Audio Engineering Society

Convention Paper 6293

Presented at the 117th Convention
2004 October 28–31 San Francisco, CA, USA

Specifying the Jitter Performance of Audio Components

Chris Travis [1] and Paul Lesso [2]

[1] Sonopsis Ltd, Wotton-under-Edge, Gloucestershire, UK
<Chris.Travis@Sonopsis.Ltd.UK>

[2] Wolfson Microelectronics plc, Edinburgh, UK
<Paul.Lesso@WolfsonMicro.com>

ABSTRACT

The question of sample-clock quality is a perennial one for digital audio equipment designers. Yet most chip makers provide very little information about the jitter performance of their products. Consequently, equipment designers sometimes get burnt by jitter issues. The increasing use of packet-based communications and class-D amplification will throw these matters into sharp relief. This paper reviews various ways of characterizing and quantifying jitter, and refines several of them for audio purposes. It also attempts to present a common, unambiguous terminology. The focus includes wideband jitter, baseband jitter, jitter spectra, period jitter, long-term jitter and jitter signatures. Comments are made on jitter transfer through phase-locked loops and on the jitter susceptibility of audio converters.

1 INTRODUCTION

Clocks tick at the heart of every digital audio product. Jitter on clocks that are applied to audio converters (analog-to-digital and digital-to-analog) can degrade audio performance. To make sense of this situation, designers need a framework for thinking about jitter. To make progress, the industry needs some good ways of characterizing and quantifying jitter performance. It also needs a common terminology. This paper aims to contribute in all of these areas.

At the time of writing, a new project has been proposed to the AES Standards Committee with the short title 'Jitter Performance Specification'. It is likely to be allocated to AES SC-02-01, the Working Group on Digital Audio Measurements. The authors encourage all interested parties to participate in this project. Details can be found at <www.aes.org/standards>. The authors also invite feedback on the present paper by direct email.

2 OVERVIEW

The emphasis in this paper is on audio sample clocks and on the clocking chains from which they are derived. It is not on interface-specific aspects of jitter.

We have chosen to focus at this stage on components (e.g. chips) rather than equipment (e.g. mixing desks). Progress with the former will hopefully spawn progress with the latter in due course.

There is considerable disagreement in the industry on how low jitter must be for its effects to be inaudible [1,2,3]. Some of this may be due to the inappropriate use of *period* jitter as a measure of sample clock quality (section 3.5.1). Further research is needed in this area. The authors hope that the present paper is entirely complementary to work on audibility aspects.

Currently, there is little in place that helps designers to predict their jitter-related performance degradation. This is a problem. For example, the use of conventional phase-locked-loop techniques to lock to timestamps in packet-based audio interfaces can easily produce clocks that are too jittery for use in professional products [4]. Technologies are available that solve the problem, but unless designers know early-on that they must use them, redesign may be required. An aim of this paper is to help designers avoid such uncertainty and expense.

The core of this paper is section 3, which looks at ways of characterizing and quantifying jitter. It refines the *wideband* and *long-term* jitter measures for audio use, and introduces a new measure called *baseband* jitter. It sets out some guidelines for plotting jitter spectra, and unifies the frequency-domain and time-domain views through plots that the authors refer to as jitter signatures. Numeric examples are included to keep things concrete. Section 4 covers clocking chains and jitter transfer. Section 5 presents a unified qualitative treatment of the jitter susceptibility of audio converters, discussing four distinct aspects. The paper's key points are reiterated in section 6. An appendix explores reasons for preferring jitter spectra (s/rtHz) to phase noise spectra (dBc/rtHz), and illustrates the effect of clock division.

3 CLOCK JITTER

3.1 What is jitter?

3.1.1 One definition

The following general definition of jitter may be useful:

> Jitter is the dynamic deviation of event instants in a stream or signal from their ideal positions in time, excluding modulation components below 10 Hz.

Box 1. A general definition of jitter [1].

This definition has been adapted from the one in [5]. The latter uses the expression "short-term", implicitly defining it as referring to modulation frequencies "greater than or equal to 10 Hz". This seems awkward, and is at odds with the expression "long-term jitter", which has found use in various circles. Additionally, the definition in [5] limits its scope to "timing signals". Our adapted definition sidesteps these problems.

Modulation components below 10 Hz comprise wander rather than jitter. Some specs say the demarcation point can actually depend on context. At-least one standard states that the cutoff "is usually specified at 1 Hz" [6]. These contradictions are not a problem in practice, but they do underline the need to be clear about bandwidths when using highpass and bandpass measures of jitter.

3.1.2 Multiple measures

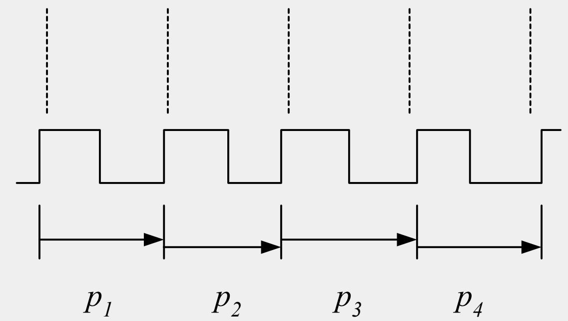
The jitter-free form of a signal is uniquely defined, but the range of possible jittered versions is infinite. How might we sensibly rank the versions? The answer depends entirely on context. This is an important point. To underline it, here are a couple of examples.

Consider first a synchronous state machine, where excessive clock jitter will lead to incorrect operation. The failure mechanism responds directly to the times between consecutive rising edges of the clock, so it is

---

[1] Some jitter definitions use the word "non-cumulative" [7]. This may be intended to make it clearer that the effects of frequency offsets do not count as jitter. A mathematical view is that such effects correspond to modulation by components below 10 Hz, and so are already adequately covered.

natural to quantify the jitter in the same way. This is peak-to-peak [2] "period jitter", illustrated in figure 1.



$$\text{instantaneous period jitter} = pjit_n = p_n - \text{mean}(p)$$

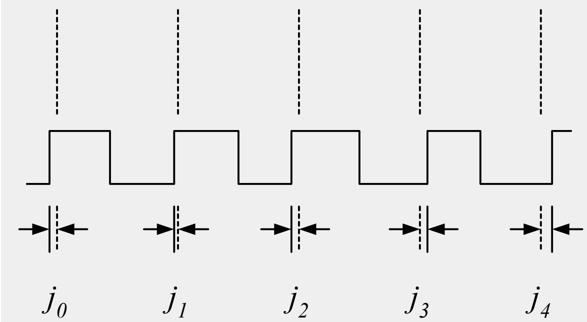
$$\text{peak.to.peak period jitter} = \max(pjit_n) - \min(pjit_n)$$
$$= \max(p_n) - \min(p_n)$$

$$\text{RMS period jitter} = \sqrt{\sum (pjit_n)^2}$$

Figure 1. Period jitter.

