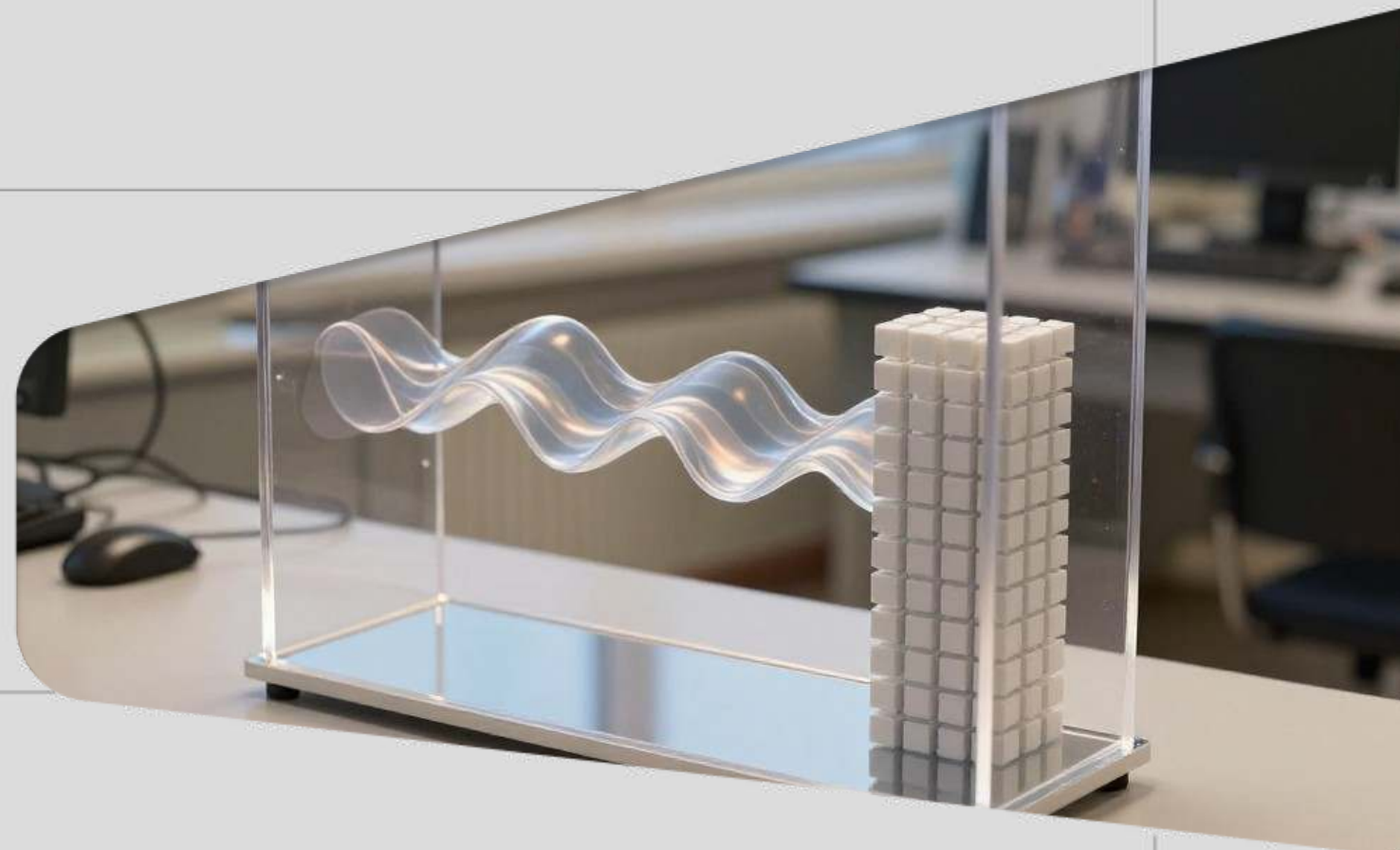


Представление звуковой информации в компьютере

Звук превращается в числа: от параметров волны до форматов, хранения и передачи.



