
ДИСЦИПЛИНА: ОСНОВЫ ЗВУКОЗАПИСИ И ЦИФРОВАЯ ОБРАБОТКА АУДИО
ИНФОРМАЦИИ

Форматы аудиофайлов и публикация контента

Академическая лекция

WAV, MP3, AAC и требования платформ



ПЛАН ЛЕКЦИИ

Структура изучения форматов
аудиофайлов и их публикации



ОСНОВНЫЕ ТЕМЫ ЗАНЯТИЯ

- 01** Основы цифрового аудио
Сэмплирование, квантование, терминология (Sample Rate, Bit Depth)
- 02** Формат WAV (Lossless)
Технические спецификации, преимущества и недостатки
- 03** Форматы MP3 и AAC (Lossy)
Алгоритмы сжатия, CBR vs VBR, профили и области применения
- 04** Сравнение форматов
Сводные таблицы, визуализация соотношения "размер/качество"
- 05** Экспорт и требования платформ
Спецификации Spotify, Apple Music, YouTube, SoundCloud
- 06** Особенности публикации и итоги
Метаданные, контроль громкости (LUFS), Quality Control



- **Частота дискретизации (Sample Rate)**

Количество измерений амплитуды сигнала в секунду (Гц).
Определяет максимальную воспроизводимую частоту.

- **Глубина бит (Bit Depth)**

Количество бит для записи каждого сэмпла. Определяет динамический диапазон (разницу между тихим и громким).

- **Битрейт (kbps)**

Объем данных, передаваемый или обрабатываемый за одну секунду (килобит в секунду). Ключевой параметр lossy-сжатия.

- **Контейнер vs Кодек (PCM)**

Контейнер (WAV, MP4) хранит данные и метатеги. Кодек (PCM, AAC) определяет алгоритм кодирования аудио.

- **Lossless vs Lossy**

Lossless (без потерь) сохраняет 100% данных. Lossy (с потерями) удаляет неслышимые частоты для сжатия.

Визуальный глоссарий



Sample Rate

Шаг по оси X (время) → 44.1 kHz = 44,100 сэмплов/сек



Bit Depth

Шаг по оси Y (амплитуда) → 16-bit = 65,536 уровней громкости



Продвинутые концепции

Dithering: добавление слабого шума при снижении битности (24→16) для маскировки искажений квантования.

LUFS: единицы измерения воспринимаемой громкости, стандарт для стриминговых платформ (-14 LUFS).



Связь параметров

Битрейт (PCM) = Sample Rate × Bit Depth × Channels

Пример: 44.1 kHz × 16 bits × 2 (Stereo) = 1411 kbps (стандарт CD-Audio).



• Теорема Найквиста-Шеннона

Частота дискретизации должна быть минимум в два раза выше максимальной частоты сигнала (например, 44.1 кГц для аудио до 22 кГц).

• Aliasing (Алиасинг)

Искажения, возникающие при нарушении теоремы Найквиста. Частоты выше предела Найквиста отражаются в слышимый диапазон.

• Квантование и шум

Округление аналогового значения амплитуды до ближайшего дискретного уровня. Разница между реальным и округленным значением создает шум квантования.

