

# Кодирование звуковой информации

С начала 90-х годов персональные компьютеры получили возможность работать со звуковой информацией. Каждый компьютер, имеющий звуковую плату, микрофон и колонки, может записывать, сохранять и воспроизводить звуковую информацию.

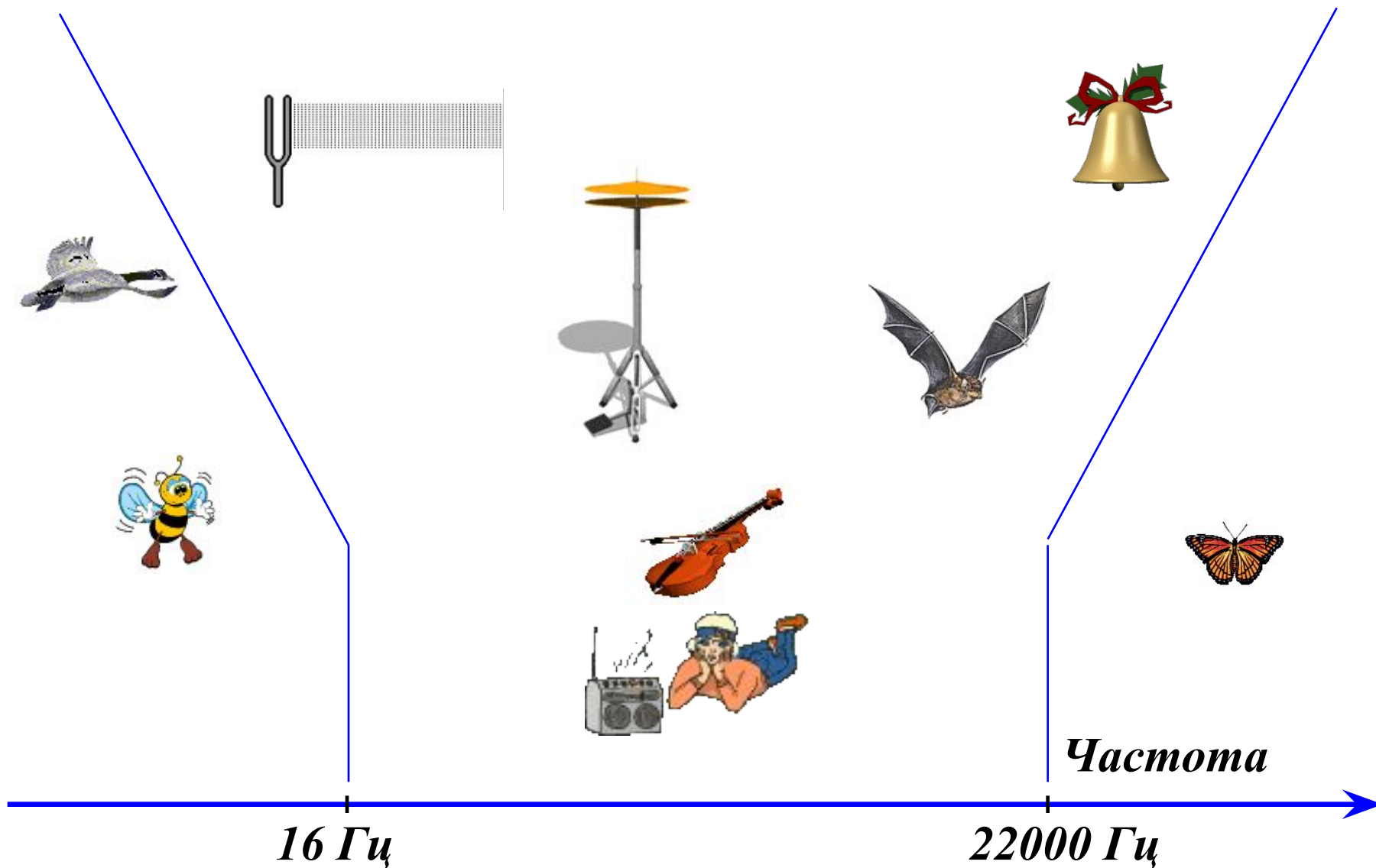
С помощью специальных программных средств (редакторов аудиофайлов) открываются широкие возможности по созданию, редактированию и прослушиванию звуковых файлов.

Создаются программы распознавания речи и появляется возможность управления компьютером при помощи голоса.

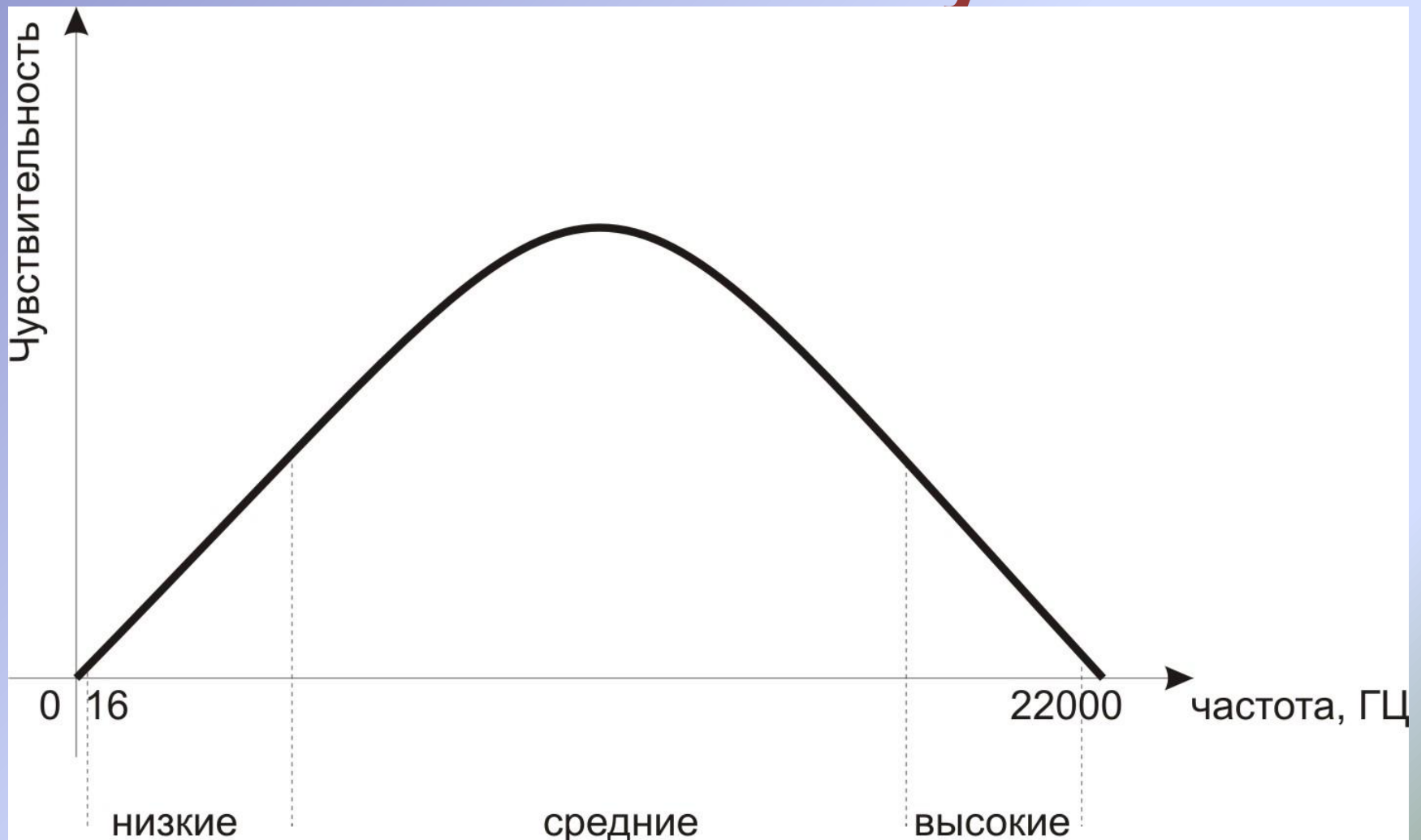
У всех источников звука имеются колеблющиеся части, которые приводят в колебательное движение частицы окружающей среды (воздуха) → распространяющаяся звуковая волна вызывает колебательное движение барабанной перепонки уха человека, которое воспринимается мозгом как звук → не все источники колебаний являются источниками звука → звук – механические колебания в частотном диапазоне от 16 Гц до 22000 Гц

Упругие волны в воздухе с частотой от 16 Гц до 20000 Гц вызывают у человека звуковые ощущения. Волны с частотой меньше 16 Гц называют *инфразвуковыми*, а с частотой больше 20 000 Гц - *ультразвуковыми*.