Почему звук нельзя «сохранить как волну»

Компьютер работает с дискретными данными, а звуковая волна в природе непрерывна. Поэтому нужны аналого-цифровое преобразование и обратное воспроизведение, а качество записи задают дискретизация, разрядность и число каналов.



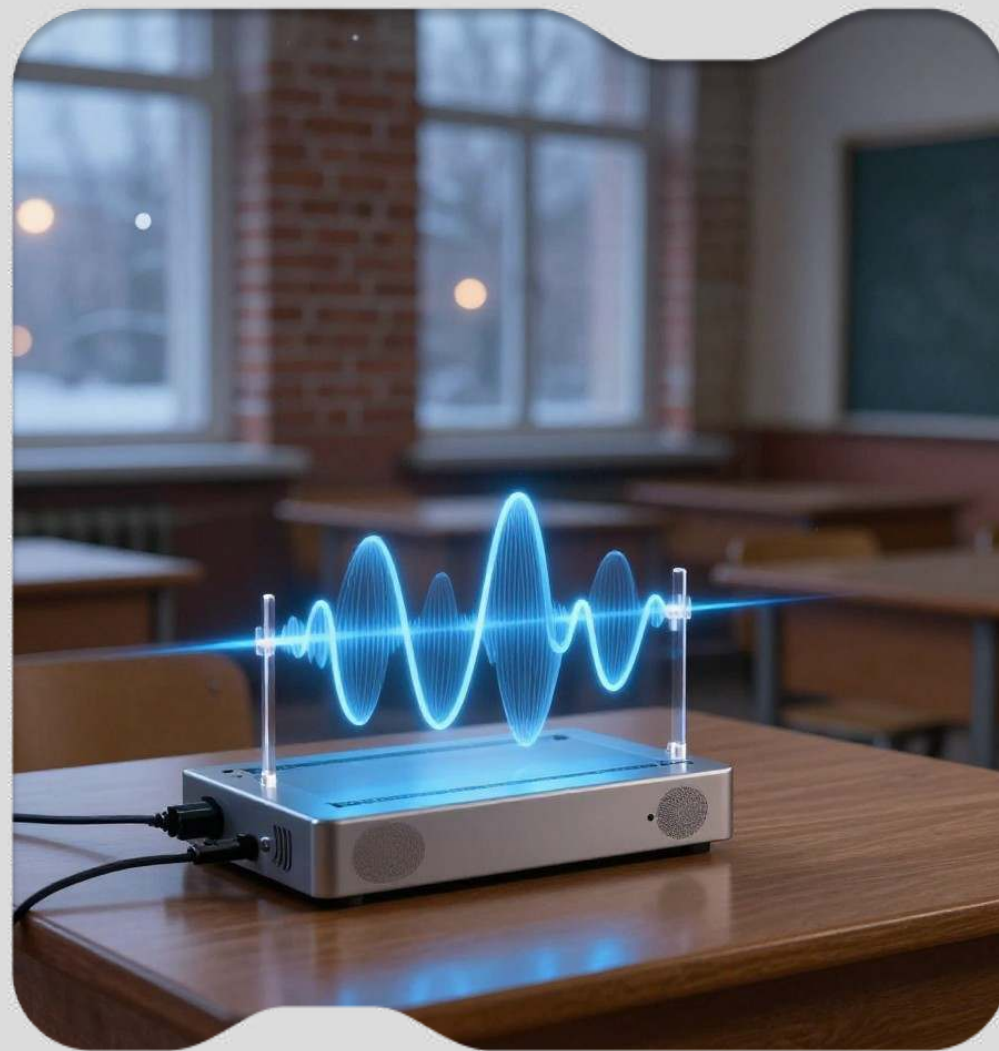
Физическая природа звука

● Звук возникает из механических колебаний среды и распространяется как волна. Мы воспринимаем не сам «движок» частиц, а изменения давления и скорости колебаний в воздухе, воде или твёрдых телах.