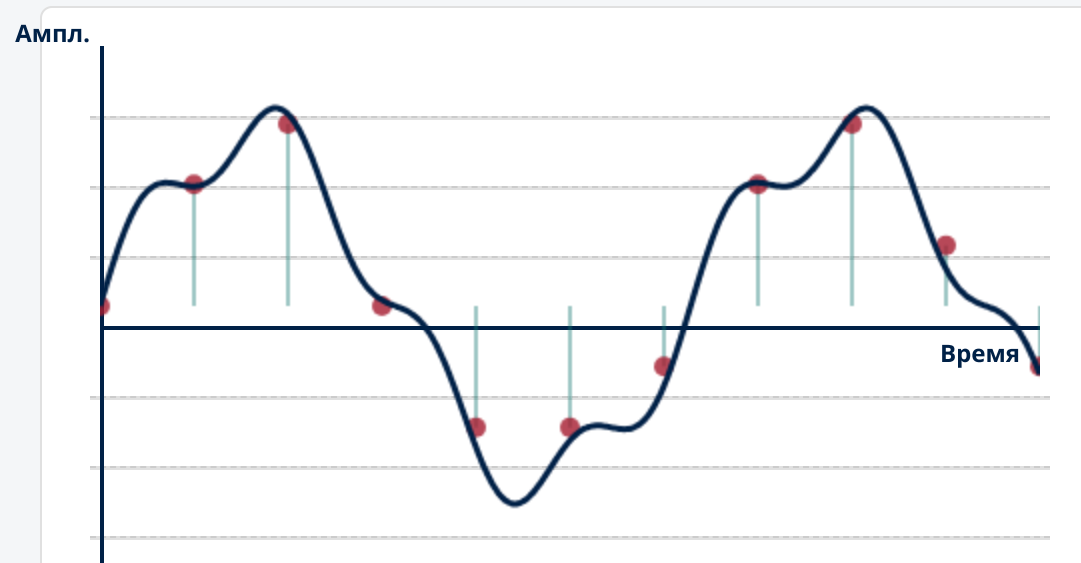
• Dithering

Добавление псевдослучайного шума низкого уровня перед снижением битности (например, 24 → 16 бит). Декоррелирует ошибку квантования, превращая гармонические искажения в белый шум.

Формула динамического диапазона:

Динамический диапазон $\approx 6.02 \times N + 1.76$ дБ (где N - битность)

Процесс оцифровки аналогового сигнала



- Непрерывный аналоговый сигнал (бесконечное разрешение)
- Сэмплирование (дискретизация по оси времени X)
- Квантование (округление по оси амплитуды Y)



РЕКОМЕНДУЕМЫЕ ПАРАМЕТРЫ

Сценарий	Sample Rate	Bit Depth
Музыка / Мастеринг	44.1 / 48 kHz	24-bit
Видео / Кино (Post)	48 / 96 kHz	24-bit
Полевая запись	48 / 96 / 192 kHz	32-bit float




Важно: Headroom и True Peak


При экспорте для стриминговых платформ (lossy кодирование) оставляйте запас по уровню **-1 dBTP (True Peak)**. Это предотвратит межсемпловый клиппинг (intersample clipping) при конвертации WAV в MP3/AAC алгоритмами платформы.

Расчет размера файла (WAV / PCM)

$$\text{Размер (Байт)} = \text{Sample Rate} \times (\text{Bit Depth} / 8) \times \text{Каналы} \times \text{Секунды}$$

Пример расчета (Трек: 3 минуты, Стерео)

 CD-качество (44.1 kHz, 16-bit)
 $44100 \times (16/8) \times 2 \times 180 \text{ сек} = 31,752,000 \text{ Байт}$
 $\approx 30.3 \text{ МБ}$

 Студийный мастер (48 kHz, 24-bit)
 $48000 \times (24/8) \times 2 \times 180 \text{ сек} = 51,840,000 \text{ Байт}$
 $\approx 49.4 \text{ МБ}$

Переход с 16-bit на 24-bit увеличивает размер файла на 50%.
Для lossy форматов (MP3/AAC) размер зависит только от битрейта.

WAV: общая характеристика



- **Контейнер формата**

Использует RIFF (Resource Interchange File Format) / WAVE структуру для организации данных.

- **Используемые кодеки**

Чаще всего несет некомпьютеризированный PCM (Pulse-Code Modulation). Расширение BWF применяется для профессионального вещания.

- **Формат без потерь (Lossless)**

Обеспечивает 100% сохранение исходного аудиосигнала без артефактов компрессии.

- **Стандарт индустрии**

Идеален для редактирования, сведения и является стандартом де-факто для финального мастеринга.

- **Метаданные (BWF/iXML)**

Broadcast Wave Format позволяет внедрять BEXT (Broadcast Extension) и iXML блоки для таймкодов и съемочной информации.

Структура файла WAV



Контейнер (RIFF/WAVE)

Внешняя оболочка, организующая данные в чанки (chunks).



Аудио данные (Кодек PCM)

Сырые аудиосэмплы (fmt chunk + data chunk). Lossless - без сжатия.



Метаданные (BEXT/iXML)

Дополнительные чанки с информацией: таймкод, имя сцены, трека и дубля.



Идеально для производства

Ключевое преимущество: WAV предоставляет прямой доступ к сэмплам, что делает его самым быстрым и точным форматом для нелинейного редактирования (NLE/DAW).