# Источники звуков



# Спектр частот, которые способно воспринимать человеческое ухо

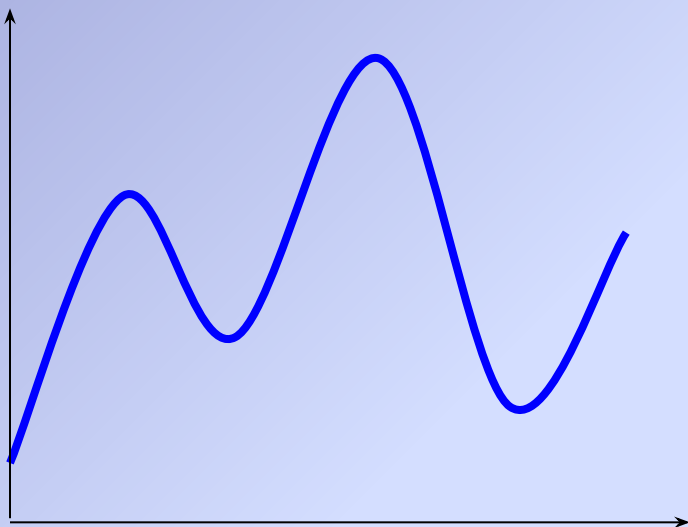


Звук представляет собой звуковую волну с непрерывно меняющейся амплитудой и частотой.

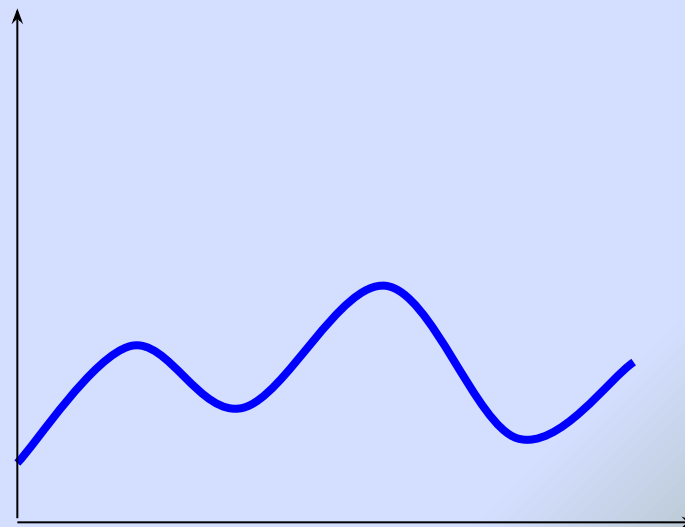
Чем больше амплитуда сигнала, тем он громче для человека, чем больше частота сигнала, тем выше тон.

# Звуки различной громкости

---



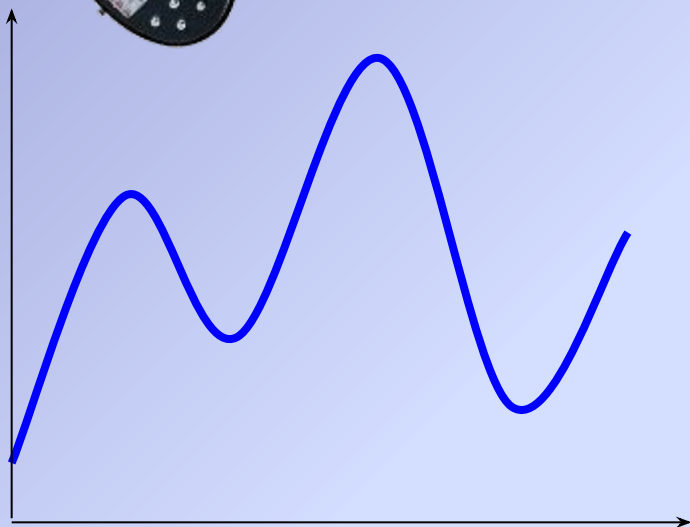
**Громкий звук**



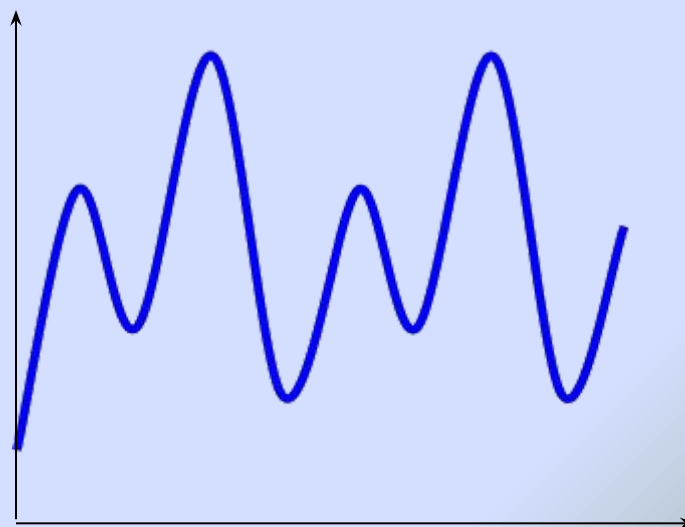
**Тихий звук**



# Звуки различной высоты



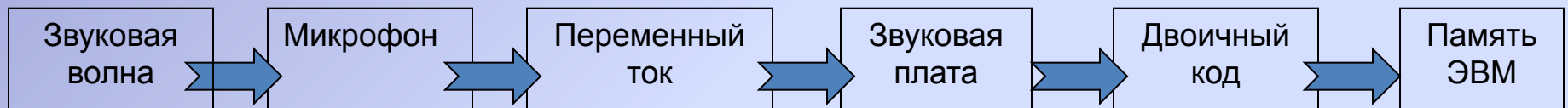
**Низкий звук**



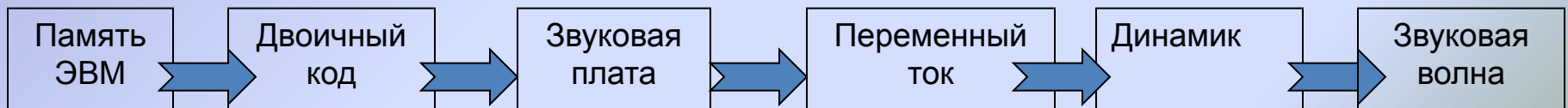
**Высокий звук**

Для того чтобы компьютер мог обрабатывать звук, непрерывный звуковой сигнал должен быть превращен в последовательность электрических импульсов (двоичных нулей и единиц).