● Высота тона определяется частотой: чем она больше, тем тон выше. Громкость связана с амплитудой, а тембр формируется спектральным составом — тем, какие частоты присутствуют сильнее всего.

● Для человека слышимый диапазон обычно лежит примерно от 20 Гц до 20 кГц, но с возрастом верхняя граница часто уменьшается. Поэтому один и тот же сигнал может звучать по-разному для разных слушателей.

От звуковой волны к цифровому коду



Дискретизация по времени

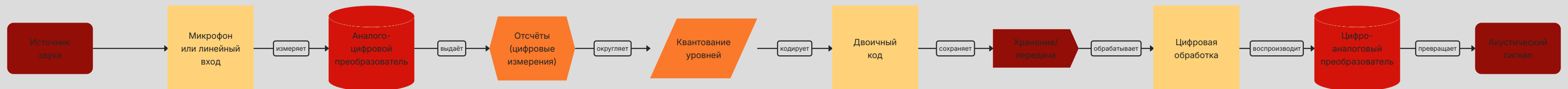
Аналоговый сигнал измеряют через равные промежутки времени: в каждый момент получают одну «выборку». Так синусоида превращается в ряд точек, которые приближают исходную волну. Важно выбрать шаг отсчёта.

Восстановление формы при воспроизведении

Дальше по точкам выполняют восстановление формы при воспроизведении: чем точнее измерения, тем меньше ошибок формы. На результат влияют точность измерения и согласованность параметров, чтобы сигнал не искажался.

Преобразование звука: от микрофона до акустики

Логическая цепочка этапов, влияющих и на качество, и на объём данных



Дискретизация и частота отсчетов

.01

Частота дискретизации показывает, сколько раз за одну секунду измеряют значение сигнала. Это задаёт «разрешение по времени» и напрямую влияет на то, удастся ли корректно отразить быстрые изменения звука.

• .02

По теореме Котельникова частота дискретизации должна быть не ниже удвоенной максимальной частоты полезного сигнала. Тогда снижается риск искажений при преобразовании «в цифру» и обратно.

• .03

На практике CD-аудио использует 44,1 кГц, а для видео и многих профессиональных задач часто применяют 48 кГц. В студийной работе нередко выбирают и более высокие значения, когда это оправдано требованиями.

