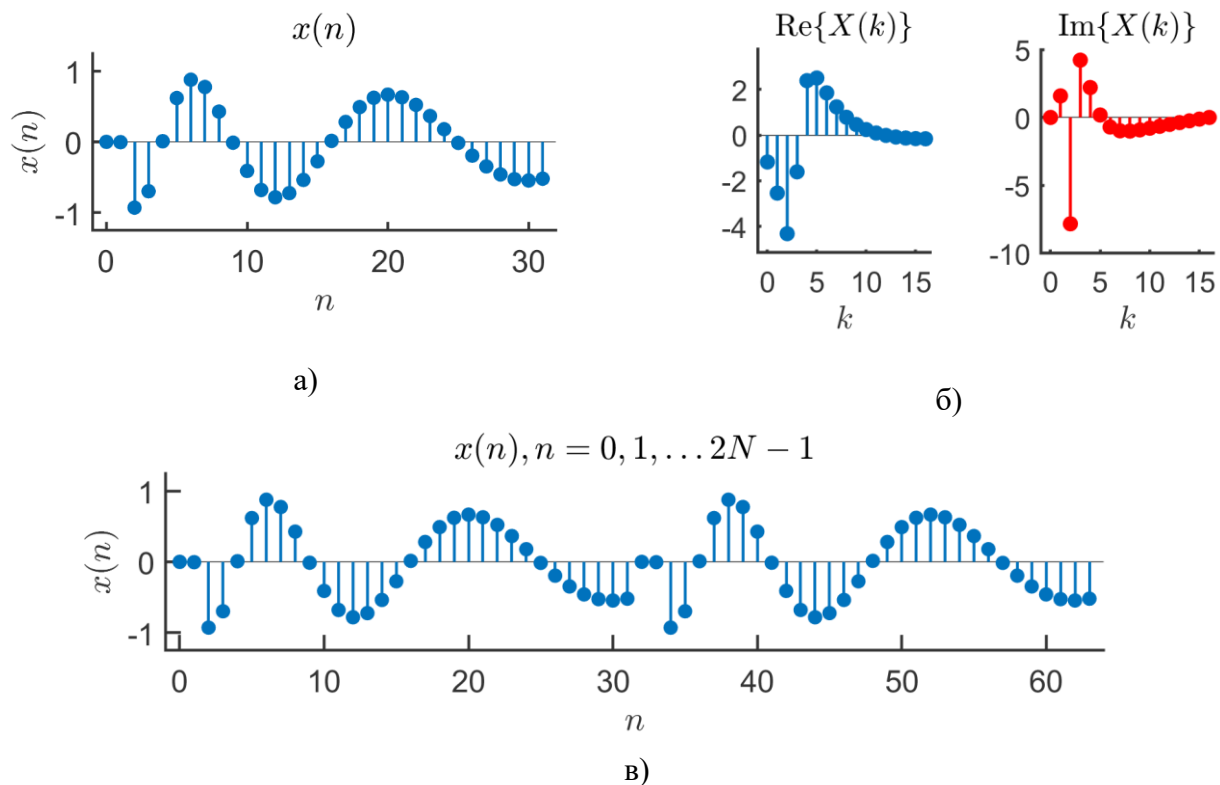


Рис. 1.19 – Периодическое продолжение дискретного сигнала:
 а) временное представление сигнала; б) ДПФ сигнала;
 в) периодическое продолжение сигнала

По этой причине «чистый» синус (согласно модели ДПФ) представляется, как показано на рис. 1.20 (без последнего нулевого отсчета в конце). Так происходит потому, что именно в этом случае ДПФ-расширение сигнала тоже является синусом (рис. 1.21).