### Схема кодирования звука

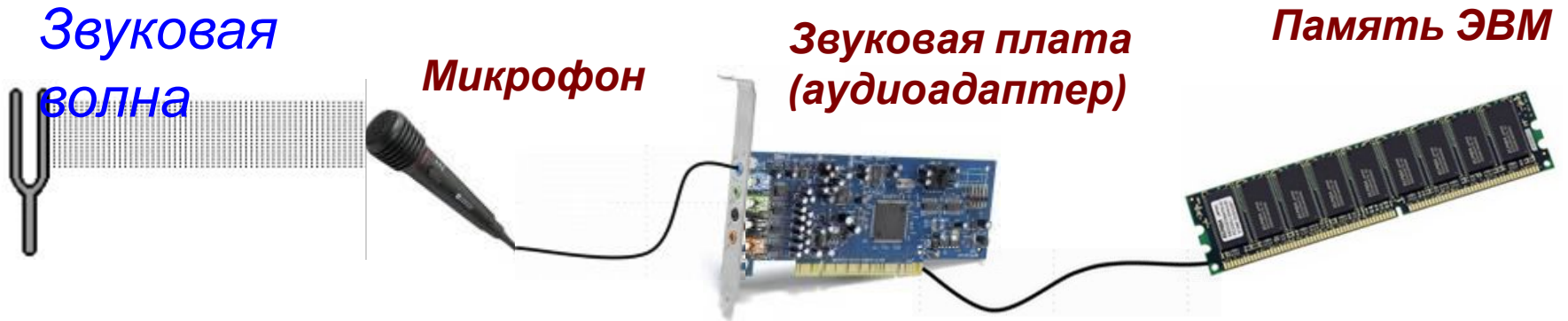


### Схема декодирования звука



# Схема преобразования звуковой волны в двоичный код

---



# Схема воспроизведения звука, сохранённого в памяти ЭВМ

---



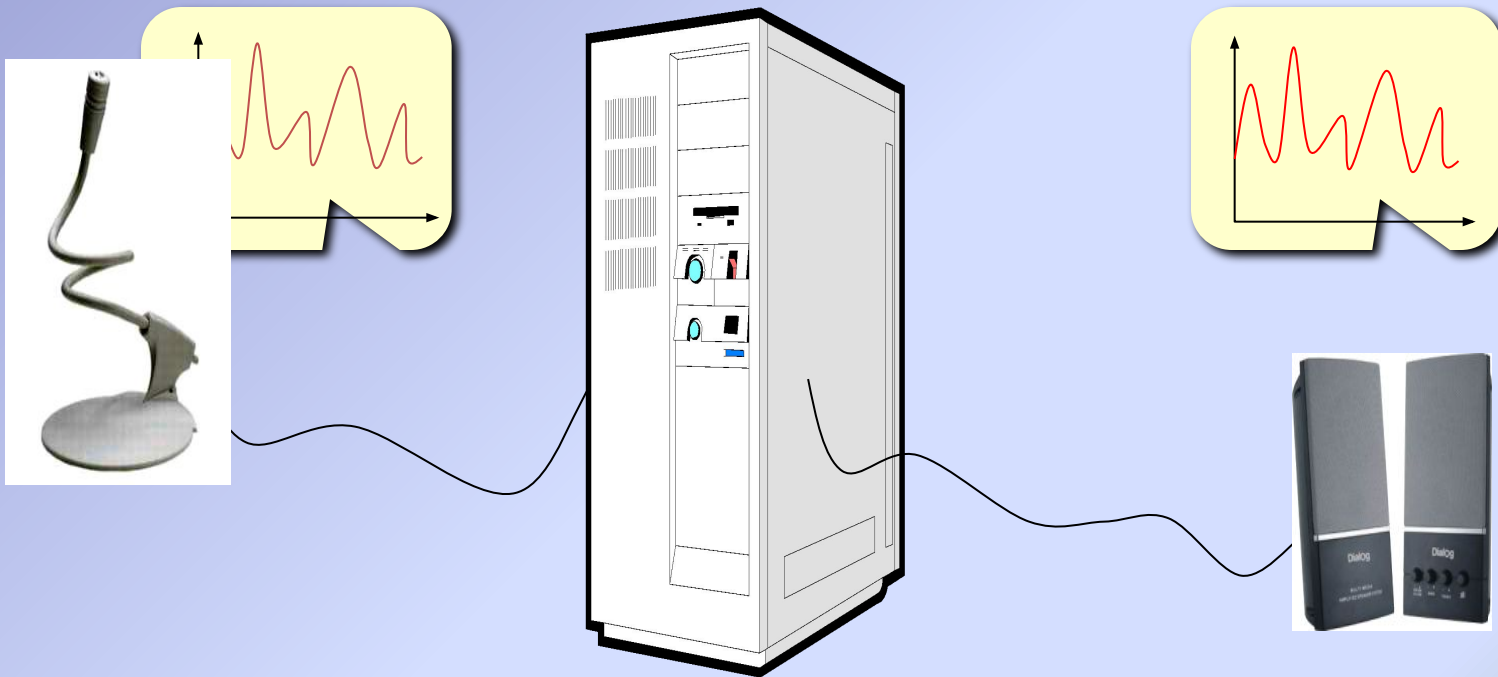
# Оцифровка (перевод в цифровую форму)

аналоговый сигнал

цифровой сигнал

1011010110101010011

аналоговый сигнал

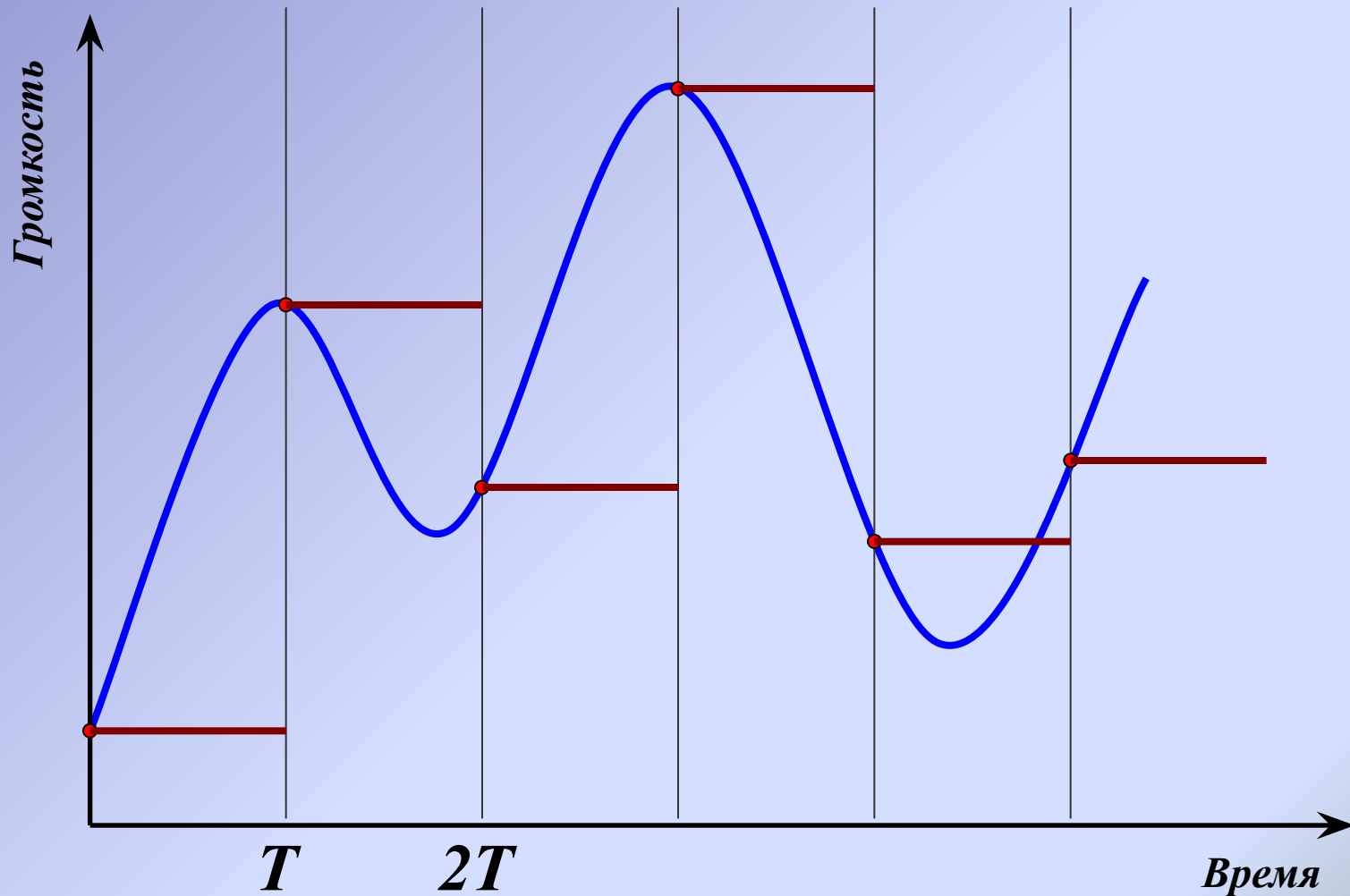


Непрерывная звуковая волна разбивается на отдельные маленькие временные участки, причем для каждого такого участка устанавливается определенная величина амплитуды.

Таким образом, непрерывная зависимость амплитуды сигнала от времени  $A(t)$  заменяется на дискретную последовательность уровней громкости. На графике это выглядит как замена гладкой кривой на последовательность “ступенек”.

