File formats, Bit Rate, Sample Rate, Sample Rate Conversion

This form explains the core technical concepts that determine audio quality and file size: **bit rate**, **sample rate**, **bit depth**, and **sample rate conversion (SRC)**. It describes how these parameters affect frequency response, dynamic range, and artifacts, and compares major **audio file formats** (uncompressed, lossless, and lossy). A comprehensive comparison table of **audio and video file formats** is provided for practical reference. New sections address **audible differences between lossy formats at different bitrates** and **practical workflows for converting between formats without losing quality**.

---

1. Introduction

When audio is recorded digitally, an analog waveform is converted into a series of discrete numbers. Three key parameters control this process:

Parameter	What It Controls	Unit
Sample rate	Frequency range (how high a frequency can be captured)	Hertz (Hz) or kHz
Bit depth	Dynamic range (precision of each sample)	bits
Bit rate	Data volume per second (quality + file size)	kbps or Mbps

2. Sample Rate

2.1 Definition

Sample rate is the number of samples taken per second when digitizing audio. It determines the **maximum frequency** that can be accurately reproduced.

2.2 Nyquist–Shannon Sampling Theorem

The maximum reproducible frequency is exactly **half the sample rate**, called the **Nyquist frequency**:

$$f_{\text{max}} = \frac{f_s}{2}$$

Where f_s is the sample rate.

Sample Rate	Nyquist Frequency	Typical Use
44.1 kHz	22.05 kHz	CD audio, most consumer music
48 kHz	24 kHz	Professional video, DVD
96 kHz	48 kHz	High-resolution audio, studio work
192 kHz	96 kHz	Ultra-high-definition audiophile recordings

2.3 Practical Implications

- **44.1 kHz** extends just beyond human hearing (~20 kHz), sufficient for professionally released music.
- Higher sample rates (96 kHz, 192 kHz) capture more detail but increase file size significantly. Many engineers find no perceivable quality improvement above 48 kHz for final delivery.

---

3. Bit Depth

3.1 Definition

Bit depth is the number of bits used to represent each audio sample. It controls **dynamic range** (difference between loudest and quietest sounds) and amplitude precision.

3.2 Dynamic Range Formula

$$\text{Dynamic Range} \approx 6.02 \times \text{bit depth} + 1.76\ \text{dB}$$

Bit Depth	Levels	Dynamic Range	Typical Use
16-bit	65,536	~96 dB	CD audio, streaming deliverables

Bit Depth	Levels	Dynamic Range	Typical Use
24-bit	16,777,216	~144 dB	Recording/mixing/mastering (gold standard)
32-bit float	~4.3 billion	~1528 dB (theoretical)	Internal DAW processing, prevents clipping

- **24-bit** is recommended for recording and mixing due to lower noise floor and higher dynamic range.
- **32-bit float** offers virtually unlimited headroom during processing.

---

4. Bit Rate

4.1 Definition

Bit rate is the amount of digital audio data processed per second, measured in **kilobits per second (kbps)**. It represents audio quality and file size.

4.2 Calculation for Uncompressed Audio

For PCM (uncompressed) audio:

$$\text{Bit Rate} = \text{Bit Depth} \times \text{Sample Rate} \times \text{Channels}$$

Example: CD-quality stereo audio (16-bit, 44.1 kHz, 2 channels):

$$\text{Bit Rate} = 16 \times 44{,}100 \times 2 = 1{,}411{,}200 \text{ bps} = 1{,}411 \text{ kbps}$$

4.3 Bit Rate vs. Bit Depth vs. Sample Rate

Concept	Controls	Measured In	Primary Impact
Bit rate	Data volume per second	kbps	Quality + file size (especially for compressed audio)
Bit depth	Amplitude precision per sample	bits	Dynamic range, noise floor
Sample rate	Time resolution (samples per second)	Hz/kHz	Frequency range

- **Higher bit rate** = better quality but larger file (especially for lossy formats like MP3).
- **Uncompressed files** (WAV, AIFF) have the highest bit rates; **lossy compressed files** (MP3, AAC) have the lowest.