ОСНОВНЫЕ ПАРАМЕТРЫ ФОРМАТА

Частота дискретизации	От 8 kHz до 384 kHz
Битовая глубина (Bit Depth)	8, 16, 24-bit, 32-bit float
Количество каналов	Моно, Стерео до Multichannel
Порядок байтов	Little-endian



Совместимость с ПО

Формат WAV обладает максимальной совместимостью со всеми современными цифровыми звуковыми рабочими станциями (DAW) и системами нелинейного монтажа (NLE). Является отраслевым стандартом для обмена и архивирования аудиоданных.

Ограничения размера и RF64

Лимит формата RIFF: ~4 ГБ

Классический контейнер WAV использует 32-битные указатели размера файла, что физически ограничивает максимальный размер файла значением 4 Гигабайта.

Решение проблемы: формат RF64



Спецификация EBU (European Broadcasting Union)

RF64 использует 64-битные указатели размера, позволяя создавать файлы, превышающие ограничение в 4 ГБ.



Длительные записи и многоканал

Критически важно для многоканальных записей в высоком разрешении (напр. 96kHz/24-bit/5.1) или многочасовых сессий.

Большинство современных DAW автоматически переключаются в режим RF64.

WAV: преимущества и недостатки



ПРЕИМУЩЕСТВА



Максимальное качество

Формат без потерь (lossless) сохраняет 100% оригинальных аудиоданных без искажений и артефактов сжатия.



Универсальная совместимость

Является индустриальным стандартом. Поддерживается практически всеми DAW, плеерами и операционными системами.



Точность редактирования

Позволяет производить монтаж с точностью до сэмпла. Требует минимальных ресурсов процессора при воспроизведении.

НЕДОСТАТКИ



Большой размер файлов

Несжатые данные занимают значительно больше места на накопителе по сравнению с lossy-форматами (MP3, AAC).



Неудобство доставки

Высокий битрейт делает формат неподходящим для прямого интернет-стриминга из-за жестких требований к каналу связи.



Отсутствие экономии трафика

При передаче по сети, подкастинге или публикации на массовых платформах требует неоправданных затрат трафика.

★ Вывод: WAV — идеальный формат для создания, мастеринга и архивирования, но не для конечного потребления пользователями.



- **MPEG-1/2 Layer III**

Официальное название стандарта. Разработан институтом Фраунгофера.

- **Lossy формат**

Сжатие с потерями. Часть исходных данных удаляется для уменьшения размера файла.

- **Психоакустическая модель**

Основа сжатия: удаляются частоты, которые человеческое ухо не воспринимает из-за эффекта акустической маскировки.

- **Широкая совместимость**

Поддерживается практически всеми устройствами, операционными системами и автомобильными магнитолами.

- **Метаданные (ID3)**

Теги ID3v1/ID3v2 для хранения информации об артисте, альбоме, обложке.

- **Типичные битрейты**

От 128 kbps (минимально приемлемо) до 320 kbps (наивысшее качество).

Таймлайн и значение

Эволюция MP3

1993 год

Опубликован стандарт MPEG-1 Audio Layer III.

Конец 1990-х

Революция в цифровой музыке, файлообменные сети (Napster).

2017 год

Истечение срока действия патентов (формат стал свободным).



Историческое наследие

Формат, изменивший музыкальную индустрию. До сих пор де-факто стандарт для подкастов.



Широкая поддержка

Главное преимущество MP3 сегодня — это 100% совместимость с любым устройством воспроизведения, выпущенным за последние 25 лет.



- **Основы алгоритма**

Использует **MDCT** (модифицированное дискретное косинусное преобразование) и **квантование** для удаления избыточной и неслышимой информации. Для финального сжатия применяется **код Хаффмана**.

- **Методы оптимизации**

Joint Stereo: кодирование разницы между каналами вместо полного разделения. **Lowpass фильтры:** обрезка частот выше 16-18 kHz (зависит от битрейта).