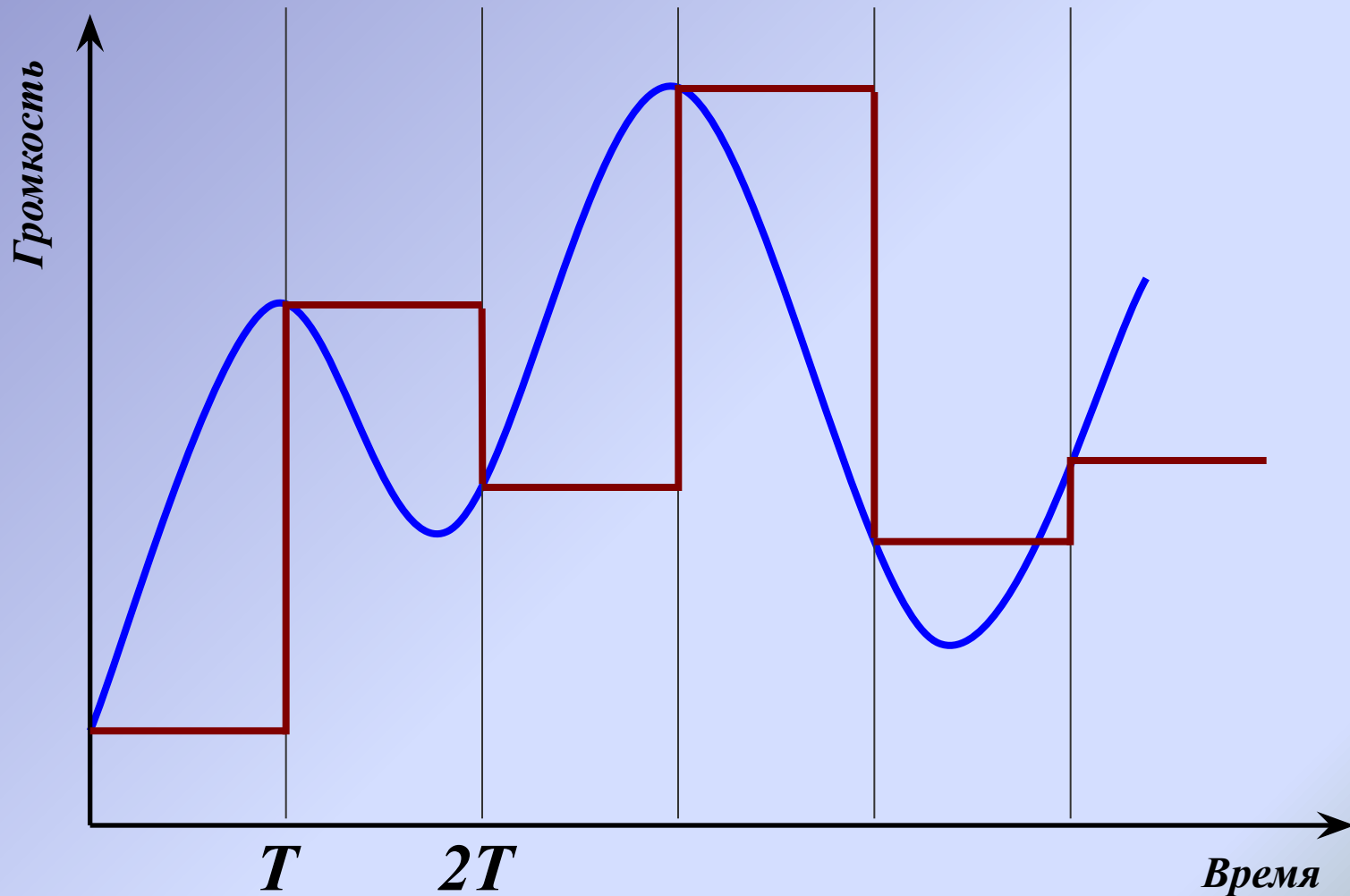
Процесс преобразования непрерывного аналогового сигнала в дискретный (прерывистый) называется **временной**

