MP3 vs AAC vs FLAC vs CD

By John Atkinson • Posted: Mar 8, 2008
As Wes Phillips recently reported on this website, CD sales are down and legal downloads of audio files are up. *Stereophile* has been criticized more than once for not paying enough attention to the subjects of MP3 and other compressed file formats, such as AAC, and for offering no guidance *at all* to readers about how to get the best sound quality from compressed downloads.

These criticisms are correct. We don't.

The reason is simple: Although they are universally described in the mainstream press as being of "CD quality," MP3s and their lossy-compressed ilk do *not* offer sufficient audio quality for serious music listening. This is not true of lossless-compressed formats such as FLAC, ALC, and WMA lossless—in fact, it was the release of iTunes 4.5, in late 2003, which allowed iPods to play lossless files, that led us to welcome the ubiquitous Apple player to the world of high-end audio. But *lossy* files achieve their conveniently small size by discarding too much of the music to be worth considering.

In the past, we have discussed at length the reasons for our dismissal of MP3 and other lossy formats, but recent articles in the mainstream press promoting MP3 (examined in Michael Fremer's "The Swiftboating of Audiophiles") make the subject worth re-examining.

Lossless *vs* Lossy
The file containing a typical three-minute song on a CD is 30–40 megabytes in size. A 4-gigabyte iPod could therefore contain just 130 or so songs—say, only nine CDs' worth. To pack a useful number of songs onto the player's drive or into its memory, some kind of data compression needs to be used to reduce the size of the files. This will also usefully reduce the time it takes to download the song.

Lossless compression is benign in its effect on the music. It is akin to LHA or WinZip computer data crunchers in packing the data more efficiently on the disk, but the data you read out are the same as went in. The primary difference between lossless compression for computer data and for audio is that the latter permits random access within the file. (If you had to wait to unZip the complete 400MB file of a CD's content before you could play it,

you would rapidly abandon the whole idea.) You can get reduction in file size to 40–60% of the original with lossless compression—the performance of various lossless codecs is compared here and here—but that increases the capacity of a 4GB iPod to only 300 songs, or 20 CDs' worth of music. More compression is necessary.

The MP3 *codec* (for COder/DECoder) was developed at the end of the 1980s and adopted as a standard in 1991. As typically used, it reduces the file size for an audio song by a factor of 10; *eg,* a song that takes up 30MB on a CD takes up only 3MB as an MP3 file. Not only does the 4GB iPod now hold well over 1000 songs, each song takes less than 10 seconds to download on a typical home's high-speed Internet connection.

But you don't get something for nothing. The MP3 codec, and others that achieve similar reductions in file size, are "lossy"; *ie*, of necessity they eliminate some of the musical information. The degree of this degradation depends on the data rate. Less bits always equals less music.

As a CD plays, the two channels of audio data (not including overhead) are pulled off the disc at a rate of just over 1400 kilobits per second. A typical MP3 plays at less than a tenth that rate, at 128kbps. To achieve that massive reduction in data, the MP3 coder splits the continuous musical waveform into discrete time chunks and, using Transform analysis, examines the spectral content of each chunk. Assumptions are made by the codec's designers, on the basis of psychoacoustic theory, about what information can be safely discarded. Quiet sounds with a similar spectrum to loud sounds in the same time window are discarded, as are quiet sounds that are immediately followed or preceded by loud sounds. And, as I wrote in the February 2008 "As We See It," because the music must be broken into chunks for the codec to do its work, transient information can get smeared across chunk boundaries.

Will the listener miss what has been removed? Will the smearing of transient information be large enough to mess with the music's meaning? As I wrote in a July 1994 essay, "if these algorithms have been properly implemented with the right psycho-acoustic assumptions, the musical information represented by the lost data will not be missed by most listeners.

"That's a mighty big 'if.'"

And while lossy codecs differ in the assumptions made by their designers, all of them discard—permanently—real musical information that would have been audible to some listeners with some kinds of music played through some systems. These codecs are not, in the jargon, "transparent," as can be demonstrated in listening tests (footnote 1).