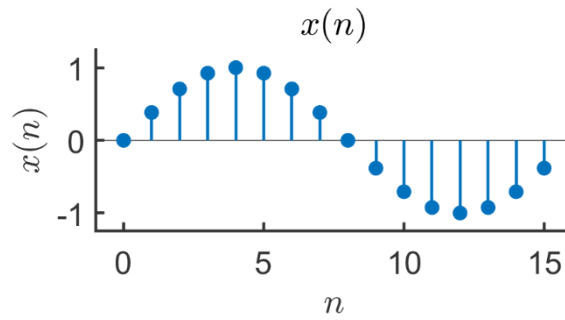


Рис. 1.20 – Дискретный синус

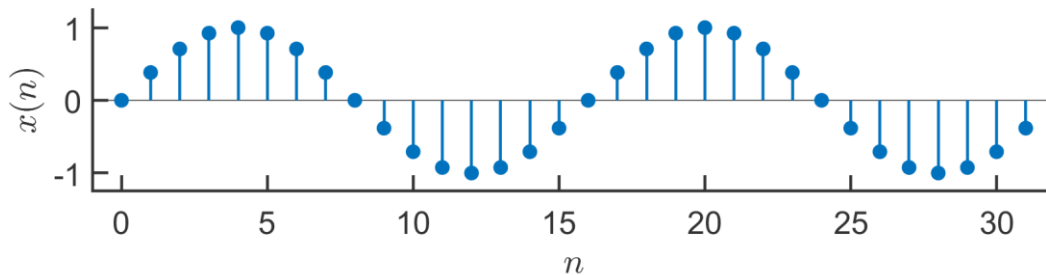


Рис. 1.21 – Дискретный синус: ДПФ расширение

Рассмотрим вопрос анализа спектра сигнала при помощи ДПФ. Остановимся на двух ситуациях. В первом случае во входном сигнале содержится в точности целое число периодов синусоидального сигнала. Тогда согласно модели сигнала ДПФ имеет место периодическое продолжение исследуемого фрагмента сигнала, формируя таким образом бесконечный непрерывный периодический сигнал, как показано на рис. 1.22.

Рис. 1.23 отражает ситуацию, когда в сигнале нет целого числа периодов. Разрывы, которые образуются на краях интервала наблюдения приводят к появлению новых компонент в спектре сигнала. Появляются паразитные боковые «лепестки» в спектре сигнала, происходит снижение разрешающей способности по частоте. Обратите внимание, что боковой лепесток на 12 дБ ниже основного лепестка и что боковые лепестки имеют спад только 6 дБ/октаву. Такая ситуация неприемлема для большинства задач анализа спектра.