5. Sample Rate Conversion (SRC)

5.1 Definition

Sample rate conversion alters the number of digital audio samples per second to match a different desired rate. It changes the **time-domain resolution** of the signal.

Example: Converting from 352.8 kHz (mixing resolution) to 44.1 kHz (CD delivery).

5.2 Two Stages of Downsampling

1. **Low-pass filtering**: Removes high-frequency components above the new Nyquist frequency to prevent **aliasing**.
2. **Decimation (downsampling)**: Samples are discarded to achieve the target sample rate.

5.3 Aliasing and the Nyquist Frequency

Aliasing is a distortion phenomenon where high-frequency content **folds back** into lower frequencies when a signal is downsampled incorrectly.

- Frequencies above the **Nyquist frequency** (half the sample rate) cause aliasing.
- **Anti-aliasing low-pass filters** must be applied before downsampling to remove frequencies above Nyquist.

5.4 Filter Types and Artifacts

Linear-Phase Filters

- Maintain consistent phase across all frequencies → **no phase distortion**
- Introduce **pre-ringing** (echo before transients) and **post-ringing**
- Can cause **transient smearing** (blurred attack on drums, plucked strings)

Minimum-Phase Filters

- Introduce **phase distortion** and more **post-ringing**
- **No pre-ringing**
- Preferable for transient-rich content; more "analogue" feel

5.5 Filter Steepness Trade-off

- **Steeper filters** preserve more high-frequency content but cause **more ringing**
- **Gentler filters** reduce ringing but attenuate high frequencies earlier
- Infinite steepness would mean infinite ringing (physically impossible)

5.6 Quality Levels for SRC

Quality	Speed	Artifacts	Use Case
Fast	Quick	Higher aliasing, more distortion	Preview, rough drafts
Best	Slow	Minimal artifacts	Final masters, distribution

5.7 When SRC Matters

- Poor SRC can introduce **audible artifacts** (aliasing, transient smearing, phase shifts)
- Well-implemented SRC with proper filtering is **reversible** and minimally destructive
- Always record and mix at high resolution (e.g., 96 kHz/24-bit), then convert once at the end for delivery

---

6. Audio File Formats

6.1 Three Categories

Category	Compression	Quality	File Size	Examples
Uncompressed	None	100% accurate	Largest	WAV, AIFF, PCM, DSD
Lossless	Compressed (ZIP-like)	Identical to original	~50–60% of original	FLAC, ALAC
Lossy	Compressed (data removed)	Some quality loss	~10% of original	MP3, AAC, WMA, OGG

Key details:

- **WAV**: Default on Windows; compatible with all players; very large files
- **AIFF**: Apple's equivalent of WAV; supports metadata tags (WAV does not)
- **FLAC**: Open-source; 9 compression levels; no Apple Music/iTunes support
- **ALAC**: Apple's lossless; ~60% of original size; limited hardware support

- **MP3**: Highest bitrate 320 kbps; 128 kbps sounds "off" to most listeners
- **AAC**: Better quality than MP3 at same bitrate; wider frequency range (up to 96 kHz)
- **OGG**: Smallest file format; used by Spotify; limited hardware compatibility

6.3 Format Recommendations

Purpose	Recommended Format
Recording/mixing/mastering	WAV or AIFF (uncompressed, 24-bit/48 kHz or higher)
HD streaming/archiving	FLAC or ALAC (lossless)
General music library/portable	AAC or MP3 320 kbps (lossy)
Audiophile high-res	FLAC, ALAC, or DSD

---

7. Audible Differences Between Lossy Formats at Different Bitrates *(NEW)*

7.1 What Listeners Actually Hear at Different Bitrates

Technical paper mentioned bitrates but didn't explain what listeners actually hear. This section addresses audible artifacts at each bitrate level.