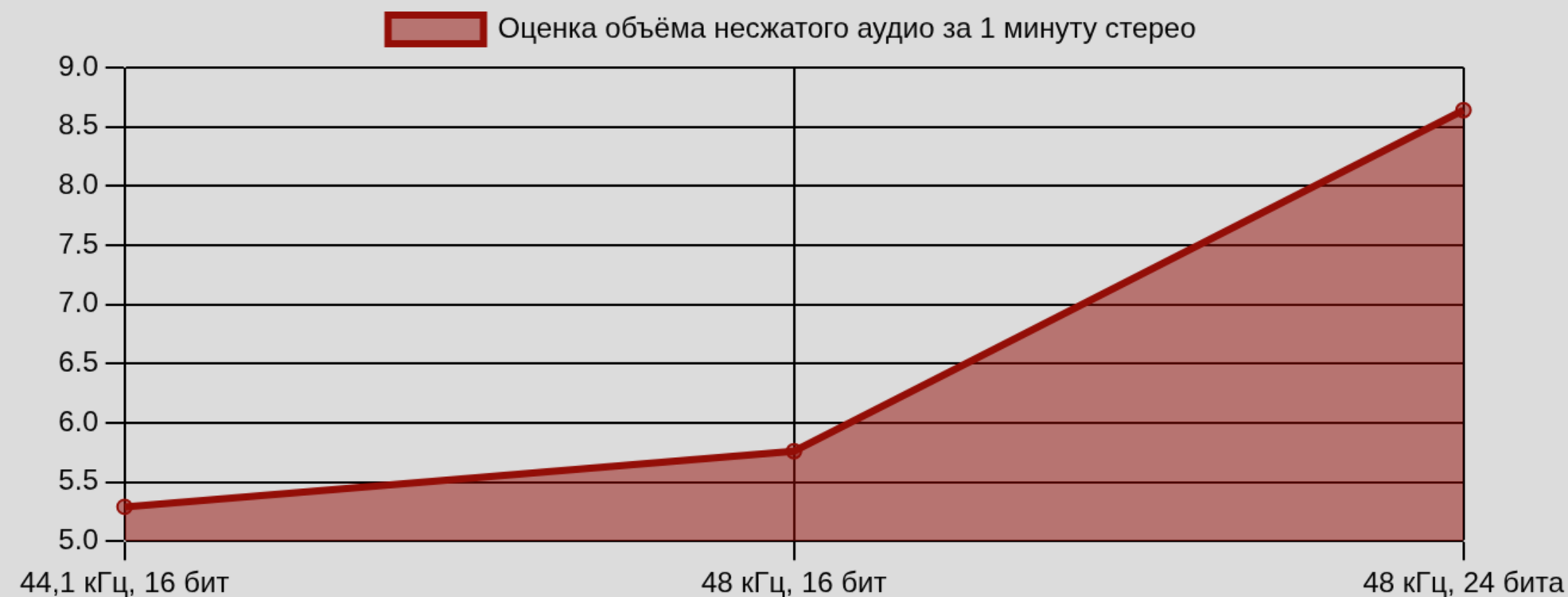
16 бит

При разрядности 16 бит число уровней становится достаточно большим, чтобы лучше сохранять нюансы тихих фрагментов записи.

Обобщённое описание из курса по представлению цифрового аудио (математическая связь разрядности и числа уровней: 2^N).

Разрядность задаёт число разрешённых уровней квантования: 16 бит соответствуют 65 536 уровням, уменьшая шум квантования и повышая точность передачи динамики.

Как параметры влияют на объём аудиоданных



Расчёт по формуле объёма несжатого PCM: частота × разрядность/8 × число каналов × длительность (для 1 минуты стерео), без учёта заголовков и округлений.

При одинаковой длительности и числе каналов рост частоты дискретизации или разрядности линейно увеличивает объём несжатого файла. На размер также влияет формула: частота × разрядность × каналы.

Увеличение 48 кГц до 44,1 кГц уже заметно меняет объём, а переход с 16 до 24 бит даёт самый большой прирост. Поэтому выбор разрядности критичен для «тяжёлых» файлов.

Сравнение популярных форматов аудио

Форматы отличаются тем, используют ли они сжатие, и как это влияет на качество и сценарии применения.

Если важна работа с исходником и точность — выбирают WAV/FLAC. Для повседневного обмена и потоков обычно выгоднее форматы с потерями, где качество оптимизируют под ограниченный размер и битрейт.

Формат	Тип сжатия	Качество	Типичное применение
WAV	Без сжатия	Максимальная точность	Монтаж и обработка
MP3	С потерями	Компромисс качество/размер	Музыка и обмен файлами
AAC	С потерями	Эффективнее при том же битрейте	Стриминг и мобильные платформы
FLAC	Без потерь	Сохранение оригинала	Архивирование и хранение

Технические характеристики форматов из общедоступной документации и справочных материалов спецификаций (общепринятое описание).

Стерео, моно и многоканальный звук



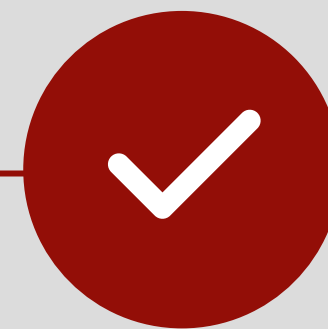
Моно (1 канал)

Моно использует один канал и часто подходит для речи. Оно упрощает запись и воспроизведение, а также помогает сохранять разборчивость без сложного пространственного эффекта для слушателя.