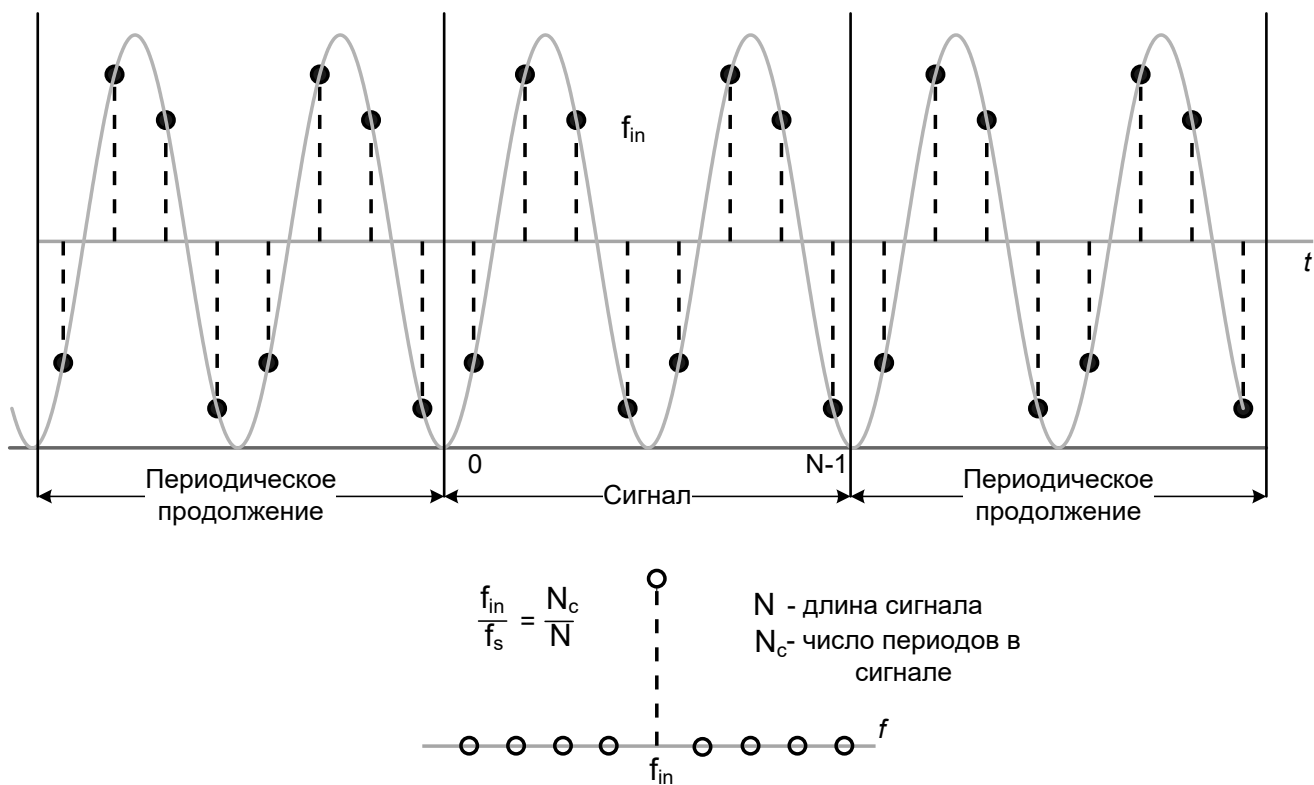


Рис. 1.22 – ДПФ синусоидального сигнала с целым числом периодов

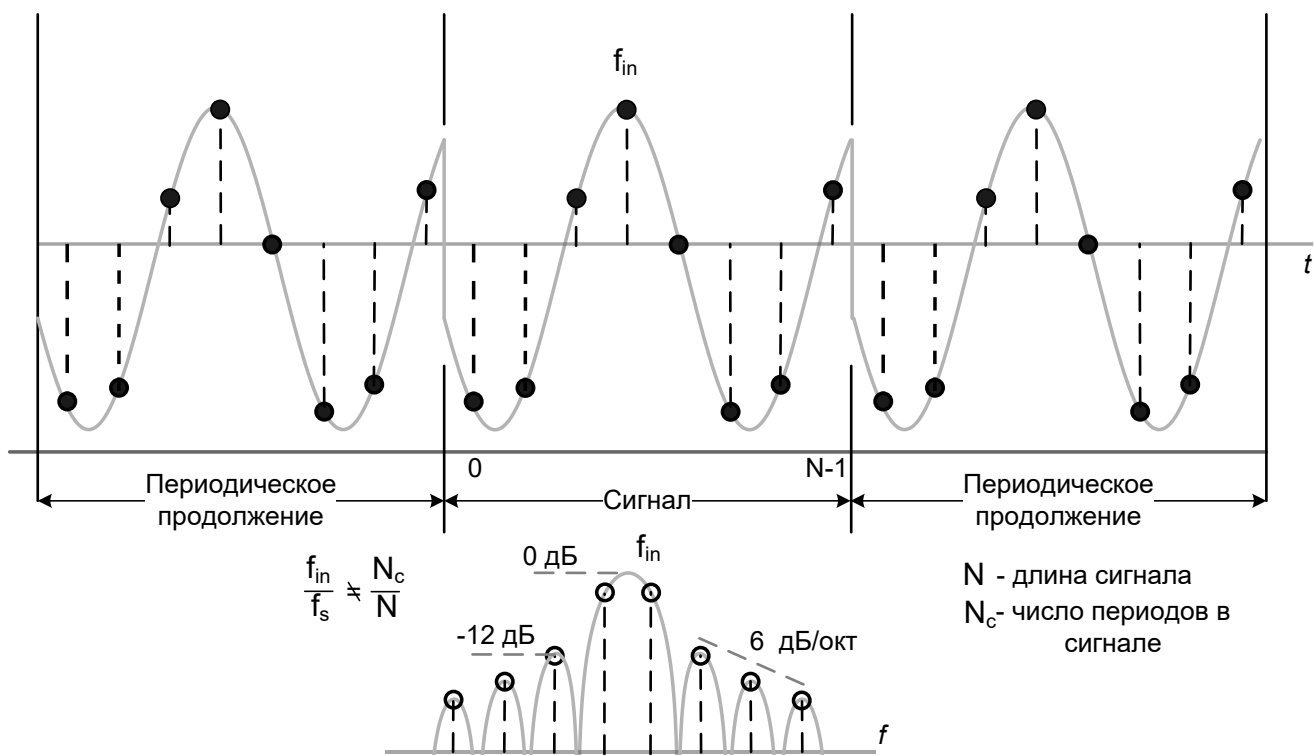


Рис. 1.23 – ДПФ синусоидального сигнала с нецелым числом периодов

Для устранения эффекта разрыва сигнала на концах интервала наблюдения необходимо входные отсчеты умножать на соответствующую оконную функцию. Применение оконной функции уменьшает амплитуду сигнала в начале и конце интервала наблюдения.

Алгебраический взгляд на ДПФ

В данном разделе мы хотим продвинуться в понимании алгебраической и вычислительной структуры ДПФ. Наше понимание проистекает из двух новых путей интерпретации ДПФ:

- 1) как произведение матрицы \mathbf{W} на вектор (сигнал);
- 2) как полином от W_N .

Мы видели, что ДПФ является *линейным* преобразованием вектора длины N в вектор такой же длины. Из курса линейной алгебры нам известно, что линейное преобразование представимо матрицей. Поэтому вместо того, чтобы говорить о функции, которая преобразует N отсчетов сигнала $x(0), x(1), \dots, x(N-1)$ в такое же количество частотных отсчетов $X(0), \dots, X(N-1)$, мы можем говорить о произведении $N \times N$ матрицы \mathbf{W} на N -мерный вектор $[x(0), x(1), \dots, x(N-1)]$, которое дает N -мерный вектор $[X(0), X(1), \dots, X(N-1)]$:

$$\mathbf{X} = \mathbf{W}\mathbf{x}. \quad (1.63)$$

