

«Кодирование и обработка звуковой информации»

Звуковая информация

- Звук представляет собой распространяющуюся в воздухе, воде или другой среде волну (колебания воздуха или другой среды) с непрерывно меняющейся **амплитудой** и **частотой**. Человек воспринимает звуковые волны с помощью слуха в форме звука различной **громкости** и **тона**. Чем больше амплитуда звуковой волны, тем громче звук, чем больше частота колебаний, тем выше тон звука.



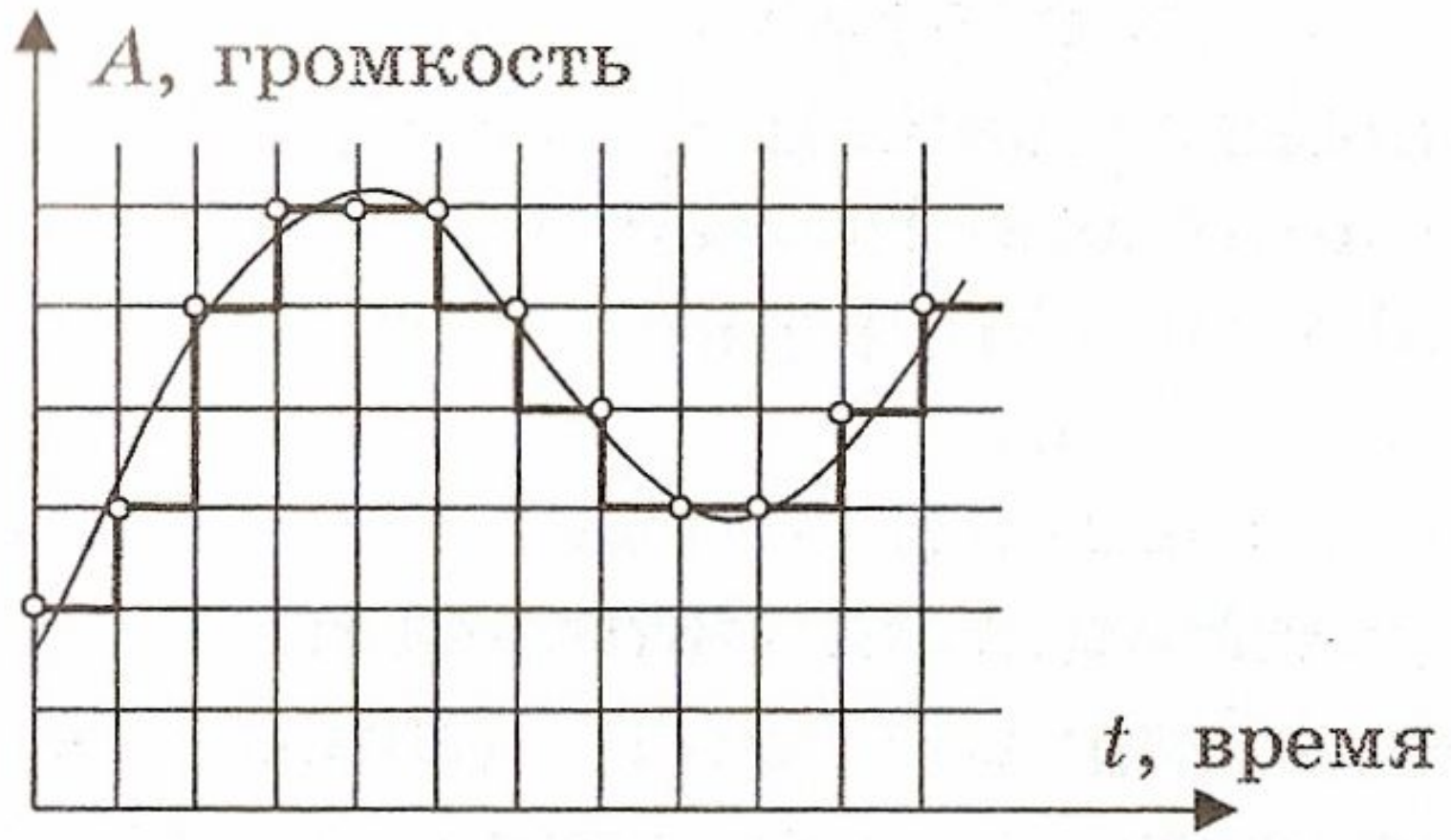
- Человеческое ухо воспринимает звук с частотой от 20 колебаний в секунду (низкий звук) до 20000 колебаний в секунду (высокий звук).
- Человек может воспринимать звук в огромном диапазоне амплитуд, в котором максимальная амплитуда больше минимальной в 10^{14} раз (в сто тысяч миллиардов раз).
- Для измерения громкости звука применяется специальная единица **децибел** (дБ). Уменьшение или увеличение громкости звука на 10 дБ соответствует уменьшению или увеличению амплитуды звука в 10 раз.

