

Представление звука в ЭВМ

Аналоговые и дискретные сигналы

Аналоговый сигнал

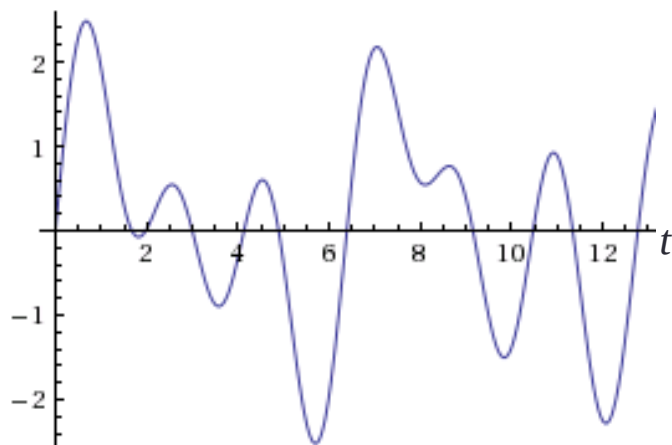
- Сигнал, изменяющийся плавно, непрерывно.
- Может принимать любые промежуточные значения.
- Невозможно отличить сигнал от шума.
- Большинство физических параметров, как их воспринимает человек — аналоговые.

Цифровой сигнал

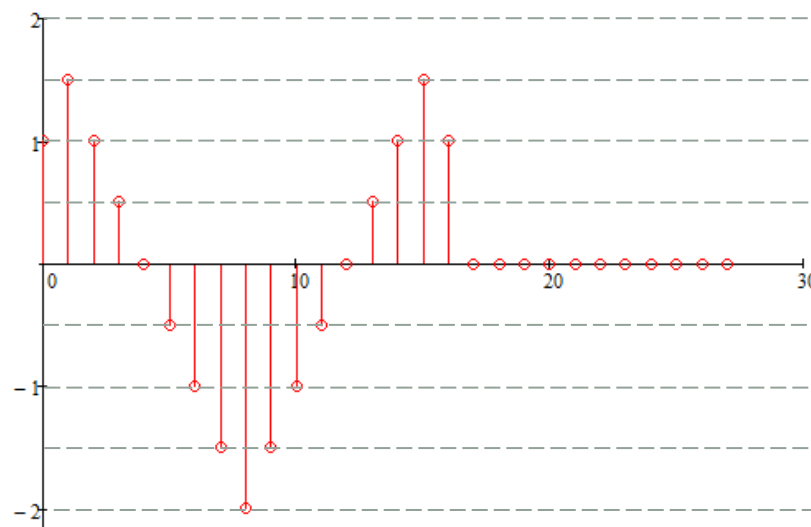
- Может находиться только в фиксированном числе разных состояний.
- Значения сигнала известны только в определённые моменты.
- Из-за дискретных состояний легче выявлять и исправлять шум.
 - Можно делать точные копии.
- Компьютеры работают только с цифровым сигналом.