A good thing about period jitter is that it is easy to measure with a sufficiently fast scope (oscilloscope). Perhaps for this reason, period jitter is what many engineers think of first when the J word is mentioned.

Our second example centres on a phase-locked loop (PLL) for use alongside analog-to-digital converters (ADCs) and/or digital-to-analog converters (DACs). Let's say we want to see whether double buffering of the associated audio data will suffice, in normal operation. We lock the PLL to an effectively jitter-free reference, and look at its output using a scope triggered from the same reference. Setups such as this, with no explicit bandwidth limits, measure "absolute jitter/wander" [3]. Depending on the PLL's loop bandwidth, the results may be dominated by sub-10Hz components. This is not a problem in the given context.



$$\text{instantaneous absolute jitter} = j_n$$
$$\text{mean}(j) = 0$$

$$\text{peak.to.peak absolute jitter} = \max(j_n) - \min(j_n)$$

$$\text{RMS absolute jitter} = \sqrt{\sum (j_n)^2}$$

Figure 2. Absolute jitter/wander.

Absolute jitter/wander is also useful for quantifying certain algorithmic processes such as packet stuffing. This can be through analysis rather than measurement, so avoiding wrinkles related to observation interval.

The key point is that there are various different numeric measures of jitter, each one providing a different one-dimensional view of the same underlying multidimensional phenomenon. Each measure is appropriate in some situations but not in others.

It follows that specs such as "Jitter 200 ps RMS" are practically meaningless. Jitter specs should always identify what measure of jitter they are referring to, as in "Period jitter 200 ps RMS" for example. Note also that naming a jitter component, process or mechanism (e.g. "stuffing jitter") is not the same as identifying a jitter measure. Sometimes both are necessary (e.g. "absolute stuffing jitter").

---

[2] Strictly speaking, negative-peak period jitter would be a better measure than peak-to-peak in this first example. In practice, people tend to assume that symmetry applies.

[3] Care is necessary because the term "absolute jitter" has been the subject of a number of conflicting usages. Some people have sought to redefine it as an asymptote of interval jitter (section 3.5.2), failing to notice that this causes a disparity of typically 3 dB. If it becomes appropriate to coin a new term, one reasonable candidate is "fullband jitter/wander".

3.1.3 Peak-to-peak and RMS

In the examples just given, it was the peak-to-peak jitter (or the peak jitter) that mattered. This is quite typical of situations that involve a hard limit or error mechanism. But peak-to-peak measurements can be problematic. When the jitter is noise-like, your results depend on the observation interval. In telecommunications circles it is common to take RMS measurements instead (root mean square) and to compute corresponding peak-to-peak values for a given bit error rate [8].

For quantifying sample-clock jitter, RMS measures are in-any-case the natural choice. There is no sharp failure as the jitter increases, but instead a progressive rise of modulation products, scaling with jitter power. We will use RMS measures throughout the rest of the paper.

3.1.4 Jitter as a signal

One-dimensional measures of jitter are important, but they can obscure the bigger picture. A key step in understanding jitter is to start thinking of it as a signal in its own right. In figure 2 for example, the jitter is a stream of values, one per rising edge of the clock. In other words, it is a sampled signal. Its bandwidth can extend to half the clock frequency, but no further.

Hence we can analyze and characterize jitter in the same ways that we might analyze and characterize any signal. For example, given appropriate tools we can examine a section of the jitter waveform (i.e. instantaneous jitter versus time). Such "modulation-domain" analysis is increasingly well-catered for in modern test equipment.

Pictures can of-course be more telling than numbers. For example, eye diagrams give a very intuitive view of interface jitter, and histograms can allow random and deterministic components to be distinguished [4]. But the most consistently useful 2D representation of jitter is the frequency domain view. This is largely because jitter transfer (PLLs), jitter tolerance (receivers) and jitter susceptibility (converters) are all strong functions of frequency.

[4] This kind of analysis has been taken to rarefied levels in serial communications, to predict bit error rates. However, there are concerns about its validity. Alternative approaches that include a frequency-domain aspect may be more robust.

3.2 In the frequency domain

The majority of oscillators have jitter spectra that are dominated by the same three characteristic features. These are shown diagrammatically in figure 3.

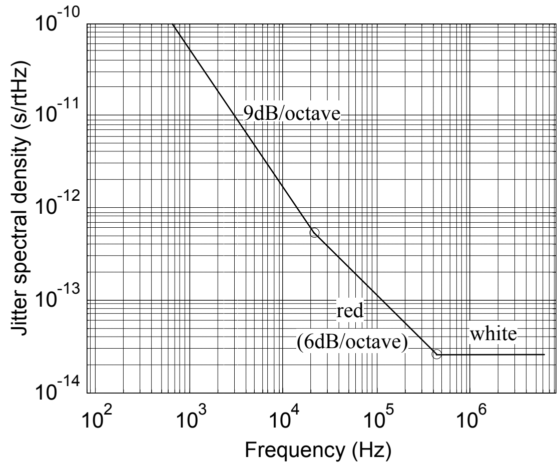


Figure 3. Asymptotes of oscillator jitter spectra.

The most important feature is a "red" noise component, i.e. one with a 6dB/octave negative slope. This comes from the phase-integrating behaviour that is central to the definition of oscillation. For example, white noise at the input of a VCO (voltage-controlled oscillator) emerges as red jitter. At some point, the red component descends into a white floor. This floor reflects the path between the active device(s) and the measurement point. It can also show limitations of the measurement setup. The third feature is an increase of slope to 9dB/octave below some frequency. It is related to flicker noise in the active device(s).

The level of the 9dB/octave feature can vary greatly between different circuits. The red jitter component is more fundamental, and therefore shows less variation. Indeed, to a first approximation all low-Q oscillators have the same red jitter! Medium-Q LC oscillators are maybe ~30 dB better and crystal oscillators are maybe ~60 dB better. Individual circuits will generally show further features, such as spectral humps or lines due to interference, e.g. via the supply rails.

In audio systems, free-running oscillators are normally crystal-based, and therefore basically beyond reproach [5].

[5] On the other hand, many recent metal-can clock generators incorporate a factory-programmable frequency synthesis PLL with low-Q oscillator. Jitter can be higher than expected!

Low-Q oscillators are certainly used, but only in PLLs. Figure 4 shows the idealized intrinsic jitter spectrum of a fairly conventional PLL incorporating a low-Q VCO. (The word "intrinsic", used in this way, underlines that this is what you get when the PLL is locked to an effectively jitter-free reference.)