Bitrate	What You Hear	Artifacts Present	Recommendation
64 kbps	Noticeably dull/muffled; high frequencies cut off around 11 kHz	Aggressive high-frequency roll-off; severe smearing; "telephone quality"	Avoid for music; OK for podcasts/speech only
128 kbps	Acceptable for casual listening; muddy highs; pre-echo on transients	High-frequency roll-off near 16 kHz; pre-echo (ghost of loud transients before actual sound); hi-hat smearing	Minimum acceptable for music; not for critical listening
160 kbps	Good quality; minor artifacts only on complex transients	Slight high-frequency muddiness; occasional pre-echo	Good balance for most listeners

Bitrate	What You Hear	Artifacts Present	Recommendation
192 kbps	Very good quality; nearly transparent for most listeners	Minimal artifacts; subtle high-frequency loss	Sweet spot for mobile/commuting
256 kbps	Excellent quality; transparent for most equipment	Very subtle artifacts only on complex material	Recommended for casual audiophiles
320 kbps	Indistinguishable from CD for most listeners in real-world use	Nearly imperceptible differences; residual ~-30 dBFS (3% of original amplitude)	Best lossy quality; transparent for most

7.2 Key Findings from Blind Listening Tests

The "Great MP3 Experiment" (Coding Horror):

- **128 kbps CBR**: Clearly lower quality with extremely high statistical confidence
- **192 kbps VBR**: Optimal "aural bang for the byte" — even dogs can't hear the difference from CD
- **320 kbps CBR**: Nobody can hear the difference from CD in controlled tests

Null Test Results (WAV vs. 320 kbps MP3):

- Residual signal: True peak **-47.76 dBFS**, RMS **-64.29 dBFS**
- This is at or below the noise floor of many playback systems
- Average residual: **-30 dBFS** (~3% of original signal amplitude)
- **97% of the removed data is stereo width information**, not mono content

7.3 Which Instruments/Frequencies Are Most Affected

Lossy compression uses **psychoacoustic masking** to remove data:

Artifact Type	What It Affects	Most Noticeable On
High-frequency roll-off	Treble above encoder cutoff	Cymbals, hi-hats, string harmonics, sibilance ("s" sounds)
Pre-echo	Echo before loud transients	Drum hits, plucked strings, piano attacks

Artifact Type	What It Affects	Most Noticeable On
Transient smearing	Softened attack on percussive sounds	Hi-hats, snare drums, castanets, pizzicato strings
Stereo width reduction	Spatial information between channels	Wide-panned instruments, reverb tails, orchestral recordings
Quantization noise	Background noise shaped by masking	Quiet passages, reverb tails, sustained notes

Hi-hats are the classic test case: At 128 kbps, hi-hat "smearing" is clearly audible (missing spectral values). Quality improves at 192 kbps but high frequencies remain muddy. At 256 kbps, artifacts are not audible.

7.4 When Lossy Compression Becomes Audible

Listener Type	Equipment	Audibility Threshold
Casual listener	Phone speakers, cheap earbuds	~192 kbps (can't hear difference above this)
Enthusiast	Good headphones, consumer speakers	~256 kbps (subtle differences above this)
Audiophile	High-end studio monitors, reference headphones	320 kbps vs. lossless is measurable but often imperceptible
Trained professional	Reference monitors, quiet room, analytical listening	Can detect 320 kbps vs. WAV in A/B tests, but not in real-world use

Critical factors affecting audibility:

- **Age and hearing damage**: High-frequency hearing loss makes artifacts harder to hear
- **Playback equipment**: Boosted treble makes encoding errors more visible; attenuated treble masks them
- **Listening environment**: Noisy environments (cars, gyms) mask compression artifacts
- **Volume level**: Lower volumes can enhance ability to detect subtle details

7.5 Format Quality at Same Bitrate

Not all formats are equal at the same bitrate:

Format	Quality at 128 kbps	Quality at 192 kbps	Notes
MP3	Noticeable artifacts	Very good	Older encoder; needs 256+ kbps for transparency
AAC	Good (muddy highs)	Excellent	Better than MP3 at same bitrate; 160 kbps = transparent
Opus	Audible errors	Transparent	Best quality at low bitrates; YouTube standard at 130 kbps
Vorbis	Good	Excellent	Spotify uses this; similar to AAC

Opus breakthrough: 128 kbps Opus encodes "the wildest transients without audible distortion" — equivalent to MP3 at ~192-256 kbps.