# Временная дискретизация



# Временная дискретизация

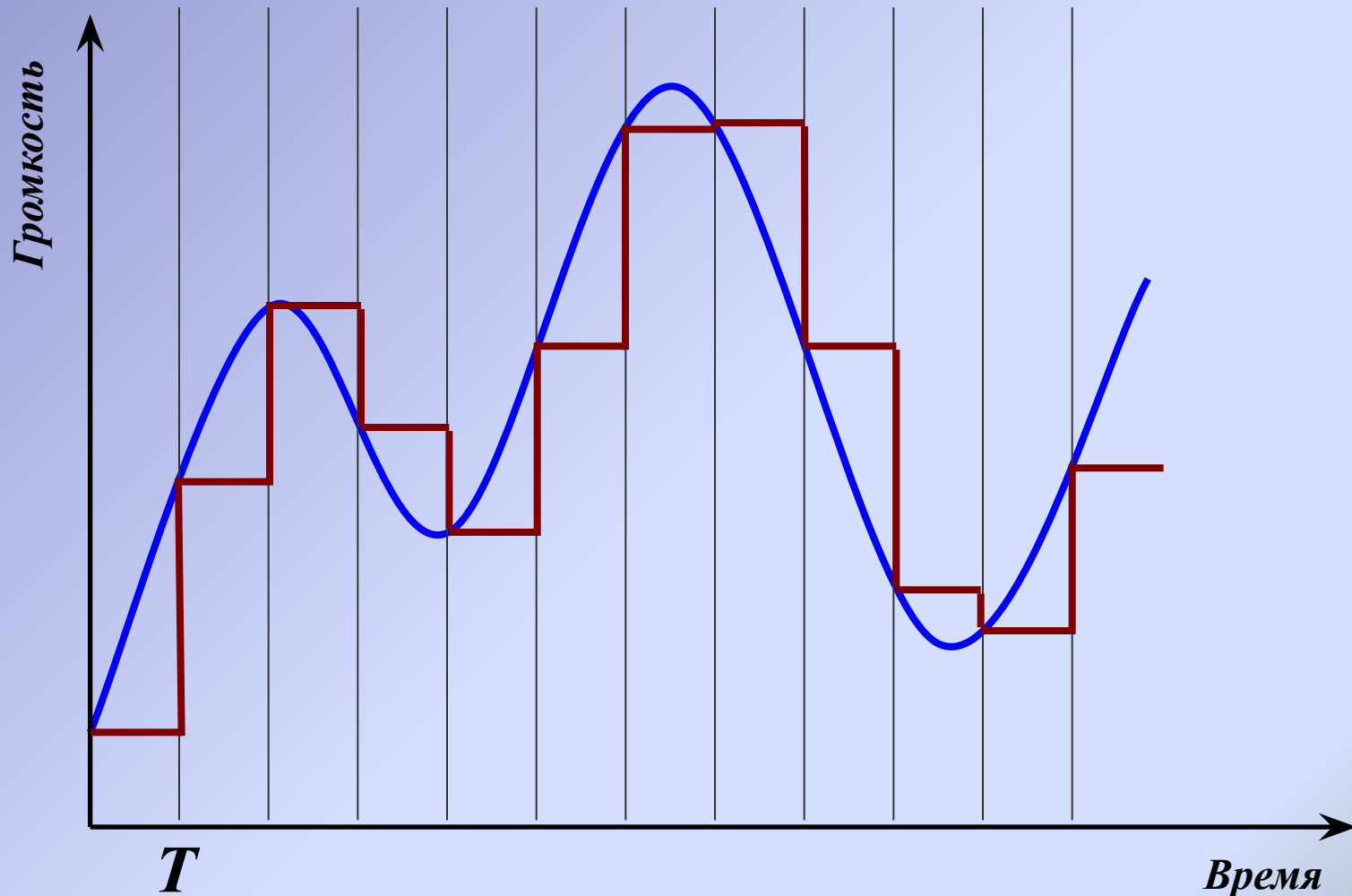
---





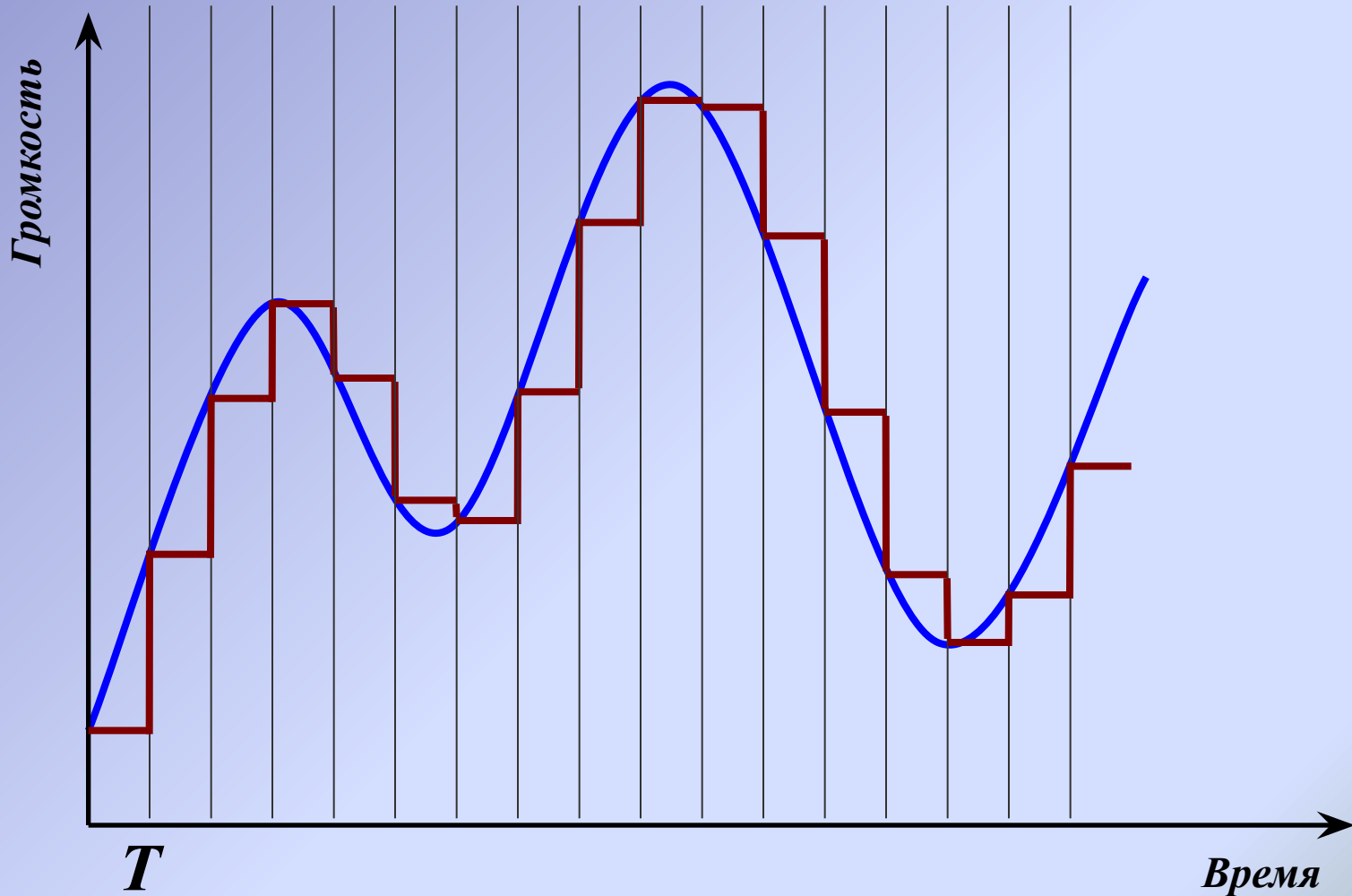
# Временная дискретизация

---



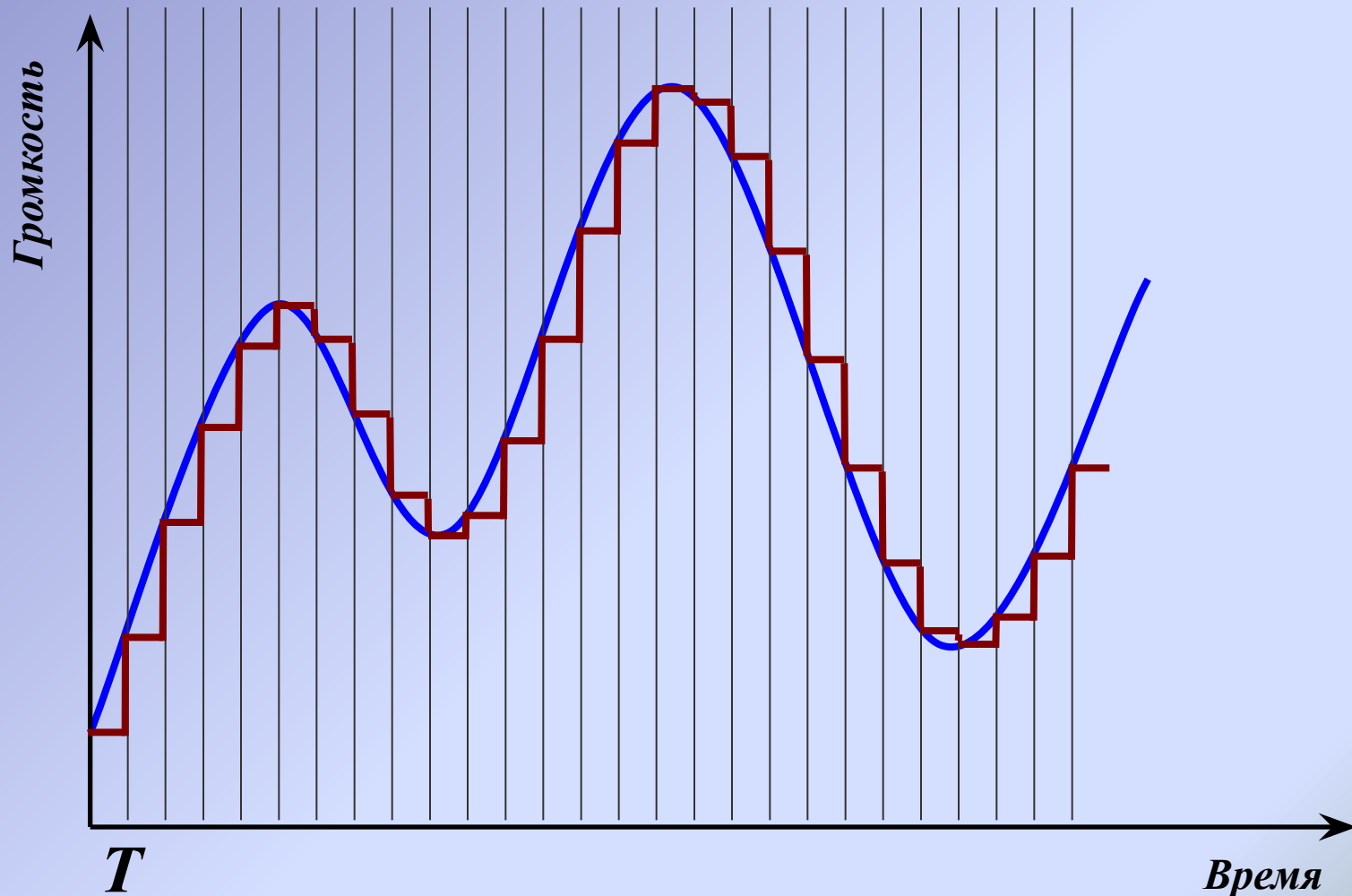