So to us at *Stereophile*, the question of which lossy codec is "the best" is moot. We recommend that, for serious listening, our readers use uncompressed audio file formats, such as WAV or AIF—or, if file size is an issue because of limited hard-drive space, use a lossless format such as FLAC or ALC. These *will* be audibly transparent to all listeners at all times with all kinds of music through all systems.

Putting Codecs to the Test
Do I have any evidence for that emphatic statement?

For an article published in the March 1995 issue of *Stereophile*, I measured the early PASC, DTS, and ATRAC lossy codecs and put four of the test signals used for that article on our *Test CD 3* (Stereophile STPH006-2). For the present article, I used two of those signals, tracks 25 and 26 on *Test CD 3*. But first, to set a basis for comparison, I used that most familiar of test signals: a 1kHz tone.

The spectrum of this tone, played back from CD, is shown in fig.1. The tone is the sharply defined vertical green line at the left of the graph. There are no other vertical lines present, meaning that the tone is completely free from distortion. Across the bottom of the graph, the fuzzy green trace shows that the background noise is

uniformly spread out across the audioband, up to the 22kHz limit of the CD medium. This noise results from the 16-bit Linear Pulse Code Modulation (LPCM) encoding used by the CD medium. Each frequency component of the noise lies around 132dB below peak level; if these are added mathematically, they give the familiar 96dB signal/noise ratio that you see in CD-player specifications.
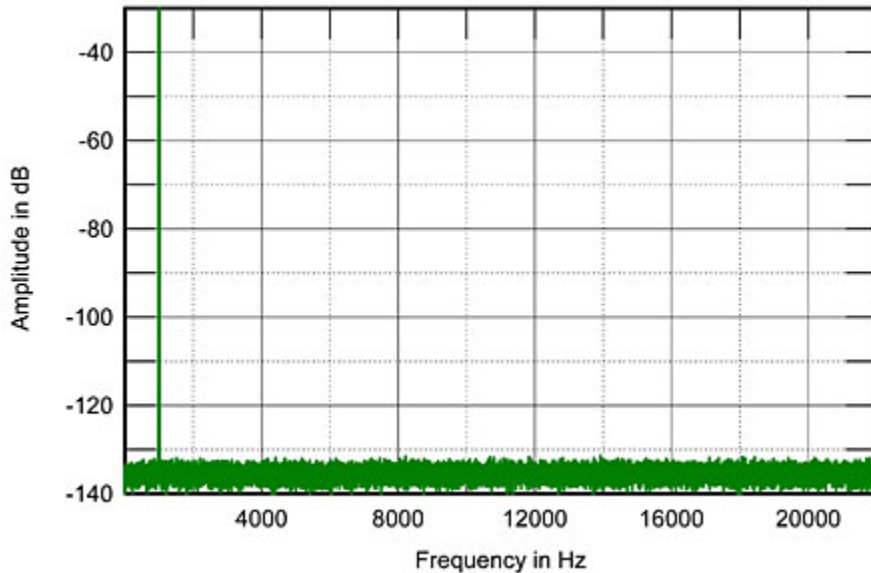

Fig.1 Spectrum of 1kHz sinewave at –10dBFS, 16-bit linear PCM encoding (linear frequency scale, 10dB/vertical div.).

