

## ВВЕДЕНИЕ

Благодаря прогрессу в области микроэлектроники, многие области техники, которые прежде были аналоговыми, становятся цифровыми. В качестве примеров можно привести цифровое телевидение, цифровую фотографию, разработку цифровых видеокамер, цифровых магнитофонов, цифровых автоответчиков и т.д. В тоже время появляются новые области техники, такие как мобильная телефония, Интернет-телефония, видеоконференции через Интернет и т. д., которые используют цифровые методы обработки сигналов. Относительно недавно *мультимедиа* компьютерные технологии стали находить применение в самых разных областях человеческой деятельности: в образовании (например, дистанционное обучение), в научных исследованиях (например, для моделирования процессов), в медицине (например, для обучения врачей). Ярким примером использования мультимедиа являются многопользовательские игры через Интернет.

Согласно википедии *“мультимедиа – это комплекс программных и аппаратных средств, позволяющих пользователю, работать в интерактивном режиме с разнородными данными (текст, данные, видео и др.), организованными в виде единой информационной среды”*. В более узком смысле мультимедиа – это программное и аппаратное обеспечение, предназначенное для создания, хранения и передачи информации, представленной в различных цифровых форматах. Хотя термин мультимедиа данные в самом общем смысле включает многие типы данных, мы будем рассматривать только такие мультимедиа сигналы как речь, изображения, аудио и видео сигналы.

Первая проблема на пути использования цифровой техники для передачи и хранения таких мультимедиа сигналов как речь, изображения, аудио и видео сигналы состоит в том, что эти сигналы исходно являются аналоговыми. Таким образом, первое, что необходимо сделать для того, чтобы поместить такие сигналы в цифровое устройство – преобразовать их в цифровую форму, чтобы получить поток мультимедиа данных.

К сожалению, число битов, требуемое для хранения речи, изображений, аудио и видео информации, преобразованных в цифровую форму, очень велико. Это обстоятельство существенно снижает эффективность функционирования соответствующих цифровых систем. Рассмотрим несколько примеров.

Одна фотография, полученная с помощью 4 Мегапиксельной (Mpixels) цифровой фотокамеры требует 12 Мбайт памяти для хранения. Другими словами это означает, что только 21 фотография, полученная с помощью этой камеры, может быть сохранена на компактфлэш памяти (CF) размером 256 Мбайт.

Видеопоследовательность, содержащая 30 кадров (1 сек стандартного фильма) с размером кадра  $480 \times 720$  точек занимает  $480 \times 720 \times 3 \times 30 \approx 31$  Мбайт памяти. Это означает, что только 21 сек такого фильма может быть записана на компакт диск (CD) емкостью 650 Мбайт.

Передача цветного изображения размером  $352 \times 288$  точек (каждая точка занимает 3 байта и изображение занимает, соответственно, 2.4 Мбита) по стандартному телефонному каналу, используя модем со скоростью 33600 бит/сек требует около 1 минуты (72 сек).

Передача 1 минуты речевого сигнала дискретизованного с частотой 8 кГц при квантовании каждого отсчета в 16 битовое целое число при использовании того же модема потребует около 4 минут.

Таким образом, вторая задача, которую необходимо решить – это сжатие мультимедийных данных для того, чтобы передавать их быстрее и хранить эффективнее.

Отметим, что в рассмотренных примерах, для передачи оцифрованных мультимедиа сигналов использовались стандартные телефонные сети. Эти сети были разработаны для передачи аналоговых сигналов и для передачи цифровых данных нам необходим модем,

который преобразует цифровые данные в аналоговые сигналы для передачи по этим сетям. Для ускорения передачи по таким каналам необходимо улучшать (ускорять) модемы. Однако, сейчас не только эти сети используются для передачи цифровых данных. Появляются новые широкополосные каналы связи. Например, так называемые цифровые сети интегрального обслуживания (ISDN – integrated service digital networks) включают один или несколько цифровых телефонных каналов со скоростью передачи 64 Кбит/сек. В этих каналах униполярные цифровые данные (например, поток 0 и 1 из компьютера) преобразуются в биполярные цифровые сигналы (0 заменяется на  $-1$ ). Передача по каналу ведется биполярными импульсами. У получателя биполярные импульсы опять преобразуются в униполярный поток. Относительно недавно появились так называемые модемы для цифровых абонентских линий (DSL – digital subscriber line modem). Они применяются в тех же кабельных линиях (скрученные пары), которые используют стандартные телефонные сети. Разница в том, что они используют не полосу 4 кГц, соответствующую стандартному телефонному каналу, а всю полосу соответствующей кабельной линии. Они обеспечивают скорость передачи более 1 Мбит/сек. Увеличиваются и емкости запоминающих устройств, таких как CD и DVD. Может показаться, что вторая проблема будет полностью решена в будущем. Это не совсем так. Фактически, не имеет значения, как увеличиваются полосы каналов и емкости дисков. Проблема сжатия мультимедийных данных с тем, чтобы хранить их эффективнее и передавать быстрее, все равно сохраняется.

Системы сжатия мультимедийных данных можно разделить на два больших класса: системы сжатия без потерь и системы сжатия с потерями. Системы сжатия без потерь сохраняют исходный файл, т.е. после сжатия и восстановления восстановленный файл совпадает (как говорят, bit by bit) с исходным файлом. Сюда относятся такие методы как: кодирование кодом Хаффмена, арифметическое кодирование, кодирование, основанное на алгоритмах Зива-Лемпела-Велча и Зива-Лемпела и т. д.

Системы сжатия с потерями устраняют избыточность из мультимедийных данных ценой внесения некоторых искажений в исходный файл. Теория сжатия с потерями исследует обмен (компромисс) между степенью сжатия и внесенными искажениями. К этому классу относятся системы сжатия, основанные на преобразованиях (дискретное преобразование Фурье, дискретное косинусное преобразование, вэйвлетная фильтрация, дифференциальное кодирование).