Стерео (2 канала)

Стерео разделяет сигнал на левый и правый каналы. Благодаря этому появляется ощущение расположения источников в пространстве, а музыка и эффекты звучат объёмнее, чем в моно.



Многоканальный звук (5.1/7.1)

Форматы 5.1 и 7.1 применяют в кино, играх и домашнем кинотеатре. Они дают точнее позиционирование эффектов вокруг слушателя, поэтому сцены воспринимаются реалистичнее.

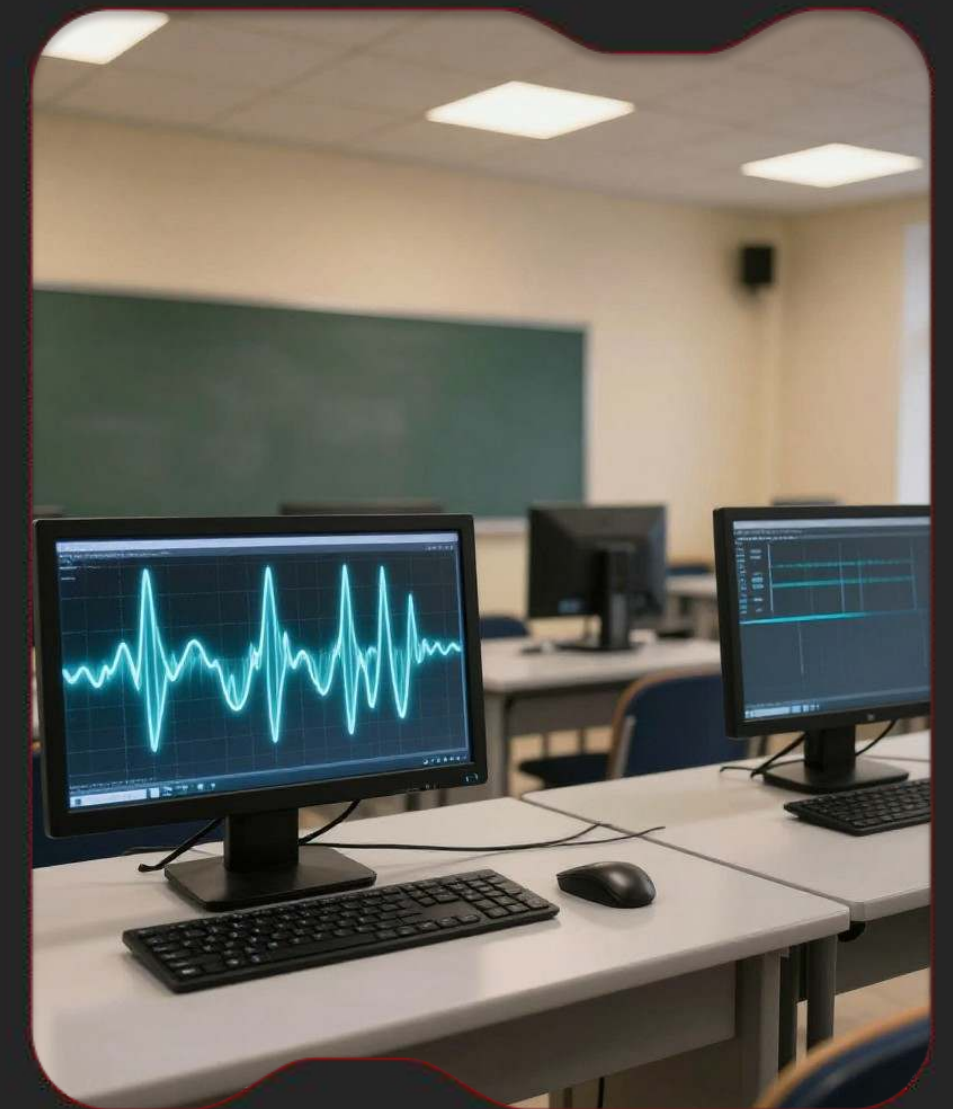
Волновое представление и спектр

Запись во времени: осциллограмма

Одну и ту же запись можно смотреть по-разному: как волну во времени или как набор частот. На осциллограмме видно, как меняются значения амплитуды, паузы и динамика сигнала по секундам.