Простейший пример 2-х точечного ДПФ:

$$\begin{bmatrix} X(0) \\ X(1) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \end{bmatrix}.$$

В общем случае матрица ДПФ имеет вид:

$$\mathbf{W} = \begin{bmatrix} W_N^0 & W_N^0 & W_N^0 & \dots & W_N^0 \\ W_N^0 & W_N^1 & W_N^2 & \dots & W_N^{N-1} \\ W_N^0 & W_N^2 & W_N^4 & \dots & W_N^{2(N-1)} \\ W_N^0 & W_N^3 & W_N^6 & \dots & W_N^{3(N-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ W_N^0 & W_N^{N-1} & W_N^{2(N-1)} & \dots & W_N^{(N-1)(N-1)} \end{bmatrix},$$

где $W_N = e^{-j\frac{2\pi}{N}}$ – корень N -й степени из единицы. Поэтому степени W_N можно сокращать по модулю N . Например,

$$\mathbf{W}_4 = \begin{bmatrix} W_4^0 & W_4^0 & W_4^0 & W_4^0 \\ W_4^0 & W_4^1 & W_4^2 & W_4^3 \\ W_4^0 & W_4^2 & W_4^4 & W_4^6 \\ W_4^0 & W_4^3 & W_4^6 & W_4^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & W_4^1 & W_4^2 & W_4^3 \\ 1 & W_4^2 & W_4^0 & W_4^2 \\ 1 & W_4^3 & W_4^2 & W_4^1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{bmatrix}$$

Матрица ДПФ является симметричной (её элементы симметричны относительно главной диагонали). Она также обладает рядом других симметрий. Например, любые две строки матрицы ортогональны, а квадрат длины любой строки (сумма квадратов элементов строки) равно в точности N . Таким образом, матрица ДПФ является *ортогональной*. Но не *ортонормированной*.