Аналоговые и дискретные сигналы

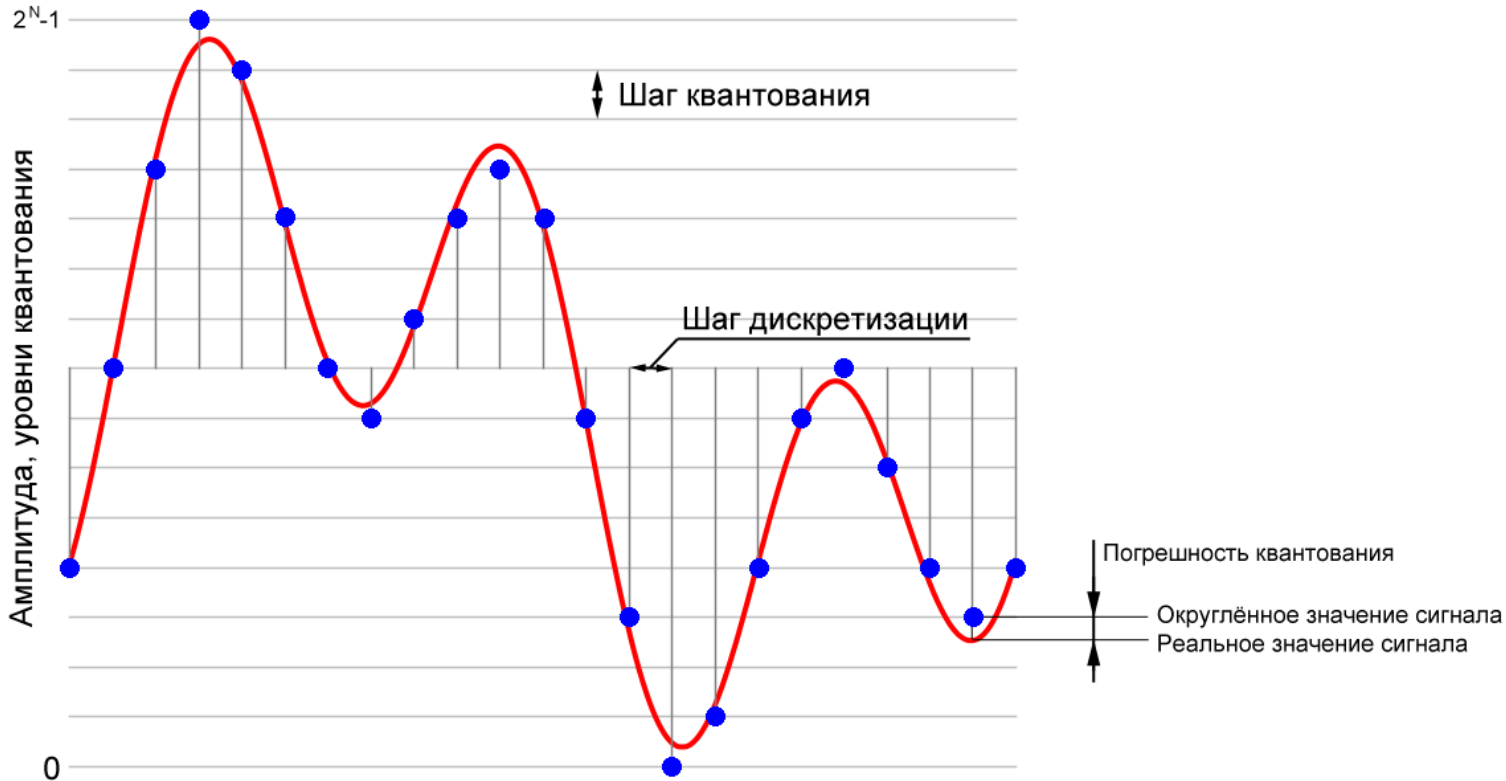
Аналоговый сигнал



Цифровой сигнал

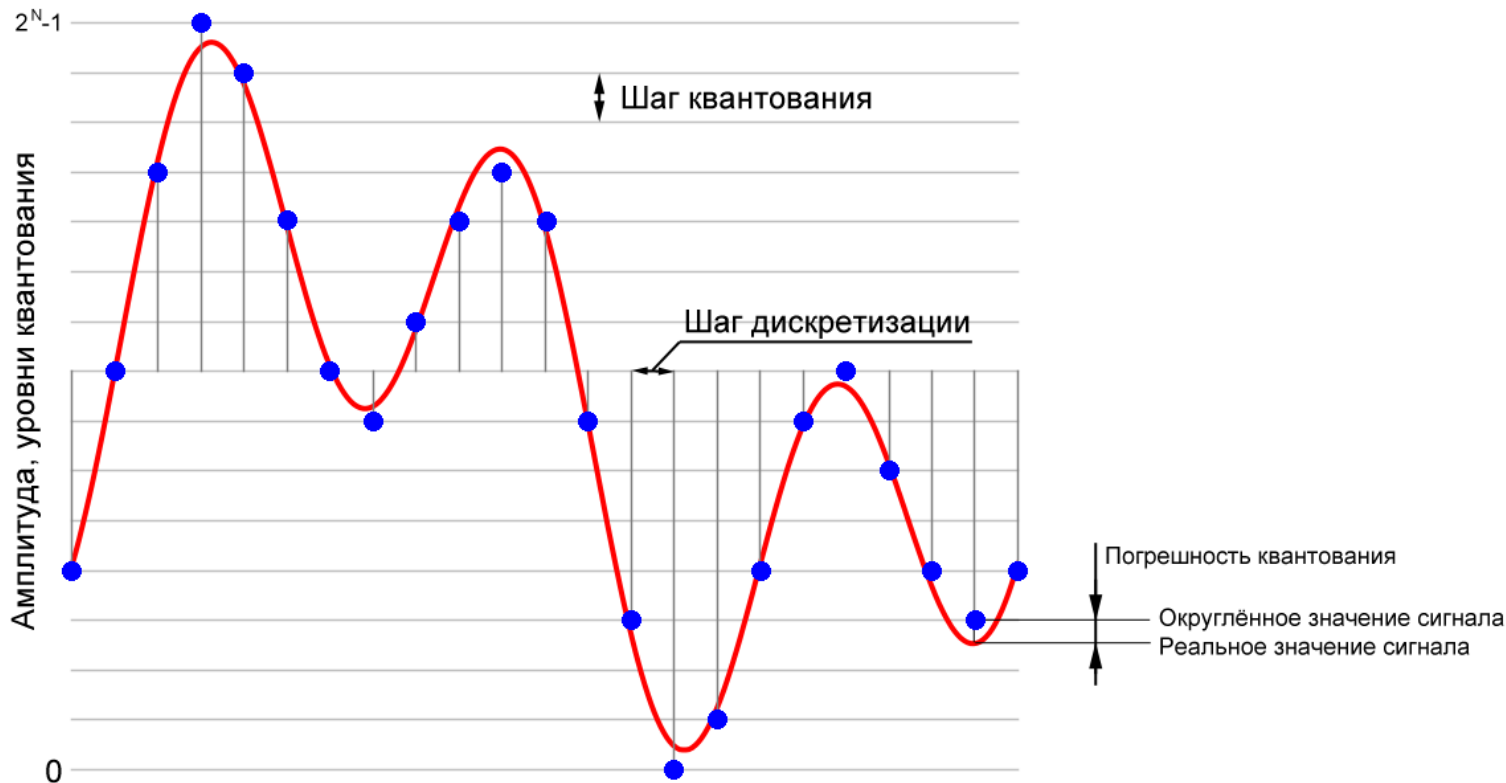


Аналогово-цифровое преобразование



- Непрерывная звуковая волна разбивается на отдельные маленькие временные участки, в начале каждого из которых определяется уровень громкости.
- Уровни громкости могут принимать только значения из фиксированного набора.

