RECORDING ACADEMY™
PRODUCERS & ENGINEERS WING®

RECOMMENDATIONS FOR HI-RESOLUTION MUSIC PRODUCTION

CONTRIBUTORS

- Leslie Ann Jones and Phil Wagner: Quality Sound Matters Committee Co-Chairs
- Leslie Ann Jones: Hi-Res Music Production Recommendations Subcommittee Chair
- Bill Gibson: Document Author
- Marc Finer: Additional Contributions
- Chuck Ainlay: Committee Member
- Bob Ludwig: Committee Member
- Rick Plushner: Committee Member
- Phil Wagner: Committee Member
- Maureen Droney: Managing Director, Recording Academy Producers & Engineers Wing

INTRODUCTION

Recent research from Music Watch, the Consumer Technology Association and DEG has shown that consumers are interested in better quality audio and that many are willing to pay for it. The sound of digital audio has improved vastly from its early days. It is now possible for consumers to hear studio quality music—recorded music with sound quality equal to what the artists, producers, mix engineers, and mastering engineers worked to achieve. Providing the listener with the best audio quality gives them the best possible experience, which results in more engaged consumers, increased satisfaction in the creative community, and an improved financial environment for all who are involved in creating, recording, producing, and promoting music.

Sales of Hi-Resolution Audio downloads from companies such as HD Tracks and Pro Studio Masters have become a healthy business with many subscription streaming services focusing on the delivery of higher quality audio, whether CD-quality or true Hi-Resolution using MQA. And now, most record companies require the delivery of hi-res masters because they recognize new opportunities to monetize better-sounding music.

Although there is growing interest in hi-res audio within the music industry digital supply chain, we currently lack hi-res production standards. This leads to confusion and inefficiencies. The Recording Academy Producers & Engineers Wing has created "Recommendations for Hi-Resolution Music Production" to help increase efficiency and transparency in the production process. It is important that we all work together to ensure the consistent quality of recorded masters, from their inception, through the supply chain, and ultimately to consumers.

This set of recommendations is not a tutorial on digital concepts. However, when it leads to a more thorough understanding of hi-resolution audio, some aspects of digital recording are briefly explained.

BACKGROUND: PURPOSE OF COMMITTEE

Historically, the audio industry has seen an ebb and flow of new ideas, from the lineage of analog—cylinder recorders, wire recorders, records (78s), magnetic tape recorders, microgroove 33 1/3 and 45 rpm discs, prerecorded open reel tape, 8-track cartridges, and cassette tape—to digital tape, CDs, and files. Some of these technologies have improved sound quality and some have increased consumer convenience. Our industry has adjusted its use of technology based on the consumer's desire for quality or convenience. This has been demonstrated by the transitions from LPs (quality) to cassettes (conveniences) to CDs (quality) to MP3s (convenience). Hi-resolution audio represents an opportunity to fill a growing demand for high-quality experiences for the modern recorded music consumer. Supporting this cyclical return to an emphasis on quality is the rapid growth and development of virtual and augmented reality content, which combines high-quality immersive video and audio.

It is now possible for consumers to hear music at its best rather than after it has been converted to conventional uncompressed digital resolutions (CD-quality) or worse, after it has been converted to inferior compressed consumer formats (MP3, AAC, and so on).

The Committee conducted an extensive series of interviews with active recording professionals from all genres and specialties to determine helpful suggestions and proven strategies for producing, recording, mixing, and mastering hi-resolution music. Those conversations formed the basis for this set of recommendations, which was created to:

- Promote an environment where the best possible audio quality is consistently delivered to consumers
- Assist record labels, online distributors, and aggregators with their master delivery requirements

There remains some resistance to the idea that hi-res audio is better than other commercially released audio formats. Many producers are unaware of the sonic benefits offered by recording in hi-res, so they begin all recording sessions at their equipment's default sample rate and bit depth, which in some cases is still 44.1 kHz/16-bit.

During our interviews, we encountered cases where the master mix existed solely in MP3 format, limiting the quality of all subsequent consumer offerings to the lesser quality of the source MP3. The fact that this could happen today points at two immediate concerns:

1. The need for education regarding master recordings
2. The importance of agreeing about what is acceptable for final master delivery

While producers and engineers may recognize and appreciate the benefits of hi-res audio, sometimes the artists they work with don't know about, or are unconcerned about, format options. They may not realize that distribution exists specifically for hi-res formats, and that the distribution of hi-res recordings represents new and enhanced streams of income. This lack of knowledge highlights the need for education among the creative community. Some labels and many independent, self-released artists are unaware that these distribution channels exist, so they don't request hi-res masters and therefore miss out on an entire stream of potential income.

Our industry is evolving. The quality of music recordings being delivered to consumers is evolving as well. And, this committee recognizes the unique approach that each recording may require. We respect and understand that each creative professional relies on inspiration through his or her personal workstyle. Our goal in developing this document is to recommend ways for audio professionals to effectively incorporate hi-res audio into their creative workflow.

WHO ARE THESE RECOMMENDATIONS FOR?

These recommendations are for anyone who is actively recording, producing, and delivering music. Unfailingly, in our discussions and interviews the following questions would come up:

- What about the rock community? Does hi-res audio really matter if you're recording a rock band?
- What about the urban hip-hop community? Will hi-res audio make a difference and will the audience care?

- What about electronic dance music?
- Can people really hear the difference, and do consumers care?
- And so on…

Breaking this document into sections according to genre would be inefficient and unnecessary. Recording hi-resolution audio makes a difference for anyone recording, producing, or listening to music. It raises the bar. We're on the verge of advancement—a new era of excitement about music and the way it sounds.

If you care about how your music sounds, this document is for you.

WHAT IS HI-RESOLUTION AUDIO?

The term "Hi-Resolution Audio" is defined by sample rates, bit depth, and file formats. However, and more importantly, the function of hi-res audio is to provide the consumer with a studio quality listening experience that reflects what artists, producers, tracking engineers, mix engineers, and mastering engineers hear in the studio. These music creators work tirelessly to achieve the best possible audio quality. And, although consumers care passionately about the convenience of their listening experience, current research indicates that they also care about audio quality.

By the simplest definition, hi-res audio is lossless audio that is better than CD-quality in both sample rate and bit-depth. Since the Red Book specification for CD-quality audio requires a 44.1 kHz sample rate, 16-bit PCM format, the lowest possible format to be considered hi-res audio is 48 kHz/20-bit PCM. However, even though 48 kHz/20-bit audio quality is technically hi-res audio, the recommended minimum-resolution for a recording project is 48 kHz/24-bit PCM. At 48/24, all modern computers are capable of recording and playing back large sessions with high track counts, along with a substantial number of plug-ins. The conclusion drawn by this committee is that tracking, mixing, and mastering at a resolution of 96 kHz/24-bit or 96 kHz/32-bit float PCM is preferred, providing true hi-res audio while imposing less burden on the computer CPU and allowing for higher track counts than when recording at 192/24, 192/32-bit float, or higher.

There are two common digital audio formats for recording, mixing, and mastering hi-res audio: Pulse Code Modulation (PCM) and Direct Stream Digital (DSD).

Recommended Formats	Minimum Quality	Preferred Quality
PCM	48 kHz/24-bit	96 kHz/24-bit or better
DSD	2.8224 MHz	5.6448 MHz or better

HI-RESOLUTION AUDIO AND PROCESSING POWER

Audio recorded at a higher sample rate and bit depth provides better quality. However, even with the power of modern computers, there may be a trade-off between the highest possible audio quality and the practical number of audio tracks or voices that a computer can realistically process. Audio resolutions of 192 kHz/24-bit and higher are preferred but, at those resolutions, the user may experience a prohibitive reduction in track count potential.

The balancing act between audio resolution and computing power is a fundamental concern, especially when all processing is being handled natively by a single computer. Modern card-based DAWs can deliver large numbers of tracks at higher resolutions because the cards handle much of the processing. In addition, an increasing number of modern outboard interfaces absorb part of the audio processing requirements and, therefore, extend the recording system's capacity beyond the capacity of the CPU alone. It's worth noting that even some card-based systems that were developed prior to our current move to hi-resolution audio, struggle to support high track counts at 96/24 and above.

To work successfully in hi-res, each engineer or producer should know the limitations of his or her system and work within those boundaries. For example, if you are mixing internally within the digital audio workstation (DAW), 32-bit floating point digital signal processing can provide better results than 24-bit processing. But, if the CPU and supporting peripherals can't support the processing demands of a specific audio setting, a compromise in track count or audio resolution might be required.

Recommendations:

- Keep current. Update computers, software, interfaces, and drives as frequently as is practical.
- Increase processing capacity substantially by incorporating additional cards and interfaces that share the processing requirements with the CPU.
- Set your system default preferences to record at 96/24 unless intentionally set otherwise.

THE REAL DIFFERENCE BETWEEN 44.1/16, 48/24, 96/24, 192/24 AND BEYOND

Is there truly a noticeable difference between MP3s and 192/24 files? Absolutely, but everyone owes it to themselves to listen and compare. In most cases the differences between CD-quality and 192/24 are at least noticeable, and frequently, they are stark. Skillfully mixed and mastered music with a wide dynamic range benefits dramatically from a hi-res workflow. For recordings such as symphonic film scores, classical music, or other recordings that feature acoustic instruments, hi-res audio is a perfect fit—the increased audio quality can be appreciated by virtually anyone who hears it. In the experience of this committee and the audio professionals we interviewed (including numerous rock, pop, and urban producers and engineers whose work is aggressive and powerful), recording, mixing, and mastering at resolutions 96/24 or better results in a final product that is both sonically superior and faithful to the sound of the final mastered mix.

THE RESULT OF INCREASING THE PCM SAMPLE RATE

Many hi-res fans speak in subjective terms that can be less than convincing. Phrases like, "buttery lows," "transparency," and "it's much more open" are difficult to support with mathematics and technical specifications. However, there are certain aspects of hi-res audio that are easily explained and quantifiable. Sample rate and bit depth are two of those aspects.

The sample rate is the number of times per second that the voltage created by the audio waveform is measured and stored. Because of reasons we'll discuss later in this document, there isn't a theoretical downside to increasing the sample rate above 96 or 192 kHz or beyond. However, there is a practical limitation imposed by the computing system. If the system can't keep up with data flow demands, severe track count limitations will be imposed and the entire system will become slow, inefficient, and unreliable. On the other hand, as stipulated by the Nyquist-Shannon Theorem, there is a limitation on the minimum sample rate. The Nyquist-Shannon Theorem states that a frequency can only be accurately reconstructed if it is sampled at least twice per cycle. Therefore, the highest frequency that can be sampled accurately is half the sample rate.