7.6 Practical Recommendations

Use Case	Recommended Bitrate	Format
Podcasts/audiobooks (voice only)	64–96 kbps	AAC or MP3
Casual music listening (commute, gym)	192 kbps	MP3 or AAC
Critical music listening at home	256–320 kbps	AAC or MP3
Audiophile collection	320 kbps or lossless	FLAC/ALAC preferred
Archiving/mastering	Never use lossy	WAV/AIFF/FLAC only

Key takeaway: You can't restore quality that's already compressed away. Always archive in lossless format, then create lossy copies for distribution.

8.3 Step-by-Step Conversion Workflows

Workflow A: High-Quality Master Conversion (WAV → MP3/AAC)

1. **Keep original as backup**: Never overwrite your master WAV file
2. **Open in converter**: Use iZotope RX, r8brain, or SoX
3. **Set quality to "Best"**: Never use "Fast" for final masters
4. **Choose encoder settings**:
    - MP3: LAME 3.100+, 320 kbps CBR or V0 VBR
    - AAC: iTunes AAC LC or FFmpeg AAC, 256 kbps
5. **Enable dither** if converting from 24-bit to 16-bit
6. **Verify conversion**:
    - Check spectral analysis (should match original up to Nyquist)
    - Perform null test if possible (residual should be -30 dBFS or lower for 320 kbps)
7. **Listen to critical sections**: Focus on transients (drums) and high frequencies (cymbals)

Workflow B: Batch Conversion for Library (Large Collections)

1. **Use SoX command line** or Foobar2000 batch converter
2. **Set 保守 quality**: Use high-quality settings (don't sacrifice quality for speed in batch)
3. **Convert to VBR (variable bitrate)**:
    - MP3: V0 (highest quality VBR, ~245 kbps average)
    - AAC: 256 kbps VBR
4. **Preserve metadata**: Ensure ID3 tags transfer correctly
5. **Sample rate**: Convert to 44.1 kHz for music, 48 kHz for video/audio

Workflow C: Sample Rate Conversion Only (e.g., 96 kHz → 44.1 kHz)

1. **Use professional SRC tool** (iZotope RX or r8brain)
2. **Select "Best" quality**: Critical for sample rate changes
3. **Apply only once**: Never convert sample rate multiple times
4. **Verify with spectrum analyzer**: Ensure no aliasing artifacts above 22.05 kHz
5. **Check transient response**: Compare drum hits before/after

8.4 How to Verify Conversion Quality

Method	**What to Check**	**Pass Criteria**
Spectral analysis	Frequency content up to Nyquist	No artificial roll-off below expected cutoff

Method	What to Check	Pass Criteria
Waveform comparison	Transient shape, peak amplitude	Transients preserved (not softened)
Null test	Subtract converted from original	Residual ≤ -30 dBFS for 320 kbps MP3
A/B listening test	Switch between files repeatedly	Cannot reliably identify converted file in blind test
Metadata inspection	Sample rate, bit depth, channels	Matches intended output parameters

Spectral analysis tips:

- 320 kbps MP3 should show nearly identical spectral detail to WAV
- Slight roll-off near upper hearing threshold (18–20 kHz) is normal
- 128 kbps MP3 shows clear cutoff around 16 kHz
- 64 kbps MP3 shows aggressive cutoff around 11 kHz

8.5 Batch Conversion Best Practices

1. **Convert from original masters only**: Never convert already-converted files (generation loss)
2. **Use consistent settings**: Same encoder, bitrate, sample rate across entire batch
3. **Preserve metadata**: ID3 tags, album art, track numbers
4. **Organize output**: Create separate folders by bitrate/format
5. **Verify random samples**: Check 5–10% of batch with spectral analysis and listening