# Временная дискретизация

---



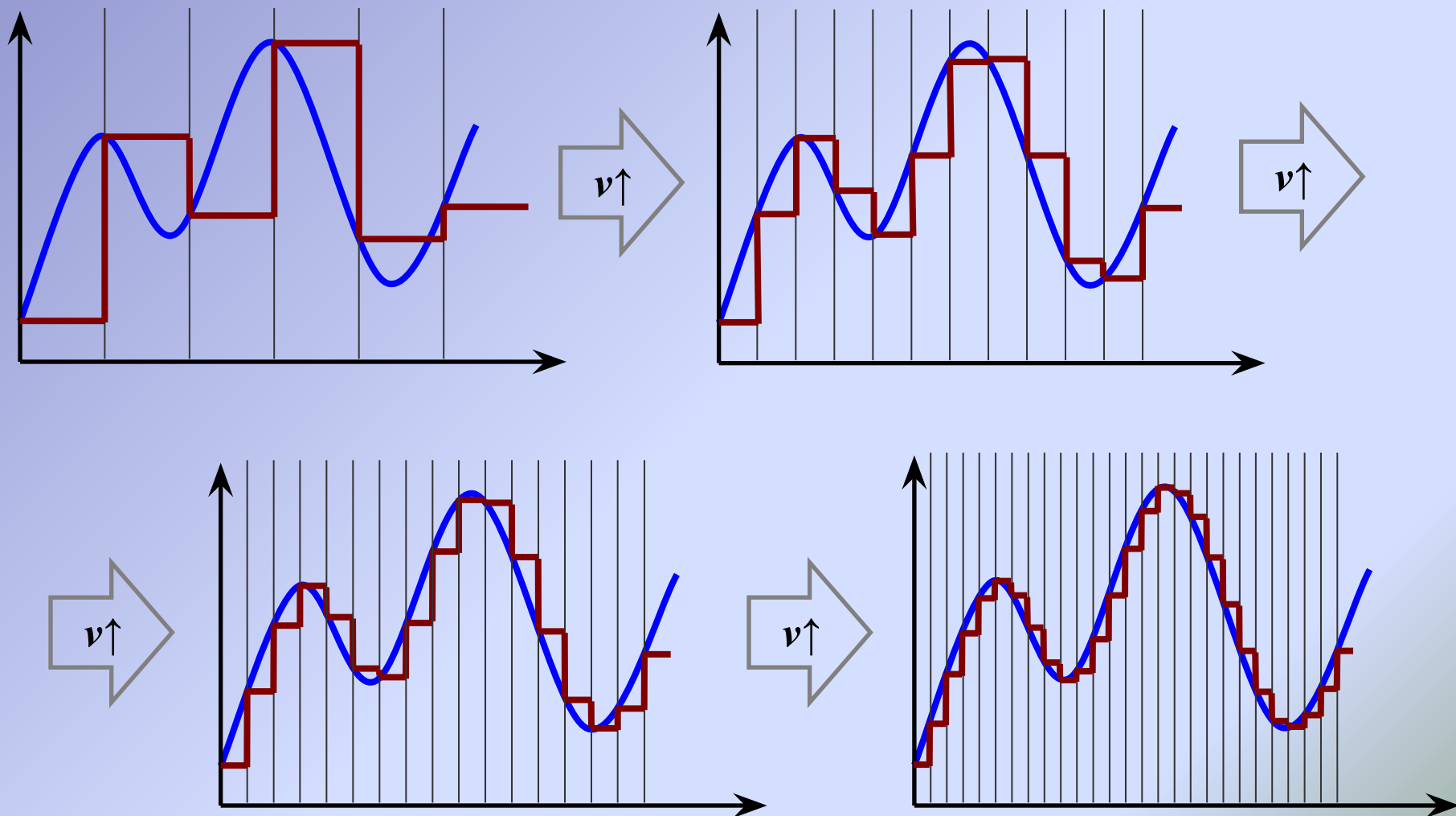
# Временная дискретизация

---



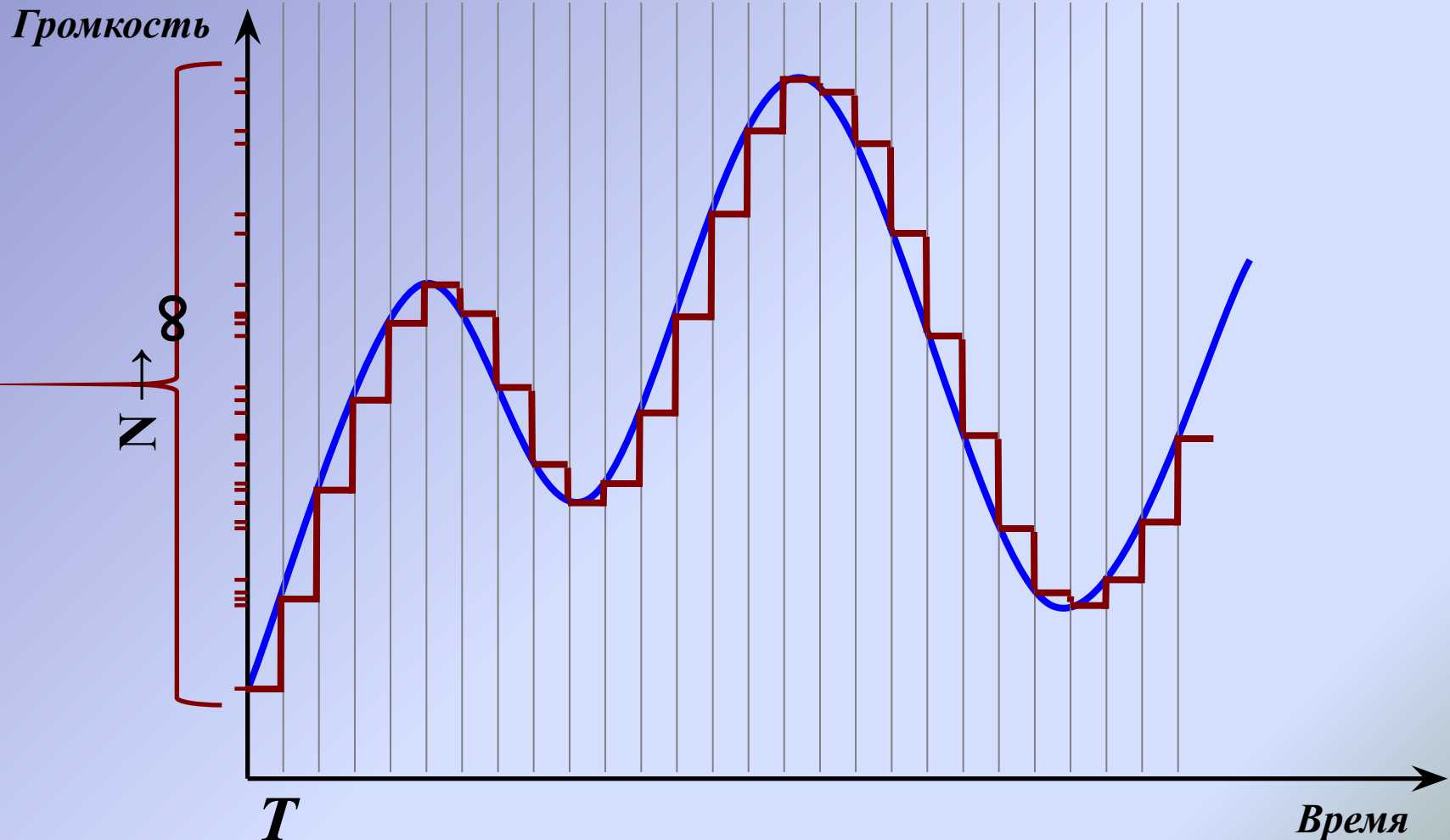
Количество измерений уровня звукового сигнала за 1 секунду называют ***частотой дискретизации***.