Аналогово-цифровое преобразование



Частота дискретизации — число отсчётов в секунду.

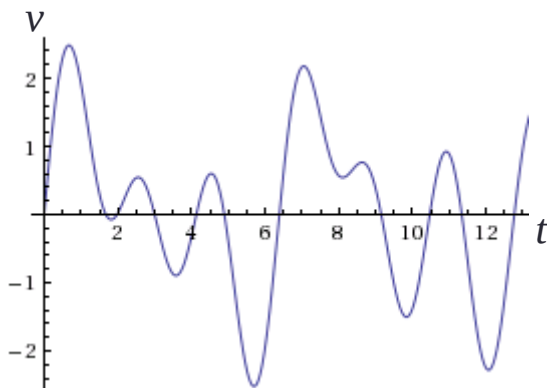
- Шаг дискретизации равен $1/F$ секунды, где F — частота дискретизации (Гц).

Глубина кодирования — число бит, выделенных на запись одного отсчёта.

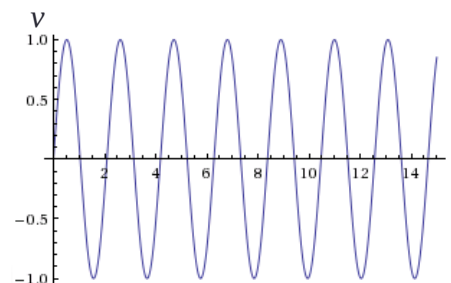
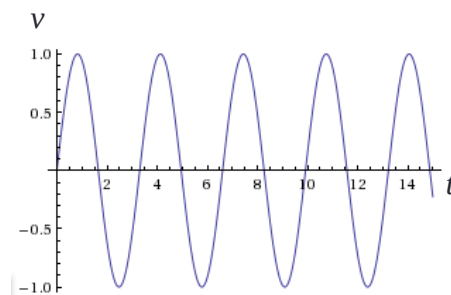
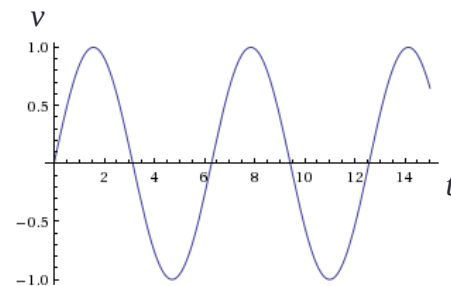
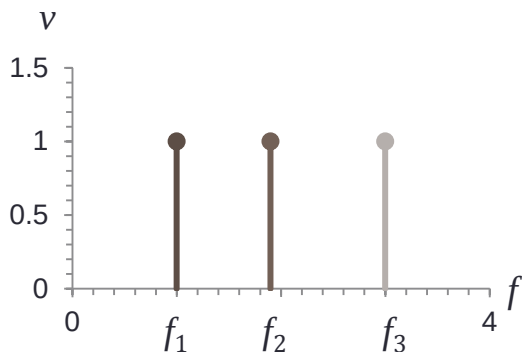
- Число уровней квантования равно 2^N , где N — глубина кодирования (бит).

О спектре сигнала

Любой аналоговый сигнал можно разложить на простые колебания (гармоники):



Спектр – диаграмма распределения простых колебаний в сигнале.

