naim

world class sound...

Asynchronous USB

January 2013

Naim Audio Ltd • Southampton Road • Salisbury • Wilts SP1 2LN • England

Introduction

A decade ago the outboard digital-to-analogue converter (DAC) – once a popular means of upgrading a CD player – had virtually disappeared as an audiophile product category. Today, with burgeoning interest in computer audio, the outboard DAC is enjoying a new lease of life.

But with the computer as a music source comes the need for a new audio interface. The S/PDIF and AES/EBU digital interfaces which used to suffice for an outboard DAC are not generally available on a computer, and where an S/PDIF output is provided it is usually via the Toslink optical interface and may not be capable of working at up to 192kHz sampling rate.

The obvious interface to use with a computer music source is USB – Universal Serial Bus. Widely employed to connect all sorts of peripheral, it is as universal as its name suggests and capable – in 2.0 or 'Hi-Speed' form – of a maximum data rate of 480Mbits per second, which is ample to handle the 9.2Mbits/sec required by 24-bit, 192kHz stereo.

To achieve the highest audio quality via USB, though, demands that the interface be operated in the optimum way. USB supports three different modes of data transmission – adaptive, synchronous and asynchronous – but only the last of these is compatible with state-of-the-art audio. First-generation USB-equipped DACs typically used one of the other modes, and were often restricted to 16-bit resolution and 48kHz maximum sampling rate. As a result, they delivered poor sound quality. To understand why asynchronous mode is superior we first have to understand jitter and how it affects sound quality.

Jitter

When, during recording, an analogue audio signal is converted to digital or, during replay, the digital signal is converted back to analogue, timing errors – known as jitter – can add unwanted distortion. It isn't enough that the analogue-to-digital converter at the recording end of the chain accurately converts the signal amplitude at each sampling point to the appropriate digital code, or that the digital-to-analogue converter at the replay end performs the reverse process with equivalent precision. It is also vital that the time period between successive samples, during signal capture or reconstruction, be consistent to a very high degree of accuracy, otherwise the time displacement will modify the signal waveform, as the simple example in Figure 1 shows. In fact the timing error gives rise to frequency intermodulation distortion in the output signal.
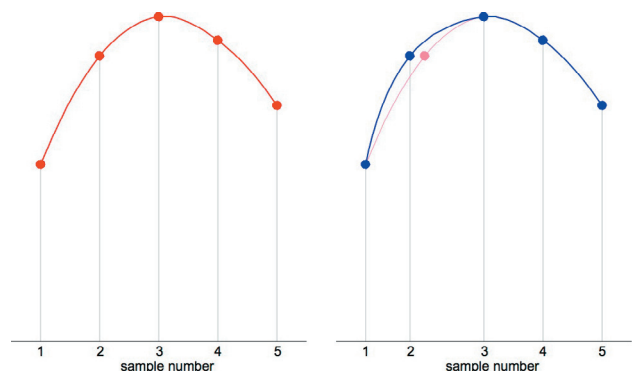

Figure 1

Figure 1. A simple illustration of how jitter introduces distortion. The left diagram shows five successive samples, equally spaced as they should be, defining the desired output waveform (red trace, with the dots showing the sample points). The right diagram shows how if one of the samples is displaced in time with respect to the others – here it's the second sample, which occurs a quarter of a sampling period too early – the output waveform (blue trace) is altered even though the sample values remain exactly the same

At the CD sampling rate of 44.1kHz (44,100 samples per second) the sampling period – the time between each successive sample – is 22.7 millionths of a second (22.7μs or 0.0000227s). At 192kHz sampling rate the sampling period reduces to 5.2μs. In order for jitter not to introduce significant distortion, the variation in sampling period must be limited to tiny fractions of these figures, ideally to within 40 picoseconds (40ps), ie 40 millionths of a millionth of a second (0.00000000004s).

We can make an estimate of the required timing accuracy by taking the example of a full-scale 10kHz sine wave and calculating what sample timing error results in a signal amplitude error equal to 1 least significant bit (LSB), which is the inherent amplitude resolution of the digital signal, ignoring the linearising effect of dither noise. For 16-bit encoding, the maximum rate of change of amplitude of a full-scale 10kHz sine wave – which occurs at the zero-crossing point – is equivalent to 2059 LSB per microsecond. To ensure that the amplitude error due to jitter is limited to 1 LSB or less therefore requires that the sample timing error be limited to a maximum of 1/2059μs or 486 picoseconds (486ps). In practice we'd like to achieve, say, 10 times better than that, ie 48.6ps peak, equivalent to 34.3ps rms.