# Зависимость качества звука от частоты дискретизации

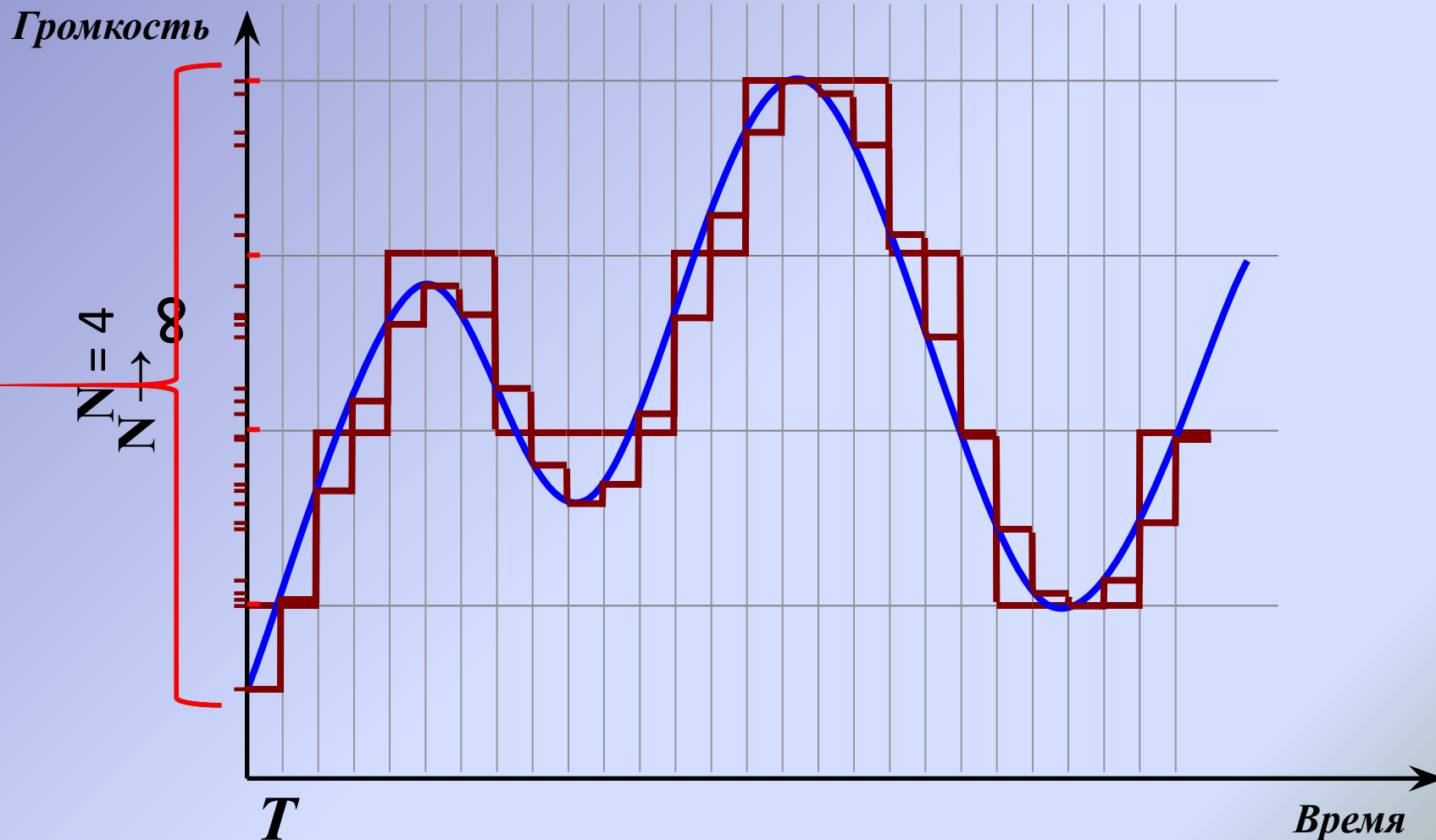


# Количество уровней громкости при дискретизации по времени

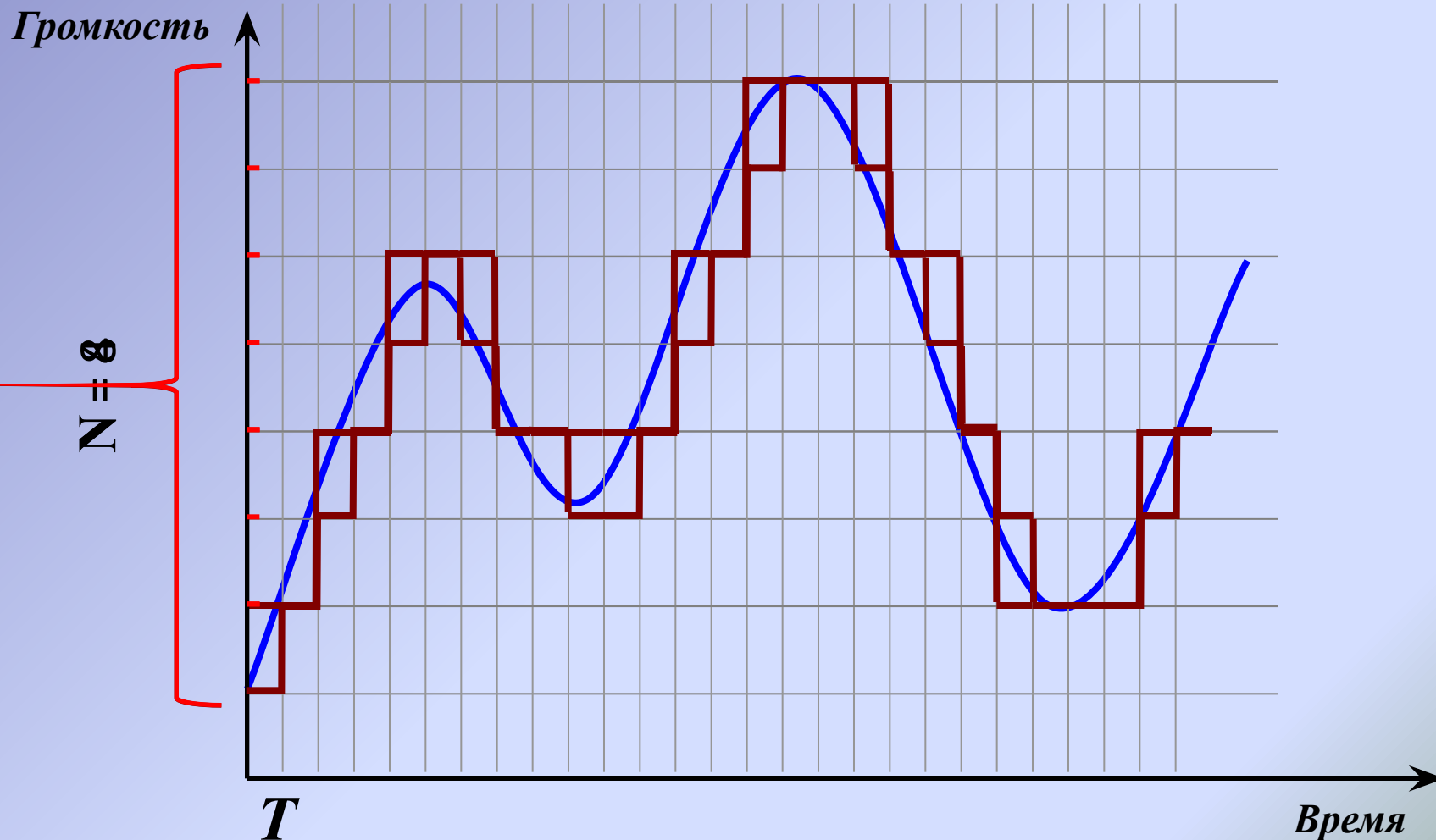
---



# Изменение качества звука при дискретизации по уровню



# Изменение качества звука при дискретизации по уровню

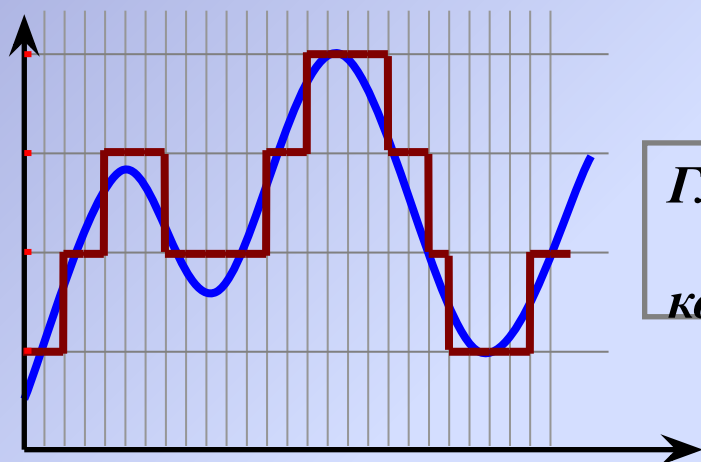
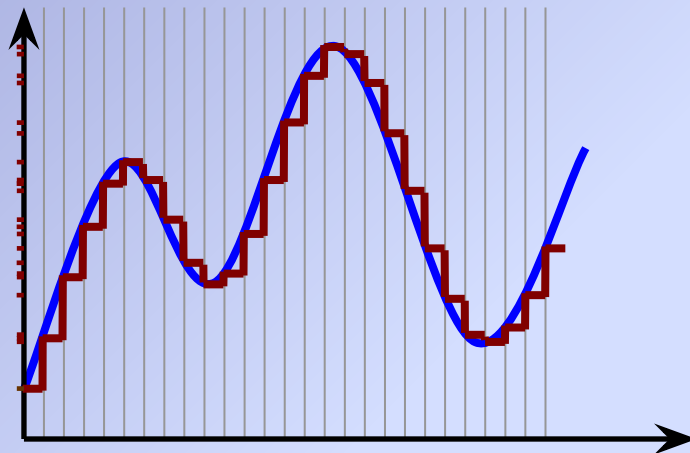