Спектральный взгляд: спектрограмма и смысл частот

Если разложить звук по частотам, появляется спектр. Спектральный анализ показывает гармоники, шумовые компоненты и их распределение, а также помогает распознавать тембр, что особенно важно для речи, музыки и аудиопоиска.



Сжатие аудио: с потерями и без потерь

.01

Сжатие без потерь уменьшает размер файла за счет рационального хранения данных и не меняет исходное содержимое. Пример — FLAC и ALAC, которые сохраняют точность записи при последующем восстановлении.

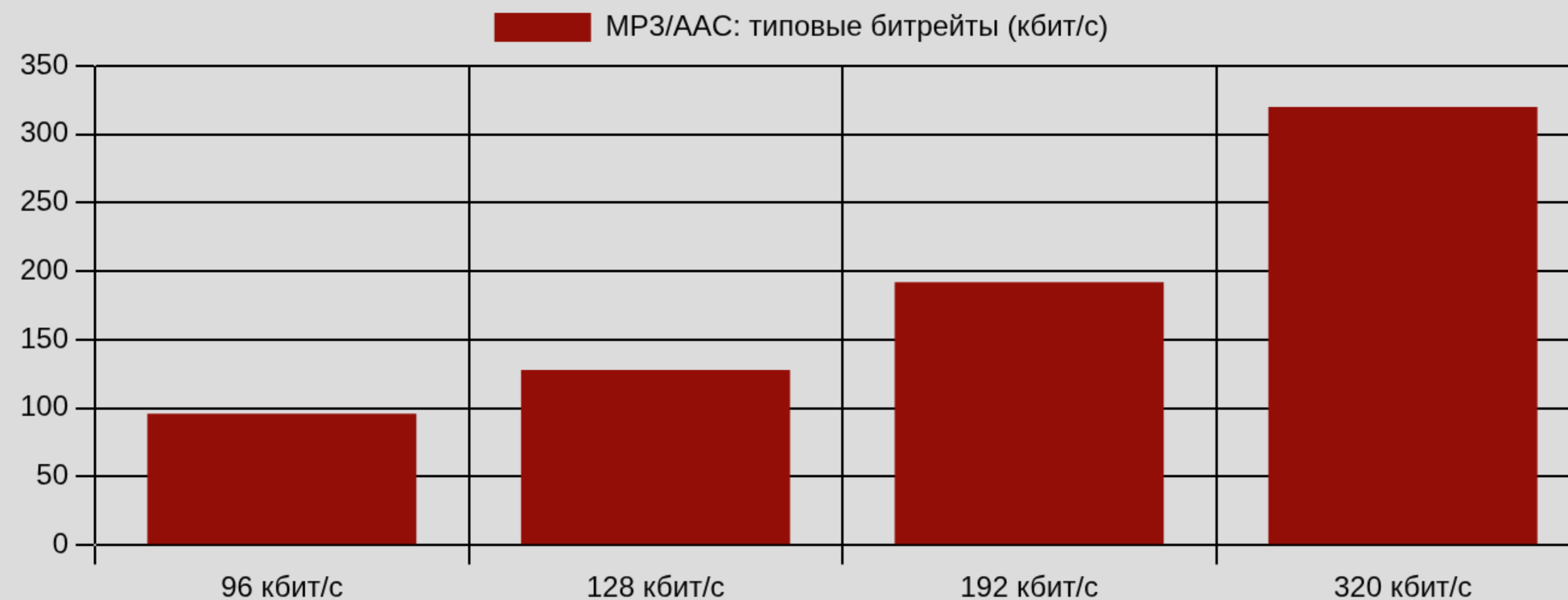
.02

Сжатие с потерями экономит место иначе: часть данных удаляется как менее заметная для слуха. В MP3 и AAC используют психоакустические эффекты, поэтому качество зависит от битрейта и характера материала.