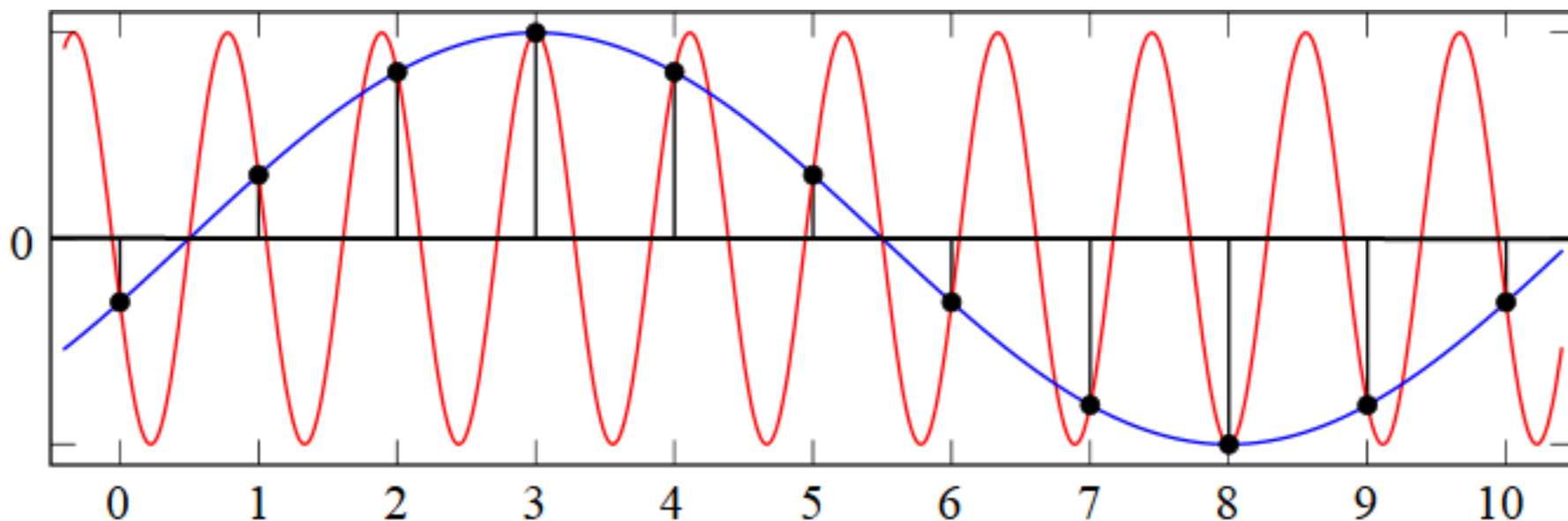


Теорема Котельникова-Найквиста-Шеннона

- Любой аналоговый сигнал может быть восстановлен с какой угодно точностью по своим дискретным отсчётам, взятым с частотой $F > 2 F_c$, где F_c — максимальная частота, которой ограничен спектр реального сигнала.
 - Это верно для бесконечно продолжающегося сигнала.
 - В реальности сигналов с ограниченным спектром почти не бывает.
 - Поэтому, частоту дискретизации берут с запасом.
- Если максимальная частота в сигнале равна или превышает половину частоты дискретизации, то способа восстановить сигнал из дискретного в аналоговый не существует.

Алиасинг

- Если частота дискретизации недостаточна, возникает эффект, известный как Aliasing – высокочастотный сигнал проявляет себя как низкочастотный.
- Перед оцифровкой желательно отфильтровать высокочастотные сигналы, чтобы избежать алиасинга.

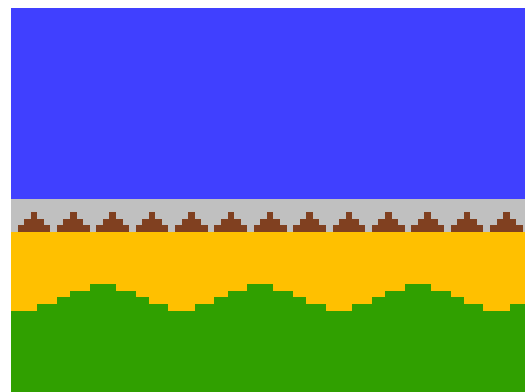


Алиасинг - не только в звуке

Фото



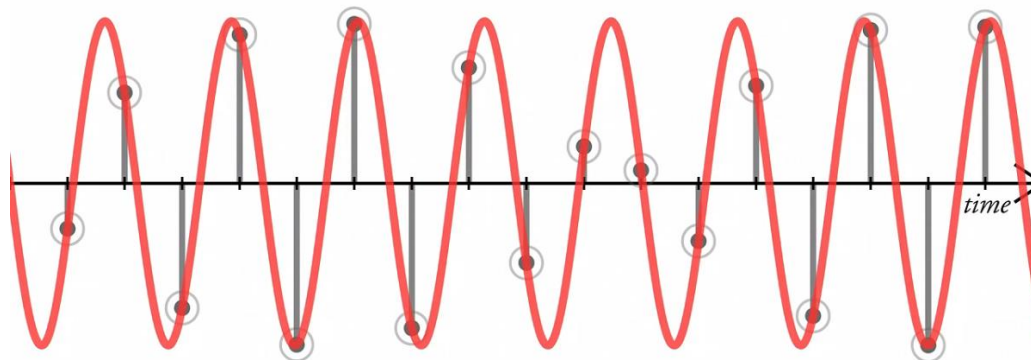
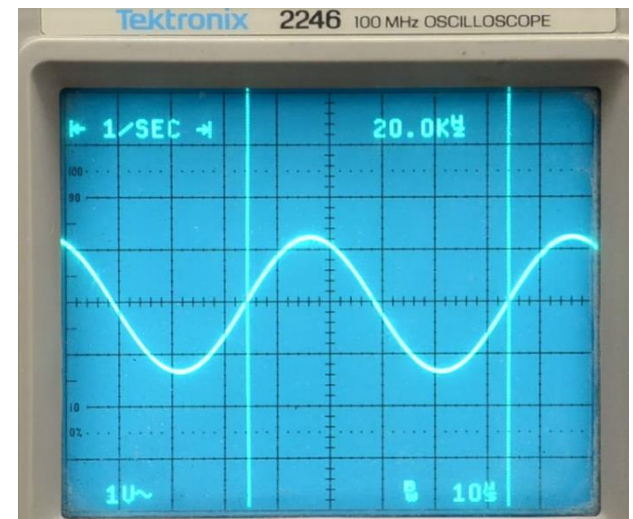
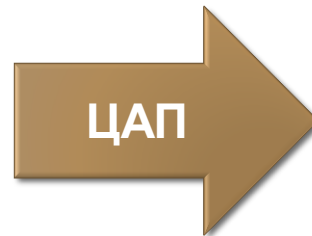
Видео



Влияние частоты дискретизации на качество звука

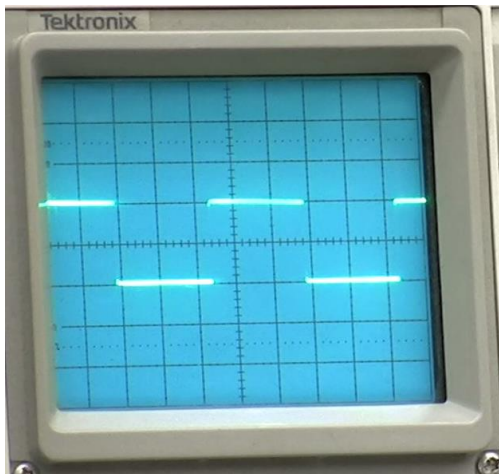
- Ограниченный по частотному диапазону аналоговый сигнал может быть **однозначно** восстановлен по цифровому сигналу.