В настоящем курсе мы рассмотрим теоретические аспекты сжатия с потерями. Подробно остановимся на технике квантования. Затем рассмотрим методы сжатия, используемые на практике. Завершим курс обзором стандартов сжатия речи, изображений, аудио и видео сигналов.

## Лекция 1,2

### АНАЛОГОВЫЕ И ДИСКРЕТНЫЕ СИГНАЛЫ

Введем некоторые определения.

Функция  $f(x)$  называется *непрерывной* в точке  $x = a$ , если  $\lim_{x \rightarrow a} f(x) = f(a)$ . Мы говорим, что функция  $f(x)$  непрерывна, если она непрерывна во всех точках ее области определения, т.е. при всех допустимых значениях аргумента  $x$ .

Будем называть множество *дискретным*, если оно содержит конечное или счетное число элементов (элементам счетного множества можно поставить в соответствие числа натурального ряда или другими словами перенумеровать их).

*Аналоговые сигналы* – это непрерывные функции непрерывного аргумента такого как время, пространство или другие непрерывные физические переменные. Мы будем в основном рассматривать непрерывные функции, заданные на некотором непрерывном отрезке времени. Например, функция вида  $x(t)$ , где  $t$  – время, принимающая вещественные значения, описывает некоторый аналоговый сигнал.

*Дискретные сигналы* могут быть дискретными по множеству значений функции или (и) по множеству значений аргумента. Если мы имеем дело с отсчетами аналогового сигнала, взятыми в заданные моменты времени, то мы называем это множество вещественных чисел *дискретным по времени* или *дискретизованным* сигналом. Если каждый отсчет также принимает значения из некоторого дискретного множества значений, то мы говорим, что имеем дело с *цифровым* сигналом.

В силу непрерывности пространства и времени такие сигналы как речевой сигнал, изображение, видеосигнал, сигналы от всевозможных датчиков (температуры, давления, скорости) и.т.д. являются непрерывными. Для того, чтобы поместить непрерывный сигнал в дискретное (цифровое) устройство такое, как, например, ЭВМ необходимо преобразовать его в дискретный сигнал.

Так называемое *аналого-цифровое* преобразование включает в себя две операции:

- Операция дискретизации или взятия отсчетов непрерывного сигнала (квантование по времени)
- Операция квантования непрерывного сигнала по уровню и связанная с ней операция кодирования.

### ДИСКРЕТИЗАЦИЯ. ТЕОРЕМА КОТЕЛЬНИКОВА

Пусть задана непрерывная функция  $x(t)$ . *Дискретизацией* называется взятие отсчетов этой функции через интервал времени  $T$ . В результате вместо функции  $x(t)$  мы получаем последовательность отсчетов  $x(nT)$ ,  $n = 0, 1, 2, \dots$ . Нас интересует, будут ли внесены искажения в исходный сигнал  $x(t)$  при выполнении дискретизации и, если будут, то каким образом они зависят от шага дискретизации  $T$ ? Ответы на эти вопросы дает теорема отсчетов, называемая по имени ее автора, теорема Котельникова.

Прежде чем перейти к формулировке теоремы напомним некоторые понятия, которые потребуются нам при ее доказательстве.

Преобразование Фурье, или спектр непрерывного сигнала  $x(t)$  определяется следующим образом

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt,$$

где  $\omega = 2\pi f$  – круговая частота. Эта функция в общем случае комплексная, она может быть представлена в виде

$$X(f) = A(f) e^{j\phi(f)},$$

где  $A(f) = |X(f)|$  называется спектром сигнала или амплитудно-частотной характеристикой сигнала, а  $\phi(f)$  – это фазовый спектр или фазо-частотная характеристика сигнала.

Существует обратное преобразование Фурье, которое позволяет восстановить сигнал  $x(t)$  по его спектру

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j\omega t} df.$$

Преобразование Фурье тесно связано с преобразованием Лапласа. Для временных функций, которые существуют только при  $t \geq 0$ , преобразование Лапласа определяется как функция комплексной переменной  $s$  следующим образом

$$L(s) = \int_0^{\infty} x(t) e^{-st} dt.$$

При  $s = j\omega$  преобразование Лапласа совпадает с преобразованием Фурье. В терминах, применяемых при обработке сигналов, преобразование Фурье вводится для того, чтобы представить сигнал в частотной области, а  $f$  – аргумент преобразования Фурье – понимается как частота.

Если  $x(t)$  – периодическая функция времени с периодом  $p$ , то она может быть разложена в ряд Фурье

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(k\omega_p t) + b_k \sin(k\omega_p t)),$$

где

$$\omega_p = 2\pi/p, \quad a_0 = \frac{2}{p} \int_{-p/2}^{p/2} x(t) dt, \quad a_k = \frac{2}{p} \int_{-p/2}^{p/2} x(t) \cos(k\omega_p t) dt, \quad b_k = \frac{2}{p} \int_{-p/2}^{p/2} x(t) \sin(k\omega_p t) dt, \quad k = 1, 2, \dots$$

Чтобы избежать использования двух слагаемых ( $\cos(k\omega_p t)$  и  $\sin(k\omega_p t)$ ) для каждого индекса  $k$  после простых вычислений мы получаем

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} A_k \cos(k\omega_p t + \varphi_k),$$

где  $\varphi_k$  – это начальная фаза. Далее, используя формулу Эйлера  $e^{j\omega t} = \cos(\omega t) + j \sin(\omega t)$  мы получаем комплексную форму ряда Фурье