Таблица 1.5. Громкость звука

Звук	Громкость, дБ
Нижний предел чувствительности человеческого уха	0
Шорох листьев	10
Разговор	60
Гудок автомобиля	90
Реактивный двигатель	120
Болевой порог	140

Временная дискретизация звука.

- Для того чтобы компьютер мог обрабатывать звук, непрерывный звуковой сигнал должен быть преобразован в цифровую дискретную форму с помощью **временной дискретизации**. Непрерывная звуковая волна разбивается на отдельные маленькие временные участки, причем для каждого такого участка устанавливается определенный уровень громкости.
- Таким образом, непрерывная зависимость громкости звука от времени $A(t)$ заменяется на дискретную последовательность уровней громкости. На графике это выглядит как замена гладкой кривой на последовательность «ступенек».



Частота дискретизации

- Для записи аналогового звука и его преобразования в цифровую форму используется микрофон, подключенный к звуковой плате. Качество полученного цифрового звука зависит от количества измерений громкости звука в единицу времени, т. е. **частоты дискретизации**. Чем большее количество измерений производится за одну секунду (чем больше частота дискретизации), тем точнее «лесенка» цифрового звукового сигнала повторяет кривую аналогового сигнала.
- **Частота дискретизации звука** — это количество измерений громкости звука за одну секунду. Частота дискретизации звука может лежать в диапазоне от 8000 до 48000 измерений громкости звука за одну секунду.

Глубина кодирования

- Каждой «ступеньке» присваивается определенный уровень громкости звука. Уровни громкости звука можно рассматривать как набор N возможных состояний, для кодирования которых необходимо определенное количество информации I , которое называется глубиной кодирования звука.
- **Глубина кодирования звука** — это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука.

- Если известна глубина кодирования, то количество уровней громкости цифрового звука можно рассчитать по формуле $N=2^I$. Пусть глубина кодирования звука составляет 16 битов, тогда количество уровней громкости звука равно:
 $N=2^I = 2^{16} = 65536$.
В процессе кодирования каждому уровню громкости звука присваивается свой 16-битовый двоичный код, наименьшему уровню громкости будет соответствовать код 0000000000000000, а наибольшему 1111111111111111.

Качество оцифрованного звука.

- Чем больше частота и глубина дискретизации звука, тем более качественным будет оцифрованный звук. Самое низкое качество оцифрованного звука, соответствующее качеству телефонной связи, будет при частоте дискретизации 8000 раз в секунду, глубине дискретизации 8 битов и записи одной звуковой дорожки (режим моно). Самое высокое качество оцифрованного звука, соответствующее качеству аудио-CD будет при частоте дискретизации 48000 раз в секунду, глубине дискретизации 16 битов и записи двух звуковых дорожек (режим стерео).