Sample Rate	Nyquist Limit
44.1 kHz	22.05 kHz
48 kHz	24 kHz
88.2 kHz	44.1 kHz
96 kHz	48 kHz
176.4 kHz	88.2 kHz
192 kHz	96 kHz

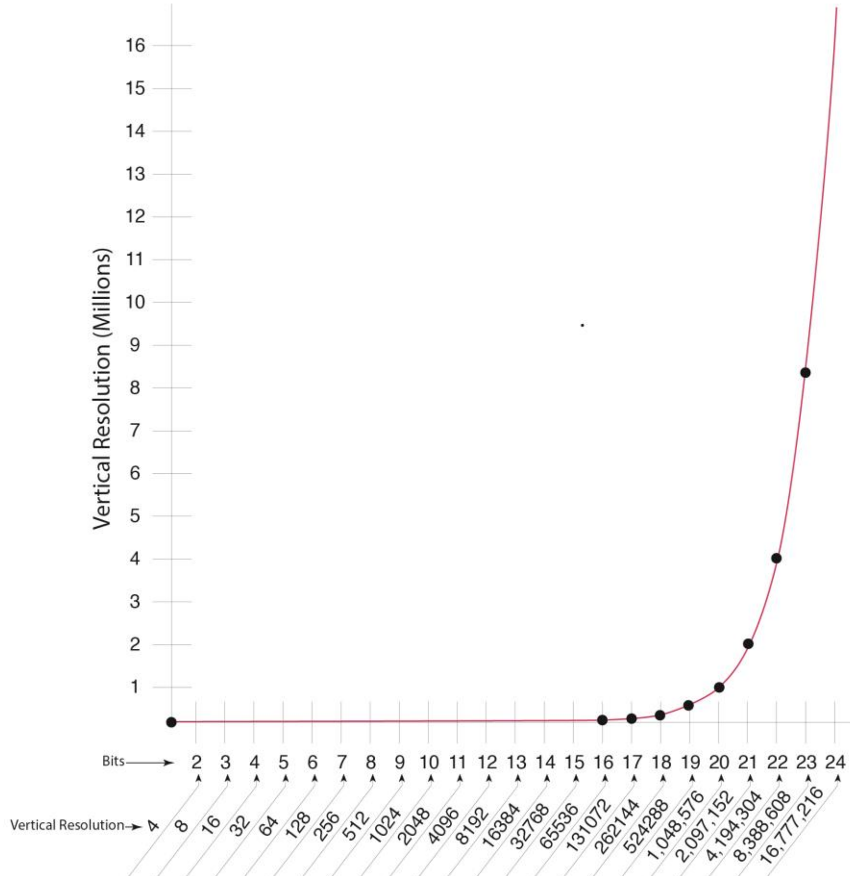
Since the range of audio that we're trying to capture is 20 Hz to 20 kHz, the Nyquist-Shannon Theorem indicates that the sample rate should be at least 40 kHz to capture an accurate recording. However, if the digital recorder is given frequencies above its ability to accurately reproduce, the inaccuracy results in a phantom waveform called an "alias," which manifests as a lower frequency that does not exist in the source audio. An anti-aliasing filter (a low-pass filter) is used to remove any alias-causing frequencies above the Nyquist limit. So, the goal of CD-quality recording, at a sample rate of 44.1 kHz, is to provide full-bandwidth audio content that extends from 20 Hz to 20 kHz. By extending above the 40-kHz sample rate that's indicated by the Nyquist-Shannon Theorem, a 44.1-kHz sample rate allows for the use of a low-pass filter with a reasonable slope.

Why is aliasing important to discuss regarding hi-res audio? At a sample rate of 44.1 kHz, early digital recorders used a steep slope on the anti-aliasing filter, which resulted in an anomaly at the onset of the slope. This anomaly manifested as a ringing or zinging sound in the high end. A gentler slope on the low-pass filter helps flatten the bump at the low-pass frequency, resulting in a smoother sound without the characteristic zing of early digital recordings. In addition, as the sample rate increases, the low-pass filter extends the onset of the slope to the upper frequencies, allowing for a gentle low-pass slope and smoother-sounding digital recordings.

THE RESULT OF INCREASING BIT DEPTH

The Red Book CD specification calls for 44,100 samples per second using 16-bit words. Whereas the number of samples defines the waveform's horizontal resolution along the x-axis, the bit-depth defines the vertical resolution along the y-axis. A binary word has just two possible values (one and zero) and the number of potential combinations is calculated with the number of bits being an exponent of 2 (2^x). Sixteen binary digits in a word provides exactly 65,536 potential vertical steps (2^{16}) from top to bottom (on the y-axis) with which to define the amplitude status (voltage) in each sample.

In contrast, the recommended sample rate and bit depth for hi-resolution audio is 96/24, or 96,000 samples per second with a 24-bit word, which has 2^{24} (16,777,216) vertical steps from maximum positive amplitude to maximum negative amplitude.



Note: Compared to the 16-bit vertical resolution of 65,536 steps, a 24-bit vertical resolution is represented by 16,777,216 steps! The exponential expansion from 16-bit audio to 24-bit audio results in a ratio of 256 steps (at 24-bits) for every 1 step (at 16-bits). Even among engineers who are convinced that higher sample rates are better, it is widely believed there is a greater sonic pay-off by choosing to record with 24-bit words than by choosing a faster sample rate.

Though the increase in bits from 16 to 24 undeniably increases the resolution of the grid with which analog waveforms can be represented in the digital domain, there is a quantifiable sonic change that occurs with these increases. The dynamic range—the distance in decibels from the quietest sound in a source to the loudest sound in a source—increases with each added bit. A recording with a lower noise floor has an increased dynamic range. Analog recordings on cassette tape have a limited dynamic range because, by virtue of the small tape size and the slow tape speed (1 7/8 ips), the noise floor is very high. Professional audio recordings made on analog reel-to-reel tape machines provide a greater dynamic range than cassette tape because the noise floor of a reel-to-reel recording is inherently lower than that of a cassette. The actual dynamic range provided by an analog tape machine is dependent on the tape formulation, electronic alignment settings, tape speed, size of the tape tracks, and the quality of the circuit design and components.

Because the inherent noise floor of CD-quality recordings is lower than cassette or reel-to-reel recorders, they have a greater dynamic range than either cassette or reel-to-reel recordings. And, since the noise floor in 192/24 recordings is still lower yet, they have a greater dynamic range than CD-quality recordings.

The calculation of the dynamic range for any bit-depth is, very simply, 6.02 time the number of bits. For a 16-bit audio recording, the dynamic range is 6.02 × 16, which equals just a little over 96 dB. The dynamic range of a 24-bit recording is 6.02 × 24, which is barely under 144.5 dB.

Considering that the best dynamic range from an analog tape machine without noise reduction is about 80-dB (equivalent to about a 13-bit recording), the dynamic ranges of both 16- and 24-bit recordings are spectacular. As a further comparison, the best analog recordings with the hottest tape formulation and Dolby SR noise reduction achieve a broadband dynamic range of about 100 dB (equivalent to a 17-bit digital recording) and the typical compact cassette provides a dynamic range of 50 to 56 dB (equivalent to a bit-depth ranging from 8 to 10 bits).

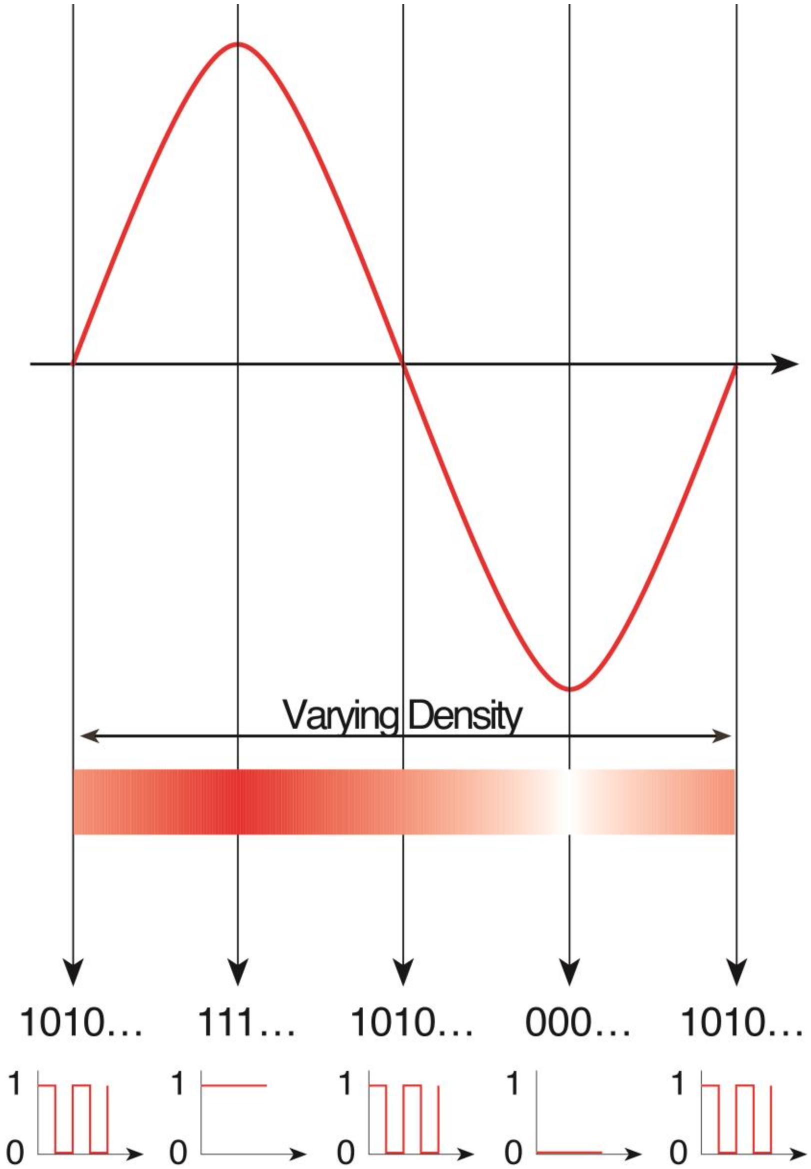
DIRECT STREAM DIGITAL (DSD)

Developed by Sony and Philips and commercially released as Super Audio CD (SACD) in 1999, Direct Stream Digital (DSD) audio recording is a pulse density modulation (PDM) system. DSD encodes the analog sound wave using a 1-bit binary bit stream at a sample rate of 2.8224 MHz. Compared to PCM recordings that quantify the voltage (amplitude) status of each sample by a 16 to 24-bit word, DSD recordings define the varying amplitude voltage by the density of ones in the continuously flowing binary bit stream. Full positive amplitude is represented by a constant series of ones; full negative amplitude is represented by all zeroes; and zero amplitude is represented by alternating ones and zeroes.

Increasing positive amplitude results in the increased density of ones, while increasing negative amplitude results in decreased density of ones or, in other words, an increase in the density of zeroes. DSD recording systems faithfully follow the analog waveform. When plotted, the varying density looks like an analog imprint with shades of gray varying back and forth from black (full positive amplitude) to white (full negative amplitude). Both PCM and DSD systems fall short regarding linearity, quantization errors, and quantization noise. Increasing the sample rate helps both systems move problematic errors and noise out of the audible spectrum.

The initial DSD sample rate of 2.8224 MHz is frequently referred to as DSD64 because its sample rate is 64 time the CD-quality sample rate of 44.1 kHz. The DSD format also uses sample rates of 5.6448 MHz (also called DSD128 or double-rate DSD), 11.2896 MHz (DSD256 or quad-rate DSD), and 22.5792 MHz (DSD512 or octuple-rate DSD). The higher sample rates are intended primarily for studio use but there are very few available multitrack mix/edit systems at this writing with Sonoma, Sadie, and Pyramix leading the way.

The DSD and PCM processes are different enough that it’s hard to make a true mathematical comparison. However, a 2.8224 MHz DSD is understood to have 120 dB dynamic range, which equates to a 20-bit PCM recording, and an accurate frequency response from 1 – 100 kHz. But, quantization noise above 25 kHz increases dramatically, requiring the use of complicated noise-shaping filters that move the noise out of the audible spectrum. So, it is generally held that A 2.8224 MHz DSD recording is similar in quality to a 96 kHz/20-bit recording and that with each increase in the sample rate, from double-, to quad-, and to octuple-rate DSD, the audio quality increases.