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \frac{A_k}{2} (\exp(jk\omega_p t + j\varphi_k) + \exp(-jk\omega_p t - j\varphi_k)) = \sum_{k=-\infty}^{\infty} c_k e^{jk\omega_p t},$$

$$\text{где } c_k = \frac{1}{2} A_k e^{j\varphi_k} = \frac{1}{2} (a_k - jb_k) = \frac{1}{p} \int_{-p/2}^{p/2} x(t) \exp(-jk\omega_p t) dt, \quad c_0 = 1/2 a_0, \quad b_0 = 0, \quad a_{-k} = a_k,$$

$$b_{-k} = -b_k.$$

Говорят, что разложение периодической функции с периодом  $p$  в ряд Фурье дает ее представление в виде суммы гармонических колебаний с частотами  $k\omega_p = 2k\pi/p$ ,  $k = 1, 2, \dots$  тогда как интеграл Фурье дает разложение непериодической функции в виде суммы бесконечно большого числа колебаний с непрерывно изменяющейся частотой или дает разложение функции в непрерывный спектр. Отметим, что непериодическая функция конечной длины  $T$  также может быть разложена в ряд Фурье. Для того, чтобы выполнить такое разложение необходимо построить периодическое продолжение этой функции с периодом  $T$ .

*Теорема Котельникова.*

Пусть задана функция  $x(t)$  и ее спектр имеет вид  $X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt$ ,  $X(f) = 0$  при  $|f| > w_0$ . Тогда эта функция полностью определяется своими мгновенными значениями в моменты, отстоящие друг от друга на  $1/(2w_0)$  секунд.

Доказательство.

Продолжим спектр  $X(f)$  периодически и получим функцию  $\hat{X}(f)$ . Тогда на интервале  $(-w_0, w_0)$ , т.е. на периоде, функция  $\hat{X}(f)$  может быть разложена в ряд Фурье с коэффициентами

$$a_k = \frac{1}{2w_0} \int_{-w_0}^{w_0} \hat{X}(f) e^{-j2\pi fk/(2w_0)} df, \quad (1.1)$$

$$\text{т.е. } \hat{X}(f) = \sum_{k=-\infty}^{\infty} a_k e^{j2\pi fk/(2w_0)}.$$

Т.к.  $X(f)$  есть преобразование Фурье для  $x(t)$ , то  $x(t)$  можно представить как обратное преобразование Фурье для  $X(f)$

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df = \int_{-w_0}^{w_0} \hat{X}(f) e^{j2\pi ft} df.$$

Рассмотрим значения функции  $x(t)$  в точках отсчета, расположенных на  $1/(2w_0)$  секунд друг от друга, т.е. пусть  $t = k/(2w_0)$ , тогда

$$x\left(\frac{k}{2w_0}\right) = \int_{-w_0}^{w_0} \hat{X}(f) e^{j2\pi fk/(2w_0)} df. \quad (1.2)$$

$$\text{Сравнивая (1.1) и (1.2), получаем, что } a_k = \frac{1}{2w_0} x\left(\frac{-k}{2w_0}\right).$$

Таким образом, если функция известна в точках отсчета  $\dots -2/(2w_0), -1/(2w_0), 0, 1/(2w_0), 2/(2w_0), \dots$ , то коэффициенты  $a_k$  определены. Эти коэффициенты определяют спектр  $X(f)$ , а  $X(f)$  определяет  $x(t)$  для всех значений  $t$ . Это значит, что существует единственная функция, не содержащая частот выше  $w_0$  и

проходящая через заданные мгновенные значения в точках отсчета, отстоящих друг от друга на  $1/(2w_0)$  секунд.

Для восстановления функции по этим мгновенным значениям заметим, что

$$X(f) = \sum_k a_k e^{j2\pi f k / (2w_0)} \text{ при } |f| \leq w_0 \text{ и } X(f) = 0 \text{ при } |f| > w_0.$$

Взяв обратное преобразование, получаем

$$\begin{aligned} x(t) &= \int_{-w_0}^{w_0} X(f) e^{j2\pi f t} df = \int_{-w_0}^{w_0} \sum_{k=-\infty}^{\infty} a_k e^{jkf 2\pi / 2w_0} e^{j2\pi f t} df = \sum_{k=-\infty}^{\infty} a_k \int_{-w_0}^{w_0} \exp(j2\pi f (k/2w_0 + t)) df = \\ &= 2 \sum_{k=-\infty}^{\infty} a_k \int_0^{w_0} \cos(2\pi f (k/2w_0 + t)) df = 2 \sum_{k=-\infty}^{\infty} a_k \sin(2\pi f (k/2w_0 + t)) / (2\pi (k/2w_0 + t)) \Big|_0^{w_0}, \\ x(t) &= 2w_0 \sum_k a_k \frac{\sin \pi(2w_0 t + k)}{\pi(2w_0 t + k)} = \sum_k x\left(\frac{-k}{2w_0}\right) \frac{\sin \pi(2w_0 t + k)}{\pi(2w_0 t + k)} = \sum_k x\left(\frac{k}{2w_0}\right) \frac{\sin \pi(2w_0 t - k)}{\pi(2w_0 t - k)}. \end{aligned}$$

Другими словами функцию  $x(t)$  можно представить в форме суммы элементарных функций вида  $\frac{\sin x}{x}$ , расположенных симметрично относительно точек отсчета и имеющих пиковое значение равное  $x(t)$  в соответствующих точках отсчета. Для восстановления функции  $x(t)$  нужно лишь создать серию импульсов вида  $\frac{\sin x}{x}$ , пропорциональных отсчетам, и сложить их.

Таким образом, из теоремы отсчетов следует, что если спектр непрерывного сигнала ограничен, то выполняя дискретизацию с частотой в два раза превышающей максимальную частоту в спектре непрерывного сигнала, мы не вносим никаких искажений и этот сигнал может быть однозначно восстановлен по полученной последовательности отсчетов.