Fig.2 shows the spectrum of this tone after it has been converted to an MP3 at a constant bit rate of 128kbps. (The MP3 codec I used for this and all the other tests was the Fraunhöfer, from one of the original developers of the MP3 technology.) The 1kHz tone is now represented by the dark red vertical line at the left of the graph. Note that it has acquired "skirts" below –80dB. These result, I believe, from the splitting of the continuous data representing the tone into the time chunks mentioned above, which in return results in a very slight uncertainty about the exact frequency of the tone. Note also that the random background noise has disappeared entirely. This is because the encoder is basically deaf to frequency regions that don't contain musical information. With its very limited "bit budget," the codec concentrates its resources on regions where there *is* audio information. However, a picket fence of very-low-level vertical lines can be seen. These represent spurious tones that result, I suspect, from mathematical limitations in the codec. Like the skirts that flank the 1kHz tone, these will not be audible. But they do reveal that the codec is working hard even with this most simple of signals.
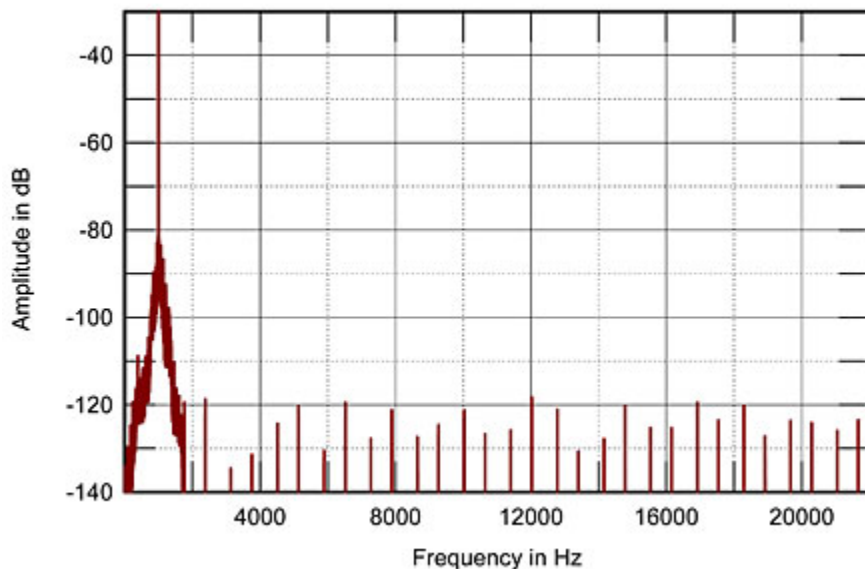
Continued Below:

Fig.2 Spectrum of 1kHz sinewave at –10dBFS, MP3 encoding at 128kbps (linear frequency scale, 10dB/vertical div.).

But what about when the codec is dealing not with a simple tone, but with music? One of the signals I put on *Test CD 3* (track 25) simulates a musical signal by combining 43 discrete tones with frequencies spaced 500Hz apart. The lowest has a frequency of 350Hz, the highest 21.35kHz. This track sounds like a swarm of bees, but more important for a test signal, it readily reveals shortcomings in codecs, as spuriae appear in the spectral gaps between the tones..

Footnote 1: Something I have rarely seen discussed is the fact is that because all compressed file formats, both lossless and lossy, effectively have zero data redundancy, they are much more vulnerable than uncompressed files to bit errors in transmission.

MP3 vs AAC vs FLAC vs CD Page 2

For reference, fig.3 shows the spectrum of the signal on the CD. Other than the well-defined green vertical lines representing the tones and the uniform background noise, the spectrum is clean. Important points to note with this graph are that 1) all musical fundamentals lie to the left of the 4000Hz (4kHz) mark; 2) the region between the next three divisions, 4kHz, 8kHz and 16kHz, is where musical harmonics and the "air" on a recording reside; and 3) the region above 16kHz—more than a quarter of the horizontal scale—will be inaudible to most adults.
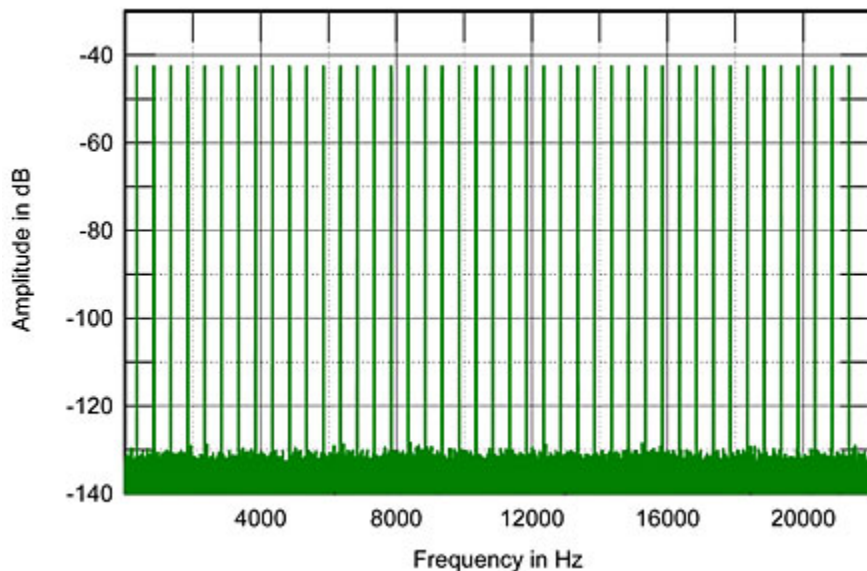
Fig.3 Spectrum of 500Hz-spaced multitone signal at –10dBFS, 16-bit linear PCM encoding (linear frequency scale, 10dB/vertical div.).