.03

Битрейт напрямую связывает качество и объем: чем он выше, тем больше данных сохраняется и тем меньше артефактов. При низких значениях заметнее «сжатые» искажения и сужение воспринимаемого спектра.

Потери качества при разных битрейтах



Набор типовых режимов кодирования MP3/AAC; конкретные потери зависят от кодека и настроек (официальные значения приведены как примеры, точные измерения не представлены).

По мере снижения битрейта увеличивается вероятность артефактов сжатия и заметнее деградация высоких частот. При росте битрейта качество приближается к исходному, но растет размер файла.

Диаграмма показывает «ступени» параметров кодирования: более высокий битрейт дает больший запас данных, поэтому критичные для тембра детали теряются реже, хотя файл становится тяжелее.



≈10 МБ

**Для стереозаписи это
число показывает,
сколько места
занимает минута
несжатого аудио при
базовых параметрах.**

Расчетная оценка по формуле: $44,1 \text{ кГц} \times 16 \text{ бит} \times 2 \text{ канала} \times 60 \text{ с}$ (без сжатия).

Для несжатого формата размер определяется произведением частоты дискретизации, разрядности, числа каналов и длительности. Поэтому при увеличении любого параметра объем растет быстрее, чем кажется.

Обработка звука в цифровой среде

Типовой пайплайн обработки в аудиоредакторах и цифровых рабочих станциях



Передача звука по сети

.01

При потоковой передаче важны не только кодек и битрейт, но и задержка: слушатель ощущает ее как «отставание» звука от изображения или реакции. Стабильное соединение помогает избежать провалов и рассинхрона.

.02

Чтобы уменьшить нагрузку на канал, аудио обычно кодируют в сжатые форматы. Чем сильнее сжатие, тем меньше данных нужно пересылать, однако при этом возрастает риск заметных артефактов при низком битрейте.

.03

В видеосвязи применяют кодеки с низкой задержкой и настройками, которые автоматически адаптируют качество под доступную скорость интернета. Идея проста: поддерживать непрерывность важнее, чем всегда держать максимум качества.

.04

На практике систему дополняют буфером и алгоритмами восстановления: они сглаживают кратковременные перебои. Поэтому итоговое качество зависит от сети так же сильно, как и от параметров кодирования звука.