Матричное представление позволяет прийти к наиболее простой интерпретации обратного ДПФ. Матрица ОДПФ должна быть обратной к матрице ДПФ:

$$\mathbf{x} = \mathbf{W}^{-1}\mathbf{X},$$

причем как следствие ортогональности обратная матрица ДПФ является транспонированной версией матрицы ДПФ:

$$\mathbf{W}^{-1} = \frac{1}{N} \mathbf{W}^\dagger,$$

где † – обозначение транспонированной сопряженно матрицы. Элементами матрицы \mathbf{W}^\dagger являются

$$(\mathbf{W}^\dagger)_{N}^{nk} = e^{j\frac{2\pi nk}{N}} = W_N^{-nk}.$$

где \mathbf{U}^\dagger обозначает не просто транспонированную, но сопряженную матрицу \mathbf{U} (эквивалентный термин эрмитово сопряженная матрица), т.е. $u_{ij}^\dagger = u_{ji}^*$.

Наконец рассмотрим ещё один способ записи базовой формулы ДПФ, который позволяет по-новому взглянуть на это преобразование. Для фиксированных N и k отбросим эти индексы, так что $W \equiv W_N^k$. Тогда ДПФ принимает форму многочлена от W с коэффициентами $x(n)$:

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

Данное представление ценно по двум причинам. Первое – устанавливается связь с многочленами, что позволяет использовать эффективные алгоритмы вычисления полиномов для ДПФ (такие как, например, схема Горнера). Заметим, что БПФ, которое впервые было введено в ЦОС, можно также рассматривать как эффективный способ умножения многочленов. Кроме того, использование схемы Горнера позволяет построить эффективный рекурсивный алгоритм вычисления ДПФ известный, как алгоритм Гёрцеля. Второе, более современный подход рассматривает ДПФ как способ аппроксимации полиномом реального спектра. Отсюда ясно, что когда реальный спектр сигнала имеет острые пики, то полиномиальная аппроксимация является неэффективной, в этом случае для описания спектра лучше использовать рациональные функции.

1.1.16 DTMF-сигналы

Важным примером использования частотного представления сигнала может служить анализ двухтональный многочастотных сигналов (англ. *Dual-Tone Multi-Frequency*), которые ещё называют DTMF-сигналами или тональным набором.

DTMF-сигналы используются для набора телефонного номера. Для передачи одного символа телефонный аппарат генерирует два тона – один низкой частоты, а другой высокой частоты. Эти два сигнала складываются, в результате чего получается DTMF-сигнал. Стандартный набор частот, используемый для получения DTMF-сигналов приведен в табл. 1.3.

Таблица 1.3 – Соответствие частот синусоид и символа при формировании DTMF-сигнала

1	2	3	697 Гц
4	5	6	770 Гц
7	8	9	852 Гц
*	0	#	941 Гц
1209 Гц	1336 Гц	1477 Гц	

Согласно приведенной таблице для того, чтобы сформировать сигнал, соответствующий символу «7» необходимо сложить синусоиды с частотами 941 Гц и 1209 Гц. Обычно длительность сигнала при кодировании одно символа равна 0,5 секунды. Если посчитать ДПФ от DTMF-сигнала, то в спектре должны быть

два четко выраженных пика, которые отвечают за частотные составляющие, кодирующие передаваемый символ. Таким образом, ДПФ может быть использован, как инструмент для декодирования DTMF-сигналов.

1.1.17 Выделение частотных компонент полезного сигнала

На практике, при анализе сигналов при помощи ДПФ часто ставится задача автоматического определения значений спектральных компонент $X(k)$, превосходящих некоторый заданный порог ϵ , и соответствующих частотных индексов k , которые часто называют *частотными бинами*. Полезно смотреть на данную задачу, как на выделение полезного сигнала на фоне шума. Компоненты, превышающие порог являются полезными, несущими информацию, а компоненты, находящиеся «под порогом» относятся к шуму.