Note: In this preceding illustration, the status of “1” is indicated by a red vertical line; a status of zero is indicated by the lack of a red line (white). As the amplitude rises above the center line, the density of ones increases in the bit stream.so the shading is darker and darker until peak amplitude is reached and the shading is dark red. And, as the amplitude decreases below the center line, the density of ones decreases in the bit stream.so the shading is lighter and lighter until the shading fades to white.

THE LIMITATIONS OF HI-RESOLUTION AUDIO

The quality of hi-res audio is dependent on the source. A 192 kHz/24-bit recording of a cassette will not sound better than that cassette. Once in the digital domain, it’s possible to apply digital processing tools to help compensate for audio quality problems or to enhance the sound quality; however, the simple transfer from cassette or any other format to 192/24 does not improve the fundamental sound quality that’s delivered from the output of the source player to the record input of the A/D converter. (For further information, see “The Pros and Cons of Upsampling” later in this document.)

Recommendations:

- Record audio at the highest practical sample rate for your system.
- Record 24-bit audio.
- Upsampling will not increase audio quality.

HI-RESOLUTION HARDWARE AND SOFTWARE CONSIDERATIONS

SYSTEM SPECIFICATIONS

Compared to recording standard-resolution audio, recording music in hi-res places greater demands on a computer system. Whether considering the core CPU, DAW software, digital audio interface, type of storage media, or delivery format, every music recordist needs to stay current. Modern computing systems are quite capable of recording hi-res audio; however, at the highest practical sample rate and bit depth (192/24 or 192/32 float) there can be increased limitations in the available track count. A small system might be limited to eight or fewer simultaneous tracks for playback and recording at 192/24. However, a larger system using interfaces or PCI cards that handle much of the processing might be virtually unlimited in track count—constrained only by bus speed capacity and the type of storage device being used.

PROCESSOR

To take advantage of current software and data requirements, use a computer that's as fast and up to date as possible. Processing power and speed has increased rapidly for many years—typically doubling in capacity every 12 to 18 months—so producers and engineers interested in recording in hi-res should upgrade their computers when warranted by technological advances and as frequently as is financially practical.

RAM

To facilitate a positive creative experience, an engineer who is serious about recording hi-res audio should use as much RAM as the computer will hold. Also, it's best to use a 64-bit operating system and software. A 64-bit OS optimizes the way RAM is accessed; therefore, it increases the amount of RAM that can be accessed.

SOFTWARE

Some software manufacturers place limitations on the number of available tracks or voices, but many manufacturers place no limitations on the track count. In that case, the only limitations are imposed by the capacity of the computer, amount and speed of RAM, interface, and storage device.

Modern DAW and plug-in software can record, process, and play back hi-res audio; however, there are still some limitations imposed that can govern the sample rate and bit depth to a lower resolution than intended. It is impractical for this document to list all the software and plug-ins that might create limitations because there are constant changes and strides forward. But, it is extremely important that each engineer confirm that any software used to record hi-res audio imposes no unacceptable or destructive band limitations.

INTERFACE

Especially when recording multitrack hi-res sessions, it's advisable to use an interface that provides additional processing power to supplement your core CPU. As mentioned previously, some manufacturers offer PCI cards that handle much of the processing and others include additional processing power in rack-mounted or desktop hardware. Either type of system offers a substantial performance bonus and a boost in speed, performance, and stability.

LATENCY

Latency is the time delay from the onset of the source audio to its eventual monitoring from the DAW output. Recording systems that are less able to keep up with the demands of digital recording can induce a delay between the DAW's audio input and output that is prohibitive to the addition of overdubs in the multitrack recording process. The delay is simply too long to allow for the overdubbing of additional audio tracks in sync with the existing tracks.

Some manufacturers provide processors within the interface that dramatically minimize the load on the computer CPU, releasing the CPU to perform basic recording tasks, leaving plug-in processing to the interface processor and helping to provide a minimal- or zero-latency recording environment.

Some modern DAW manufacturers include low-latency modes, optimizing processing to minimize latency issues. Although increases in CPU processing power and speed help minimize problematic latency issues, modern audio/MIDI interface manufacturers also take advantage of the speed and capabilities of USB 3.0, USB 3.1, and Thunderbolt 1, 2, and 3 to facilitate latency times well below 1 ms. And, some interfaces provide a path for the artist to monitor directly from the interface input for overdubbing rather than monitoring the track after it has travelled through the DAW's signal path.

When latency is a problem, reducing the size of the RAM buffer reduces the length of the latency. In contrast, during a mixdown that uses multiple processor-intensive plug-ins, the RAM buffer should be set higher. Buffer settings that are too low tend to cause freezes, crashes and error messages. Since latency issues are not problematic during most mixdowns, it is good practice to increase the RAM buffer prior to beginning the mix session. Whereas optimal RAM buffer settings for tracking might be 32 or 64 samples, during mixdown the RAM buffer might need to be 1024, 2048, or higher to keep up with processor demands.

Recommendations:

- Design the system for your specific types of recording applications. In the current computing era, a small system isn't an inferior system—it's just limited in capacity.
- In a quick paced workflow with high capacity sessions, it's important to maintain current computers, the maximum RAM the computer will hold, current software, and an up-to-date interface that augments the system's processing capacity.
- Use a 64-bit OS and software.
- For recording and overdubs reduce the size of the DAW's RAM buffer to minimize latency.
- For mixdown, increase the size of the RAM buffer to better handle the demands imposed by large, processor-intensive plug-ins.
- Use an interface with additional processing power and an accommodation for zero-latency monitoring.
- Use a USB 3.0/3.1 or Thunderbolt interface for dramatically reduced latency

HARD DISK DRIVE

Hard disk drive (HDD) speed is important in any audio recording application, but that importance elevates when recording hi-res audio, especially in a multitrack application. In a system that utilizes a tower with multiple internal drive bays it has long been recommended that the operating system and applications should be run from one internal drive, and the recording session should be run from a separate internal drive. Traditionally, the internal SATA transfer rates have exceeded the transfer rates available from external drives. It is ideal if your computer contains at least two built-in 7200 RPM (or faster) drives with SATA/eSATA interfaces.

When the recording computer, such as a laptop or the current Mac Pro, contains just one internal hard-disk or solid-state drive, recording to an external drive might be the only viable option. In this case, USB-3.0, USB-3.1, and Thunderbolt 1, 2, and 3 provide sufficient connection and transfer rates to record and playback large hi-res multitrack sessions.

SOLID STATE DRIVE
Solid State Drives are storage devices containing nonvolatile flash memory, used in place of a hard disk. They have no moving parts, offer faster read and write access, and are more durable and quieter than hard disks. They are more road-worthy than HDDs, but aren't yet well-suited for long-term storage. Whereas the life of an HDD is primarily determined by revolutions of the disk and how long it can continue spinning, the life of an SSD is determined by a finite number of read/write actions.

The upside of SSDs is their speed and small size. The downside is clearly the fact that they have a higher failure rate than HDDs. And, an SSD failure is catastrophic; when it fails, the data cannot be retrieved. Therefore, SSD drives are only considered acceptable for use as Transitional Master Backup Storage Media, providing they are "new" (i.e. not previously used) AND are "enterprise-grade" varieties including SLC (Single Level Cell) and eMLC (Enterprise Multi Level Cell). Primary Storage should be on an HDD.

For more information on Master Delivery recommendations, refer to the Producers & Engineers Wing "Delivery Recommendations for Recorded Music Projects (Including Stems and Mix Naming Conventions)" https://www.grammy.org/files/pages/deliveryrecommendations.pdf

The following two tables display: 1) Representative HDD/SSD speeds, capacities, and transfer rates and 2) Data connection interfaces.

HARD DISK DRIVES/SOLID-STATE DRIVES

Drive Type	Drive Speed (rpm)	Drive Capacity	Transfer Rate
Toshiba SATA	7200 rpm	4 TB	6 Gbps/750 MBps
Western Digital SATA	5400 rpm	2 TB	3 Gbps/375 MBps
Samsung SATA	7200 rpm	1 TB	3 Gbps/ 375 MBps
Western Digital SATA	5400 rpm	2 TB	6 Gbps/750 MBps
SSD		1 TB	400–550 MBps

DATA CONNECTION INTERFACES

Connection Type	Version	Speed
USB	2.0	480 Mbps/60 MBps
USB	3.0	5 Gbps/625 MBps
USB	3.1	10 Gbps/1250 MBps
SATA	1	1.5 Gbps/187.5 MBps
SATA	2	3 Gbps/375 MBps
SATA	3	6 Gbps/750 MBps
FireWire	400	400 Mbps/50 MBps
FireWire	800	800 Mbps/100 MBps
Ethernet	100/1000	100&1000 Mbps/12.5&125 MBps
Thunderbolt	Original	10 Gbps/1.25 GBps
Thunderbolt	2	20 Gbps/2.5 GBps
Thunderbolt	3	40 Gbps/5 GBps

Recommendations:

- Use 7200 RPM hard drives.
- Use USB-3.0 or faster data interface connections.
- When possible, use an internal system drive (fast HDD or SSD) for applications and operating system software, along with a separate internal SATA/eSATA 7200 RPM HDD for the recording session.
- When the recording computer only contains one internal drive, use a fast internal drive (HDD or SDD) for the system and applications along with an external USB-3 or Thunderbolt drive for the session files.

HI-RESOLUTION RECORDING AND MIXING CONSIDERATIONS

SETTING UP YOUR RECORDING SESSION: CHECK THE SAMPLE RATE AND BIT DEPTH

It's imperative to check the audio resolution of each new session. Some manufacturers still default to a resolution of 44.1 kHz/16-bit. Other manufacturers set their defaults to 48 kHz/24-bit. Still others create each new session at the audio resolution of the last session opened. This fundamental consideration must be checked and set correctly before the start of any new session.

Recording, mixing, mastering, and distributing audio at 96 kHz/24-bit resolution is preferred. This audio resolution provides a good balance between high-quality digital audio and the processing potential of the modern computer. Increasing the bit-depth is at least as important as increasing the sample rate.

When starting work on a project that was previously recorded by other producers and engineers, verify the sample rate of all the previously recorded audio files. Whereas almost all DAWs require that all the audio in a single session matches the session audio setting, some allow the recording, playback, and importing of multiple file formats within the same session. These DAWs perform real-time conversions for playback of audio files that don't match the session settings while playing back and/or recording new audio. Because most DAW software requires that all audio in the session conforms to the session's sample rate and bit-depth settings, and considering the extra drain that real-time conversion places on the CPU, it's important confirm that all audio clips match the session audio settings.