Ошибка, вызванная дискретизацией с частотой  $f_s < 2w_0$  носит название aliasing или перекрытие. Это название выражает тот факт, что высокие частоты в спектре сигнала “проявляются” в восстановленном сигнале как более низкие. Чаще всего анализ этого явления выполняется в частотной области. Мы сначала проиллюстрируем перекрытие во временной области. На Рис. 1.1 показана косинусоида с частотой 3 Гц. Предположим, что мы берем отсчеты этого сигнала с частотой 4 Гц. В соответствии с теоремой Котельникова минимальная частота взятия отсчетов, необходимая для однозначного восстановления сигнала из последовательности отсчетов равна 6 Гц. Таким образом, условия теоремы в нашем примере нарушены. Пунктиром на Рис. 1.1 показан восстановленный сигнал. Очевидно, что это косинусоида с частотой 1 Гц. Вместо исходного сигнала с частотой 3 Гц по отсчетам мы восстановили более низкочастотный сигнал (с частотой 1 Гц).

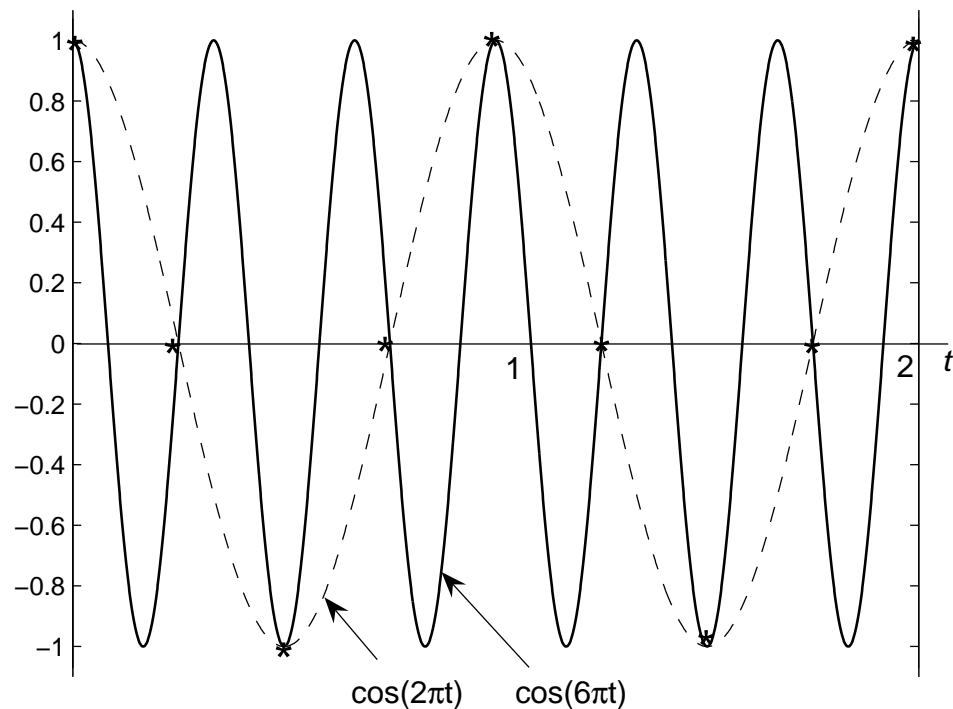


Рис. 1.1. Перекрывание во временной области

Для анализа преобразования Фурье дискретизованного сигнала, т.е. последовательности отсчетов, необходимо связать с этой последовательностью аналоговый сигнал. Далее мы будем использовать так называемую дельта-функцию Дирака  $\delta(x)$ . Напомним некоторые свойства этой функции.

$$\int_{\Delta} \delta(x) \varphi(x) dx = \begin{cases} \varphi(0), & \text{if } 0 \in \Delta \\ 0, & \text{иначе} \end{cases},$$

где  $\varphi(x)$  - это произвольная функция, а  $\Delta$  - это произвольный интервал на вещественной оси. Очевидно, что

$$\int_{\Delta} \delta(x - x_0) \varphi(x) dx = \begin{cases} \varphi(x_0), & \text{if } x_0 \in \Delta \\ 0, & \text{иначе} \end{cases}$$

и

$$\int_{\Delta} \delta(x - x_0) dx = \begin{cases} 1, & \text{if } x_0 \in \Delta \\ 0, & \text{иначе} \end{cases}.$$

Дельта-функция может также рассматриваться как предел последовательности прямоугольных импульсов ширины  $\Delta$  и высоты  $1/\Delta$  при стремлении  $\Delta$  к нулю.

Стандартный способ сопоставления аналогового сигнала последовательности отсчетов – это вычисления произведения исходного аналогового сигнала  $x(t)$  и последовательности импульсов, следующих друг за другом с интервалом  $T_s$ . В этом случае дискретизованный сигнал может быть представлен в виде

$$x_s(t) = \sum_{k=-\infty}^{\infty} x(kT_s) \delta(t - kT_s) \quad (1.3)$$

Так как функция  $\delta(t - kT_s)$  равна нулю всюду кроме точки  $t = kT_s$ , то мы можем переписать формулу для  $x_s(t)$  следующим образом

$$x_s(t) = x(t) \sum_{k=-\infty}^{\infty} \delta(t - kT_s). \quad (1.4)$$

Принимая во внимание, что функция  $\sum_{k=-\infty}^{\infty} \delta(t - kT_s)$  - периодическая, мы можем разложить эту функцию в ряд Фурье

$$\sum_{k=-\infty}^{\infty} \delta(t - kT_s) = \sum_{n=-\infty}^{\infty} c_n e^{jn2\pi/T_s}, \quad (1.5)$$

$$c_n = \frac{1}{T_s} \int_{-T_s/2}^{T_s/2} \delta(t) e^{-jn2\pi/T_s} dt = \frac{1}{T_s}.$$