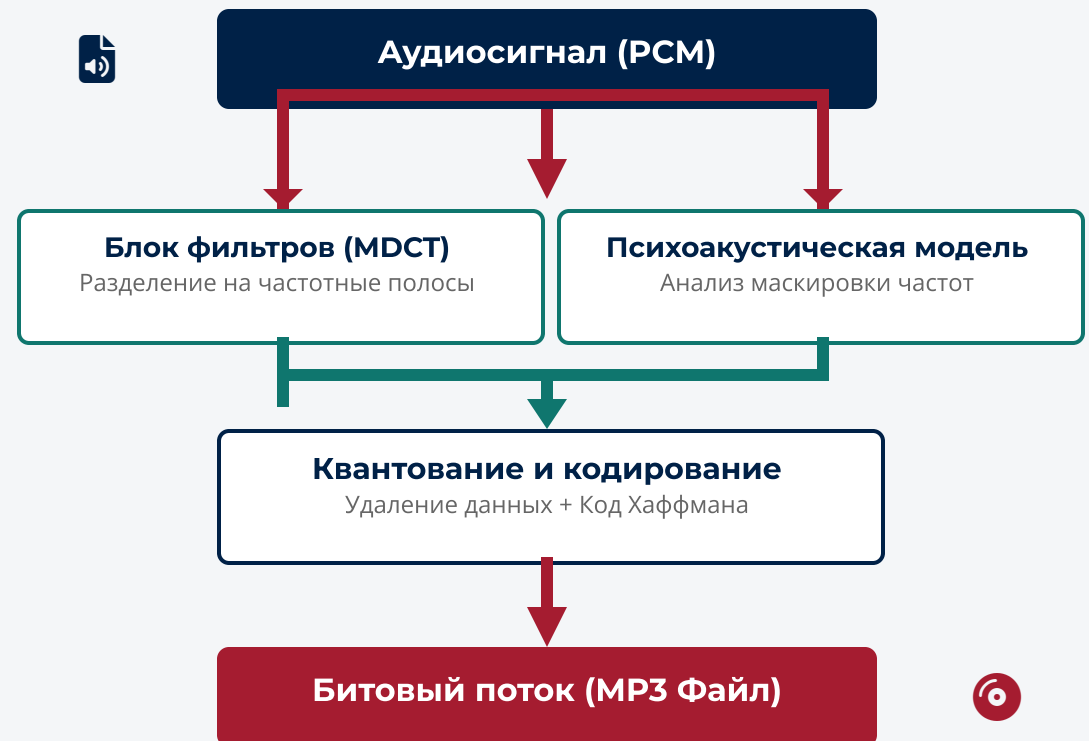
- **Артефакты сжатия**

Pre-echo: шум перед резкими транзиентами (ударами).
Зернистость: искажение высоких частот (тарелки, сибиланты) из-за грубого квантования.

- **Gapless воспроизведение**

MP3 кодирует аудио блоками по 1152 сэмпла, что добавляет "хвосты" тишины. Воспроизведение без пауз требует специальных тегов (LAME tag) для компенсации задержки.

Блок-схема кодера MP3





CBR (CONSTANT)



Постоянный битрейт

Выделяет одинаковое количество данных для каждой секунды аудио.



Предсказуемый размер

Идеально для потокового вещания (радио) и расчета размера файла.

VBR (VARIABLE)



Переменный битрейт

Динамически меняет битрейт в зависимости от сложности звука.



Лучшее качество/бит

Меньший размер файла при сохранении высокого качества звучания.

ABR (AVERAGE)



Целевой средний

Компромисс между CBR и VBR с заданным средним битрейтом.

Рекомендации по экспорту MP3:

- **320 kbps CBR:** Архивное хранение и максимальное качество lossy-мастера
- **V0 / V2 (VBR):** Повседневное прослушивание и оптимальный баланс размера/качества
- **96-128 kbps:** Разговорные форматы, подкасты, аудиокниги (предпочтительно моно)



- **Профили кодирования**

LC-AAC (Low Complexity), HE-AAC (High Efficiency с SBR), HE-AAC v2 (с Parametric Stereo). Адаптация под любые битрейты.

- **Качество vs MP3**

Обеспечивает лучшее качество звука, чем MP3, при одинаковом или меньшем битрейте. Стандарт де-факто для потери качества.

- **Поддерживаемые контейнеры**

Чаще всего встречается в контейнерах MP4 (видео/аудио), M4A (только аудио) или как чистый поток AAC.