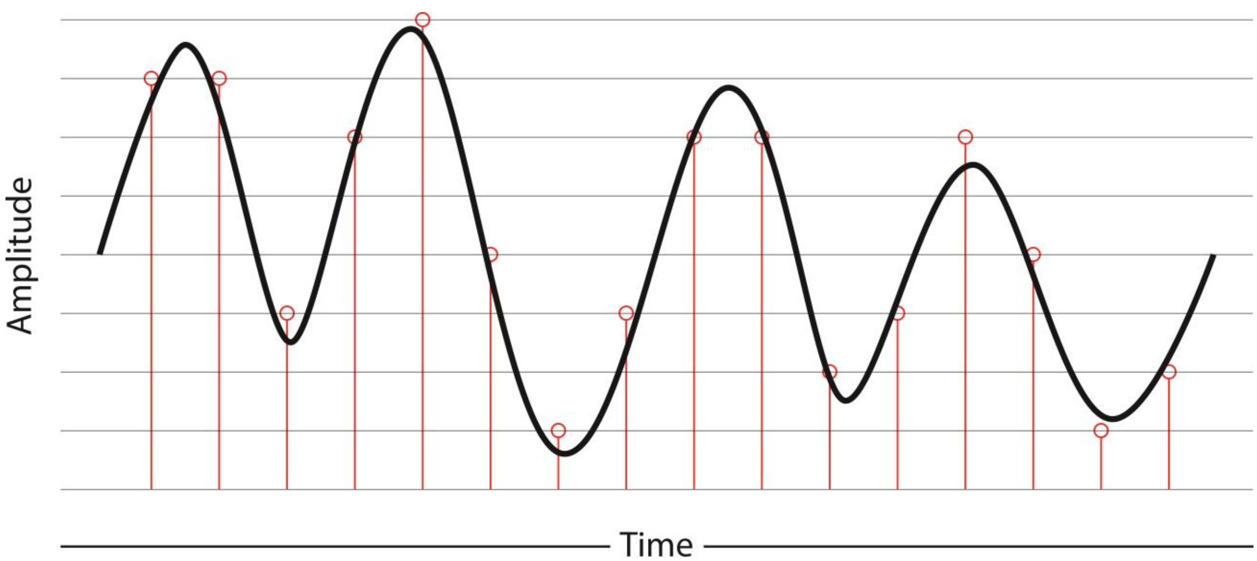
Recommendations:

- Be intentional about sample rate and bit-depth settings.
- Set up sessions for 96 kHz/24-bit audio. If track count or processing requirements demand it, record 48 kHz/24-bit audio.
- When continuing a session that someone else started, verify the sample rates of existing audio files match the new session settings.

QUANTIZATION

Quantization is the process whereby the continuously varying analog audio waveform conforms to the closest available sample point and amplitude voltage. This illustrates the origin of the stair step analogy when describing a digital rendition of an analog waveform; however, some prefer other analogies. When the analog waveform is quantized, it is forced to conform to the available digital grid. When the digitally plotted and quantized voltage isn't precisely equal to the voltage of the continuously varying analog wave form, quantization errors occur. Therefore, compared to CD-quality audio, using higher sample rates and longer digital words such as 96 kHz/24-bit results in fewer quantization errors and a much more accurate conversion from analog to digital and then back to analog.

Quantization errors—inaccuracies and digital artifacts—are most problematic in extremely quiet audio recordings. These inaccuracies and artifacts contribute to the noise floor.



Note: The low-level (quiet) waveform in the preceding graph cannot be accurately plotted by the available grid. The digital recording process can only quantize the amplitude voltage to the closest available grid point. And, these inaccuracies, called “quantization errors,” are most extreme when converting quiet waveforms like the one represented here because there are fewer bits available to accurately represent the analog wave.

Recommendations:

- Use the highest practical sample rate when recording digital audio for the most faithful capture and reproduction of the source wave form.
- Use 24-bit words during recording for greater accuracy and fewer quantization errors.

DITHER

At especially low signal levels, digital recording systems perform very poorly because there simply aren’t enough bits available to accurately record and play back an audio source. When any sound is represented by the bottom bits in a digital word, the resulting sound is a distorted and static-like rendition of the source audio. The sonic integrity of low-level signals—most common during the fadeout of a commercial song and during the quietest passages of classical recordings such as those of a symphony, choir, or string quartet—can be maintained by the proper use of dither.

Dither is simply white noise added to the program source at very low levels. The actual noise color—the frequency balance—can be shaped depending on the source audio and desired results. Though it seems ironic to add noise to an otherwise noiseless system, dithering helps improve the perceived accuracy of low-level passages in a digital recording and can help decrease the apparent source noise. In fact, dither doesn’t simply mask the artifacts; it removes them, similar to the way that a soft-focus filter smooths out a pixelated image.

Increasing the bit depth increases dynamic range. For most sources, this increased dynamic range keeps the intended audio signal much farther from the noise floor. Therefore, 24-bit audio is freer from noise than 16-bit audio and is less dependent on the dithering process.

The choice to use dither is typically left for the mastering stage of an audio production—it is best to rely on the mastering engineer to choose the best strategy to implement dither.

Recommendations:

- When tracking audio, don’t use dither.
- Reserve dithering for the mastering process where its use can be selected and applied for the best results.
- Record 24-bit audio files to benefit from its increased dynamic range and decreased noise floor.

32-BIT FLOATING-POINT AUDIO RESOLUTION

Most current professional digital workstations can process audio at 32-bit floating point or better. Incorporating 32-bit float mode into an audio production workflow provides a means of extending headroom for internal DSP tasks. However, although there are manufacturers working to increase the native playback standard for consumer devices, current analog-to-digital and digital-to-analog converters rarely operate at more than 24 bits. Until there are consumer devices capable of playing back higher bit-depth files, 32-bit float audio files are not appropriate for consumer distribution.

The benefits from recording 32-bit floating-point audio are realized during internal processing (in the box) where it provides several advantages:

1. It allows the meaningful 24-bit signal to be scaled so the noise floor follows the signal level.
2. It provides headroom.
3. It helps avoid clipping and rounding errors when internally processing the 24-bit audio that's recorded by the DAW.
4. It also helps avoid the addition of unnecessary dithering noise that could be introduced by any plug-in that renders a new version of the source audio.
5. It provides headroom when rendering virtual instruments to audio tracks in the same way that it provides headroom for other internal processing tasks.

When tracking, it is always important to avoid distorting the analog input. Distortion at the console input stage is not affected by selecting 32-bit float. Selecting 32-bit float only provides additional headroom when processing audio internally within the host computer.

The one disadvantage to recording 32-bit floating-point audio is the increased demand it places on the recording system. Although recording 32-bit float audio isn't fundamentally helpful during basic tracking and extremely simple mixing sessions, even modest sessions include some sort of internal processing such as clip gain, clip effects, or other processing action in which a newly render version of the source audio is created. In each of these scenarios, 32-bit float provides additional headroom and improved results. Therefore, if the DAW is capable, it's always appropriate to record and mix 32-bit floating-point audio.

Most professional DAWs provide an option to switch to 32-bit float from 24-bit mode after tracks have been recorded—a valuable feature in those cases where a session was tracked at 24-bit but then during mixing and mastering it makes sense to switch to 32-bit float.

Use 32-bit float audio settings during mastering, where final levels are set and internal processing such as EQ, compressions, limiting, and so on are applied. It's important to be aware that the consumer will receive 16- or 24-bit audio, not 32-bit floating-point files. The format that they receive is dependent on the formats delivered by the mastering engineer, which are currently limited to MP3, CD-quality, or any of the hi-resolution formats mentioned in this document.

Recommendations:

- When the audio tracks will be processed internally within the DAW, use 32-bit float audio.
- Be sure to add dither to a bounce or to the output of a 32-float DAW when creating 24bit or 16 bit files from the session.
- 24-bit audio settings are appropriate for sessions where there will be no internal processing applied by plug-ins or virtual instruments, and where there will be no processing performed by clip gain, clip effects, and so on.
- If the tracking computer can't handle the load of tracking large sessions in 32-bit float mode, track at 24-bits and then switch to 32-bit float for processing tasks that take advantage of the additional headroom provided in 32-bit float mode.
- When routing the outputs of a digital system to an analog mixing desk and then re-recording to the analog inputs of a separate workstation, record in 24-bits if there will be no further processing.

OPTIMIZING PROCESSING RESOURCES DURING MIXDOWN

A modern mixdown demands a lot from the computer processer. High track counts are common and the additional processing requirements imposed by the instantiation of multiple plug-ins is formidable. Relative to recording CD-resolution multitrack sessions, recording and mixing in hi-res raises processing requirements dramatically. Therefore, it is increasingly more important to be organized during the mixing process and to find ways to get the most out of all available processors. Tracks should be arranged and structured in a way that lets the mix engineer work quickly and efficiently.

Assign groups of like tracks to a stereo aux (subgroup) fader so the group level can be quickly and easily controlled. If, for example, your session contains 20 to 30 layered backing vocal tracks, they are much easier to control in the mix if all those tracks are assigned to a single stereo aux fader that is then routed to the main mix output. If each of those tracks includes compression, limiting, EQ, or other plug ins, the CPU might be struggling to keep up. In such an instance, consider bouncing all those tracks to a stereo submix and then deactivating the source tracks. This releases any processing power that was previously in use by the individual tracks in that group for use in other tasks. In a hi-res environment, printing these types of subgroups and then deactivating the source tracks, provides substantial relief to the CPU and enables a more efficient and trouble-free mixing experience.

STEMS

The term “stem” has become widely used in the audio community. It is commonly used to reference any bounced or recorded mix variation or submix. Therefore, although the process described in the previous section involves bouncing a group of vocal tracks with the intent to deactivate the source tracks and regain processing power for other mix tasks, the result is also referred to as a “backing vocal stem.”

Printing stereo or surround stems that represent the primary mix groupings, such as drums, backing vocals, strings, brass, and so on, helps the mix engineer fold the mix down to a manageable number of faders. However, in addition to the stereo and/or surround mix, many record labels, artist managers, and artists also specify other specific stems as a part of the mix deliverables. Stems are also routinely required for film and video game projects.

The most convenient time to print stems is immediately after completion of the final mix. Organize mixes and stems into separate stereo and surround folders and be sure to conform to any contractual specifications for delivery. Do not delete the source tracks. They might be needed for future remixes or for changes and/or updates to the deliverable files.

For more information regarding stems, refer to the Producers & Engineers Wing “Delivery Recommendations for Recorded Music Projects (Including Stems and Mix Naming Conventions): *https://www.grammy.org/files/pages/deliveryrecommendations.pdf*

MIX LEVELS

Master mix levels should be conservative enough to leave room for the mastering engineer to craft and render the final mastered versions. Preferred pre-mastering peak levels range from -3 to -6 dBFS, depending on the master recording’s bit depth and the type of music that has been recorded, and the mastering engineer’s preference. The dramatic increase in amplitude resolution provided by 24-bit and 32-bit float digital audio allows the mix engineer to leave more headroom for the mastering engineer to work with, without sacrificing audio quality.

DOCUMENTATION

It’s extremely important that engineers thoroughly document each recording and/or mixing project. It is highly recommended that both paper and electronic documentation accompany all

Master deliverables and Backup/Safety media. Paper documents should also be scanned and digitally stored within a metadata folder that is included in the session drive.

Documentation should include:

1. Tracking sheets, engineer notes, set-up notes, sketches of microphone placement, and any other pertinent session or mixing data
2. The provenance of all audio used within the session
3. All known technical and creative contributors, such as producers, performers, engineers, and songwriters
4. Use of outboard processors along with notes regarding relevant details about specific control settings, presets used, patch bay connections, signal routing, timings of necessary changes during the mix, and so on
5. Technical specifications such as digital audio resolution, storage formats, file locations, tape formulation, and so on

NAMING FILES: MIXES, STEMS, AND VERSIONS

Considering the wide range of digital file types, along with current expectations for multiple mix versions and stems, it has become very important to utilize a naming convention that clearly states the relevant aspects of each audio file. Mix versions and Stem files should contain all relevant information in their file names, and the naming protocol must be easy to understand at a glance. It's also crucial to establish a hierarchy for Project, Song, and Mix folders.