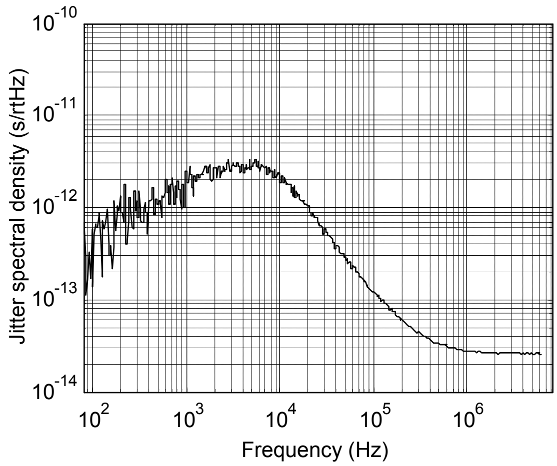


Figure 4. Idealized PLL jitter spectrum.

Below the PLL's natural frequency, feedback moderates the VCO's jitter contribution. This creates a plateau or a broad peak in the jitter spectrum. The details depend on such things as feedback order and flicker-noise effects. The peak is obvious in figure 4. The jitter of this PLL, measured via a 100 Hz highpass filter, is 320 ps RMS. 88% of that jitter is in the decade centred on 7 kHz. Hence the peak is the dominant feature in this spectrum. It will be dominant even in heavily divided versions of the clock, and even in the preamble jitter of any derived AES3 or SPDIF signals [6].

The size of the peak or plateau depends strongly on the bandwidth of the PLL. This is illustrated in figure 5. Taking the loop bandwidth below roughly 1 kHz pushes the PLL's 100Hz highpass jitter above ~3 ns RMS. Because most of this jitter is in or near the audio band, it has a direct effect on converters. As we will see later, at this level it would reduce the D+N performance (distortion-plus-noise performance) at 20kHz to only ~71 dB. These numbers are broadly representative of all conventional PLLs that use low-Q oscillators, and of all converter ICs that don't incorporate asynchronous sample rate converters.

[6] You can easily tell which region of a smooth spectrum will dominate the integrated jitter. Just sit a line with 3dB/octave negative slope onto the spectrum, and see where it touches.

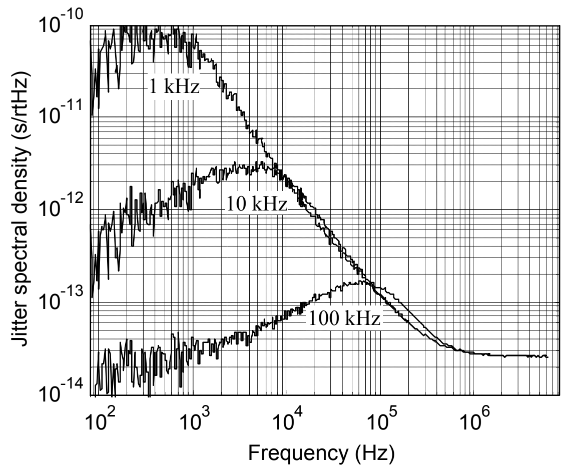


Figure 5. Spectra for different PLL bandwidths.

3.3 Plotting jitter spectra

Jitter spectra can be presented in many different ways. When homing in on some-or-other problem in the lab, engineers will use whichever way suits their purpose. But when the desire is for a representation that enables comparison with other devices, it helps to follow some common guidelines.

We remind readers that you have to choose between two quite-different approaches when plotting any spectrum. If it is discrete frequency components (spectral lines) that are of interest, you scale the Y axis so that their levels are directly readable. The result can be called a "line magnitude" plot, and it has become the default for audio spectra. Its Y-axis units might sensibly indicate "per line" (e.g. "dBFS/line") [7]. If, on the other hand, it is the continuous parts of the spectrum that are important, you normalize to unit bandwidth, e.g. per root Hertz. This yields a "spectral density" plot. It is normally this that we want for jitter spectra.

We note in passing that attempts are occasionally made to combine the two approaches in a single trace. Reference [9] does this with some success. But the present authors suspect that doing so generally raises

[7] Stating "dBFS/line" rather than "dBFS" (current practice) would serve as a useful reminder that you cannot naively read dynamic range from audio spectra. Attempting to do so is perhaps the single most-common mistake that people make in audio measurement. Even better would be to also state the effective noise bandwidth, e.g. "dBFS/line (ENB=5Hz)".

more problems than it solves. This is partly for mathematical reasons, but also by the KISS principle.

To enable reasonable comparisons to be made across different clock rates, absolute units should be used. Relative units such as UI/rtHz may be appropriate in some situations, but they are not generally so. Similarly the use of dBc/rtHz is deprecated. Further justification for this stance is given in the appendix.

Jitter spectra should normally be plotted on log axes. Common features such as those shown in figures 3 to 5 would be lost on linear axes. Jitter analysis tools that can plot only on linear axes are hardly worth having.

Direct plotting of FFT results on a log frequency axis typically gives extreme congestion in the upper decades. This "ink-blot" effect impedes the proper interpretation of results. There is also information loss, because hundreds of bins sit behind each plotted pixel or point. This is not a problem if the emphasis is on the discrete frequency components. You simply make sure that the rendering algorithm picks peaks. But such behaviour is inappropriate when the emphasis is on continuous spectral density and integration over frequency bands.

A solution that is both well-founded and practical is to apply 1/Nth-octave smoothing. The authors tentatively suggest a minimum of 1/24th-octave smoothing. More may well be appropriate, but almost any amount will eliminate the ink-blot effect. (The "round" amounts are 1/96, 1/48, 1/24, 1/12, 1/6, 1/3 and 1.) The smoothing also executes a frequency-dependent rescaling of peaks. After the rescaling, the height of each distinct peak (above its base) reflects its significance relative to the continuous part of the spectrum.

When using a log X axis the lowest two or three bins should be discarded, because they may contain leakage from DC [10]. After 1/Nth-octave smoothing it can be good to discard more points, i.e. the ones that have significantly lower percentage-bandwidth resolution than the rest of the data. Alternatively a "rectangular" plotting scheme can be used, as a visual reminder of the lower resolution in the bottom couple of octaves.

It is very useful to have a jitter integration capability built into the tool that plots the jitter spectra. If this is controllable via two vertical cursors, the user can extract highpass and bandpass jitter results, and can measure individual peaks. The combination of a jitter spectrum with associated numeric results is very powerful.

3.4 Measurements over frequency bands

3.4.1 Wideband jitter

The "highpass" jitter numbers reported in section 3.2 were derived by integrating the jitter spectral densities from 100 Hz up. One way of thinking about this is that you sum the squares across each 1Hz band, and then take the square root of the total. It is just the same as calculating noise from the plots of noise spectral density that are provided in op-amp datasheets. The units may seem strange to start with, but you get used to them.