Тон 20 kHz, оцифрованный с частотой дискретизации 44kHz

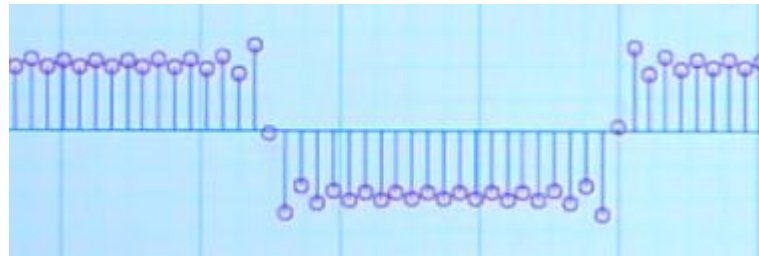


Какой сигнал не является ограниченным в частотном диапазоне

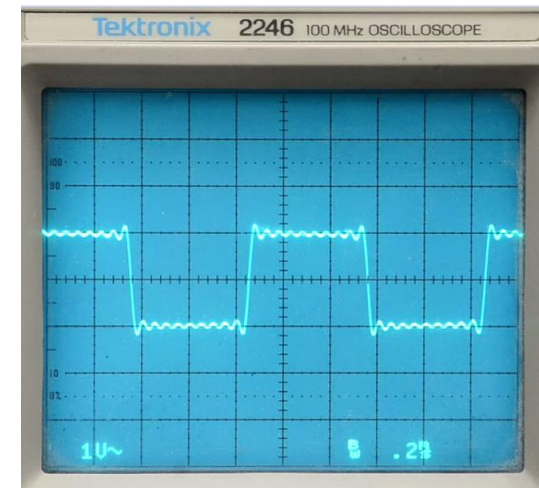
Входной сигнал
квадрат 1 kHz



Цифровой сигнал
с частотой дискретизации 44 kHz



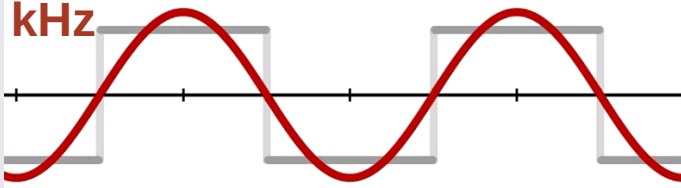
Восстановленный
аналоговый сигнал



Как получается квадрат из синусоид

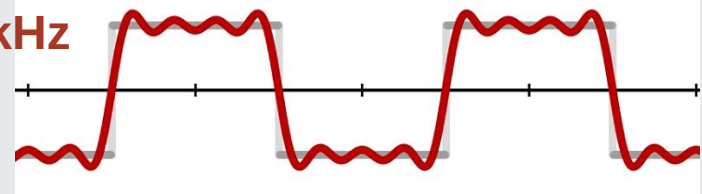
$$y(t) = \frac{4}{\pi} \sin(\omega t)$$

1 kHz



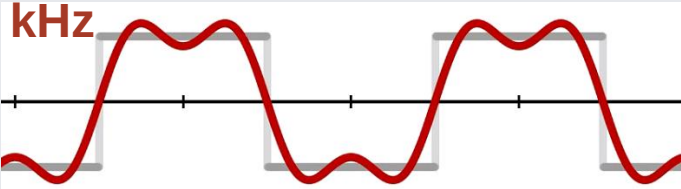
$$y(t) = \frac{4}{\pi} \sin(\omega t) + \frac{4}{3\pi} \sin(3\omega t) + \frac{4}{5\pi} \sin(5\omega t) + \frac{4}{7\pi} \sin(7\omega t)$$

7 kHz



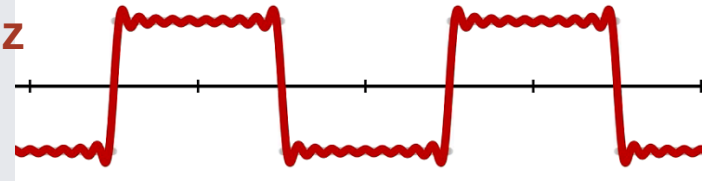
$$y(t) = \frac{4}{\pi} \sin(\omega t) + \frac{4}{3\pi} \sin(3\omega t)$$

3 kHz



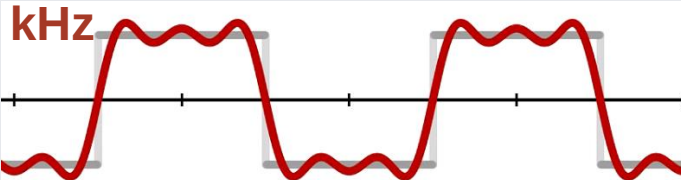
$$y(t) = \frac{4}{\pi} \sin(\omega t) + \frac{4}{3\pi} \sin(3\omega t) + \frac{4}{5\pi} \sin(5\omega t) + \frac{4}{7\pi} \sin(7\omega t) + \frac{4}{9\pi} \sin(9\omega t) + \frac{4}{11\pi} \sin(11\omega t) + \frac{4}{13\pi} \sin(13\omega t) + \frac{4}{15\pi} \sin(15\omega t) + \frac{4}{17\pi} \sin(17\omega t) + \frac{4}{19\pi} \sin(19\omega t)$$