Another way to consider the effect of jitter is to use spectrum analysis to look at the frequency dependent intermodulation distortion that it introduces. For a given value of jitter the resulting distortion worsens as signal frequency increases, so it is common to test for jitter in the analogue output of a digital audio system using a full-scale high-frequency tone, typically at a quarter of the sampling frequency. Figure 2a shows a high-resolution spectral analysis of such a signal, in this case 12kHz at a sampling frequency of 48kHz. Figure 2b overlays spectra showing the result of adding sinusoidal jitter of three different severities at three different frequencies, while Figure 2c shows the raising of the noise floor that results from jitter that is random rather than sinusoidal in nature. Figure 3 shows the result of adding a high level of sinusoidal jitter to a musical signal, the peak near 400Hz in the spectrum being jitter-induced intermodulation distortion that is very apparent on listening.
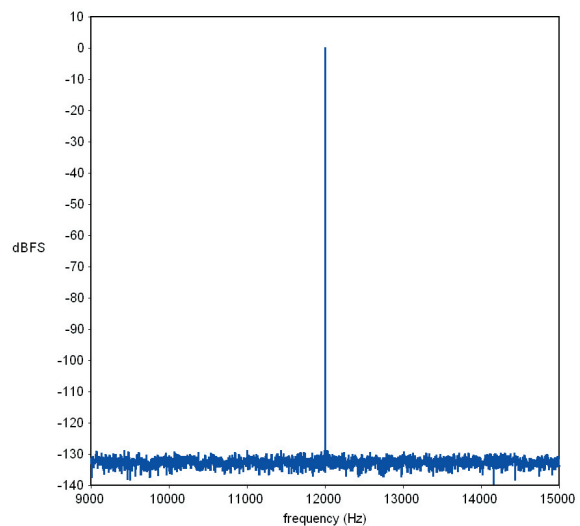

Figure 2a

Figure 2a. High-resolution spectrum analysis of a high frequency tone, as typically used to test for jitter. Here the sampling rate is 48kHz and the tone at 12kHz. Note that the only peak above the noise floor is due to the test tone itself
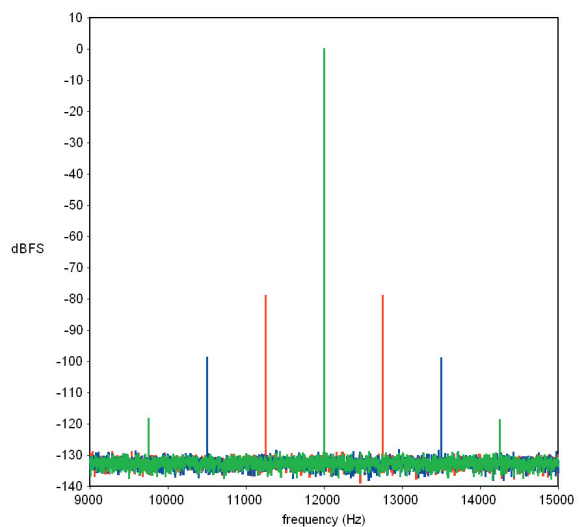

Figure 2b

Figure 2b. A simulation of how the spectrum of Figure 2a changes when sinusoidal jitter is added. Three plots are overlaid here, showing the results for 3ns peak jitter at 750Hz (red trace), 300ps jitter at 1.5kHz (blue trace) and 30ps jitter at 2.25kHz (green trace). In each case a pair of sidebands appears in the spectrum, either side of the main peak at 12kHz, due to the frequency dependent intermodulation distortion introduced by the jitter. Note how, even with this 16-bit test signal, the twin peaks at 30ns are clearly visible above the noise floor, illustrating the great sensitivity of this measurement method

When viewing an output jitter test spectrum, then, what we hope to see is a low noise floor and only low-level peaks due to sinusoidal jitter components, preferably with amplitudes of -120dBFS or lower.
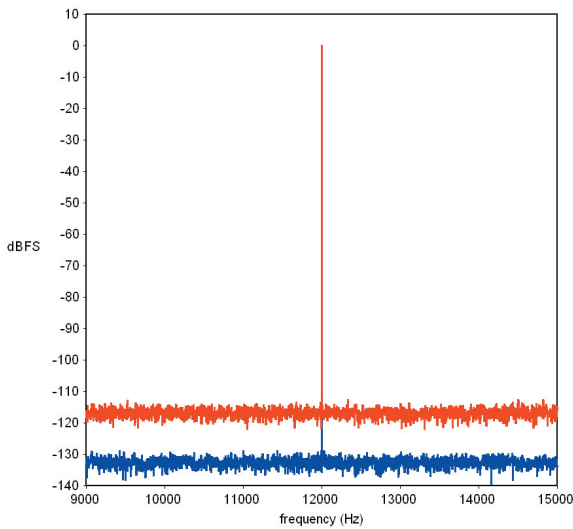