Fig.4 shows the spectrum of this demanding signal as preserved by lossless coding, in this case the popular FLAC codec (at its slowest "8" setting). To all intents and purposes, it is identical to the spectrum of the original CD. The lossless coding is indeed lossless, which I confirmed by turning the FLAC file back to WAV (LPCM) and doing a bit-for-bit comparison with the signal used to generate fig.3. The bits were the same—the music will also be the same!
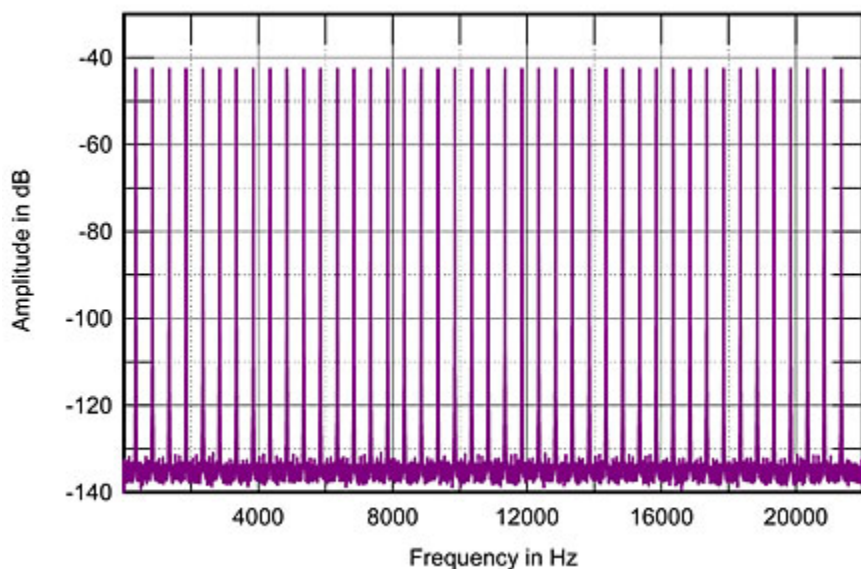

Fig.4 Spectrum of 500Hz-spaced multitone signal at –10dBFS, FLAC encoding (linear frequency scale, 10dB/vertical div.).

How did the MP3 codec running at 128kbps cope with the multitone signal? The result is shown in fig.5. The dark red vertical lines represent the tones, and none are missing; the codec has preserved them all, even those at the top of the spectrum that will be inaudible to almost every listener. But the background noise components, which on the CD all lay at around –132dB, have all risen to the –85dB level. With its limited bit budget, the codec can't encode the tones without reducing to almost half the 16 bits of CD resolution. Even with the