As we have seen, it can be informative to integrate jitter across a wide band. The resulting numbers are useful as a general measure of jitter performance. However, they would be rather more useful if everyone used the same highpass corner frequency. What frequency might be appropriate? For parts targeting the audio markets, the authors suggest 100 Hz. The justification for this is in the psychoacoustics. The audibility threshold for sinusoidal jitter shows a steep slope in the region from a few hundred hertz down to roughly 130 Hz. The slope divides regions in which the detection mechanisms are fundamentally different. The threshold can be predicted using psychoacoustic models and masking theory [1], but more-direct results are also available [11].

100Hz highpass filters are already provided in some telecoms jitter testsets, though for different reasons. Questions of filter order will need further discussion, but a first-order roll-off would certainly be inadequate because some jitter spectra have 9dB/octave slope. Perhaps a range of roll-off rates can be allowed, e.g. from third-order to infinite.

As for terminology, the existing term "wideband jitter" seems suitable. Box 2 shows some detailed wordings.

wideband jitter (100Hz corner)	10 ns RMS
700Hz highpass jitter	1 ns RMS

Box 2. Terminology - Examples.

A reminder of the 100Hz corner frequency is included. When an unusual corner frequency is used, a different form of words should be adopted, e.g. as shown.

3.4.2 Baseband jitter

Wideband jitter is a general measure. Can we find a sharper measure that homes in on sample-clock issues in particular? Because jitter susceptibility depends on converter architecture, efforts to do this tend to become converter-specific. But one aspect is common across most ADC and DAC chips: Jitter components in or near the audio band directly modulate the audio signal, with most of the modulation products falling in-band.

A bandpass measurement can capture the relevant jitter. The lower band edge can be 100 Hz, as before. For the upper edge, the authors suggest 40 kHz. This isolates all of the frequencies that can cause sub-20kHz products when the audio signal is a 20kHz sinewave. Curiously, 40kHz lowpass filters are also already provided in some telecoms jitter testsets.

The authors suggest the name "baseband jitter" for this new measure. Box 3 shows some wordings.

baseband jitter (100Hz-40kHz)	8 ns RMS
12kHz-20MHz bandpass jitter	50 ps RMS

Box 3. Terminology - More examples.

These are hopefully both descriptive and unambiguous. As before, a reminder of the band edges is included, and a different wording is offered for unusual ranges.

3.5 Measurements of time intervals

3.5.1 Period jitter

Period jitter was introduced in section 3.1.2. Unlike wideband jitter and baseband jitter, it can be measured directly in the time domain, i.e. without filter hardware. You simply use a scope, and examine the waveform one period after the trigger point. Many scopes can plot period jitter histograms and extract RMS values.

If you look not one but N periods after the trigger point, you measure N-period jitter. If it is known that a clock will be divided e.g. by four before use, its 4-period jitter may well be of more interest than its period jitter.

period jitter	50 ps RMS
4-period jitter	75 ps RMS

Box 4. Terminology - Yet more examples.

There is no hardware filter when measuring period jitter but there is an implicit filter. It arises because each edge is measured with respect to its predecessor. Modulations than take many clock periods to unfold will hardly register. Comparing figures 1 and 2 we find that $pjit_n = j_n - j_{n-1}$. In filter language, this is a two-tap FIR filter with coefficient values of 1 and -1. Figure 6 shows its frequency response. It is highpass, with a peak gain of 6 dB, a roll-off of 6dB/octave.

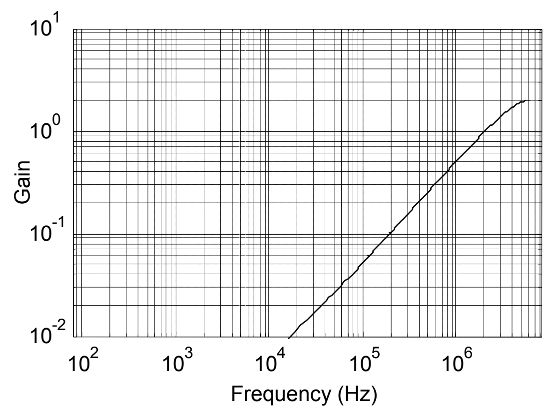


Figure 6. Period jitter implicit filter.

We saw in section 3.1.2 that period jitter is entirely appropriate for some purposes. We see here that it is entirely *in*appropriate as a general measure [14]. This is because it is basically blind to low-frequency jitter. To drive this point home, table 1 compares it with wideband jitter for the three spectra of figure 5.

	wideband jitter (100Hz corner)	period jitter (1ms interval)
100kHz loop	90 ps RMS	92 ps RMS
10kHz loop	320 ps RMS	92 ps RMS
1kHz loop	2760 ps RMS	92 ps RMS

Table 1. Period jitter values for figure 5.

We can only guess how many times it has happened that people hear the effects of *baseband* jitter, measure the *period* jitter, and reach erroneous conclusions.

3.5.2 Long-term jitter

As we will see, it can be useful to make N-period jitter measurements with very large N. Modern digital scopes are excellent for such measurements. It is usually better to express the interval in seconds rather than periods. The name "N-period jitter" then becomes awkward. Some people favour the term "accumulated jitter", but this interferes with the language that is normally used to describe how jitter builds up in a cascade of PLLs. Another possible candidate is the telecoms term TIE (timing interval error), and its relation, TIErms [5,15]. The authors find two problems with this. Firstly, TIE is defined in a way that assumes perfect knowledge of the mean frequency. This turns out to be subtly at odds with the measurement method that we have outlined. (The clash relates to the RMS values of non-zero-mean result sets, and to the question of unbiased estimation.) The second problem is simply that the terms themselves are ungainly. For these and other reasons, the authors suggest the alternative name "interval jitter".

Figure 7 shows the frequency response associated with a 1ms interval jitter measurement. It is a comb response, so its capture of spectral lines will be hit-and-miss. But it should measure the smoother parts of the spectrum fairly well. The choice of 1 ms for the interval gives us a loose approximation to the 100Hz highpass response of wideband jitter measurements.

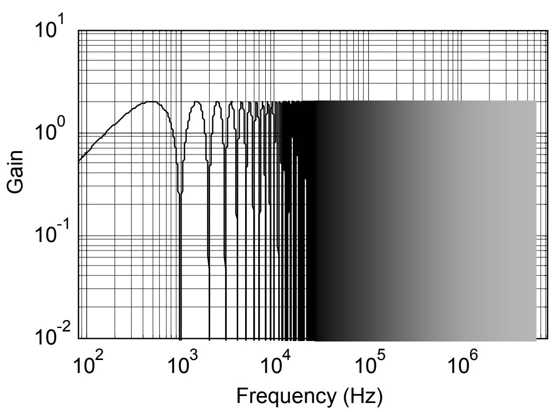


Figure 7. 1ms interval jitter implicit filter.