8.6 Common Conversion Mistakes to Avoid

Mistake	Consequence	Solution
Converting sample rate multiple times	Cumulative SRC artifacts, audible degradation	Convert once at end using high-quality SRC
Using "Fast" quality setting	Higher aliasing, more distortion, pre-echo	Always use "Best" quality for masters

Mistake	Consequence	Solution
Converting lossy → lossy (e.g., MP3 → AAC)	Compound quality loss, new artifacts	Always convert from original lossless master
Using 128 kbps for masters	Audible quality loss, unprofessional	Use 320 kbps or lossless for distribution
Not enabling dither (24→16 bit)	Quantization distortion, harsh noise floor	Enable triangular dither when reducing bit depth
Ignoring encoder version	Older encoders (iTunes 1990s) produce worse quality	Use modern encoders (LAME 3.100+, iTunes AAC)

8.7 Metadata Preservation During Conversion

What metadata survives conversion:

Format	Metadata Types Preserved	Notes
WAV → FLAC	Artist, album, title, year, genre, artwork	Use FLAC with Vorbis comments
WAV → AIFF	Full metadata support	AIFF supports embedded artwork
MP3 (ID3 v2.3)	Artist, album, title, year, genre, artwork	Most compatible version
AAC/M4A	Artist, album, title, year, genre, artwork	Uses iTunes metadata format
FLAC → MP3	Transfers with most converters	Verify ID3 tags after conversion

9. Video File Formats (Containers)

Video files are **containers** that hold video, audio, and subtitle streams encoded with various **codecs**.

9.1 Video Container Comparison

Format	Extension	Developer	Primary Use	Supported Codecs	Browser Support
MP4	.mp4	MPEG Group	Streaming/distribution	H.264, H.265, AAC	Universal
MOV	.mov	Apple	Apple/pro video editing	H.264, H.265, ProRes, AAC	Safari/limited
AVI	.avi	Microsoft	Legacy PC video	Various (MPEG, DivX, etc.)	Limited
MKV	.mkv	Open Source (Matroska)	HD video with multiple tracks	H.264, H.265, VP9, AAC, Opus	Limited (plugin)
WebM	.webm	Google	Web video (HTML5)	VP8, VP9, H.264, AAC, Opus	Universal (HTML5)
FLV	.flv	Adobe	Flash video/web	Sorenson, H.263, MP3	Limited (Flash)
WMV	.wmv	Microsoft	Windows media	WMV, VC-1, WMA	Limited
MPEG	.mpeg/.mpg	MPEG Group	DVD/TV broadcast	MPEG-1, MPEG-2	Limited

Format	Extension	Developer	Primary Use	Supported Codecs	Browser Support
TS	.ts	Broadcast	Broadcast/HDTV	H.264, H.265, MPEG-2	Limited
MXF	.mxf	SMPTE	Professional broadcast/post	ProRes, DNxHD, XAVC	None

Key details:

- **MP4**: Widely used for streaming; supports 3D graphics and interactivity
- **MOV**: Apple's QuickTime framework; high quality and flexible for professional editing
- **MKV**: Open-source; unlimited video/audio/subtitle tracks in one file
- **AVI**: Traditional Microsoft format; large file sizes
- **WebM**: Web-optimized, open-source; HTML5 video standard [table]

---

10. Practical Guidelines for Audio Clients

10.1 Recording and Mixing

- Use **24-bit/48 kHz** or **24-bit/96 kHz** WAV/AIFF for recording/mixing
- Keep **sample rate constant** throughout the session to avoid SRC artifacts
- Use **32-bit float** internally in your DAW to prevent clipping

10.2 Mastering and Distribution

- Master at the **highest resolution** you recorded (e.g., 96 kHz/24-bit)
- Perform **sample rate conversion once** at the end, using high-quality (Best) settings
- For streaming: deliver **WAV 24-bit/48 kHz** or **FLAC** if the platform accepts lossless
- For CD: **WAV/AIFF 16-bit/44.1 kHz**