masking of this noise in the presence of the tones implied by psychoacoustic theory, this degradation most certainly will be audible on music. Yes, this kind of signal is very much a worst case, but this result is *not* "CD quality."
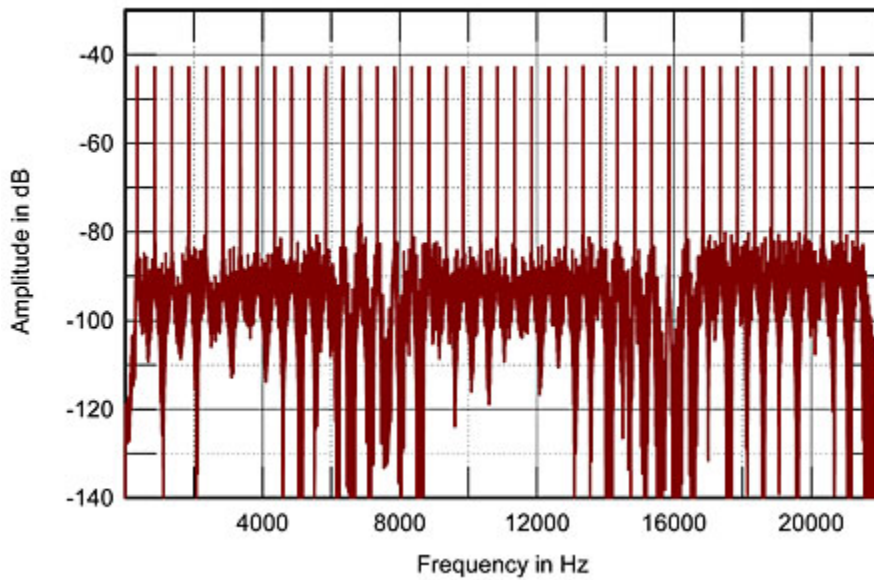

Fig.5 Spectrum of 500Hz-spaced multitone signal at –10dBFS, MP3 encoding at 128kbps (linear frequency scale, 10dB/vertical div.).

How about other lossy codecs? I looked at how the iTunes AAC codec (a version of MPEG 4, a later development than MP3) performed on this test, running at the same 128kbps. The result is shown in fig.6. At first it looks very similar, to fig.5, but there are significant differences. Note that almost all the tones above 18kHz are missing and that those above 16kHz are increasingly rolled off. The designers of the codec obviously decided not to waste the limited bit budget by encoding information that would most probably not be heard even from the CD. Instead, they devoted those resources to a more accurate depiction of the musically significant regions at lower frequencies. You can see in this graph that, below 4kHz, the noise level is 10–20dB lower than with the MP3 codec (though perhaps more "granular"). In effect, in the frequency region that is musically most important, an AAC file with this test signal has 2–3 bits more resolution than an MP3 file with the same bit rate. The AAC noise floor is higher than the MP3 noise floor between 8kHz and 18kHz, but given the physics of human hearing, this is insignificant.
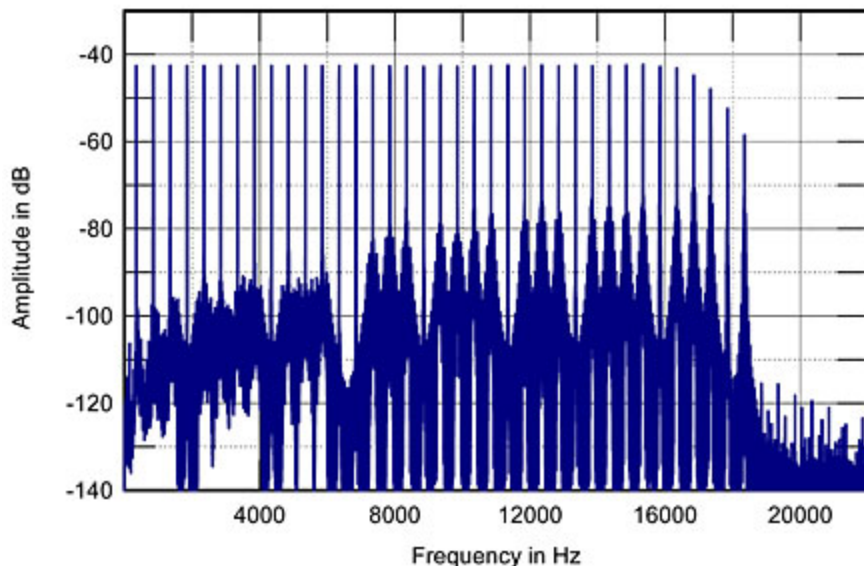
Fig.6 Spectrum of 500Hz-spaced multitone signal at –10dBFS, AAC encoding at 128kbps (linear frequency scale, 10dB/vertical div.).