Мы учли, что внутри интервала  $(-T_s/2, T_s/2)$  находится только одна дельта-функция (она соответствует  $k = 0$ ). Подставляя (1.5) в (1.3) получаем

$$x_s(t) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} x(t) e^{jn2\pi/T_s}.$$

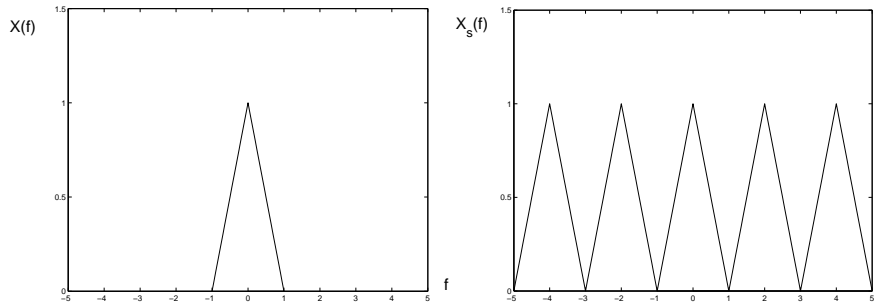
Так как умножение на  $\exp(j\alpha t)$  во временной области соответствует сдвигу преобразования Фурье, то мы получаем

$$X_s(\omega) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} X\left(\omega - \frac{2\pi n}{T_s}\right)$$

или

$$X_s(f) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} X\left(f - \frac{n}{T_s}\right).$$

Таким образом, спектр дискретизованного сигнала представляет собой последовательность копий спектра исходного сигнала  $x(t)$ , следующих друг за другом с интервалом  $1/T_s$  Гц. На Рис. 1.2 показан спектр  $X_s(f)$  для случаев  $T_s = 1/(2w_0)$  и  $T_s > 1/(2w_0)$ .





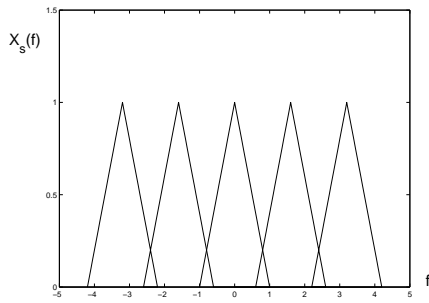


Рис.1.2 Спектр исходного и дискретизованного сигнала для случаев  $T_s = 1/(2w_0)$  и  $T_s > 1/(2w_0)$  секунд.

Как следует из Рис. 1.2 если отсчеты берутся слишком редко  $T_s > 1/2w_0$ , то копии спектра исходного сигнала накладываются друг на друга, вызывая специфическую ошибку.

Для того, чтобы восстановить сигнал  $x(t)$  по последовательности отсчетов достаточно профильтровать эту последовательность идеальным фильтром нижних частот с частотой среза равной  $w_0 = 1/(2T_s)$  Гц.

Другими словами преобразование Фурье восстановленного сигнала  $x_R(t)$  выражается следующим образом

$$X(f) = X_s(f)T_s H_L(f),$$

где

$$H_L(f) = \begin{cases} 1, & \text{if } |f| \leq w_0 \\ 0, & \text{иначе} \end{cases}.$$

Импульсный отклик идеального фильтра нижних частот представляет собой обратное преобразование Фурье от его частотной характеристики  $H_L(f)$

$$h_L(t) = \int_{-\infty}^{\infty} H_L(f) e^{j2\pi ft} df = 2 \int_{-w_0}^{w_0} \cos(2\pi ft) df = \frac{1}{T_s} \frac{\sin(\pi t/T_s)}{\pi t/T_s}.$$

Дискретизованный сигнал  $x_s(t)$  представляет собой сумму взвешенных дельта-функций. Проходя через идеальный фильтр нижних частот, каждая из этих дельта-функций генерирует сдвинутую и взвешенную копию импульсного отклика. Таким образом, выходной сигнал  $x_R(t)$  представляет собой сумму взвешенных и сдвинутых копий импульсного отклика идеального фильтра нижних частот

$$x_R(t) = \sum_{k=-\infty}^{\infty} x(kT_s) \frac{\sin(\pi(t - kT_s)/T_s)}{\pi(t - kT_s)/T_s}.$$

Однако на практике мы имеем дело с непрерывными сигналами конечной длительности и соответственно бесконечным спектром. Возникает вопрос: как следует применять теорему Котельникова в этом случае? На практике под  $w_0$  понимается только эффективная ширина спектра, в которой сосредоточена определенная часть (90-95%) энергии сигнала, т.е. мы находим  $w_0$  из следующего уравнения

$$\frac{\int_{-w_0}^{w_0} |X(f)|^2 df}{\int_{-\infty}^{\infty} |X(f)|^2 df} = \alpha,$$

где  $\alpha \approx 0.9 - 0.95$  называется *эффективной полосой* при заданном пороге  $\alpha$ .

На практике для борьбы с перекрытиями спектра (aliasing) применяются различные методы. Например, можно выполнять предфильтрацию аналогового сигнала конечной длительности в аналоговой области подходящим низкочастотным фильтром. Этот фильтр уменьшает высокочастотные составляющие в спектре сигнала и тем самым уменьшает ошибку, связанную с перекрытием спектров. Другая возможность – это выбирать период дискретизации  $T_s \ll 1/(2w_0)$  секунд. Однако, так как наша цель – представить аналоговый сигнал в компактной форме, то мы не должны выполнять его дискретизацию со слишком большой частотой. С другой стороны во многих цифровых приложениях мы имеем дело не с аналоговым сигналом, а с сигналом, дискретизованным с частотой  $f_s \gg 2w_0$  Гц (oversampled signal) и фактически выполняем прореживание сигнала (downsampling), эквивалентное уменьшению частоты дискретизации, в зависимости от требуемого качества восстановленного сигнала. В этом случае стандартный подход состоит в предфильтрации входного (oversampled) сигнала цифровым фильтром с последующим прореживанием с целью уменьшения частоты дискретизации.

