Кодирование звука.

Звук, который мы слышим каждый день - речь, музыка, шумы - представляет собой **непрерывные волны давления воздуха**. Компьютер же умеет работать только с **дискретными (цифровыми) значениями**, представленными нулями и единицами. Поэтому перед хранением или обработкой звуковую волну необходимо **оцифровать**, то есть превратить в последовательность чисел.

Процесс преобразования аналогового звука в цифровой сигнал называется **аналогово-цифровым преобразованием (АЦП)**. Чтобы понять, как хранится звук и почему одни аудиофайлы «весят» больше других, нужно разобраться в ключевых параметрах: **частоте дискретизации**, **разрядности**, **количестве каналов** и **битрейте**.

Частота дискретизации

Когда АЦП оцифровывает звук, он делает «замеры» громкости звука тысячи раз в секунду. Количество таких измерений называется **частотой дискретизации**.

Обозначается в герцах (Гц) или килогерцах (кГц).

Примеры частот:

Область применения	**Частота**
Телефонная связь	8 000 Гц
Радио/старая аудиотехника	22 050 Гц
Музыка (CD-качество)	44 100 Гц
Видео, фильмы	48 000 Гц
Студийная запись	96 000–192 000 Гц

Чем выше частота дискретизации, тем точнее цифровая копия повторяет оригинальную звуковую волну.

Закон Найквиста–Шеннона

Чтобы корректно записать звук без искажений, частота дискретизации должна быть **в 2 раза выше максимальной частоты слышимого звука**.

Человеческое ухо слышит до ~20 кГц → откуда и взялось стандартное **44,1 кГц**, что покрывает до ~22 кГц.

Если частота низкая - звук становится «глухим» и «телефонным».

Разрядность

Разрядность - это **точность измерения громкости звука**. Каждый измеренный отсчёт - это число, и разрядность показывает, сколько бит используется для его хранения.

Типичные значения:

Разрядность	Применение	Качество
8 бит	старые игры, телефон	низкое, слышен шум
16 бит	CD-качество	стандарт высокой чёткости
24 бит	студийная запись	лучший динамический диапазон

Разрядность определяет:

- **диапазон громкости** (динамический диапазон),
- количество уровней - например:
  8 бит → 256 уровней громкости
  16 бит → 65 536 уровней

Чем выше разрядность - тем меньше искажения и шумы квантования.

Уровень громкости аудио

Уровень громкости измеряется в **децибелах**, это логарифмическая шкала.
0 дБ - это порог слышимости для человека (самые тихие звуки).
140 дБ - порог боли.

Для аудиозаписи важны два параметра:

Максимальная громкость (0 dBFS)
В цифровом звуке это "ceiling" - предел, выше которого звук начинает искажаться (клиппинг).

Минимально различимый уровень (шумовой порог)
Зависит от разрядности. Чем она ниже, тем выше уровень шума.

Таблица уровней громкости реальных звуков

Звук	**Уровень, дБ**	**Комментарий**
Порог слышимости	0 дБ	человек едва слышит
Шёпот на расстоянии 1 м	20–30 дБ	очень тихо
Библиотека	40 дБ	приглушённая тишина
Обычный разговор	50–60 дБ	комфортная громкость
Городская улица	70 дБ	шумно
Пылесос	80 дБ	громко
Музыка в наушниках (громко)	90 дБ	длительное воздействие вредно
Концерт, клуб	100–110 дБ	опасно при длительном воздействии

Сирена скорой	120 дБ	очень громко
Реактивный самолёт рядом	130–140 дБ	боль, повреждение слуха
Порог боли	140 дБ	максимум восприятия

Эта шкала важна, потому что цифровая система должна уметь передать как тихие 20 дБ, так и громкие 100 дБ - без искажений.

Что такое динамический диапазон

Динамический диапазон - это разница между:

максимальным уровнем звука, который можно записать без искажений,

минимальным уровнем, который можно отличить от шума системы.

Измеряется в дБ.

Пример:
Если динамический диапазон системы = 96 дБ, то она способна различать громкость в пределах 96 дБ от самых тихих звуков до предельно громких.