The degradation is dependent on bit rate—the higher the bit rate, the bigger the bit budget the codec has to play with and the fewer data must be discarded. I therefore repeated these tests with both lossy codecs set to 320kbps. The file size is three times that at 128kbps, though still significantly smaller than a lossless version, but are we any closer to "CD quality"?

Fig.7 shows the spectrum produced by the MP3 encoder running at 320kbps. (This is the format used by Deutsche Grammophon for its classical downloads.) Again, all the tones are reproduced correctly, and the noise has dropped by around 6dB or so at higher frequencies and up to 15dB at lower frequencies. But it is still not quite as low as AAC at 128kbps below 1kHz or so.
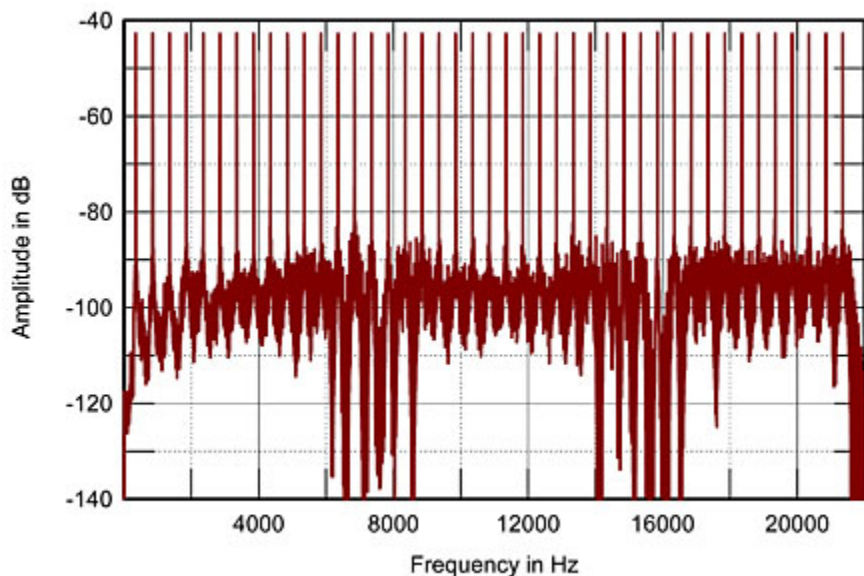

Fig.7 Spectrum of 500Hz-spaced multitone signal at –10dBFS, MP3 encoding at 320kbps (linear frequency scale, 10dB/vertical div.).

AAC at 320kbps now encodes all the tones, even the inaudible ones at the top of the audioband (fig.8). The noise floor is quite high above 18kHz, but—and it's a big "but"—the noise-floor components have dropped to below –110dB below 16kHz, and to below –120dB for the lower frequencies. Though some spectral spreading can be seen at the bases of the vertical lines representing the tones, it is relatively mild. Given the bigger bit budget at 320kbps, the AAC codec produces a result that may well be indistinguishable from CD for some listeners some of the time with some music. But the spectrum in fig.8 is still not as pristinely clean as that of the original CD in fig.3.
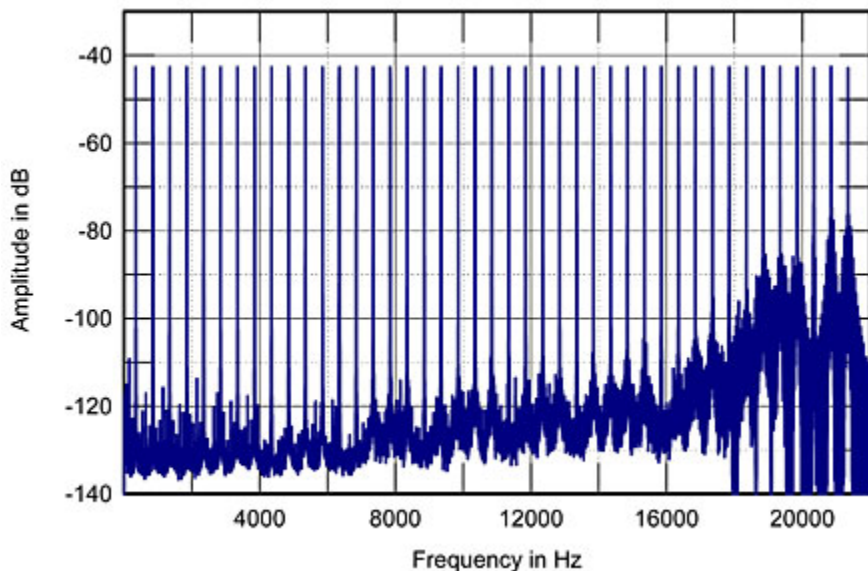