## КВАНТОВАНИЕ

Вторая операция, которая необходима для преобразования аналогового сигнала в цифровой, называется *квантованием*. Мы говорим, что квантование это процедура преобразования отсчетов сигнала в целые значения, которые, очевидно, имеют конечную длину двоичного представления. Простейший пример квантования – это обычное арифметическое округление до ближайшего целого. Любое вещественное значение, скажем,  $x$  может быть округлено до ближайшего целого, скажем,  $y$ , при этом внесенная квантованием ошибка равна  $x - y$ . Введем формальное определение.

Пусть функция принимает значения из множества  $X$ , которое может быть как дискретным так и непрерывным и пусть  $X^n$  это множество значений функции  $\mathbf{x} = (x_1, \dots, x_n)$ , соответствующих  $n$  последовательным моментам времени. Тогда операция *квантования* представляет собой операцию замены вектора  $\mathbf{x} = (x_1, \dots, x_n)$  из множества  $X^n$  аппроксимирующим вектором  $\mathbf{y} = (y_1, \dots, y_n)$  из аппроксимирующего множества  $Y$ . Нас будет интересовать случай, когда может быть бесконечное множество  $X^n$  аппроксимируется с помощью конечного множества  $Y$ , содержащего  $M$  аппроксимирующих последовательностей  $\mathbf{y}_1, \dots, \mathbf{y}_M$ .

Введем в рассмотрение функцию  $d_n(\mathbf{x}, \mathbf{y})$ , каждое значение которой будем считать величиной ошибки аппроксимации  $\mathbf{x}$  с помощью  $\mathbf{y}$

$$d_n(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n d(x_i, y_i),$$

где  $d(x, y)$  – некоторая неотрицательная функция, задающая величину ошибки и называемая *критерием качества*. Критерий качества является количественной мерой искажений, возникающих при замене  $x$  на  $y$ . Наиболее широко используемый критерий качества – квадратичный, т.е.  $d(x, y) = (x - y)^2$ . Функция  $d_n(\mathbf{x}, \mathbf{y})$  называется *средней величиной ошибки* при замене  $\mathbf{x}$  на  $\mathbf{y}$ .

*Скорость квантования*, т.е. число битов необходимое для представления вектора  $\mathbf{x}$ , отнесенное к размерности этого вектора  $n$  определяется по формуле

$$R = \frac{\log_2 M}{n} \text{ бит/отсчет.}$$

Пример. Пусть аппроксимирующее множество  $Y$  содержит 4 вектора длины 2:

$$\mathbf{y}_1 = (0.2, 0.3);$$

$$\mathbf{y}_2 = (0.3, 0.1);$$

$$\mathbf{y}_3 = (0.1, 0.5);$$

$$\mathbf{y}_4 = (0.0, 0.4).$$

Предположим, что мы хотим проквантовать вектор  $\mathbf{x} = (0.18, 0.25)$ . Ближайший к  $\mathbf{x}$  (в смысле квадратичного критерия качества) вектор – это  $\mathbf{y}_1 = (0.2, 0.3)$ . Средняя ошибка квантования равна  $1.45 \times 10^{-3}$ . Скорость квантования составляет  $R = (\log_2 4) / 2 = 1$  бит/отсчет.

## СКАЛЯРНОЕ КВАНТОВАНИЕ

Скалярное квантование – это процесс замены значения  $x$  непрерывной случайной величины  $X$  ближайшим значением  $y$  из некоторого дискретного множества значений  $Y = \{y_1, \dots, y_M\}$ . Другими словами в этом случае размерность квантуемой величины  $n = 1$ , а скорость скалярного квантования  $R = \log_2 M$  бит/отсчет.

Пусть  $[A, B]$  – область значений входных величин квантователя или интервал числовой оси, где могут находиться значения  $x$  квантуемой величины  $X$ . Мы можем определить скалярный квантователь как множество подинтервалов интервала  $[A, B]$  или множество квантов  $\Delta_i = [t_{i-1}, t_i]$ ,  $i = 1, 2, \dots, M$ , где  $t_{i-1}$  – левая и  $t_i$  – правая границы кванта. Ширина  $i$  – го кванта равна  $t_i - t_{i-1}$ . При этом  $x_0 = A$ ,  $x_M = B$ . Число квантов равно  $M$ , кванты не пересекаются и покрывают весь интервал  $[A, B]$ , т.е.

$$\Delta_i \cap \Delta_j = \emptyset \text{ и } \bigcup_{i=1}^M \Delta_i = [A, B].$$

Аппроксимирующее значение  $y_i \in \Delta_i$ . Общее определение скалярного квантования сводится к округлению до ближайшего целого если  $\Delta_i = [i - 1/2, i + 1/2]$  и  $y_i = i$  для всех целых  $i$ .

## РАВНОМЕРНОЕ СКАЛЯРНОЕ КВАНТОВАНИЕ

Скалярное квантование называется *равномерным* если все кванты одинаковы по ширине. Аппроксимирующее значение для  $i$  – го кванта  $y_i$  – это число из подинтервала  $\Delta_i$ , на которое заменяется квантуемая величина, если она находится в кванте  $\Delta_i$ . При равномерном квантовании  $y_i$  выбирается посередине кванта.