Note that there is an inbuilt 3dB boost. One way of thinking about this is as due to the power summation of jitter at both ends of the interval.

The "long-term jitter" of a clock or device is understood to mean its interval jitter at some fairly large interval. It is a useful measure, but it would be rather more useful if everyone used the same large interval. What interval might be appropriate? In the light of points just made, the authors suggest 1ms. Box 5 shows some wordings.

long-term jitter (1ms interval)	14 ns RMS
2.5us interval jitter	500 ps RMS

Box 5. Terminology - Further examples.

To explore the relationship between long-term jitter and wideband jitter, we ran calculations for the three spectra of figure 5. Table 2 presents the results.

	wideband jitter (100Hz corner)	long-term jitter (1ms interval)
100kHz loop	90 ps RMS	127 ps RMS
10kHz loop	320 ps RMS	455 ps RMS
1kHz loop	2760 ps RMS	4160 ps RMS

Table 2. Long-term jitter values for figure 5.

The results show good agreement (and demonstrate the already-mentioned $\sqrt{2}$ boost). Hence long-term jitter can be used to estimate wideband jitter. But some care is needed, as we will see in the next section.

3.6 Jitter signatures

To get any insights through interval jitter measurements, you have to measure at a number of different intervals and look at the trends. One good approach is to plot RMS interval jitter as a function of interval on log axes. The authors find it convenient to refer to such plots as jitter signatures. The left-most point is the period jitter, and if the X axis extends far enough the long-term jitter can also be read off. As before, to enable comparisons across clock rates, absolute units should be used.

Note that some scopes can capture and display plots of TIErms against interval. The only significant difference between these and jitter signatures is that the former are linear plots and the latter are by definition on log axes. This difference is actually *very* significant when the X-axis runs from nanoseconds to milliseconds.

Every complete jitter signature is a mathematical dual of its associated jitter spectrum [16]. You can move between the two notionally without losing information.

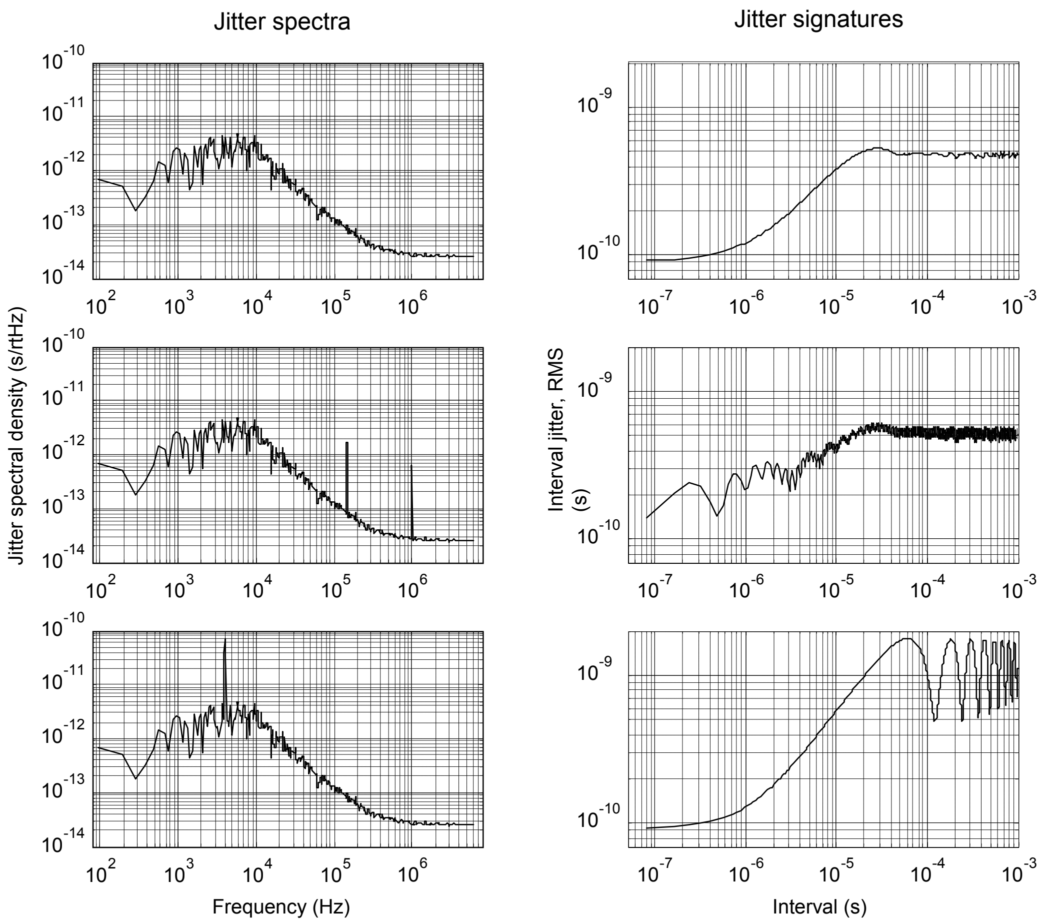


Figure 8. Jitter spectra/signature pairs.

At-least one manufacturer of jitter analysis tools has made use of this fact in its products.

Figure 8 shows several simple jitter spectra and their corresponding jitter signatures. In each case the continuous part of the spectrum is that of the PLL introduced in section 3.2. Referring to the top signature, the short-term jitter reflects the spectrum's white floor. As you measure progressively longer intervals, the jitter rises up, and then levels off. Notice that in this case the asymptotes are just as clear as their counterparts in the frequency domain.

The signatures get much busier when the spectra have distinct lines. The middle spectrum includes spurs at 150 kHz and 1 MHz. These give rise to some flourishes in the jitter signature, but their effects are largely swamped at longer intervals. The bottom spectrum is very different, with one spur dominating over the continuous parts. Because it falls at a low frequency, this spur leaves the period jitter unaffected. But it does make a big difference at longer intervals, even bringing the meaning of long-term jitter into question.

Incidentally, if you continue to increase the interval, the curves start to rise again. This is due to wander, either from the reference that the PLL is locked to, or in the scope's internal clock. (At *extremely* large intervals, things get weird. Instead of becoming more repeatable as you gather more data, the numbers start to diverge. People who analyze the month-to-month stability of rubidium clocks have to use e.g. Allan variance instead, or stay in the frequency domain [12].)

Our exploration of jitter spectra/signature pairs has cast some light on interval jitter. We conclude this section by observing that the spectra seem to be more generally useful than the signatures.