The origin of the following system for naming and organizing folders and files is derived from a document created by the Recording Academy Producers & Engineers Wing Delivery Specifications Committee and subcommittees. The folder hierarchy and naming convention included here relies on the following standardized order and abbreviation of terms. It's not necessary to include all terms in every folder or file name (e.g. AI_SongTitle_MI01_Master_96k24.wav).

Artist Initials (AI). The artist initials are usually two letters taken from the first and last name of the artist. Identify the artist by this 2 letter name code and use it consistently throughout a project. When labeling Parent folders, use the artists full name or an abbreviated version of the name. Capitalize the first letter of each name and remove all spaces.

Song Title. List the song title or a useful abbreviation thereof. Capitalize each word or word fragment. Titles should contain no spaces, punctuation, or diacritical markings (accents) so that names are universally file compatible. Song title names should be less than 15 characters. Lengthy titles may be routinely abbreviated by other programs when imported.

Mixer Initials (MI) and Mix Revision Number (01). After the song title, list the mix or stem identifier. "Master" is the example shown above. This may also be "Snare+Rim" or any descriptive identifier for the audio file. Capitalize each word or word fragment so title contains no spaces.

Mix Version or Stem Name (Master). This describes what the audio file is.

Sample Rate and Bit Depth. Sample rate and bit depth at which the audio file was created is listed after the mix version or stem type—the sample rate is followed by the bit depth. The single letter "k" is sufficient to abbreviate "kilohertz." For example, a song recorded at a sample rate of 96 kHz and a bit depth of 24 bits is abbreviated, "96k24."

File Extension. Typically, generated during file creation (e.g. ".bwf," ".wav," or ".aiff," which is sometimes ".aif."). If you have the option to show or hide the file extension, it should always be shown. Only one period should be used in the title and should only be placed before the file extension.

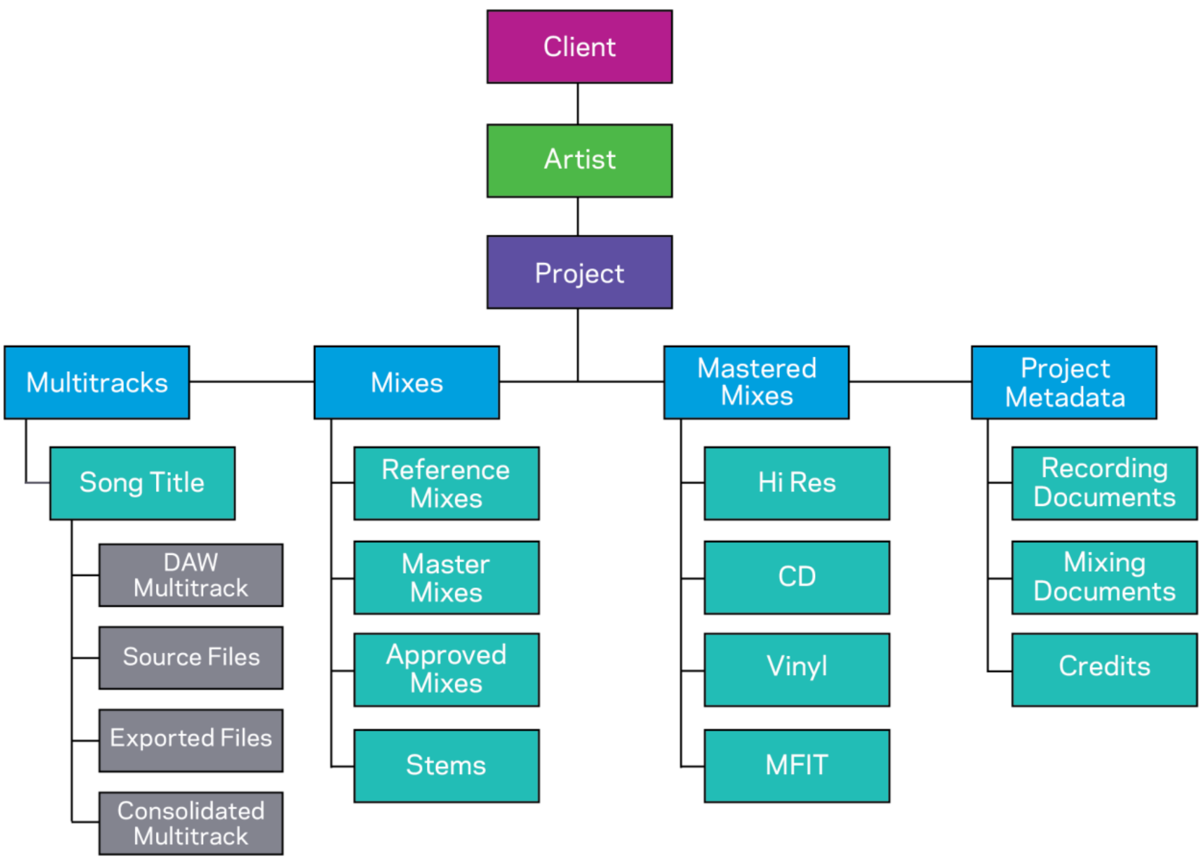
The file name length should not exceed 255 characters. Illegal characters include, but are not limited to: *backslash (/), question mark (?), left angle bracket (<), right angle bracket (>), forward slash (\), colon (:), semi colon (;), pipe (|), single quote ('), double quote ("), asterisk (*), percent (%), pound sign (#), ampersand (&), left bracket ({) and right bracket (}), dollar sign ($), exclamation mark (!), at sign (@), equal sign (=), and blank spaces.*

Note: For comprehensive guidelines on file and folder naming, refer to the Producers & Engineers Wing "Delivery Recommendations for Recorded Music Projects (Including Stems and Mix Naming Conventions)": *https://www.grammy.org/files/pages/deliveryrecommendations.pdf*

FOLDER HIERARCHY AND NAMING

Creating an easy-to-follow hierarchy of folders is as important as creating descriptive and easy-to-understand names for the files that are in those folders. The following chart displays the recommended folder hierarchy for the preservation and organization of all files during the recording, mixing, and mastering phases of a project.

Note: When a date is used in the file or folder name, the International Organization for Standardization (ISO) date format should be used without hyphens (YYYYMMDD). List the date the file was created. For example, May 8, 2017 would be "20170508".



Note: When it is appropriate for the project, additional folders can be added to customize the hierarchy. For example, in some cases the addition of a "Song Version" folder would be helpful.

FOLDER DEFINITIONS

Client. This is the main client folder for the hierarchy tree based on ownership of the recording. This could be a record label or the artist or artist's label. If the client folder is a label, all artists under this label would be contained in this client folder.

Suggested naming convention: Client name

Artist. This artist folder would contain all subsequent folders and files regarding all projects associated with this artist. It is recommended to use the full artist name to title this folder.

Suggested naming convention: Full Artist Name

Project. Within the artist folder would be a separate project folder for each project of that artist. The project folder could be an album, a single, a live recording, a surround mix etc.

Suggested naming convention: Project Title

Multitracks. This folder is the parent multitrack folder. Each project folder would contain a multitrack folder.

Suggested naming convention: Multitracks

Song Title. The multitracks folder would contain a folder for each song title of the project. The song title folder would contain the DAW multitrack for that song.

Suggested naming convention: Artist Initials_Song Title

Song Version (not pictured in chart A). – This is an optional folder and used only if a song title has more than one recorded version of a song for the project (e.g. explicit and non-explicit). All multitrack files pertaining to each version of the song should be contained within its own version folder. These folders would be contained within the song title folder.

Suggested naming convention: Artist Initials_Song Title_Version Name

DAW Multitrack. This folder would contain the DAW Multitrack session file and audio.

Suggested naming convention: ArtistInitials_SongTitle_DAWPlateformVersionMT(multitrack)_SampleRateBitDepth_Date

Source Files. Contained within the song title folder, this folder would contain the original files from an outside source that are imported into the DAW Multitrack folder. This would include source files received for the project, tuned vocal files, overdubs from outside sources etc. While this folder might contain duplicate files of the DAW Multitrack folder it would retain them in their original format.

Suggested naming convention: Source Files

Exported Files. This folder would contain exported files or sessions specifically set up for vocal tuning, musicians to use for overdubs etc. These could be stem sessions with a click, specialized track mix stem, and vocal stem. This would allow you to keep these sessions in case they're needed for future use.

Suggested naming convention: Exported Files

Consolidated Multitrack Delivery. Once the project is finished, delivery of the consolidated multitrack Master files maybe requested by the record label. Those consolidated Master multitrack sessions/audio files would be placed in this folder.

Suggested naming convention: ArtistInitials_SongTitle_DAWPlateform(&)Version(&)MT(multitrack)_SampleRateBitDepth_Date

Mixes. This folder would contain all versions of the master mixes for the project organized by mix version and type of mix.

Suggested naming convention: Mixes

Reference Mixes. This folder would contain any reference mix that you print. It is suggested that all reference mixes sent to the client and artist are kept in case they're needed for future reference.

Suggested naming convention: Reference Mixes

Master Mixes. This folder would contain all Master Mix versions and all recall mix versions. Mixes would be separated into folders by version type such as Master Mix, Vocal Up, No Lead Vocal and Instrumental. Master mixes should be organized into folders as to version type such as Master Mixes, Vocal Up, Instrumental etc.

Suggested naming convention: Master Mixes

Approved Mixes. This folder would contain only those mixes approved by the client for mastering or release. These approved mixes would be duplicated from the master mixes folder. This folder makes it easy for the mastering engineer to easily locate and identify the approved mixes that are to be mastered.

Suggested naming convention: Approved Mixes

Stem Mixes. This folder would contain any stem audio or stem multitrack sessions created from a Master Mix that is delivered to the client.

Suggested naming convention: Stems

Mastered Mixes. This folder contains the mastered mixes from the mastering engineer. The following sub-folders are recommended within the Master Mixes folder:

Suggested naming convention: Mastered Mixes

Hi Res. This folder contains the Hi Res mastered files. 96k 24 bit or higher is recommended if the project has been recorded and mixed at 96k 24 bit or higher resolution. Mastered files should retain original mix resolution.

CD. This folder contains the 44.1k 16 bit DDP file or CD audio files.

Vinyl. This folder contains the audio files used to create the vinyl master

MFIT. This folder contains the Mastered For iTunes audio files.

Project Metadata/Documentation. This folder would contain any documentation for the project including possible additional sub-folders for:

Recording Documents and Notes. Including but not limited to, notes for equipment used, signal path, photos of mic setup and location, vocal chain documentation, etc.

Mixing Documents and Notes. Including but not limited to, mixing recall / setup information, client mix notes, etc.

Credits and Other Metadata. This would contain Metadata documentation for each song and for the overall project, including the performers, producers, engineers, musicians, songwriters etc. This folder would also contain any codes such as ISRC, ISNI, UPC etc. associated with the project.

Suggested naming convention: Documentation

Recommendations:

- To regain processing capacity, create submixes/stems within your mix and then deactivate the source tracks.
- Create a separate folder of mixes that includes stems for all major instrument and vocal groups, a stereo mix, a stereo mix without vocals, and any other stereo or surround mixes and stems that are part a of a specified set of deliverable files or that might realistically be required for future use.
- Review contractual agreements regarding the resolution and organization of all mixes and stems.
- The mix engineer should work in a 24-bit or better environment when mixing via the analog domain.
- Any time internal processing is done in the session computer, 32-bit float should be used.
- The mix engineer should provide documentation including the provenance of the audio contained in the recording.
- Conform naming and file structure to the recommendations include in the previous sections about folder hierarchy and naming.