Как разрядность определяет динамический диапазон

Каждый бит разрядности даёт примерно **6 дБ динамического диапазона**.

Формула:

Динамический диапазон $\approx 6 * $ количество бит

Примеры:

Разрядность	Динамический диапазон	Описание
8 бит	≈ 48 дБ	слышимый шум, «зерно», характерно для старых игр
16 бит	≈ 96 дБ	достаточно для музыки, речи - стандарт CD
24 бит	≈ 144 дБ	превышает динамику человеческого слуха; используется в студии
32 бит float	до 1500 дБ (теоретически)	используется для редактирования, не для реального воспроизведения

Почему низкая разрядность даёт шум?

Когда громкость звука записывается в виде цифрового числа, компьютер округляет значение до ближайшего уровня.
Если уровней мало (8 бит → 256 уровней), округление становится заметным → слышен **шум квантования**.

При 16 бит (65 тысяч уровней громкости) ошибка округления минимальна, шум почти не слышен.
При 24 бит - вообще исчезает.

Взаимосвязь разрядности и качества звучания

16 бит (CD)

- динамический диапазон 96 дБ
- достаточно для музыки
- шумы практически не слышны
- стандарт для большинства аудиоформатов

24 бит (студийное качество)

- позволяет записывать очень тихие детали (дыхание, реверберацию)
- улучшенный запас громкости
- меньше риск клиппинга
- используется в профессиональных DAW

8 бит

- заметный шум
- «плоский» звук
- часто используется для ретро-эффектов или там, где качество не важно

Количество каналов

Канал - это независимая звуковая дорожка.

- **1 канал** - моно
- **2 канала** - стерео
- **5.1, 7.1** - системы окружающего звука

Стерео занимает в 2 раза больше памяти, чем моно, потому что данных - вдвое больше.

Битрейт (Bitrate) - насколько «тяжёлый» звук

Битрейт - это **количество бит, которое требуется для хранения одной секунды звука**.

Для несжатого аудио (WAV) битрейт определяется формулой:

Битрейт = F * R * C,

где

F - частота дискретизации,

R - разрядность,

C - число каналов.

Пример: WAV 44.1 кГц, 16 бит, стерео:
44100 * 16 * 2 = 1,411,200 бит/с = 1411,2 кбит/с.

То есть **1 секунда WAV почти всегда > 176 КБ**.

Для MP3/AAC битрейт указывается напрямую (128 kbps, 320 kbps…). MP3 хранит звук намного компактнее за счёт сжатия.

Расчёт объёма звукового файла

Основная формула:

I = F * R * t * C,

где
I - объём (бит),

F - частота дискретизации,

R - разрядность,

C - число каналов.
t - время в секундах.

Пример расчёта

Дан звук:

- 44 100 Гц
- 16 бит
- стерео
- 10 секунд

Ищем битрейт:

44100 * 16 * 2 = 1,411,200 бит/с.

Умножаем на длительность:
1,411,200 * 10 = 14,112,000 бит.

Перевод:

14,112,000 / 8 = 1,764,000 байт,
1,764,000 / 1024 = 1722 КБ,
1722 / 1024 = 1.68 МБ.

Итог: файл будет занимать ~1.7 МБ.

Это очень много для 10 секунд - потому что WAV **не использует сжатие**.

Сжатое и несжатое аудио

Тип	**Примеры**	**Особенности**
Несжатое	WAV, AIFF	Чистый звук, огромный размер
С потерями	MP3, AAC, OGG	Малый размер, возможные искажения
Без потерь	FLAC, ALAC	Качество = WAV, но размер меньше

MP3 может быть в 10–15 раз меньше WAV благодаря удалению информации, которую человек плохо воспринимает.

Почему нельзя бесконечно повышать частоту и разрядность?

1. Ухо всё равно слышит только до 20 кГц.
2. Размер файла растёт огромными темпами.
3. Оборудование (микрофоны, колонки) физически не может воспроизвести сигнал выше определённого диапазона.

4. Выгода минимальна: разницу между 96 и 192 кГц слышат только приборы, а не люди.