4 JITTER TRANSFER

A clean-and-simple clocking arrangement might have a free-running crystal oscillator connected directly to local ADCs and DAC chips. But such arrangements are not very common. Even in DVD players for example, the audio clocks come via a PLL (phase-locked loop), locked to 27 MHz. And in digitally connected systems, downstream devices commonly slave their audio clocks either to an incoming audio stream or to a nominated sync master. Every stage in the clocking chain can potentially contribute jitter to the derived audio clocks.

In section 3.2 we looked at the *intrinsic* jitter of PLLs, but that is only one aspect of PLL performance. Another important aspect is jitter *transfer*, i.e. how jitter on the PLL's reference input affects its clock output. Information on clocking chips' jitter transfer behavior allows equipment designers to calculate the effect of earlier stages on derived audio clocks. But chip makers do not always provide the necessary information.

PLLs can conveniently be thought of as jitter filters. Within their loop bandwidth they track their reference, but beyond it they increasingly attenuate reference jitter. Figure 9 shows this diagrammatically.

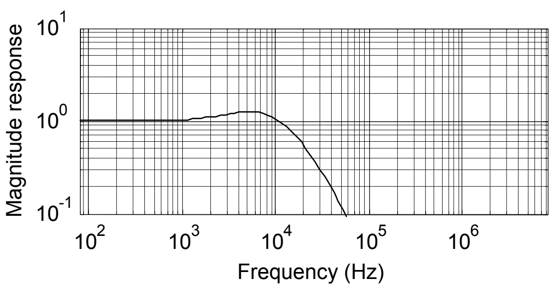


Figure 9. PLL jitter transfer function.

Pertinent characteristics include corner frequency, roll-off rate and in-band response peaking. The authors see no strong need to formalize how such information might be expressed. It may suffice, in a given context, to simply state the minimum attenuation at and above a particular frequency. In some cases the lockup time is also of interest.

Jitter transfer is predominantly a linear phenomenon, but nonlinear effects can arise, particularly at low levels. When measuring jitter transfer, methods that are blind to nonlinear effects should be avoided.

5 JITTER SUSCEPTIBILITY

5.1 General points

Audio ADCs and DACs have three important inputs; the signal input, the voltage reference, and the clock. Noise and interference on the voltage reference causes amplitude modulation, and jitter on the clock causes phase modulation. The resulting modulation products look very similar in the frequency domain. One of the authors once spent several days trying to track down a low-frequency jitter problem, only to find that it was in fact a problem of LF noise on the voltage reference.

Our emphasis here is on the phase-modulation aspect. Phase modulation of a sinusoidal signal produces sidebands at $f_{\text{signal}} \pm f_{\text{jitter}}$. Only modulation products that fall into the audio band are a direct problem. Subjecting an audio signal to 1MHz jitter, does *not* give audioband products. The frequency pairings that *do* are the straight one-to-one matches plus a 20 kHz slop.

A key point is that it is not just the basic audio signal that gets modulated. It is everything that crosses the boundary between the continuous-time domain and sampled-signal domain. This can include out-of-band interference (in ADCs), incompletely attenuated images (in DACs), and "zero-input" internal signals such as shaped quantization noise and class-D carriers.

In common ADC and DAC chips, the domain boundary is right at the pins of the chip. You can get an idea of which frequency regions might be problematic by considering the spectrum of the pin signal. You should mentally apply a 6dB/octave tilt to this, to accommodate the fact that jitter sidebands scale with signal slew rate. (Again, log axes help.) Even low-level components can cause problems if they are up at high frequencies.

The other thing to consider is the spectrum of the jitter. As we saw in section 3.2, this is typically far from flat. Often it will have a broad peak. The worst case is if this peak happens to overlap a dominant region of the tilted total signal spectrum.

Converters are sensitive to jitter in different ways and to different degrees [13,14]. The people who are best placed to find concise ways of spec'ing these things are

(More black dots means greater susceptibility.)	Reduction of dynamic range	Down-modulation of out-of-band interference	Baseband modulation noise and/or distortion	Image modulation noise and/or distortion
Common noise-shaping DACs	●		●●	●●
Noise-shaping DACs with low-order shaping and reduced images			●●	
Two-level all-digital class-D amplifiers	●●●●		●●	
Three-level all-digital class-D amplifiers	●●●		●●	
Common delta-sigma ADCs		●●	●●	
Delta-sigma ADCs with continuous-time loop filters	●		●●	

Table 3. Broad jitter susceptibilities of various types of converter.

perhaps the chip makers themselves. The authors note that for DACs and class-D amplifiers, simple plots of output signal spectrum up to 1 MHz or more can be very helpful to equipment designers. The chip makers should also consider specifying how low the clock jitter needs to be for "normal operation" of their converters. This could be done by stating a limit of acceptability for the smoothed jitter-spectral-density over a key band. Alternatively a limit on the interval jitter for some specified interval might do the job.

A selective summary of jitter susceptibility issues is presented in table 3. The four remaining parts of this section will each discuss one column of table 3.

Incidentally, table 3 deliberately omits some circuits, including single-bit DACs, upsampling DACs that don't use noise shaping, converters that incorporate asynchronous sample-rate conversion, and analog class-D amplifiers.

5.2 Reduction of dynamic range

Jitter bites equipment designers most deeply when it causes a converter that should have more than 100 dB of dynamic range to deliver e.g. only 80 dB. In such cases the jitter is interacting not with the audio signal but with an internal signal such as shaped quantization noise. Early one-bit DACs were particularly sensitive to this. More-recently the inclusion of switched-capacitor filters and the move to multi-bit designs has eased things.

Above ~200 kHz, the quantization noise is largely white at its point of injection. When you factor in the DAC's sin(x)/x frequency response and the effect of the internal switched-capacitor filter stage, its spectrum becomes more like the upper trace in figure 10 (taken from [17]). By applying the already-mentioned 6dB/octave tilt, one can estimate the region of greatest jitter sensitivity. It is typically somewhere around ~0.5 or ~1 MHz for DACs that use high-order noise shaping.

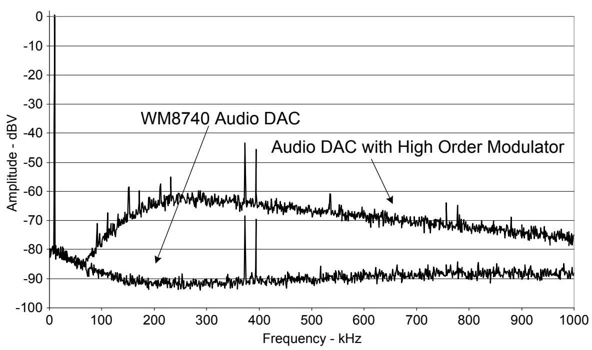


Figure 10. Out-of-band DAC spectra (from [17]).

For full dynamic range with particularly jittery clocks, equipment designers may select a DAC that instead uses low-order noise shaping. The out-of-band spectrum of one-such DAC is shown in the lower trace of figure 10.