MASTERING

If the mastering engineer is processing and manipulating the mixed masters within the computer, he or she should work in a 32-bit floating-point environment to take advantage of the extended headroom it provides and to help avoid errors, clipping, and distortion during internal digital signal processing.

The mastering engineer should provide:

Finished hi-resolution digital masters for HD downloads
Vinyl cutting masters (when required by the artist or record label)
44/16 files for CD production, normal downloads, and for MP3 aggregators
Special 44/24 or up to 192/24 Mastered for iTunes files for high-quality Apple downloads (The Mastered for iTunes applet will handle up to 192/24 files although Apple states that, "An ideal master will have 24-bit 96 kHz resolution.")

Documentation about the provenance of the mastered recording

Recommendations:

- The mastering engineer should provide hi-resolution audio files, 16-bit audio files for CD and download, and 24-bit Mastered for iTunes masters.
- The mastering engineer should provide timing, title, and technical information regarding the masters.
- The masters should be created within a 32-bit float environment when the mastering engineer is using any internal DSP.
- Whenever possible, the mastering engineer should provide documentation regarding the provenance of the recording.

BACKUP STRATEGIES

The process of tracking, mixing, and mastering in hi-resolution is very data intensive. An impressive amount of data is acquired, processed, edited, and stored during a normal production. Any session that's based on data is constantly at risk of losing that data. It is, therefore, extremely important to establish a reliable backup strategy. An effective backup plan is one that keeps data losses minimal or, ideally, inconsequential.

Many data-storage experts recommend following the 3-2-1 rule of data storage and backup. This rule states that all data should be backed up on three different drives or cloud storage sites, in two different physical locations, and that you should work from one copy. This is a simple but effective rule that can be a lifesaver when something goes wrong.

Backup solutions like Apple's Time Machine and Cloud storage solutions such as BackBlaze are useful and inexpensive. However, a bulletproof backup system takes advantage of a software utility such as ShotPutPro or Chronsync that includes a checksum, which confirms that each backup is a bit-for-bit clone of the original.

Data storage is one of the most important considerations in any high-paced recording environment. As the profile of your clients elevates, it becomes even more important to develop a bulletproof backup and archival system.

Recommendations:

- Follow the 3-2-1 Rule: Back up to three different drives in two locations and work from one copy.
- Including cloud storage in your backup routine is valuable.
- The best backup routines include a software utility that contains a checksum.

PROS AND CONS OF UPSAMPLING

Whereas upsampling lower-quality mastered audio files is unacceptable for delivery to the consumer, there can be instances where upsampling during production is acceptable.

PROS

Overdubs in Hi-Res. In the multitrack environment, when a production is being augmented or completed by overdubbing additional tracks, there is a benefit to upsampling the project to a higher resolution. Because the new tracks will be legitimately hi-res, the final product will benefit from the extended frequency content of the higher resolution audio and, depending on the density of the production, a possible increase in dynamic range. An example of this would be a scoring session where the pre-recorded tracks are delivered at 48/24 and then are upsampled to 96/24 or higher to record the orchestra. In such a scenario, the recording is legitimately hi-res if it stays hi-res through completion of the tracking, mixdown, and mastering.

Plug-ins. If new content is added at a higher resolution and if instantiated plug-ins can maintain the higher resolution, there is real value in upsampling the lower quality tracks to take advantage of the increased quality of the new tracks and instantiated plug-ins.

Outboard Hardware Processing. If new content is added at a higher resolution and then processed by high-quality full-bandwidth analog processors or digital devices that operate at hi-res audio specifications, there can be legitimate realization of increased audio quality.

CONS

Detrimental Effects of Conversion. In the opinion of many audio professionals, the conversion of digital audio from one resolution to another offers as much potential for a negative result as a positive result. Since upsampling adds samples that weren't in the original digital file, an algorithm is forced to interpolate (take an educated guess) at the time and amplitude location of the new sample. In other words, the recreated resolution is likely to be faithful to the source audio but it also might include some unintended anomalies, inaccuracies, or distortions relative to the source.

Potentially Better to Stay at the Source Resolution. Because there might be distortions or inaccuracies in the conversion process, some engineers and producers prefer not to upsample multitrack projects that originated at CD-quality, even if they're adding overdubs and have plug-ins capable of functioning in the hi-resolution range. These engineers would rather not risk the potential downside of the conversion process. However, this argument is losing some strength since modern DAWs do much more accurate conversions than their predecessors.

More Demand Placed on the Computer. Increases in sample rate and bit depth have a substantial impact on file size and they also dramatically increase demands on the CPU. Some computers and laptops aren't capable of functioning with large numbers of hi-resolution audio tracks.

Avoiding Consumer Confusion. It is important to have information available for those consumers who want to know more about provenance, the differences in hi-res audio specifications, and what it means to convert digital audio from one resolution to another. Although many audiophiles are very educated, there are also general consumers who would like to enjoy the highest audio quality but lack the knowledge and experience to feel confident that they are purchasing and enjoying all the benefits provided by hi-res audio.

Recommendations:

- Upsample a session if there will be additional hi-res overdubs recorded.
- If additional tracks will not be recorded, avoid upsampling the session.
- Maintain the integrity of each project's provenance.
- Upsampling will not make existing audio sound better.

THE PROVENANCE ISSUE

The Merriam-Webster dictionary defines provenance as, “the origin or source of something” and “the history of ownership of a valued object or work of art or literature.” As it relates to audio, provenance describes the relationship between the source audio file and the commercially delivered audio file. The provenance of an audio file answers four questions:

1. What was the format and resolution of the source recording?
2. What was the format of the mixed master?
3. What file conversions were performed on the master source?
4. What is the format and resolution of the commercially released audio file?

Labeling audio recordings as 96/24 or higher resolution when the source audio was 44.1/16 CD-quality or less is seriously problematic. This is a misrepresentation of hi-res audio and must be avoided. When it is not possible to start with hi-res master sources, the provenance must be accurately represented to the consumer.

Legitimate hi-res audio files should contain content that extends well above 20 kHz. However, files that have been upsampled from CD-quality audio will display a brick wall filter at 20 kHz because of the low-pass filter applied minimize problematic aliasing. Consumers are willing to pay extra for the highest quality audio but they deserve to get what they pay for. When a consumer purchases or streams a file labeled “Hi-Res,” that file must conform to the definition of hi-res, meaning that the file originated as hi-res and that conversions do not obscure or misrepresent the quality of the original source files.

Consumer confidence depends on reputable practices that deliver all audio files at the stated resolution. This is a serious concern for producers and engineers who might be tempted to upsample files that aren’t legitimately hi-resolution. It is extremely important to be aware that hi-res distributors and labels—and many technically adept consumers—use software to analyze the files they are sent. If the files are not truly hi-res, it will be discovered quickly. The distributor/label who has delivery requirements that specify hi-res audio will reject the files and could refuse payment to the producer and/or engineer.

HELPING THE CONSUMER AVOID CONFUSION: LABEL CODES

With the arrival of Compact Discs came the desire for a label code to help the consumer view the provenance of each CD. It was a simpler time because there were only two options: analog (A) or digital (D). A simple three-letter code was developed by SPARS that very clearly indicated the tracking, mixing, and mastering provenance in a way that was readily understood with a minimal learning curve. DDD simply meant that every stage of the recording process was digital, ADD indicated that the multitrack master was analog and that both mixing and mastering were done digitally, and so on.

With hi-res audio, it’s even more important to specify the provenance of each recording yet the variables are much more complicated. After lengthy discussions among the Digital Entertainment Group (DEG), the Consumer Technology Association (CTA), numerous major and independent labels, and The Recording Academy P&E Wing, an efficient set of labeling codes has been devised to precisely indicate the provenance of each commercially released hi-res recording.

There has been some justifiable skepticism about the release of legacy analog and digital recordings in hi-res format because the source resolution doesn’t always match the requirements to be called “hi-res.” The intent of this system of Label Codes is to provide accurate information to the consumer regarding legacy content and new releases.

To facilitate the effective release and implementation of these Label Codes, some consumer education will be required. The Master Source code always starts with MS because it tracks the provenance from the Master Source (MS) to the commercially delivered format. A label code specifies the provenance of a recording, including the file format, audio resolution (sample rate/bit depth), and any conversions or format-changing transfers.

The charts below provide an explanation of the label codes that will be used on hi-res downloads.

Code	Description
MS	Master Source
A	Analog master source
P	PCM master source
D	DSD master source
~	Transferred / Mixed To
First number	Sample rate
Second number	Bit depth

The codes in the following chart indicate the Master Source of each product. Additions can be made to the list using the nomenclature from the previous chart.

Code	Definition
MSP 48/20	Master Source is PCM 48 kHz/20-bit
MSP 48/24	Master Source is PCM 48 kHz/24-bit
MSP 88.2/24	Master Source is PCM 88.2 kHz/24-bit
MSP 96/24	Master Source is PCM 96 kHz/24-bit
MSP 176.4/24	Master Source is PCM 176.4 kHz/24-bit
MSP 192/24	Master Source is PCM 192 kHz/24-bit
MSP 192/24 ~ D5.6	Master Source is PCM 192 kHz/24-bit transferred to DSD 5.6448 MHz.
MSD 2.8	Master Source is DSD 2. 8224 MHz
MSD 5.6	Master Source is DSD 5.6448 MHz
MSD 11.2	Master Source is DSD 11.2896 MHz
MSD ~ PCM 352.8	Master Source is DSD transferred to PCM 352.8 kHz
MSA ~ P48/20	Master Source is Analog transferred to 48 kHz/20-bit PCM
MSA ~ P48/24	Master Source is Analog transferred to 48 kHz/24-bit PCM
MSA ~ P88.2/24	Master Source is Analog transferred to 88.2 kHz/24-bit PCM
MSA ~ P96/24	Master Source is Analog transferred to 96 kHz/24-bit PCM
MSA ~ P176.4/24	Master Source is Analog transferred to 176.4 kHz/24-bit PCM
MSA ~ P192/24	Master Source is Analog transferred to 192 kHz/24-bit PCM
MSA ~ D2.8	Master Source is Analog transferred to DSD 2. 8224 MHz
MSA ~ D5.6	Master Source is Analog transferred to DSD 5.6448 MHz
MSA ~ D11.2	Master Source is Analog transferred to DSD 11.2896 MHz

ANALOG SOURCE AUDIO
Considering the wide range of analog recording formats, accommodating every possible scenario in a provenance code would prove to be so cumbersome as to discourage the code's widespread use. Therefore, these codes are, by necessity, not fully comprehensive. But, they're intended to address the most likely provenance scenarios. For the purposes of the provenance code, it must be understood that MSA (Master Source Analog) implies professional tape formulation, machine alignment, and tape speed.

Substandard legacy analog sources such as cassette tape should be identified in the accompanying documentation and noted as part of the file's provenance. Given accurate information, consumers are likely to forgive a substandard source for the opportunity to own the very best rendition of a revered recording even if that source recording only offers a fraction of the maximum potential quality.

HI-RESOLUTION MUSIC AND AUDIO LOGOS
There are two logos currently used to indicate hi-resolution music and hardware.

The Hi-Res Music logo

The Hi-Res Music logo, administered by the RIAA, (Recording Industry Association of America) "…was created to help consumers find hi-resolution recordings available from digital music retailers in the U.S., Canada, and Europe for downloading or streaming. Under terms of the licensing agreement that has been issued to nearly a dozen companies, the logo must be accompanied by the name and resolution of a song's digital format."