- **Количество каналов**

Поддержка от моно (1 канал) до 48 каналов звука, что делает его пригодным для сложных систем пространственного звучания.

- **Частота дискретизации (SR)**

От 8 kHz (низкокачественная речь) до 96 kHz (high-resolution audio), перекрывая все типичные потребности.

Ключевые преимущества



Высокая эффективность

Меньший размер файла при сохранении прозрачности звучания.



Стандарт для стриминга

Используется Apple Music (256 kbps), YouTube и другими платформами.



Мобильные устройства

Аппаратная поддержка декодирования практически во всех смартфонах.



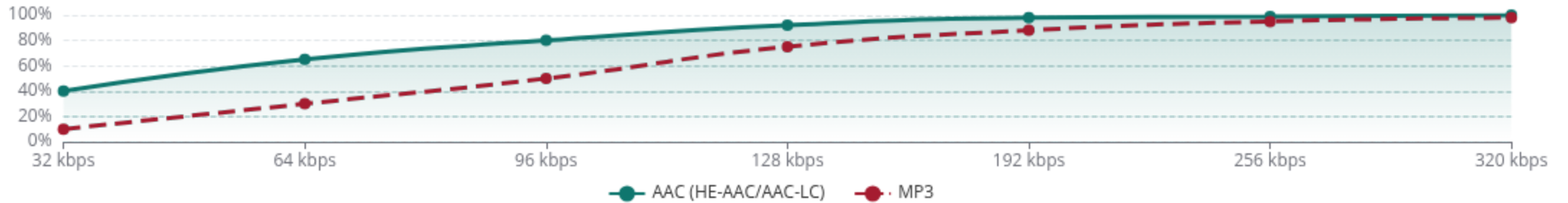
Технологическое превосходство

AAC использует более продвинутые алгоритмы обработки сигнала (например, MDCT с большим окном), что позволяет избегать типичных артефактов MP3 (pre-echo) и лучше передавать высокие частоты на низких битрейтах.

AAC: преимущества перед MP3



Сравнение качества аудио от битрейта (Субъективная оценка)



Лучшая прозрачность

Выше качество звучания на типичных битрейтах (128-256 kbps).



Высокая эффективность

Профиль HE-AAC обеспечивает приемлемое стерео при 24-64 kbps.



Стабильные ВЧ

Улучшенная обработка транзиентов, значительно меньше pre-echo.

Матрица рекомендаций: Битрейт и Применение

Битрейт (Стерео)	Оценка Качества	Целевое Применение
256 - 320 kbps	● Прозрачное (Excellent)	Музыкальный стриминг (Apple Music)
128 - 192 kbps	● Очень хорошее (Good)	YouTube видео, Качественные подкасты
96 kbps	● Хорошее (Fair)	Интернет-радио, Речевой контент
24-64 kbps (HE-AAC)	● Приемлемое (Poor)	DAB+ радио, Мобильные сети (низкий трафик)



Сравнение эффективности (Экономия 33%)

AAC-LC на 128 kbps субъективно звучит неотличимо от MP3 на 192 kbps. Это позволяет снизить размер файла на треть без ощутимой потери качества для слушателя.



- **Музыкальный стриминг**

Основной кодек для Apple Music (AAC 256 kbps, кодировщик Apple), обеспечивающий баланс между качеством и трафиком.

- **Видео-платформы**

Стандарт де-факто для YouTube (в контейнере MP4), Vimeo и большинства платформ пользовательского контента.

- **Мобильные устройства и Bluetooth**

Широко поддерживаемый кодек в экосистеме iOS и беспроводных наушниках, минимизирующий задержки и потери.

- **Цифровое радио и ОТТ**

Профили HE-AAC и HE-AAC v2 (DAB+) используются в цифровом вещании для максимизации качества на низких битрейтах.

- **Социальные сети**

Оптимизирован для коротких видео и историй, где важен небольшой размер файла при сохранении разборчивости речи.

Пайплайн публикации: от Мастера к Платформам



Исходник

WAV 24-bit / 48kHz



Транскодирование (AAC)



Apple Music
(256 kbps)



YouTube
(MP4/AAC)



Mobile / BT
(Кодеки)

Платформы обычно запрашивают lossless мастер (WAV) и сами кодируют в AAC для доставки конечным пользователям.