10.3 Archiving

- Archive masters as **uncompressed WAV/AIFF** (largest, most compatible)
- Use **FLAC** for space-efficient lossless archival (not on Apple Music)

10.4 Format Selection by Use Case

Purpose	Bit Rate	Sample Rate	Bit Depth	Format
Podcasts	64–96 kbps	44.1 kHz	16-bit	AAC or MP3
Casual music (mobile)	192 kbps	44.1 kHz	16-bit	MP3 V0 or AAC 256 kbps
Critical listening	256–320 kbps	44.1 kHz	16-bit	AAC 256 kbps or MP3 320 kbps
Audiophile library	Lossless	44.1–96 kHz	16–24 bit	FLAC or ALAC
Production/mastering	1,411+ kbps	48–96 kHz	24-bit	WAV or AIFF
Archiving	1,411+ kbps	Original	Original	WAV/AIFF or FLAC

10.5 Avoiding Common Mistakes

Mistake	Consequence	Solution
Converting sample rate multiple times	Cumulative SRC artifacts	Convert once at the end
Using 128 kbps MP3 for masters	Audible quality loss	Use 320 kbps or lossless
Mixing at 44.1 kHz then upconverting	Unnecessary SRC	Record/mix at delivery rate or higher
Using lossy formats for editing	Generation loss	Edit in WAV/AIFF, export lossy only at the end

Mistake	Consequence	Solution
Converting lossy → lossy	Compound quality loss	Always convert from original lossless master
Using "Fast" SRC quality	Audible artifacts	Use "Best" quality for masters

---

11. Conclusion

Understanding bit rate, sample rate, bit depth, and sample rate conversion is essential for maintaining audio quality throughout the production chain.

Key takeaways:

- **Sample rate** determines frequency range (44.1 kHz = CD quality)
- **Bit depth** determines dynamic range (24-bit = professional standard)
- **Bit rate** determines data volume and quality (higher = better but larger)
- **Sample rate conversion** can introduce artifacts (aliasing, ringing) if done poorly
- **Lossy compression artifacts** are audible at 64–128 kbps but minimal at 320 kbps
- **320 kbps MP3** is nearly indistinguishable from WAV for most listeners in real-world use
- Use **professional SRC tools** (iZotope RX, r8brain, SoX) with "Best" quality settings
- Use uncompressed formats (WAV/AIFF) for production; lossless (FLAC/ALAC) for archiving; lossy (AAC/MP3 320 kbps) for distribution.

Choosing the right format, bitrate, and conversion workflow at each stage ensures optimal audio quality while managing file size effectively.

References

Audioengine, "What do Bit Rate, Bit Depth, and Sample Rate mean?" (2024)
SwellAI, "How to Change Sampling Rate in Audacity"
Mastering The Mix, "Audio File Formats Explained" (2021)
Adobe, "Understanding audio bitrate and audio quality" (2025)
Adobe, "The best audio format types for audiophiles" (2025)
TRPTK, "From DSD to Download: A deep dive into sample rate conversion" (2023)
Waveroom, "The Difference Between Bit Rate, Sample Rate & Bit Depth" (2023)
Boyamic, "Bit Depth vs. Sample Rate: The Pro Audio Quality Master" (2025)
Transloadit, "Understanding different audio formats" (2024)
HiFi Sound Gear, "Does Higher Bitrate Mean Better Audio?" (2025)

Reddit r/audioengineering, "Preferred algorithm for sample rate conversion"
ToneStack, "Lossy audio compression: principles, methods, misconceptions" (2025)
Coding Horror, "Concluding the Great MP3 Experiment" (2012)
YouTube, "Best Audio Sample Rate Converter and It's FREE!" (2013)
Nottingham HiFi, "Understanding Audio Bitrate" (2025)
Steinberg Forums, "Sample Rate Conversion - what's your plugin chain" (2015)
SLDP, "LOSSLESS VS LOSSY AUDIO - Can you hear a difference?" (2025)
Triton Digital, "Choose Audio Bitrate Settings" (2026)

Moom Mastering
MoomMastering.net
Info@Moommastering.net