The Hi-Res Audio logo

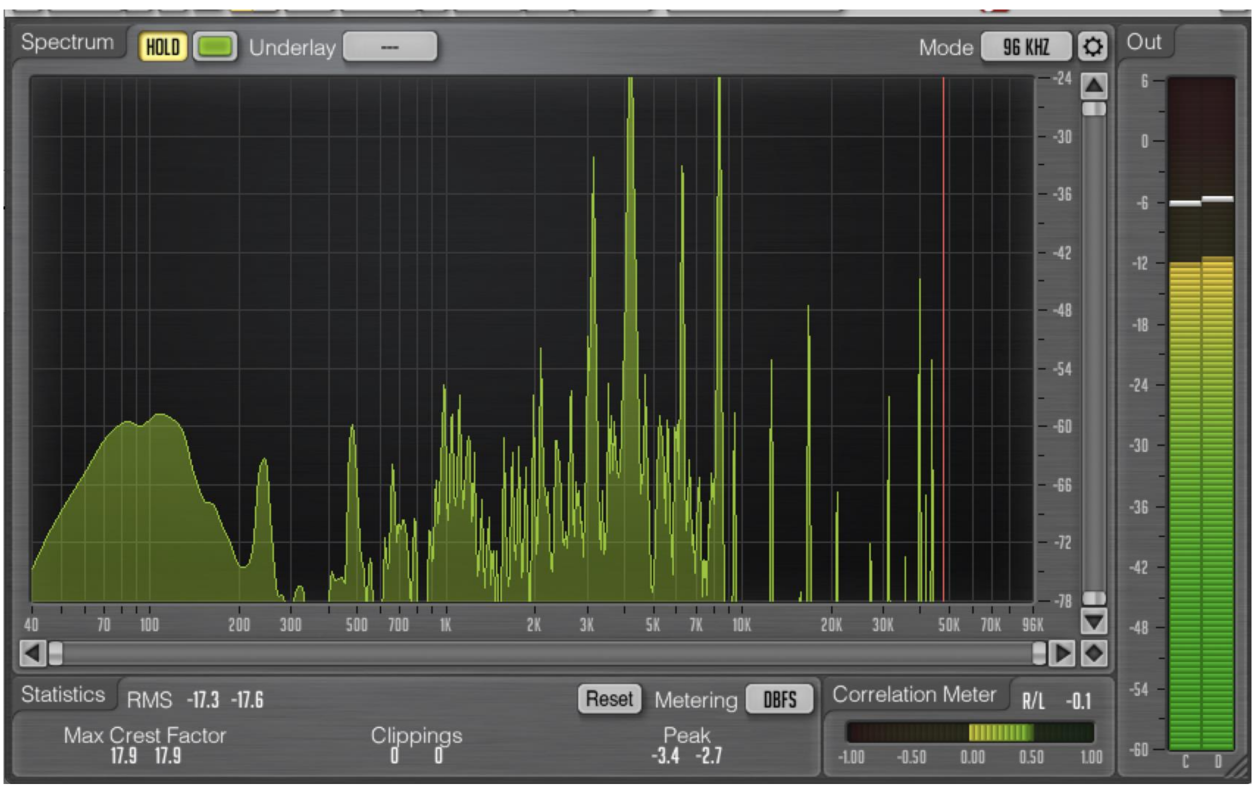
The Hi-Res Audio logo, administered by the Japan Audio Society, applies to hardware that is capable of recording, processing, and/or reproducing hi-resolution audio. As of April 2017, more than 60 companies are utilizing this logo worldwide on hundreds of different hi-res compatible devices.

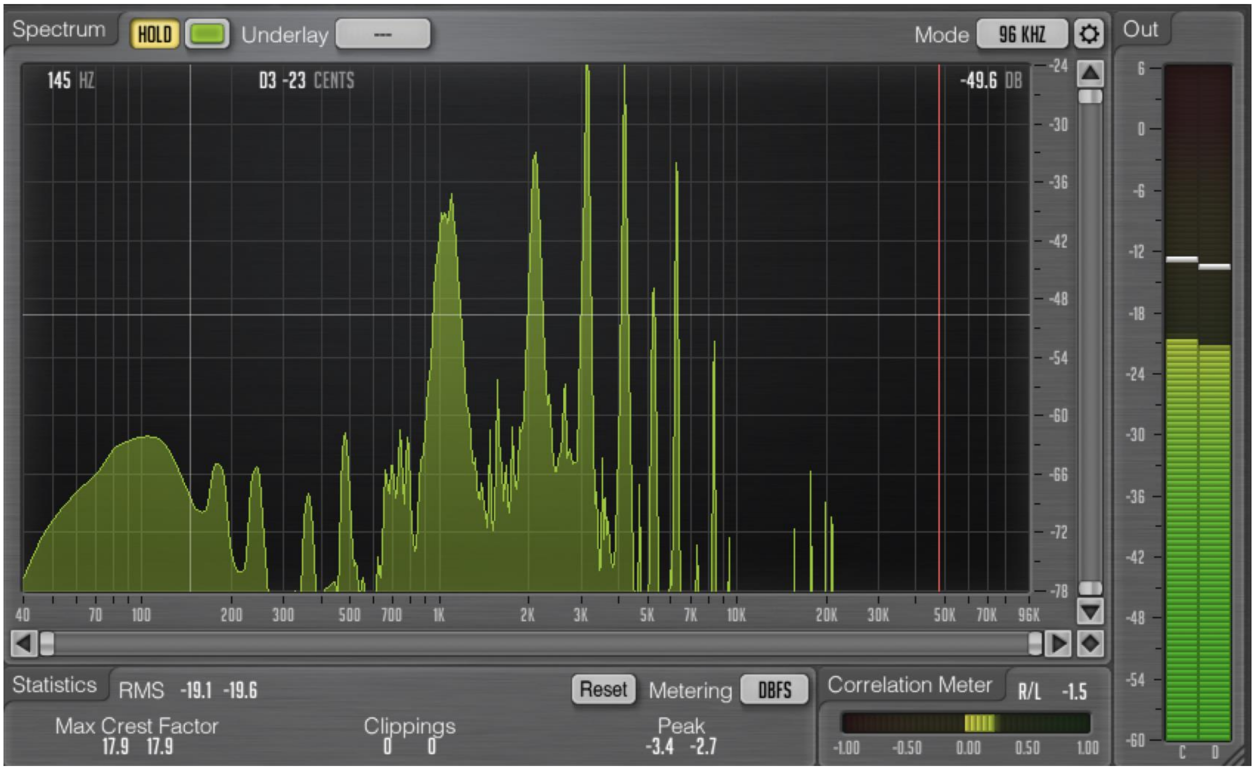
PLUG-INS

Plug-ins are inarguably interwoven into modern production and engineering. There is virtually never a modern computer-based recording that's created without their insertion at some point. They're important and ubiquitous; although, when plug-ins don't provide adequate audio quality, they become a limiting factor in the creation of truly hi-res audio. Modern plug-ins that are purely DSP-based are typically able to support audio resolutions up to and above 192/24. However, plug-ins that are modeled after classic or modern hardware are sometimes band-limited either due to lower-resolution models (48/24) or because the hardware itself imposes a band limit.

Although most modern plug-ins will function at audio resolutions of 96/24 and above, that doesn't necessarily mean that they can provide audio content that represents that resolution. For example, if you record all the multitrack sources at 96/24—and if all the microphones and processors provide content that extends into the upper bands—you should expect your mixes to contain frequency content stretching up toward 48 kHz. However, if your plug-ins are band-limited at a frequency below 48 kHz, the resulting audio quality is no longer truly representative of a 96/24 recording.

Notice the two screenshots that follow. The first shows the analysis of a 96/24 recording with frequencies represented up to 48 kHz. The second screenshot shows the same content after the insertion of a plug-in that has been band-limited at 24 kHz. The audio resolution after inserting the band-limited plug-in has been essentially changed from 96/24 to 48/24.





If you are creating a pure 96/24 production from tracking to mix to release, it is very important to be aware of the specific audio resolution of your plug-ins. Some of the industry's favorite plug-in manufacturers band-limit certain modeled plug-ins with low-pass filters set substantially below an ideal 96/24 resolution.

This brings us to a couple of very important questions:

1. Should we avoid any plug-ins modeled at 48/24 or band-limited below the session's resolution? Not necessarily, but be aware of the plug-ins you use and whether they are true 96/24 versions. All things considered, there are still plenty of uses within a hi-res mix for band-limited plug-ins, especially considering that many classic hardware processors limit the band by the nature of their design. The band-limited plug-in might accurately represent the results of passing the audio through the classic hardware. It's also important to recognize that, even if the plug-in imposes a band limitation, that limitation might still provide frequency content in the range defined as hi-res. Plus, the audio quality benefits from the increased accuracy provided by recording at 96/24, from a gentler slope at the Nyquist Frequency and from moving the low-pass cut-off up to 48 kHz

2. Should you run the entire mix through a band-limited plug-in if anything else in the mix is recorded at 96/24 or better? No. That would impose the limitations of that plug-in on your entire mix, effectively filtering anything above the band limit and decreasing the resolution of the entire recording. Since many labels now require 96/24 files, processing a complete mix through a band-limited plug-in that doesn't match the delivery specifications could lead to a refusal of payment to the engineer(s) and or producer(s).

Recommendations:

- Do not represent audio productions as "hi-res" if the source audio is CD-quality (44.1/16) or less.
- Use label codes to accurately convey the provenance of each project.
- Always research the provenance of the plug-ins used in each recording.
- If a plug-in has a resolution that is lower than the session, it can still be useful when used on a track inside the session.
- Don't process the entire mix through a plug-in that is band-limited below the session resolution.

MUSIC LABELS

Music labels have, in many ways, borne the brunt of dying consumer formats, reflected in diminishing sales and lack of consumer enthusiasm. They have also been the benefactors of each new technological advance because they were able to rerelease their complete catalogs in an entirely new and improved format. Is that a bad thing? Should they be criticized for it? Not necessarily, because there is a market demand. Posed with the option of being able to experience the studio-quality sound that hi-res audio provides for consumers, a large portion of the music-buying audience would very likely be grateful for the opportunity so long as the hi-res music is presented with an accurately documented provenance.

NEW MARKETS

Whereas the introduction of commercially released hi-res recordings represents an opportunity for record labels and online delivery services, it also represents a unique opportunity for artists, producers, engineers, songwriters, and anyone else in the business of providing hi-res content. There is so much more to this movement than rereleasing back catalogs. This is an opportunity to raise the sonic bar and to provide an exceptional listening experience for consumers who will, ideally, recognize the value in a better product and be drawn to a closer bond with artists and creative professionals.

Recommendations:

- As a member of the creative team, take advantage of the opportunity to supply new high-quality hi-res audio content.
- Strive to produce exceptional audio quality, raising the production standard and providing an exceptional experience for the consumer.

DELIVERY STANDARDS

The fact that most record labels and online music sellers require 96/24 masters reflects a definite industry shift. Music business entities recognize the commercial potential in the hi-res audio market. Noteworthy online sources have built large catalogs of hi-res content. As leaders in the hi-res audio movement, sites such as HDTracks.com, SuperHiRes.com, iTrax.com, ProStudioMasters.com, and DownloadsNow.net have quickly discovered the importance of transparency in the provenance of each product.