19 kHz



$$y(t) = \frac{4}{\pi} \sin(\omega t) + \frac{4}{3\pi} \sin(3\omega t) + \frac{4}{5\pi} \sin(5\omega t)$$

5 kHz

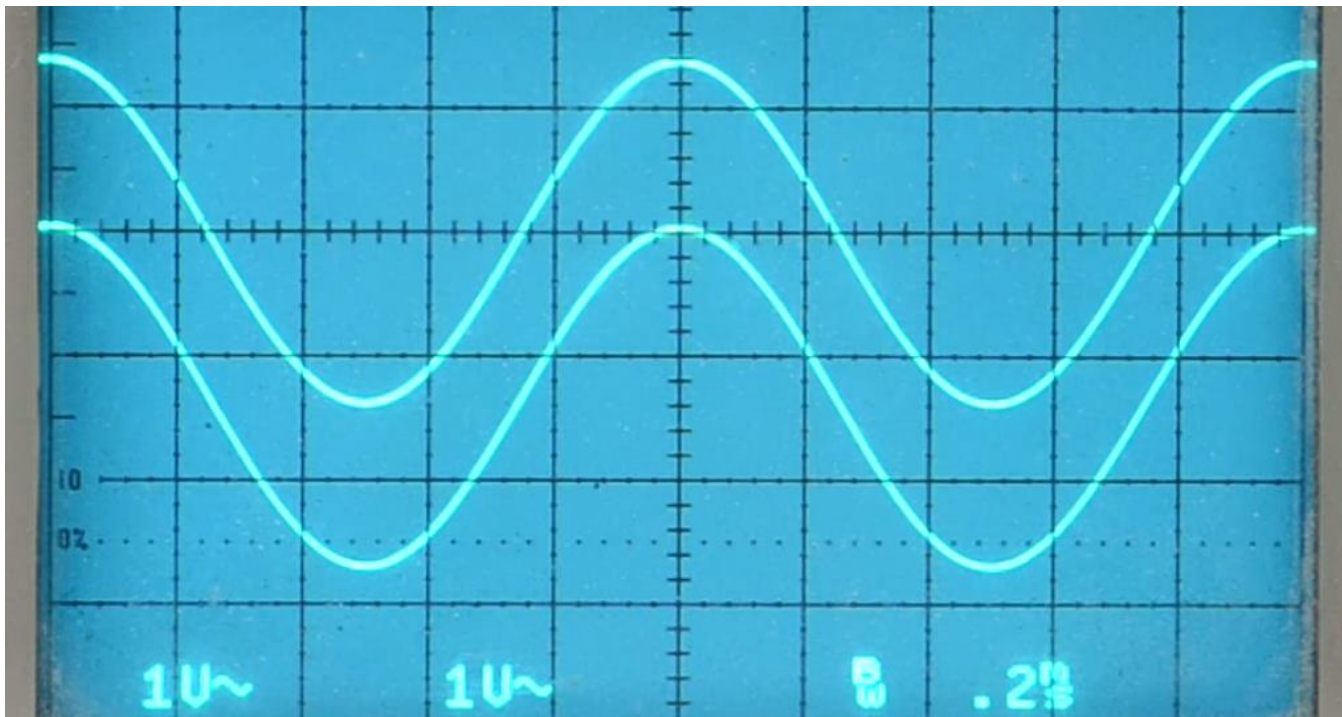


$$y(t) = \frac{4}{\pi} \sin(\omega t) + \frac{4}{3\pi} \sin(3\omega t) + \frac{4}{5\pi} \sin(5\omega t) + \frac{4}{7\pi} \sin(7\omega t) + \frac{4}{9\pi} \sin(9\omega t) + \frac{4}{11\pi} \sin(11\omega t) + \frac{4}{13\pi} \sin(13\omega t) + \frac{4}{15\pi} \sin(15\omega t) + \frac{4}{17\pi} \sin(17\omega t) + \frac{4}{19\pi} \sin(19\omega t) + \frac{4}{21\pi} \sin(21\omega t) + \frac{4}{23\pi} \sin(23\omega t) + \frac{4}{25\pi} \sin(25\omega t) + \frac{4}{27\pi} \sin(27\omega t) + \frac{4}{29\pi} \sin(29\omega t) + \frac{4}{31\pi} \sin(31\omega t) + \frac{4}{33\pi} \sin(33\omega t) + \frac{4}{35\pi} \sin(35\omega t) + \frac{4}{37\pi} \sin(37\omega t) + \frac{4}{39\pi} \sin(39\omega t) + \frac{4}{41\pi} \sin(41\omega t) + \frac{4}{43\pi} \sin(43\omega t) + \frac{4}{45\pi} \sin(45\omega t) + \frac{4}{47\pi} \sin(47\omega t) + \frac{4}{49\pi} \sin(49\omega t) + \dots$$