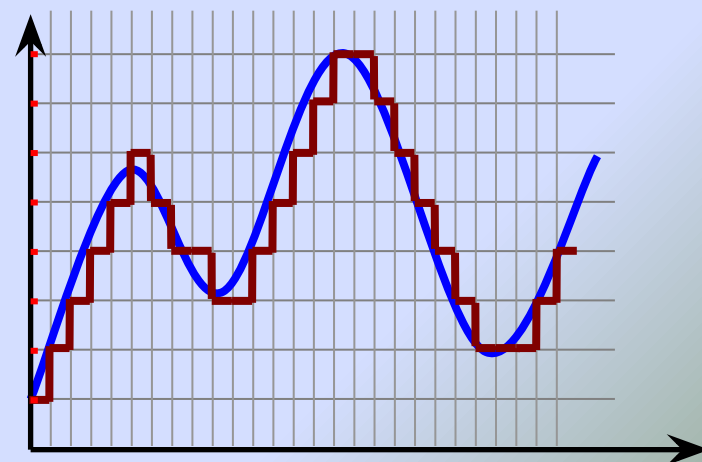
Количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука называют ***глубиной кодирования звука***.

# Зависимость качества звука от глубины кодирования

---



*Глубина кодирования* ↑



# Соответствие звуков различных характеристик некоторым источникам звука

---

*Радиотрансляция*



*8 кГц*

*8 бит*

*AudioCD*



*44,1 кГц*

*16 бит*

*DVD-Audio*



*192 кГц*

*24 бит*

# Расчёт объёма звукового файла

---

$$I = k \cdot v \cdot i \cdot t$$

Где  $I$  – размер (объём) звукового файла

$k$  – количество дорожек в записи ( $k=1$  – моно,  $k=2$  – стерео)

$v$  – частота дискретизации (в Герцах)

$i$  – глубина кодирования (в битах)

$t$  – время звучания (в секундах)

# Оценка объёма звукового файла

Определить объем памяти для хранения моноаудиофайла, время звучания которого составляет пять минут при частоте дискретизации 44 кГц и глубине кодирования 16 бит.

**Дано:**

$$v = 44 \text{ кГц}$$

$$i = 16 \text{ бит}$$

$$t = 5 \text{ мин}$$

$$k = 1$$

**Найти:**

$I$

**Решение:**

$$44 \text{ кГц} = 44000 \text{ Гц}$$

$$5 \text{ мин} = 300 \text{ с}$$

$$I = k v i t$$

$$I = 1 \cdot 44000 \text{ Гц} \cdot 16 \text{ бит} \cdot 300 \text{ с} = 26400000 \text{ бит} \approx$$

$$\approx 25781,25 \text{ Кб} \approx 25,2 \text{ Мб}$$

Ответ:  $I = 25,2 \text{ Мб}$

# Звуковые редакторы

Звуковые редакторы позволяют не только записывать и воспроизводить звук, но и редактировать его. Они позволяют изменять качество звука и объем звукового файла.

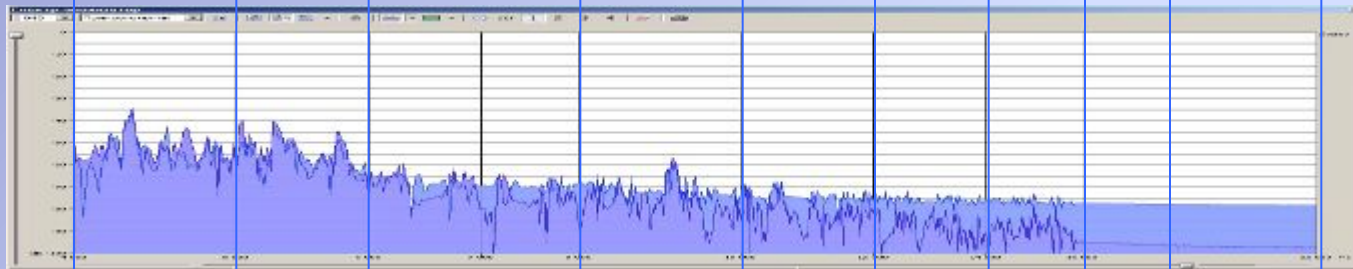
Оцифрованный звук можно сохранять без сжатия в универсальном формате **wav** или в формате со сжатием **mp3**.

- **WAV** (*Waveform audio format*), часто без сжатия (размер!)
- **MP3** (*MPEG-1 Audio Layer 3*, сжатие с потерями)
- **WMA** (*Windows Media Audio*, потоковый звук, сжатие)

# Изменение качества при сжатии звуковых файлов



*Спектрограмма несжатого звука (формат WAV)*



*Спектрограмма сжатого звука (формат mp3, битрейт 128 кбит/с)*



*Спектрограмма сжатого звука (формат WMA, битрейт 128 кбит/с)*

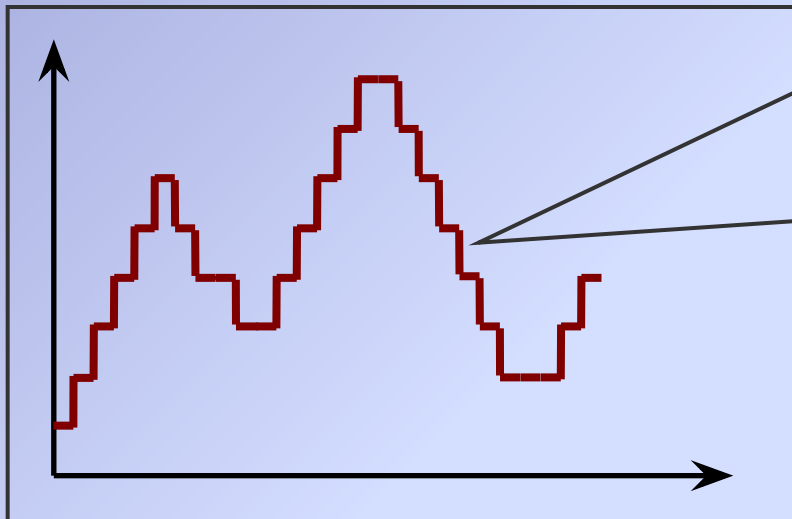
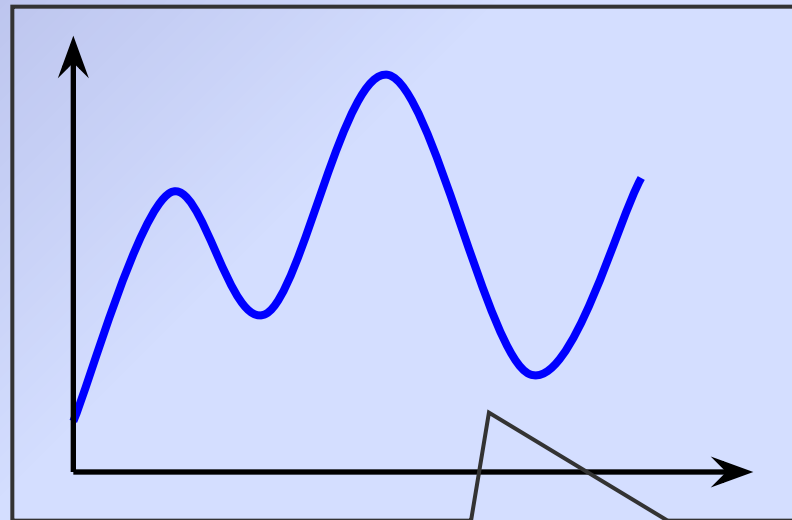
4 5 6 8 10 12 14 16 18 22,5

Частота звука, кГц



# Звук «живой» и оцифрованный

---





# Задачи

1. Оцените информационный объем моноаудиофайла длительностью звучания 20 с, если "глубина" кодирования и частота дискретизации звукового сигнала равны соответственно 8 бит и 8 кГц.

# Задачи

2. Рассчитайте время звучания моноаудиофайла, если при 16-битном кодировании и частоте дискретизации 32 кГц его объем равен 700 Кбайт.