Fig.8 Spectrum of 500Hz-spaced multitone signal at –10dBFS, AAC encoding at 320kbps (linear frequency scale, 10dB/vertical div.).

For my final series of tests, I used *Test CD 3*'s track 26, which replaces some of the tones in track 25 with silence. The spectrum of the CD original is shown in fig.9. You can see clean vertical lines representing the tones, with silence in between. You can see the random background noise below –130dB, as expected. Also as expected, encoding with FLAC gave the identical spectrum, so I haven't shown it.
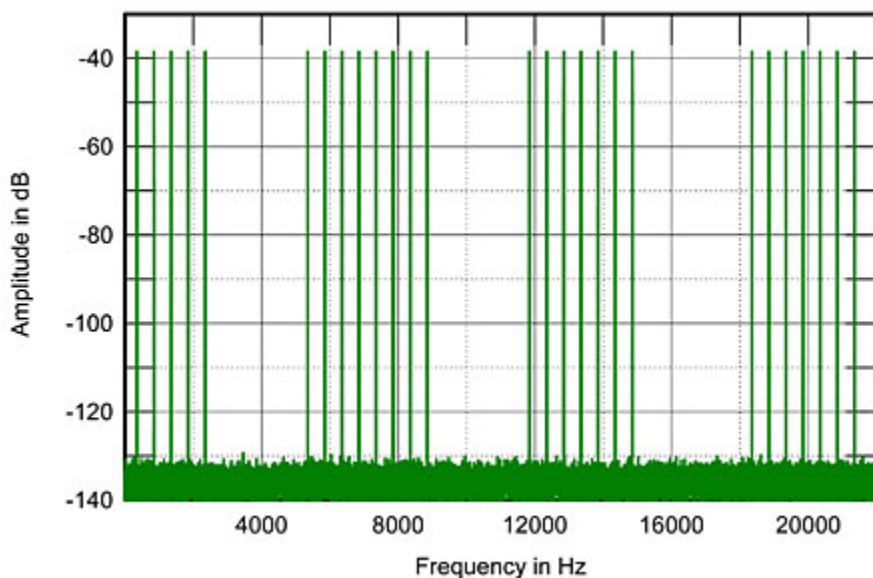
Fig.9 Spectrum of multitone signal with frequency gaps at –10dBFS, 16-bit linear PCM encoding (linear frequency scale, 10dB/vertical div.).