Влияние глубины кодирования на качество звука

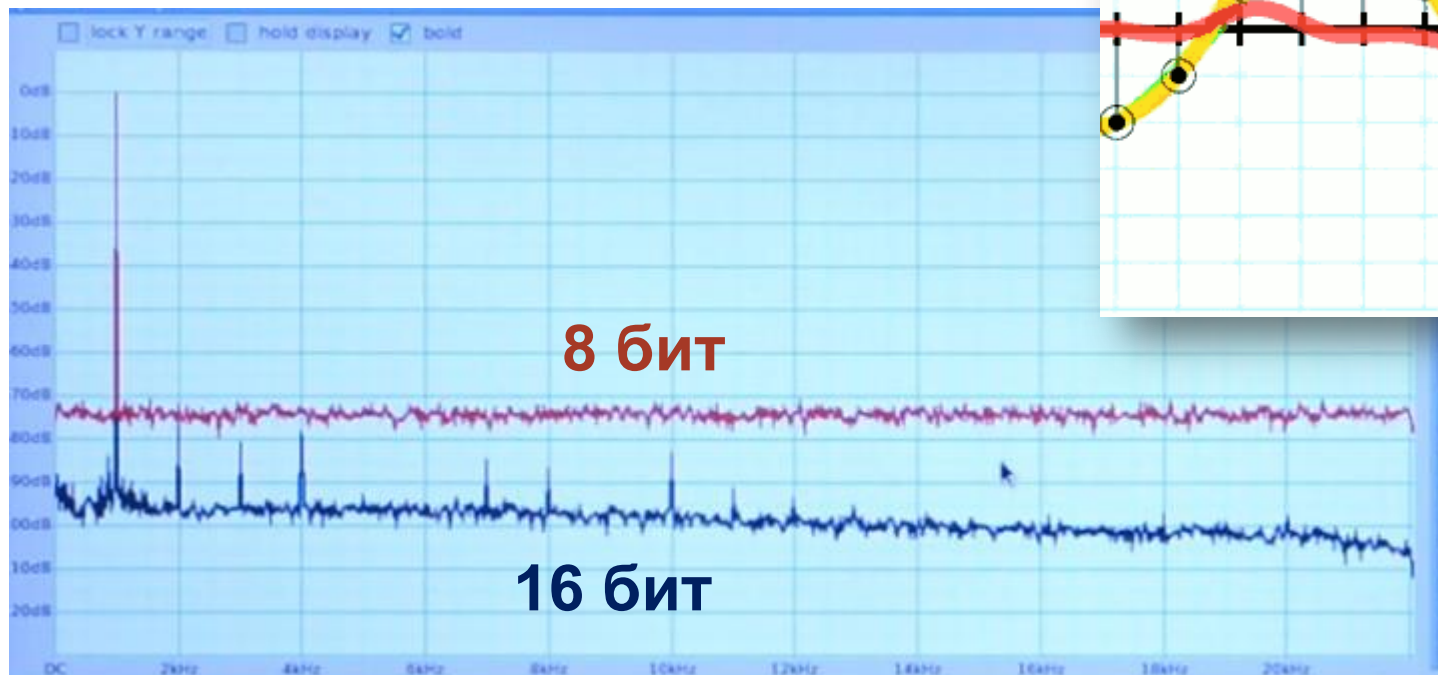
- При любой глубине кодирования выходной сигнал ЦАП будет гладкой непрерывной функцией.

Эти две синусоиды восстановлены из цифрового сигнала с глубиной кодирования 8 и 16 бит.



Влияние глубины кодирования на качество звука

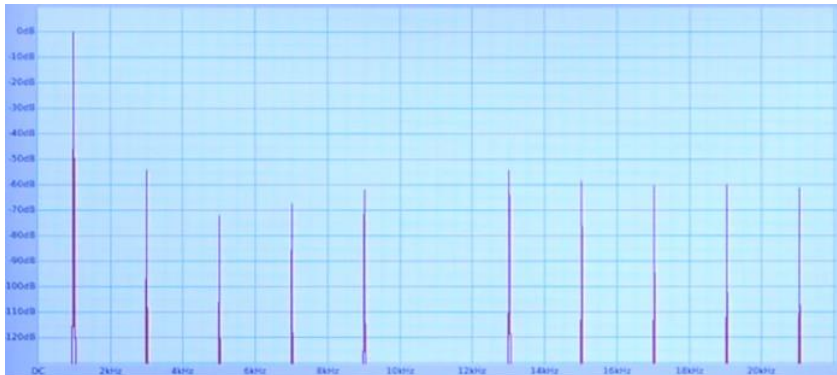
- Меньше глубина кодирования – больше шума.
 - Шум эквивалентен шуму аналоговой записи
 - Аудиокассета – 5-6 бит, в идеальном случае – 9 бит.
 - Профессиональная аппаратура – 13 бит.



Дитеринг (Dithering)

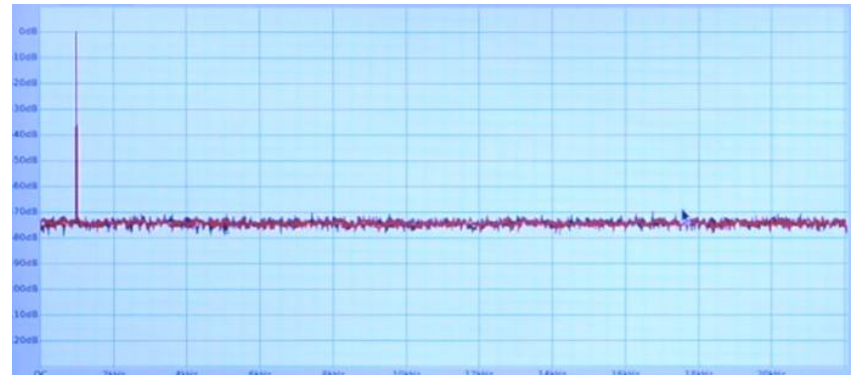
Без дитеринга

- При квантовании выбирается ближайший уровень.
- Шум квантования приводит к искажениям звука
 - При глубине 8 бит их может быть слышно