Согласно первому критерию, спектральная компонента $X(k)$ относится к полезному сигналу, если выполняется условие:

$$\frac{|X(k)|}{\max|X(k)|} > \epsilon. \quad (1.64)$$

Согласно второму критерию, считается, что нужно сравнивать амплитуду спектральной компоненты $X(k)$ не с максимальной амплитудой в спектре, а с общей мощностью сигнала:

$$\frac{|X(k)|^2}{P_x} > \epsilon, \quad (1.65)$$

где $P_x = \sum_k |X(k)|^2$.

Выбор конкретных значений порога ϵ выполняется на основании априорной информации о сигнале и шуме, с которыми приходится иметь дело при решении конкретной задачи.

1.2 Порядок выполнения работы

Возьмите вариант задания у преподавателя.

1.3 Дополнительные задания

- 1) Покажите каким образом получено выражение (1.14)?
- 2) Выведите выражение для ДПФ сигнала длины N следующего вида:

$$x(n) = \cos\left(2\pi\left(\frac{L}{N}\right)n + \phi\right).$$

3) Вычислите N -точечные ДПФ следующих сигналов:

а) $x_1(n) = \delta(n)$,

б) $x_2(n) = \delta(n - n_0)$, где $0 < n_0 < N$,

в) $x_3(n) = \alpha^n$, $0 \leq n < N$, $0 < \alpha < 1$,

г) $x_4(n) = u(n) - u(n - n_0)$, где $0 < n_0 < N$.

4) Найдите 10-точечное обратное ДПФ для

$$X(k) = \begin{cases} 3, & k = 0, \\ 1, & 1 \leq k < 9. \end{cases}$$

5) N -точечное ДПФ сигнала $X(k)$ можно интерпретировать, как N -периодическую функцию от k . Ответьте на вопрос: какие из следующих функций являются действительными (валидными) ДПФ-образами. Для каждого валидного ДПФ-образа найдите размер N и определите является ли соответствующий временной сигнал $x(n)$ действительными или нет.

а) $X(k) = j - \pi$

б) $X(k) = \sin(k/10)$

в) $X(k) = \sin(\pi k/10)$

г) $X(k) = \left(\frac{1+j}{\sqrt{2}}\right)^k$

д) $X(k) = \langle k + \pi \rangle_{10}$.

6) Для дискретной последовательности длины N :

$$x(n) = u(n) - u(n - m + 1), \quad n = 0, 1, \dots, N - 1$$

найти ДПФ $X(k)$. Построить график $|X(k)|$ для случая $m = 3$, $N = 8$.

7) Вычислите 4-точечное ДПФ сигнала

$$x(n) = 3\delta(n) - 2\delta(n - 2) + 2\delta(n - 3).$$

Постройте график амплитудного спектра сигнала $|X(k)|$.

8) Найдите формулу для ДПФ сигнала длины N следующего вида:

$$x(n) = \sin\left(\left(\frac{2\pi}{N}\right)8n + \varphi_0\right).$$

9) Определите ДПФ перечисленных конечных последовательностей, считая что все они имеют длину N , где N – четное число

а) $x(n) = \delta(n - n_0), \quad 0 \leq n_0 \leq N - 1.$

б) $x(n) = \begin{cases} 1, & n - \text{четно}, \\ 0, & n - \text{нечетно} \end{cases}$

в) $x(n) = \begin{cases} a^n, & 0 \leq n \leq N - 1, \\ 0, & \text{иначе.} \end{cases} \quad |a| < 1.$

10) Вычислите 4-х точечное ДПФ $X(k)$ для конечной последовательности

$$x(n) = \cos\left(\frac{\pi n}{2}\right), \quad 0 \leq n \leq 3;$$

11) Постройте спектрограмму речевого сигнала длительностью 3-5 секунды используя 1) окно Ханна и 2) окно Блэкмана. Какие вы опишите различие полученных спектрограмм?