Процесс квантования состоит в сравнении входной величины  $x$  с границами квантов  $t_i$  и определении номера  $i$  кванта  $\Delta_i$ , который содержит  $x$ . Номер  $i$ , записанный в двоичной форме, хранится или передается по каналу связи. Для простоты предполагаем, что число квантов  $M$  равно степени двойки и все двоичные представления номеров квантов представляют собой двоичные кодовые слова одной и той же длины  $\log_2 M$ . Такое равномерное квантование называют *квантованием с фиксированной скоростью*.

Пример равномерного скалярного квантователя с 4 квантами приведен на Рис. Интервал  $[A, B] = [0, 2]$ , границы квантов равны:  $t_0 = 0, t_1 = 0.5, t_2 = 1, t_3 = 1.5, t_4 = 2$ . Аппроксимирующие значения выбираются посередине квантов и равны, соответственно  $y_1 = 0.25, y_2 = 0.75, y_3 = 1.25, y_4 = 1.75$ .

Другой вариант равномерного квантования предполагает, что номера квантов  $i$  будут кодироваться неравномерно, т.е. длины кодовых слов могут быть различными.

Пусть  $y_i, i = 1, 2, \dots, M$  - это аппроксимирующие значения и пусть  $f(x)$  - плотность вероятности квантуемой величины  $X$ . Тогда мы можем вычислить вероятности аппроксимирующих значений (или вероятности попадания в  $i$ -й квант) по формуле

$$p_i = \int_{t_{i-1}}^{t_i} f(x) dx.$$

Используя неравномерное кодирование выходов квантователя, мы получим квантование со скоростью близкой к энтропии дискретного источника, образованного выходами квантователя

$$R = - \sum_j p_j \log_2 p_j.$$

Такой квантователь называется *квантователем с переменной скоростью*. В этом случае число квантов необязательно равно степени двойки. Наиболее часто используемый на практике равномерный квантователь такого типа представляет собой просто операцию арифметического округления, т.е. номер кванта вычисляется по формуле

$$i = \text{round}(x / \text{step}),$$

где *step* называют шагом квантования.

Очевидно, что в этом случае число квантов всегда будет нечетным,  $i$ -й квант определяется как  $\Delta_i = [i \times \text{step} - \text{step} / 2, i \times \text{step} + \text{step} / 2]$ ,  $y_i = i \times \text{step}$ .

При восстановлении по кодовому слову (двоичной записи номера  $i$ ) восстанавливается  $i$  и выбирается аппроксимирующее значение  $y_i$ .

Для того, чтобы оценить ошибку квантования необходимо выбрать критерий качества. Как уже отмечалось выше, чаще всего на практике используют квадратичный критерий, т.е.  $d(x, y) = (x - y)^2$ . Рассмотрим квантователь, приведенный на Рис. Предположим для примера, что нам необходимо проквантовать величину  $x = 1.285$ . Тогда получаем, что  $i = 3$ ,  $y_i = 1.25$ , двоичное представление номера кванта 10 (кванты нумеруются с нуля). Таким образом, мы потратили на представление  $x$  2 бита и ошибка квантования при этом составила 0.0012.

Если плотность вероятности  $f(x)$  случайной величины  $X$  - неизвестна, то мы оцениваем среднюю ошибку квантования как среднее арифметическое ошибок, полученных на экспериментальной последовательности значений случайной величины. Если  $f(x)$  - известна, то средняя ошибка квантования вычисляется по формуле

$$\varepsilon = \sum_i \int_{\Delta_i} d(x, y_i) f(x) dx.$$

Пусть  $X$  – непрерывная случайная величина с плотностью вероятностей  $f(x)$ . Тогда при квадратичном критерии качества средняя ошибка квантования вычисляется по формуле

$$\varepsilon = \int_{-\infty}^{\infty} (x - y(x))^2 f(x) dx = \sum_{i=1}^M \int_{t_{i-1}}^{t_i} (x - y_i)^2 f(x) dx \quad (1.3).$$

## НЕРАВНОМЕРНОЕ ОПТИМАЛЬНОЕ СКАЛЯРНОЕ КВАНТОВАНИЕ

Интуитивно ясно, что мы можем использовать дополнительную информацию о статистических характеристиках случайной величины для улучшения характеристик квантователя.

Рассмотрим подробнее выражение (1.3), определяющее среднеквадратическую ошибку квантования. При заданном числе квантов  $M$ , величина ошибки (1.3) зависит от  $2M - 1$  параметра, а именно от  $M - 1$  границ квантов  $\{t_1, \dots, t_{M-1}\}$  и от  $M$  аппроксимирующих значений  $\{y_i\}$ . Найдем указанные параметры так, чтобы ошибка была минимальной. Эта задача представляет собой стандартную задачу поиска минимума функции  $2M - 1$  переменных. Мы можем переписать формулу (1.3)

$$\int_{-\infty}^{\infty} x^2 f(x) dx - 2 \sum_{i=1}^M y_i \int_{t_{i-1}}^{t_i} x f(x) dx + \sum_{i=1}^M y_i^2 \int_{t_{i-1}}^{t_i} f(x) dx.$$

Дифференцируя (1.3) по  $y_i$  и по  $t_i$ , а затем приравнявая производные нулю, получаем

$$\begin{aligned} -2 \int_{t_{i-1}}^{t_i} x f(x) dx + 2 y_i \int_{t_{i-1}}^{t_i} f(x) dx &= 0, \\ -2 y_i t_i f(t_i) + 2 y_{i+1} t_i f(t_i) + y_i^2 f(t_i) - y_{i+1}^2 f(t_i) &= 0. \end{aligned}$$

Она имеет следующее решение при условии, что для всех  $x \in [A, B]$ ,  $f(x) \neq 0$ :

$$t_i = (y_i + y_{i+1}) / 2, \quad (1.4)$$

$$y_i = \frac{1}{p_i} \int_{t_{i-1}}^{t_i} x f(x) dx, \quad (1.5)$$

где

$$p_i = \int_{t_{i-1}}^{t_i} f(x) dx \quad (1.6)$$

представляет собой вероятность попадания случайной величины  $X$  в  $i$ -й квант, а формула (1.5) определяет координату центра тяжести фигуры, ограниченной функцией плотности вероятностей  $f(x)$  и границами  $i$ -го кванта. Квантование, задаваемое формулами (1.4)-(1.6), называется *оптимальным неравномерным квантованием*.