- Необходимо помнить, что чем выше качество цифрового звука, тем больше информационный объем звукового файла. Можно оценить информационный объем цифрового стереозвукового файла длительностью звучания одна секунда при среднем качестве звука (16 битов, 24000 измерений в секунду). Для этого глубину кодирования необходимо умножить на количество измерений в одну секунду и умножить на 2 (стереозвук):
 $16 \text{ битов} * 24\,000 * 2 = 768\,000 \text{ битов} = 96\,000 \text{ байтов} = 93,75 \text{ Кбайт}.$

Звуковые редакторы.

- Звуковые редакторы позволяют не только записывать и воспроизводить звук, но и редактировать его. Оцифрованный звук представляется в звуковых редакторах в наглядной форме, поэтому операции копирования, перемещения и удаления частей звуковой дорожки можно легко осуществлять с помощью мыши. Кроме того, можно накладывать звуковые дорожки друг на друга (микшировать звуки) и применять различные акустические эффекты (эхо, воспроизведение в обратном направлении и др.).

- Звуковые редакторы позволяют изменять качество цифрового звука и объем звукового файла путем изменения частоты дискретизации и глубины кодирования. Оцифрованный звук можно сохранять без сжатия в звуковых файлах в универсальном формате WAV, а также в формате со сжатием MP3. При сохранении звука в форматах со сжатием отбрасываются «избыточные» для человеческого восприятия звуковые частоты с малой амплитудой, совпадающие по времени со звуковыми частотами с большой амплитудой. Применение такого формата позволяет сжимать звуковые файлы в десятки раз, однако приводит к необратимой потере информации (файлы не могут быть восстановлены в первоначальном виде).

- На рисунке ниже изображено зафиксированное самописцем звучание 1 секунды речи. Необходимо:
 - 1) закодировать его в двоичном коде с частотой дискретизации 5 Гц и глубиной звука 4 бита;
 - 2) рассчитать информационный объем закодированного звука.



Решение:

- 1) 5 Гц — это значит, что происходит 5 измерений в 1 сек. Глубина 4 бита означает, что используются $2^4 = 16$ уровней громкости. (Результат кодирования: 1000 1000 1001 0110 0111)
- 2) для расчета информационного объема закодированного звука (V) используется простая формула:

$$V = D * I * T,$$

D — частота дискретизации (Гц);

I — глубина звука (бит);

T — время звучания (сек).

- Получаем: $V = 5 \text{ Гц} * 4 \text{ бита} * 1 \text{ сек} = 20 \text{ бит}$.

- 1) Какой объем памяти требуется для хранения цифрового аудиофайла с записью звука высокого качества (16 бит, 48 кГц) при условии, что время звучания составляет 2 минуты?
(22 Мбайта)
- 2) Оцените информационный объем моноаудиофайла длительностью звучания 1 мин, если "глубина" кодирования и частота дискретизации звукового сигнала равны соответственно:
 - а) 16 бит и 8 кГц;
(940 Кбайт)
 - 16 бит и 24 кГц.
(2,8 Мбайт)

- Определите качество звука (качество радиотрансляции, среднее качество, качество аудио-CD) если известно, что объем моноаудиофайла длительностью звучания в 10 сек равен:
 - а) 940 Кбайт;
- **(качество CD);**
- 157 Кбайт.
- **(качество радиотрансляции)**

- Рассчитайте время звучания моноаудиофайла, если при 16-битном кодировании и частоте дискретизации 32 кГц его объем равен:
 - а) 700 Кбайт;
 - **(10 сек);**
 - б) 6300 Кбайт.
 - **(1,5 мин)**

- Определите длительность звукового файла, который уместится на гибкой дискете 3,5". Учтите, что для хранения данных на такой дискете выделяется 2847 секторов объемом 512 байт.
 - а) при низком качестве звука: моно, 8 бит, 8 кГц;
- **(3 минуты);**
- б) при высоком качестве звука: стерео, 16 бит, 48 кГц.
- **(7,6 секунды).**

Домашнее задание

- Прочитать материал учебника п. 1.5 (стр. 40-44), выполнить задания в тетради 1.9 - 1.11 (стр. 44).

Итог урока

- Объясните, как частота дискретизации и глубина кодирования влияют на качество цифрового звука