It is encouraging to see trusted sites specializing in newly created hi-resolution recordings. The excitement regarding hi-res audio far supersedes the release of legacy content. Offering the consumer new music in a way that exactly reflects what the musicians, engineers, and producers heard in the studio is exciting for everyone involved, whether they're the creative team or the consumer.

History reveals that the art, craft, and science of audio recording progresses with time. And, even if—as history would indicate—today's acceptance of hi-res recordings delivers us to the next era of convenience-centered technology, we can be assured that some future technological leap will raise the audio-quality bar higher than it is today. When that happens, we owe it to ourselves, our artists, and those who march on, to leave behind the highest possible quality. Future creative professionals—and likely our future selves—will be grateful.

New technologies like Master Quality Authenticated (MQA), which enables hi-res audio streaming, make a difference in the marketplace. MQA is a particularly important enhancement for the consumer who wants the best audio quality but who also enjoys the musical experience of sites such as Apple Music and Spotify. Delivered via Tidal.com, MQA has gained acceptance from many of the most highly respected members of the professional audio community.

Recommendations:

- Track, process, and mix sessions at 96/24 or better.
- Include the session provenance with each project.

WORKING STYLES

In the digital era, a wide range of working styles has become accepted. A modern workflow might start in the most elaborate commercial studio and stay there until the mixes and stems are delivered to a world-class mastering facility where they are mastered by an award-winning engineer and delivered to the consumer at the highest possible quality. However, it is more likely that a project will start on a laptop in a bedroom and stay there all the way through the rendering of the final master. From world-class studios to home project studios to the simplest mobile iOS-based studio to any combination of recording environments, the modern workflow is extremely diverse.

MIXING TO A SEPARATE SYSTEM

Some creative teams take advantage of multiple DAWs, consoles, and recorders to capture the highest possible audio quality. The following considerations are important when choosing to mix from one system to another:

1. Mixing from the analog output of an excellent analog mix desk to the analog input of a separate mix-recorder DAW can produce excellent audio quality because it includes the character and sonic personality imprinted by the specific console. When performing an analog-to-analog transfer, it is advisable to keep the sample rate and bit-depth as high as possible for the playback and record systems. It's ideal if the source files and mix recorder are set for 96/24 or better. Since this is an analog connection, there is no need to match the digital resolution of the source output and recorder input.
2. If the mix is all in-the-box, choosing to route mix from the digital output of the mix system to the digital inputs of a separate DAW can help the efficiency of the source and record systems. However, unlike the connection of analog devices mentioned above, this is a simple digital transfer that is not meant to color, enhance, or influence the sound of the source mix. Also, since this is a digital connection, the audio resolutions of the mix and record devices must match and they must run together in sample-accurate sync. If the DAWs are connected via a 24-bit AES connection be sure to dither the playback DAW if it has a bit depth greater than 24 bits.
3. Simply recording from the analog output of the mix system, through a converter to a higher resolution mix recorder, is just another form of upsampling. In most cases, this is not a good working style and is confusing to the consumer.
4. In some instances, upsampling through an analog console is justified but it needs to be done in full disclosure to the consumer. An analog console generates its own artifacts that add personality to the sound. Components such as VCAs, transformers, and pan pots contribute crosstalk, distortion, and phase shift that frequently contribute positive

aspects to a musical sound. With a sufficient disclaimer, it is likely that consumers will understand that the most faithful way to capture these complex audio anomalies and imperfections is with hi-res audio.

Recommendations:

- Mixing to a separate hi-res system from the analog outputs of the source system to the analog inputs of the recorder can produce excellent results.
- Verify that the source system and session are legitimately hi-res.
- Verify that the mixdown system and session are legitimately hi-res.
- Confirm that any plug-in included on the final mix bus is not bandlimited and therefore does not decrease the audio resolution.
- When the source audio is not hi-res but it's being mixed through a high-quality analog console, there can still be a sonic advantage to capturing the mix with a hi-res mix recorder. However, it's imperative to add a disclaimer and label code that indicates the accurate provenance.

IN THE BOX

Mixing in the box (without leaving the host computer) keeps all the audio processing within the DAW software domain. Working in hi-res in-the-box requires a robust computer capable of keeping up with the processing demands of high track counts and hi-res plug-ins. An advantage to staying within the computer system for tracking and mixing is the ability to maintain a constant audio resolution throughout the project. As mentioned previously in this document, it is very important, no matter what process or workflow is used, to become informed about any band limits or other limitations that exist in your DAW software, digital hardware, or plug-ins.

When mixing in-the-box, DSP is constantly being used for plug-ins, clip effects, clip gain, and so on. Therefore, it is advisable to work in 32-bit float mode to take advantage of the additional headroom and accuracy that it offer.

Recommendations:

- Be certain to confirm the provenance of any plug-in used in the box.
- For large sessions that include several plug-ins, use a computer with a fast processor, the maximum amount of RAM, and fast storage devices.
- For in-the-box mixes, work in 32-bit float mode.

SUMMARY OF RECOMMENDATIONS

Hi-Resolution Audio and Processing Power

- Keep current. Update computers, software, interfaces, and drives as frequently as is practical.
- Increase processing capacity substantially by incorporating additional cards and interfaces that share the processing requirements with the CPU.
- Set your system default preferences to record at 96/24 unless intentionally set otherwise.

Sample Rate and Bit Depth

- Record audio at the highest practical sample rate for your system.
- Record 24-bit audio.
- Upsampling will not increase audio quality.

Hi-Resolution Hardware and Software Considerations

- Design the system for your specific types of recording applications. In the current computing era, a small system isn't an inferior system—it's just limited in capacity.
- In a quick paced workflow with high capacity sessions, it's important to maintain current computers, the maximum RAM the computer will hold, current software, and an up-to-date interface that augments the system's processing capacity.
- Use a 64-bit OS and software.
- Use an interface with additional processing power and an accommodation for zero-latency monitoring.
- Use a USB 3.0/3.1 or Thunderbolt interface for dramatically reduced latency

Drive Considerations

- Use 7200 RPM hard drives.
- Use USB-3.0 or faster data interface connections.
- When possible, use an internal system drive (fast HDD or SSD) for applications and operating system software, along with a separate internal SATA/eSATA 7200 RPM HDD for the recording session.
- When the recording computer only contains one internal drive, use a fast internal drive (HDD or SDD) for the system and applications along with an external USB-3 or Thunderbolt drive for the session files.

Hi-Resolution Recording and Mixing Considerations

- Be intentional about sample rate and bit-depth settings.
- Set up sessions for 96 kHz/24-bit audio. If track count or processing requirements demand it, record 48 kHz/24-bit audio.
- When continuing a session that someone else started, verify the sample rates of existing audio files match the new session settings.

Quantization

- Use the highest practical sample rate when recording digital audio for the most faithful capture and reproduction of the source wave form.
- Use 24-bit words during recording for greater accuracy and fewer quantization errors.

Dither

- When tracking audio, don't use dither.
- Reserve dithering for the mastering process where its use can be selected and applied for the best results.
- Record 24-bit audio files to benefit from its increased dynamic range and decreased noise floor.

32-Bit Floating-Point Audio Resolution

- When the audio tracks will be processed internally within the DAW, use 32-bit float audio.
- Be sure to add dither to a bounce or to the output of a 32-float DAW when creating 24bit or 16 bit files from the session.
- 24-bit audio settings are appropriate for sessions where there will be no internal processing applied by plug-ins or virtual instruments, and where there will be no processing performed by clip gain, clip effects, and so on.
- If the tracking computer can't handle the load of tracking large sessions in 32-bit float mode, track at 24-bits and then switch to 32-bit float for processing tasks that take advantage of the additional headroom provided in 32-bit float mode.

- When routing the outputs of a digital system to an analog mixing desk and then re-recording to the analog inputs of a separate workstation, record in 24-bits if there will be no further processing.

Processing Resources, Stems, Levels, and Documentation

- To regain processing capacity, create submixes/stems within your mix and then deactivate the source tracks.
- Create a separate folder of mixes that includes stems for all major instrument and vocal groups, a stereo mix, a stereo mix without vocals, and any other stereo or surround mixes and stems that are part a of a specified set of deliverable files or that might realistically be required for future use.
- Review contractual agreements regarding the resolution and organization of all mixes and stems.
- The mix engineer should generally work in a 24-bit environment when mixing via the analog domain; 32-bit float can be helpful if mixing entirely in-the-box for increased audio quality
- The mix engineer should provide documentation including the provenance of the audio contained in the recording.
- Conform naming and file structure to the recommendations include in the section covering folder hierarchy and naming.

Mastering

- The mastering engineer should provide hi-resolution audio files, 16-bit audio files for CD and download, and 24-bit Mastered for iTunes masters.
- The mastering engineer should provide timing, title, and technical information regarding the masters.
- The masters should be created within a 32-bit float environment when the mastering engineer is using any internal DSP.
- Whenever possible, the mastering engineer should provide documentation regarding the provenance of the recording.

Backup Strategies

- Follow the 3-2-1 Rule: Back up to three different drives in two locations and work from one copy.
- Including cloud storage in your backup routine is valuable.
- The best backup routines include a software utility that contains a checksum.

Pro and Cons of Upsampling

- Upsample a session if there will be additional hi-res overdubs recorded.
- If additional tracks will not be recorded, avoid upsampling the session.
- Maintain the integrity of each project's provenance.
- Upsampling will not make existing audio sound better.

The Provenance Issue

- Do not represent audio productions as "hi-res" if the source audio is CD-quality or less.
- Use label codes to accurately convey the provenance of each project.
- Always research the provenance of the plug-ins used in each recording/mix.
- If a plug-in has a resolution that is lower than the session, it can still be useful when used on a track inside the session.
- Don't process the entire mix through a plug-in that is band-limited below the session resolution.

New Markets

- As a member of the creative team, take advantage of the opportunity to supply new high-quality hi-res audio content.
- Strive to produce exceptional audio quality, raising the production standard and providing an exceptional experience for the consumer.

Delivery Standards

- Track, process, and mix sessions at 96/24 or better.
- Include the session provenance with each project.

Mixing to a Separate System

- Mixing to a separate hi-res system from the analog outputs of the source system to the analog inputs of the recorder can produce excellent results.
- Verify that the source system and session are legitimately hi-res.
- Verify that the mixdown system and session are legitimately hi-res.
- Confirm that any plug-in included on the final mix bus is not bandlimited and therefore does not decrease the audio resolution.
- When the source audio is not hi-res but it's being mixed through a high-quality analog console, there can still be a sonic advantage to capturing the mix with a hi-res mix recorder. However, it's imperative to add a disclaimer and label code that indicates the accurate provenance.

In the Box

- Be certain to confirm the provenance of any plug-in used in an in-the-box mixdown.
- For large sessions that include several plug-ins, use a computer with a fast processor, the maximum amount of RAM, and fast storage devices.
- For in-the-box mixes, work in 32-bit float mode.

SUMMARY

The members of this committee and the audio professionals we interviewed share a deep desire to do the best work possible and to share that work with listeners. Virtually every person who has contributed to the creation of these recommendations has performed detailed listening tests, comparing different equipment, file types, and audio resolutions. Some of our interviewees recorded complete sessions at different resolutions and then compared the sounds between those sessions. Many also upsampled and downsampled sessions to hear how each conversion effected the resulting sound quality. In the end, everyone felt that hi-resolution audio most accurately reflected what they heard in the studio, whether recording, mixing, or mastering.

RECORDING ACADEMY™
PRODUCERS & ENGINEERS WING®