Сравнение форматов: сводная таблица



Формат	Тип	Параметры (SR/BD/BR)	Размер (3 мин стерео)	Качество	Метаданные	Совместимость
 WAV	Lossless	SR: 44.1 - 384 kHz BD: 16/24/32-float BR: ~1411+ kbps	~50 МБ	Отличное	BEXT iXML	Мастеринг, DAW (Стандарт индустрии)
 MP3	Lossy	SR: до 48 kHz BR: 128 - 320 kbps (CBR/VBR)	~7 МБ	Хорошее	ID3 (v1/v2)	Универсальная (Старые устройства/Архив)
 AAC	Lossy	SR: 8 - 96 kHz BR: 96 - 256 kbps (LC / HE-AAC)	~6 МБ	Отличное	MP4/M4A Atoms	Стриминг, Мобильные (Apple, YouTube)

★ Вывод: Для доставки контента (стриминга) предпочтительнее AAC за счет лучшего соотношения размер/качество.



Размер файла vs Качество



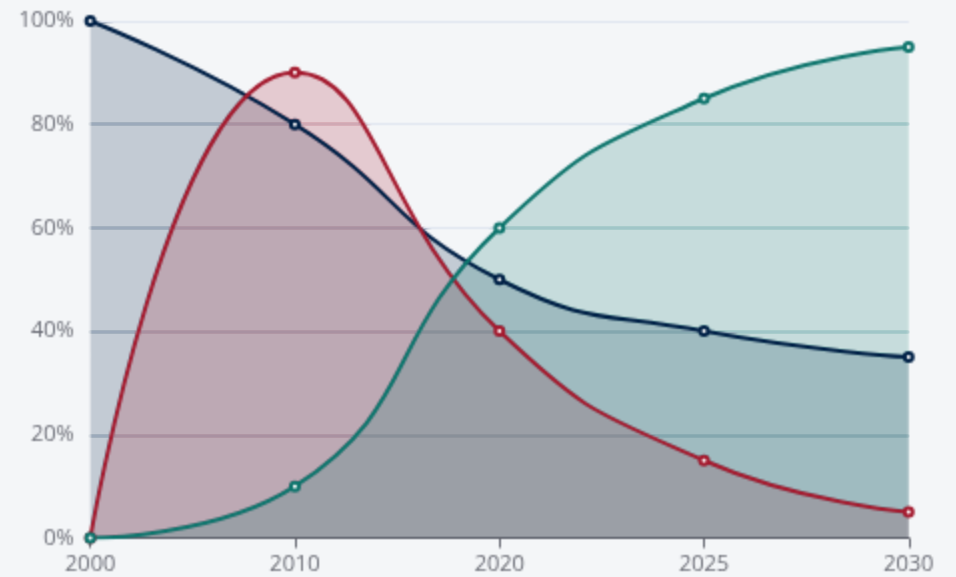
Модель трёх горизонтов (Аудиоформаты)

H1 (Горизонт 1): WAV - Текущая база (Архив/Мастеринг)

H2 (Горизонт 2): MP3 - Наследие (Снижается)

H3 (Горизонт 3): AAC/Opus - Рост (Стриминг/Мобайл)


Доля использования



○ H1: WAV ○ H2: MP3 ○ H3: AAC/Opus



ТЕХНИЧЕСКИЕ ПАРАМЕТРЫ

Платформа	Кодек	Loudness
 Spotify	Ogg Vorbis / AAC	~ -14 LUFS
 Apple Music	AAC 256 kbps	~ -16 LUFS



Apple Digital Masters

Специальный процесс контроля качества. Требуется 24-bit, 44.1-96 kHz файлы и строгий контроль межсэмповых пиков (intersample peaks) перед конвертацией в AAC.

Чек-лист экспорта мастера



Формат файла

WAV, 24-bit, 44.1 / 48 / 88.2 / 96 kHz. Никакого клиппинга.



Громкость (Loudness)

Таргет \approx -14 LUFS (интегрированная) для Spotify.



True Peak

Максимум \leq -1.0 dBTP для предотвращения клиппинга при кодировании.



Dithering

Обязательно применяйте Dither при редукции битности (например, с 32-bit float или 24-bit до 16-bit).