Figure 2c

Figure 2c. If we add random rather than sinusoidal jitter to the 12kHz test signal, this is the result – no distinct sidebands are added but the noise floor is raised (red trace). In this simulation the peak jitter is 3ns
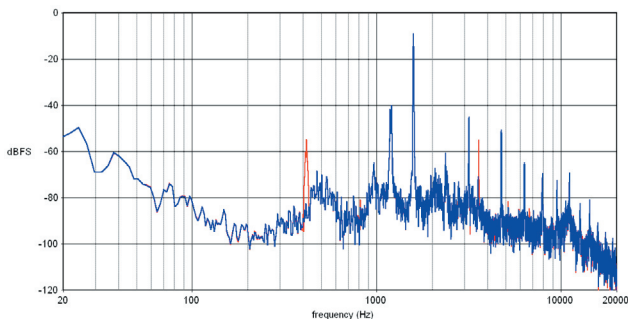

Figure 3

Figure 3. Spectrum analysis of a short section of a solo flute recording without jitter (blue trace) and with simulated 1µs peak sinusoidal jitter at 2kHz (red trace). The clearly visible intermodulation distortion at around 400Hz is not masked by the signal and is readily audible

Asynchronous mode USB

Asynchronous mode is superior to other USB data transfer modes because it allows the DAC to control the flow of audio data rather than the computer. Data is 'pulled' from the computer under control of the DAC, not 'pushed' to the DAC under control of the computer. This allows a high precision, low phase noise clock to be placed optimally, close to the digital-to-analogue converter chips, thereby ensuring the lowest possible output jitter and best possible sound quality.
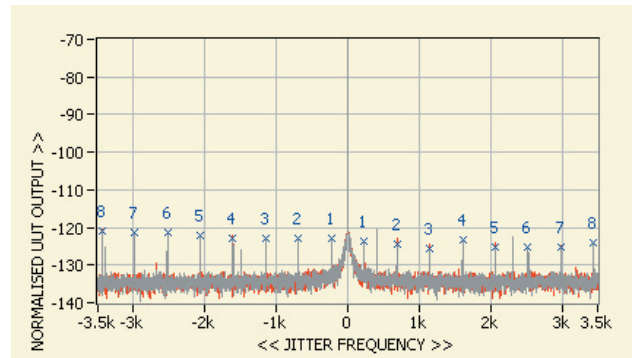

Figure 4 DAC-V1 Jitter Spectra

Figure 4 shows the output jitter spectrum from the Naim DAC-VI at a sampling rate of 44.1kHz, measured using a J-test signal (a slightly more complex jitter test signal than the single tone described above) at 11.025kHz, via the DAC's USB input. Because the DAC-VI uses asynchronous mode USB data transfer, random and sinusoidal jitter are both kept to very low levels. Total correlated jitter 121psec.
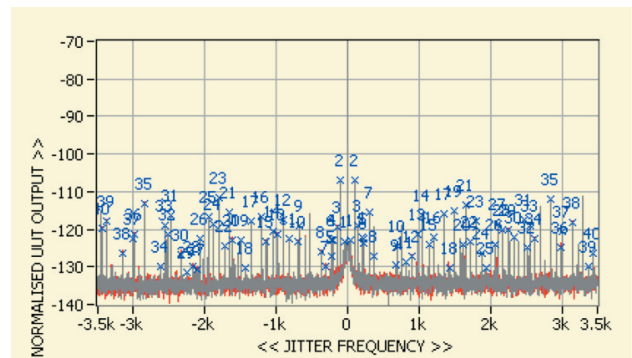

Figure 5 Unidentified DAC Jitter Spectra

Figure 5, by contrast, shows the comparable output jitter spectrum of an unidentified USB DAC that does not use asynchronous mode data transfer. The jitter performance is considerably worse, and this can clearly be heard as a degradation in sound quality.
N.B. This DAC is a good example of its type. Many more poorer examples exist than good ones, some exhibiting jitter over 1us. Total correlated jitter 451psec.