Параметры записи и их практический эффект

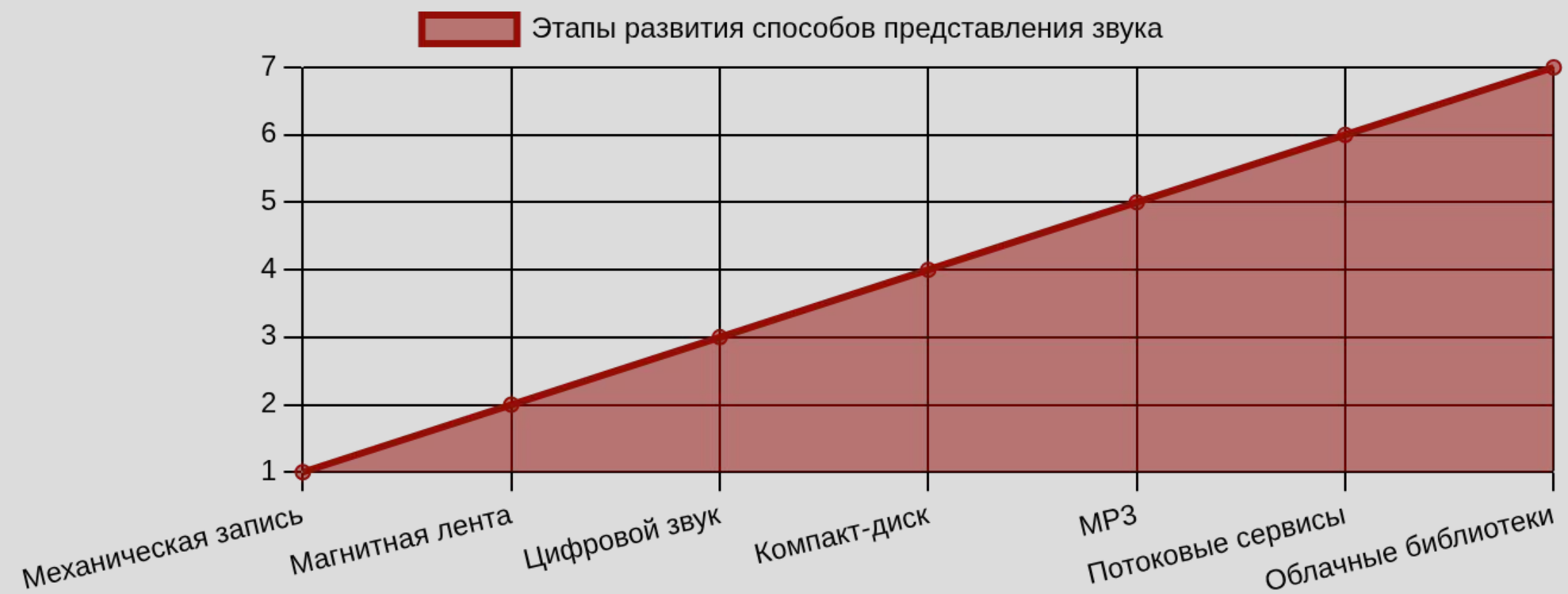
Параметр	Значение	Эффект
Частота	44,1 кГц / 48 кГц	44,1 кГц — стандарт музыки; 48 кГц — стандарт для видео и трансляций
Разрядность	16 бит / 24 бита	16 бит — базовое качество; 24 бита — студийный запас по динамическому диапазону
Каналы	моно / стерео / 5.1	моно — речь; стерео — музыка; 5.1 — кино

Сводные практические стандарты из области цифрового аудио и мультимедиа.

Выбор параметров определяет, какой именно «запас качества» получает запись: по частотному диапазону, по динамике и по каналному формату. Под задачу подбираются совместимые значения.

Большинство различий сводятся к тому, что частота влияет на детализацию по высоте, разрядность — на тишину и динамику, а число каналов — на пространственность.

Путь развития аудиотехнологий



Обобщенная историческая последовательность развития аудиосистем; конкретные даты и темпы внедрения зависят от рынка и компании.

Каждый новый этап уменьшал зависимость от конкретного носителя: звук становился все более универсальным и удобным для копирования. Параллельно снижались барьеры для прослушивания и обмена файлами.

Переход к цифровым и далее к сетевым моделям сделал звук доступным «по требованию». Технологический прогресс уменьшал трение между записью, хранением и распространением.

Ограничения и ошибки цифрового звука

.01

Если частота дискретизации слишком низкая, часть информации по частотам «складывается» и возникают искажения в виде алиасинга. В результате тембр становится менее естественным, а некоторые гармоники слышатся неправильно.

.03

Потери кодека добавляют артефакты сжатия, а шумы оборудования — дополнительные составляющие, не связанные с исходным звуком. В совокупности такие ошибки ухудшают разборчивость речи и детальность музыки.

.02

Малая разрядность усиливает шум квантования: тихие детали теряются, а динамика сжимается. Это особенно заметно в паузах и при тихом фоновом звучании, где мозг ожидает более гладкую картину.

.04

Эти эффекты проявляются даже без специальной «тренировки слуха»: при слишком низких параметрах качество падает заметно. Поэтому при выборе настроек важно балансировать точность, размер и требования к обработке.

Звук становится «ПОНЯТНЫМ» компьютеру числами

В компьютере звук представляют числами, полученными после дискретизации и квантования аналоговой волны. Качество определяется частотой, разрядностью, каналами и кодеком, а формат выбирают по задаче и ограничениям по объему и скорости.