Золотое правило: Платформы сами применяют нормализацию и кодирование. Отправляйте максимально качественный WAV (24-bit) с запасом по True Peak, а не пережатый MP3.

Экспорт: YouTube, SoundCloud, подкасты



Платформа	Формат загрузки	Требования к аудио
 YouTube	Оригинал: WAV или FLAC Транскодирование: <ul style="list-style-type: none">• AAC (в контейнере MP4)• Opus (для некоторых видео)	Громкость: ≈ -14 LUFS Пик (True Peak): ≤ -1.0 dBTP <i>Превышение -14 LUFS приведет к автоматическому понижению громкости платформой.</i>
 SoundCloud	Оригинал: WAV / AIFF (16 или 24-bit) Стриминг: <ul style="list-style-type: none">• MP3 (128 kbps)• Opus (высокое качество для GO+)	Главное: Избежать клиппинга <i>Оставляйте запас (headroom) не менее -1 dBTP для компенсации артефактов при агрессивном транскодировании в MP3 128 kbps.</i>
 Подкасты	Мастер: WAV Дистрибуция (RSS): <ul style="list-style-type: none">• MP3 64-96 kbps (моно, голос)• MP3 128-192 kbps (стерео, музыка)	Громкость: -16 LUFS (Apple Podcasts) Пик: ≤ -1.0 dBTP Обложка: 1400×1400 - 3000×3000 px (JPEG/PNG)

★ Практика: Экспорт в 24-bit WAV с параметрами -14 LUFS и -1 dBTP покрывает требования большинства видео и музыкальных платформ.



• Контроль качества (QC)

Проверка пиков/межсэмплых пиков (ISP), фазовая коррекция, удаление DC offset, проверка тишины.

• Метаданные и идентификация

Тегирование ID3/BWF/iXML. Обязательно: Artist, Title. Коды: ISRC (трек) и UPC/EAN (релиз).

• Визуальное оформление (Арт)

Квадрат 3000×3000 px, цветовой профиль sRGB, формат JPG/PNG, размер файла < 10 МБ.

• Именованние файлов (Нейминг)

Шаблон: Artist_Title_Version_SR-BD

Пример: Band_SongName_Master_48k-24b.wav

• Экспорт версий и Громкость

Создание Мастера (WAV) и дистрибутива (AAC/MP3). Строгая проверка Loudness (≈ -14 LUFS) и True Peak (≤ -1 dBTP).

Pipeline: Мастер → Дистрибуция



1. Создание Мастера (WAV)

Экспорт в 24-bit/48kHz. Контроль громкости (-14 LUFS). Применение дизеринга при даунсэмплинге.



2. Проверка (QC) и Метаданные

Анализ на клиппинг (True Peak). Заполнение тегов ISRC/UPC. Проверка обложки (sRGB, 3000px).



3. Дистрибуция / Публишинг

Загрузка Lossless на стриминги/агрегаторы. Транскодирование платформой (AAC/Ogg) для доставки слушателям.

** Качественный исходный файл (WAV) гарантирует лучший результат после транскодирования платформой.*



• Производство и архив

Использовать WAV 24-bit (в некоторых случаях 32-bit float для полевой записи) для максимального качества.

• Дистрибуция и стриминг

AAC-LC 256 kbps для музыки; 96-128 kbps для речевого контента.

• Совместимость (Наследие)

MP3 320 kbps (архив) или V0/V2 для высокой эффективности и совместимости со старыми устройствами.

• Параметры экспорта

Целевая громкость: -14 LUFS, True Peak ≤ -1.0 dBTP. Доставлять на платформы в формате 24-bit WAV.

• Чек-лист публикации

Проверка метаданных, обложки (арт), Quality Control (QC), и контроль транскодинга.

Решения по сценарию



Студия / Мастеринг

WAV 24-bit / 48 kHz (Минимум) — Lossless, без сжатия.



Подкасты

MP3 96-128 kbps (Моно) или 192 kbps (Стерео).



Стриминг (Spotify, Apple)

Мастер в WAV, платформа сожмет в AAC 256 kbps / OGG.



Финальный чек-лист

1. Проверяйте уровни: -14 LUFS, True Peak ≤ -1.0 dBTP.
2. Всегда сохраняйте lossless-мастер (WAV) перед сжатием.
3. Тестируйте финальный файл на разных устройствах!