The situation with power DACs (i.e. all-digital class-D amplifiers) is very different. The high jitter sensitivity of the early one-bit DACs re-emerges. Designs that use two-level pulse-width modulation demand special care. Three-level modulation eases things because it has lower carrier energy at low audio levels. Some designs sidestep these issues by incorporating asynchronous sample rate conversion.

Common delta-sigma ADCs are at the other extreme. Their feedback loop is completely contained within the sampled-signal domain. As a consequence they tend to keep their dynamic range even with very jittery clocks. A few designs use a continuous-time loop filter instead. Their jitter susceptibility depends on the internal DAC in the feedback path.

5.3 Out-of-band interference

We have said that switched-capacitor delta-sigma ADCs tend to keep their dynamic range. But this not the case if their input carries significant out-of-band interference. Jitter down-modulates some of that interference into the audio band, where it stays even after its parent has been removed by the digital anti-alias filter [18,19]. It can be noise-like or tonal, steady or varying, depending on the parent interference and the jitter. The weakness exists because the sampling is done up-front.

The ADC designs that use continuous-time loop filters avoid this weakness. Their sampling is done within the delta-sigma loop, so down-modulation is moderated by the action of feedback.

5.4 Baseband modulation

The jitter performance differences that we have seen relate entirely to signal components that are above the audio band. When you look at phase modulation of the basic audio signal, the differences dissolve.

Unlike the impairments discussed in previous sections, the ones produced by baseband modulation scale with the audio signal. (More specifically, they scale with the slew rate of the audio signal.) This makes them rather less apparent when listening and measuring.

Here is the maths of it: Modulating a 20kHz sinewave with e.g. 1ns RMS of sinusoidal jitter would produce sidebands each $-20.\log(2\pi.20e^{3}.10e^{-9}/\sqrt{2}) = 61$ dB down [1,2]. For jitter frequencies up to 40 kHz, one of the sidebands falls in the core audio band. The total effect is found by integrating the jitter spectral density from a chosen low frequency (100 Hz) up to 40 kHz. This is our definition of baseband jitter (section 3.4.2). So 1 ns RMS of baseband jitter reduces the D+N performance, with a 20kHz sinewave, to ~81 dB. From this one can easily find the D+N limit for other cases. For example, with 10 ns of baseband jitter it is 61 dB at 20 kHz and 61+20+6 = 87 dB at 1 kHz. (Note that these numbers do not say much about audibility. In reference [2], with music as the test signal, none of the subjects found baseband jitter below 20 ns to be audible.)

We emphasize that these results apply to every current converter chip that does not incorporate asynchronous sample-rate conversion. Furthermore, in some systems baseband modulation is the only practically-significant jitter susceptibility mechanism. For those systems, the baseband jitter measure introduced earlier in this paper is the only appropriate one-dimensional jitter measure that the authors know of.

5.5 Image modulation

In upsampling DACs, the digital interpolation filter attenuates near images of the baseband audio signal. Typically, it outputs at eight times the audio base rate. This pushes the first images of a 48kHz-sampled signal up to the 384kHz region, as in figure 10 for example.

The primary reason for attenuating the images is to ease the analog-domain filtering. But there is also a secondary benefit. Interpolating the waveform reduces its susceptibility to jitter. More clock edges get used, so the modulation products get spread over a wider band.

For white jitter, the theoretical benefit is 3 dB per doubling of rate. This number goes up when you consider real-world jitter spectra and the effects of any switched-capacitor filter stage.

Stopping the interpolation at eight times the base rate gives us converters that are fine in many applications, but are less-good with more-jittery clocks. The DACs that have the lowest jitter susceptibility are the ones that interpolate to higher rates. The out-of-band spectrum of one-such DAC is shown in the lower trace of figure 10.

The results of image modulation are very similar to those of baseband modulation. Whether one of these mechanisms dominates over the other depends in part on the jitter spectrum. The spectra from some modern clocking solutions show densities that are higher in the 384kHz region than in the baseband. In such cases, with common DACs, image modulation dominates.

6 SUMMARY OF KEY POINTS

Jitter spectra, when plotted properly, can be very useful. They can be even more useful when accompanied by numbers for RMS wideband jitter (100Hz corner) and RMS baseband jitter (100Hz-40kHz). The makers of audio clocking chips should consider providing this information in their product datasheets.

Wideband jitter is appropriate as a general measure of clock quality. Period jitter is not. Baseband jitter focuses on an aspect that is relevant to converter clocks. RMS long-term jitter (1ms interval) is easy to measure and gives an estimate of the wideband jitter (times $\sqrt{2}$). Jitter signatures are plots of interval jitter vs. interval. They can build understanding and aid intuition.

Different types of converter are sensitive to jitter in different ways. Table 3 presents a broad summary. Sometimes jitter causes a reduction of dynamic range. The makers of audio converter chips should consider illuminating this, e.g. by stating limits of acceptability for smoothed jitter-spectral-density over key bands.

DACs that have low-order shaping and reduced images are less susceptible to jitter than more-common designs. Most all-digital class-D amplifiers are more susceptible, especially ones that use two-level modulation.

All present-day converter chips that do not include asynchronous sample rate conversion are equally susceptible to baseband jitter.

7 ACKNOWLEDGEMENTS

This work was enabled and supported by Peter Frith at Wolfson Microelectronics. The authors acknowledge helpful technical contributions from Anthony Magrath, Bruno Putzeys and Steve Harris.

8 REFERENCES

[1] Julian Dunn, "Jitter and Digital Audio Performance Measurements", from 'Managing the Bit Budget', the proceedings of the AES UK Conference, London, May 1994, pp 148-166.

[2] Eric Benjamin and Benjamin Gannon, "Theoretical and Audible Effects of Jitter on Digital Audio Quality", AES Preprint 4826, presented at the AES 105th Convention, San Francisco, September 1998.

[3] Bob Katz, "Mastering Audio: The Art and the Science" (chapter 19), ISBN 0-240-80545-3, published by Focal Press, 2002.

[4] Julian Dunn, "Sample clock jitter and real-time audio over the IEEE1394 high-performance serial bus", AES Preprint 4920, presented at the AES 106th Convention, Munich, May 1999.

[5] ITU-T Recommendation G.810, "Definitions and terminology for synchronization networks".

[6] American National Standard T1.523-2001, "Telecom Glossary 2000", <www.atis.org/tg2k>.

[7] ANSI INCITS 172-2002, "American National Standard Dictionary of Information Technology".

[8] "Converting between RMS and Peak-to-Peak Jitter at a Specified BER", Maxim Application Note HFAN-4.0.2, December 2000.