Используя формулы (1.3)-(1.5), получаем, что оптимальная среднеквадратическая ошибка равна

$$\varepsilon = \sum_{i=1}^M \int_{t_{i-1}}^{t_i} x^2 f(x) dx - y_i^2 \int_{t_{i-1}}^{t_i} f(x) dx.$$

Это квантование впервые было описано С.П. Ллойдом и позднее популяризировано Дж.Максом. В частности Дж. Макс выполнил полную оптимизацию неравномерного квантования для случая, когда случайная величина  $X$  является гауссовской, т.е. ее функция плотностей вероятностей равна

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-(x-m)^2/(2\sigma^2)\},$$

где  $m$  – математическое ожидание, а  $\sigma^2$  – дисперсия случайной величины  $X$ .

На Рис 1.3 показана некоторая функция плотности вероятностей и оптимальное неравномерное квантование (границы квантов и аппроксимирующие значения) при  $M = 4$ . Как следует из Рис.1.2  $t_i - y_i = y_{i+1} - t_i$  и значение  $y_2$  является координатой центра тяжести заштрихованной фигуры.

Метод решения уравнений (1.3)-(1.6) представляет собой итеративную процедуру. Мы начинаем вычисления с равномерного квантователя с числом квантов  $M$ , вычисляем значения  $y_i$  по формуле (1.5) и  $t_i$  по формуле (1.4). Мы продолжаем вычисления до тех пор пока  $t_i$ ,  $y_i$  и  $\varepsilon$  не перестают меняться.

При оптимальном неравномерном квантовании длины квантов, вообще говоря, различны. Они малы в области высоковероятных значений  $x$  и велики в области маловероятных  $x$ . Часто, однако, желательно иметь квантование, для которого все кванты имеют одинаковые длины (в случае, когда  $A = -\infty$  и  $B = \infty$ , т.е. первый и последний квант имеют бесконечные длины). Такое квантование удобно с точки зрения технической реализации. Кроме того, такое квантование может быть более эффективным в сочетании с использованием неравномерного кодирования. Пусть число квантов равно  $M$ , а все кванты, кроме быть может двух крайних, имеют одинаковую длину  $\Delta$  и аппроксимирующие значения берутся в середине кванта:  $y_i = t_i - \Delta/2$ . Тогда ошибка  $\varepsilon$  зависит только от двух параметров: от длины кванта  $\Delta$  и одной из границ, скажем  $t_1$ . Если искать указанные два параметра из условия, чтобы ошибка  $\varepsilon$  была минимальной, то получится квантование, которое называется *оптимальным равномерным квантованием*.

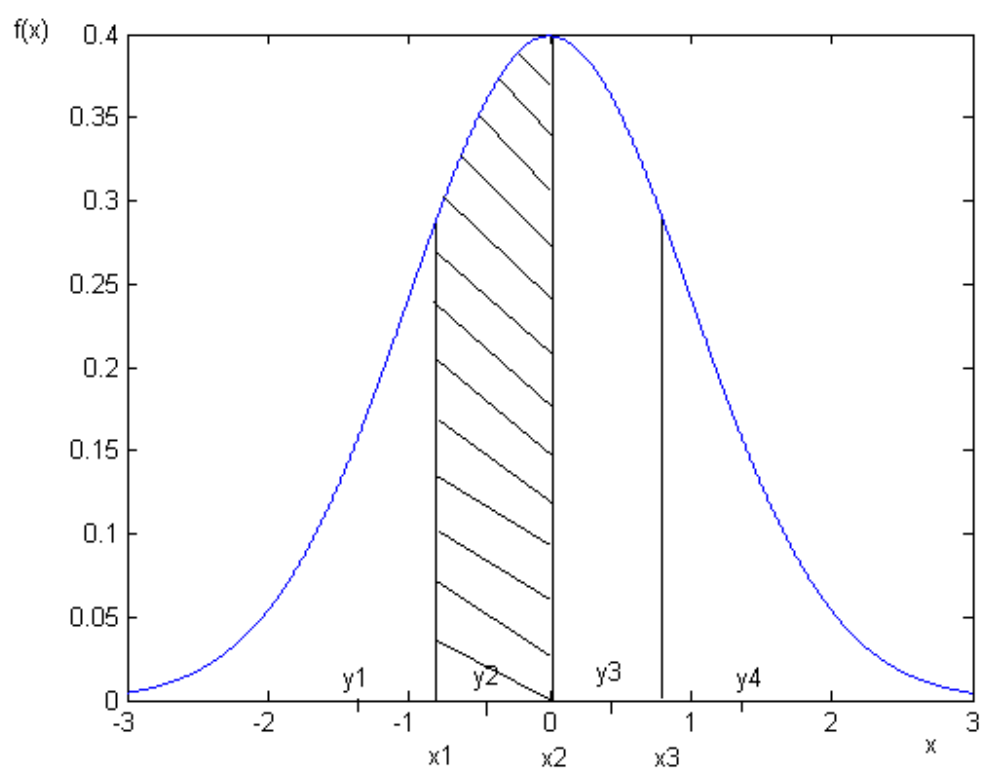


Рис.1.3