MP3 at 320kbps gave the spectrum shown in fig.10. All the tones are present, but if you look closely, you can see some extra ones, at low levels. The noise also leaks into the spaces between the groups of tones. AAC at 320kbps gave the spectrum in fig.11. Again, there is much more noise and less resolution above 18kHz, where it doesn't really matter. Again, the noise around the groups of tones is lower than with MP3 at the same bit rate. Some low-level spurious tones can be seen in the spaces between the groups of tones; though there are more than with MP3, these are all lower in level. The noise floor between the groups is also higher in level than with MP3, but is still low in absolute terms.
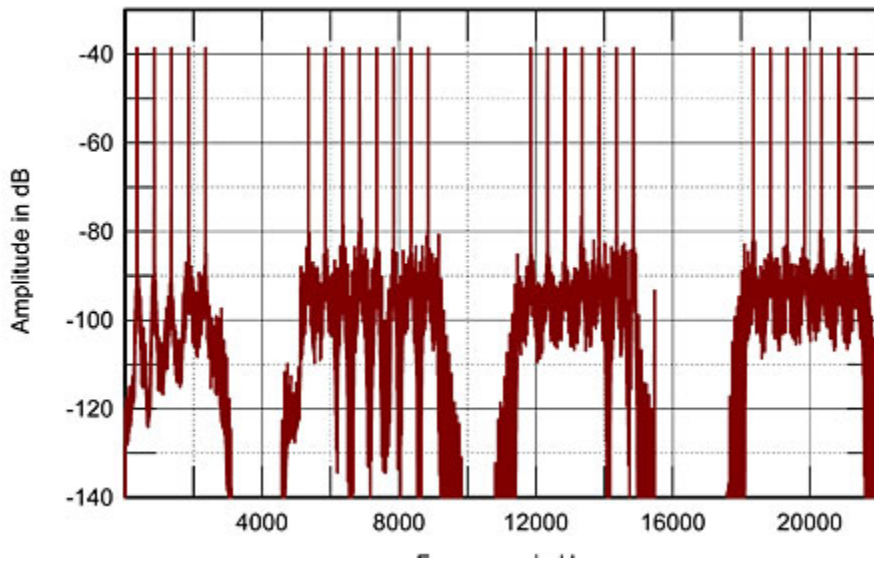

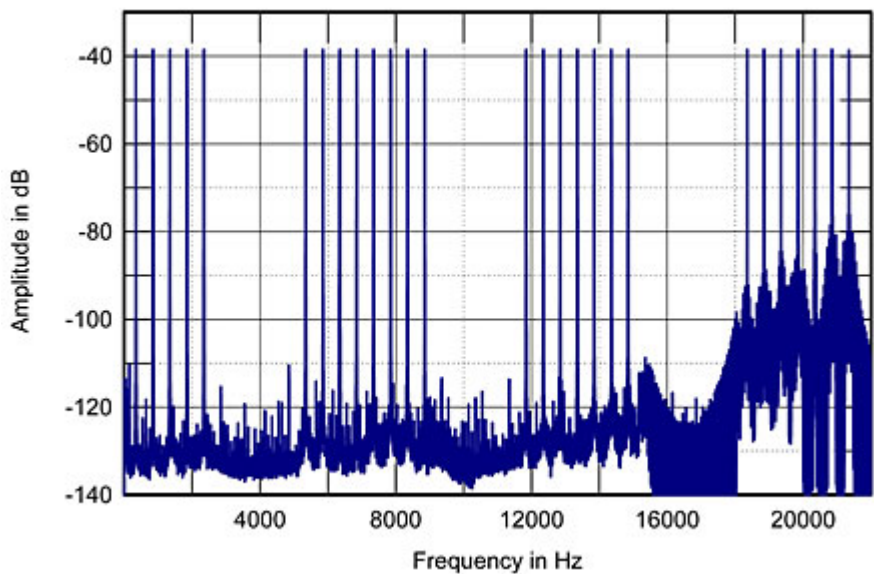Fig.10 Spectrum of multitone signal with frequency gaps at –10dBFS, MP3 encoding at 320kbps (linear frequency scale, 10dB/vertical div.).

Fig.11 Spectrum of multitone signal with frequency gaps at –10dBFS, AAC encoding at 320kbps (linear frequency scale, 10dB/vertical div.).

What does all this mean?
Basically, if you want true CD quality from the files on your iPod or music server, you must use WAV or AIF encoding or FLAC, ALC, or WMA Lossless. Both MP3 and AAC introduce fairly large changes in the measured spectra, even at the highest rate of 320kbps. There seems little point in spending large sums of money on superbly specified audio equipment if you are going to play sonically compromised, lossy-compressed music on it.

It is true that there are better-performing MP3 codecs than the basic Fraunhöfer—many audiophiles recommend the LAME encoder—but the AAC codec used by iTunes has better resolution than MP3 at the same bit rate (if a little noisier at the top of the audioband). If you want the maximum number of files on your iPod, therefore, you take less of a quality hit if you use AAC encoding than if you use MP3. But "CD quality"? Yeah, right!

Article taken from Stereophile.com

http://www.stereophile.com/features/308mp3cd