[9] Eric Benjamin, "Effects of DAC Nonlinearity on Reproduction of Noise Shaped Signals", AES Preprint 3778, presented at the AES 95th Convention, New York, October 1993.

[10] Julian Dunn, "Analog-to-Digital Converter Measurements", Audio Precision Tech Note 24, <www.audioprecision.com>.

[11] Armin Kohlraush, Ralf Fassel, and Torsten Dau, "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers", J. Acoust. Soc. Am. 108(2), August 2000.

[12] Charles Greenhall, "Spectral Ambiguity of the Allan Variance", IEEE Trans. Inst. and Meas., 47(3), June 1998.

[13] Steve Harris, "The Effects of Sampling Clock Jitter on Nyquist Sampling Analog-to-Digital Converters, and on Oversampling Delta-Sigma ADCs", J. Audio Eng. Soc., 38(7), July 1990, pp 537-542.

[14] Bruno Putzeys, and Renaud de Saint Moulin, "Effects of Jitter on AD/DA Conversion. Specification of Clock Jitter Performance", AES Convention Paper 6122, presented at the AES 116th Convention, Berlin, May 2004.

[15] Stefano Bregni, "Clock Stability Characterization and Measurement in Telecommunications", IEEE Trans. Inst. and Meas. 46(6), December 1997.

[16] Masashi Shimanouchi, "An Approach to Consistent Jitter Modeling for Various Jitter Aspects and Measurement Methods", ITC 2001, IEEE.

[17] David Hossack, Peter Frith, Julian Hayes and Angus Jackson, "Design and Evaluation of an Audio DAC with Non-Uniformly Weighted Dynamic Element Matching", from 'Silicon for Audio', the proceedings of the AES 16th UK Conference, London, April 2001, pp 82-87.

[18] John Siau, "Jitter and Its Effects", Benchmark Media Systems, Digital Application Note #1, <www.benchmarkmedia.com/appnotes-d>.

[19] Mike Story, "Audio Analog-to-Digital Converters", J. Audio Eng. Soc., 52(3), March 2004, pp 145-158.

[20] Salvatore Levantino, Luca Romanò, Stefano Pellerano, Carlo Samori, and Andrea Lacaita, "Phase Noise in Digital Frequency Dividers", IEEE J. Solid State Circuits, 39(5), May 2004, pp 775-784.

[21] Venceslav Kroupa, "Low-noise microwave-frequency synthesizers", IEE Proceedings - H, 130(7), December 1983, pp 239-243. Reprinted in 'Direct Digital Frequency Synthesizers', edited by Venceslav Kroupa, ISBN 0-7803-3438-8, published by the IEEE Press, 1999.

[22] "IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology", October 1988.

[23] Sander Gierkink, "Control Linearity and Jitter of Relaxation Oscillators" (chapter 2), thesis, University of Twente, June 1999, ISBN 90-36513170, <http://doc.utwente.nl/14152>.

APPENDIX

In engineering for radio-frequency communications, oscillator stability is often expressed in terms of single-sideband phase noise $\mathcal{L}(f)$, in units of dBc/Hz. This is the sideband-to-carrier power ratio, directly measurable using a spectrum analyzer. Conversion to jitter spectral density is straightforward. In words, you reverse out of dB in the usual way (i.e. $10^{(x/20)}$), multiply by $\sqrt{2}$ to accommodate both sidebands, and divide by the radian rate of the carrier (i.e. $2\pi f_c$).

$$\text{Jitter spectral density in s/rtHz} = 10^{\left(\frac{\text{Phase noise in dBc/Hz}}{20}\right)} \times \frac{\sqrt{2}}{2\pi f_c}$$

To interpret or compare phase noise, you need to know the carrier frequency f_c. Quite often, phase noise results are presented without stating f_c, which at-best is a pain and at-worst makes the results useless to others.

This awkwardness of dBc/Hz shows up more obviously when you divide clocks. Figure 11 shows first the phase noise spectra and then the jitter spectra of an original clock and a divided version.

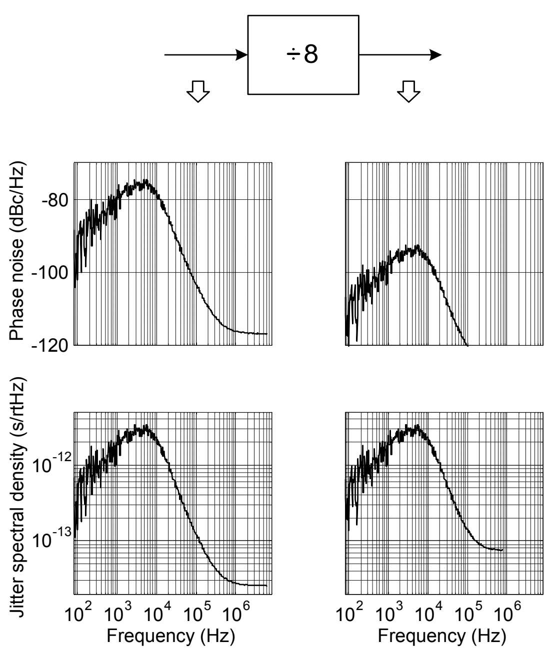


Figure 11. The effect of clock division.

Division by eight reduces the phase noise by 18 dB, simply by definition. But, assuming a perfect divider, the absolute jitter/wander should remain unaffected. This is apparent in the lower two spectra of figure 11. The spectral density of the white floor rises by 9 dB because it has to fit into one-eighth of the bandwidth, but the spectral features show identically. (Of course, real dividers are not perfect. Reference [20] addresses some real-world issues with dividers.)

The above considerations argue for using absolute units on the Y axis. Similar points have been made in the radio-frequency community since 1983 [21] or earlier.

The IEEE has formalized the maths and nomenclature of frequency and time metrology, e.g. in reference [22]. For expressing frequency stability they recommend absolute measures; the fractional frequency error y(t), and its power spectral density $S_y(f)$. For phase stability they again define absolute measures; the time error x(t) and its power spectral density $S_x(f)$. $\mathcal{L}(f)$ is notable by its absence! Our jitter spectral density $\mathcal{J}(f)$ is simply the square root of the IEEE's measure $S_x(f)$, above ~10 Hz. A good summary of such things and more is given in chapter 2 (19 pages) of reference [23].

Figure 11 prompts us to consider what division does to our one-dimensional measures of jitter. Wideband jitter is normally unaffected by division. Indeed, that is part of its utility. Note though that spurs at fractions of the clock rate can alias to DC, and so remove themselves from the measurement. Things are a little different with bandpass measures such as baseband jitter. Division will inevitably alias some energy from high frequencies down into the 100Hz-40kHz band. As we saw in section 3.2, dominant features are often at lower frequencies. Hence this aliasing is normally insignificant, as in figure 11.