С дитерингом

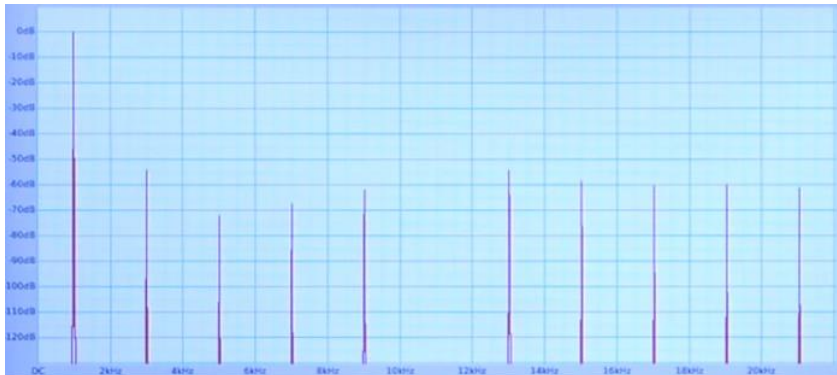
- Выбирая уровни квантования более хитро можно сделать шум равномерным.



Дитеринг (Dithering)

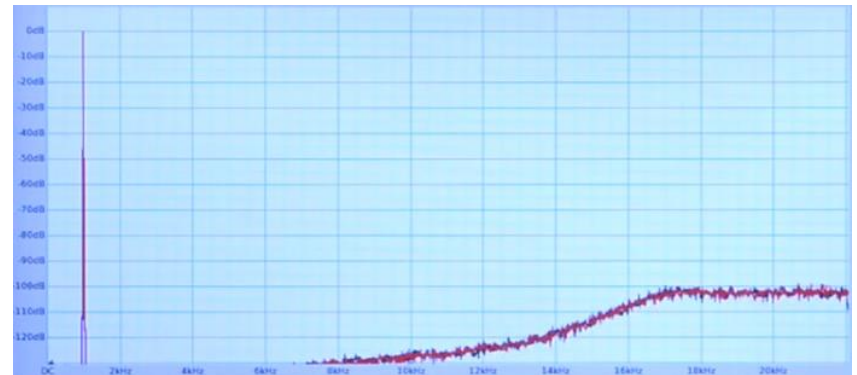
Без дитеринга

- При квантовании выбирается ближайший уровень.
- Шум квантования приводит к искажениям звука
 - При глубине 8 бит их может быть слышно



С дитерингом

- Выбирая уровни квантования более хитро можно сделать шум равномерным.
- Или убрать его в высокие частоты.



Ухудшение качества звука при цифровой обработке

- Ограничение частотного диапазона
 - Удаляются высокочастотные сигналы.
 - В местах резкого изменения сигнала возникают волнообразные искажения.
 - Ошибка лежит высокочастотной области.
 - Это происходит только один раз. При повторной обработке искажения не накапливаются.
- Дискретизация по времени
 - Определяет частотный диапазон звука.
 - Не приводит к потерям.
- Квантование по уровням
 - Вносит шум квантования.
 - Можно уменьшить и убрать из диапазона слышимых частот с помощью дитеринга.
- Если над звуком будут проводится вычисления, лучше иметь запас по уровням
 - Иначе можно добавить ещё шумов.

Объём информации для хранения звука

Для хранения звука требуется

$$\begin{aligned} I &= k * N * F * t \text{ бит} = \\ &= k * N * F * t / 8 \text{ байт,} \end{aligned}$$

где t – продолжительность звука в секундах,

N – глубина кодирования в битах,

F – частота дискретизации в Гц,

k – число каналов (1 - моно, 2 – стерео, 4 – квадро и т.п.)

*Например, при $F = 44\text{кГц}$, глубине кодирования **16 бит** на отсчёт и длительности звука **4 минуты** для хранения **двух** каналов требуется*

$$I = 2 \cdot 44000 \cdot 16 \cdot 4 \cdot 60 / 8 / 1024 / 1024 \approx 40,28 \text{ Мбайт.}$$

Сжатие звуковой информации

- Сжатие без потерь

- Исходный цифровой сигнал может быть восстановлен без каких-либо искажений.
- Возможно сжатие до половины исходного размера.
- Форматы: ape, flac, m4a, wma (некоторые варианты)...

- Сжатие с потерями

- Из звука удаляются компоненты, которые «мало» заметны человеку. Восстановление исходного сигнала невозможно.
- Можно сжимать в 7 раз при приемлемом качестве.
- Можно сжимать в 200 раз, если нужен только голос.
- Форматы: mp3, ogg, aac, wma, ...

Другие подходы к хранению музыки

- **MIDI** (*Musical Instrument Digital Interface*) — стандарт на формат обмена данными между электронными музыкальными инструментами.
- Кодировать с привязкой во времени:
 - нажатие клавиш (ноты),
 - настройку громкости,
 - выбор тембра, темпа, тональности и др.,
 - позволяет управлять освещением, пиротехникой и т.п.
- Описывает как формат передачи данных, так и электрические интерфейсы.
- Воспроизведение может осуществляться на основе семплов — коротких звуковых фрагментов.