Sigma-delta idle tones

Marcel van de Gevel, 11 April 2021

0. Introduction

A well-known technique to reduce the requirements on the resolution of audio DACs is noise shaping/sigma-delta modulation. A negative feedback loop around a coarse quantizer running at a high clock rate (usually at least a couple of megahertz for an audio converter) changes a digital input signal with many bits into a massively oversampled signal with only a few or even just a single bit. The quantization error caused by the coarse quantizer is shaped by the feedback loop: it is kept small in the audio band, where the loop has a large loop gain, and it is large in the ultrasonic region, where the loop has almost no loop gain.

Sigma-delta modulators, particularly single-bit sigma-delta modulators, can produce some strange artefacts at low signal levels, such as tones that get frequency modulated by the signal. In section 1, the simplest and worst of all types of sigma-delta modulator is analysed. The results are extended to higher-order sigma-delta modulators in section 2. Some countermeasures are discussed in section 3 and a technique to find out whether there are any idle tone issues in section 4. Section 5 is about noise modulation and has only little to do with the rest of the document.

Traditionally noise shaping loops are called sigma-delta or delta-sigma modulators when they are drawn as ordinary feedback loops and are called noise shapers when they are drawn as error correction loops, although these are essentially equivalent. To keep things simple, the loops in this document are all drawn as sigma-delta modulators with a feedback architecture.

Also for simplicity, the two possible levels at the output of a single-bit sigma-delta modulator will be interpreted as -1 and +1 rather than 0 and 1.

1. First-order single-bit undithered sigma-delta modulator

A first-order single-bit undithered sigma-delta modulator is shown in figure 1. The block marked z^{-1} is a unit delay block, in practice this will be a register clocked at the massively oversampled rate. The unit delay and the summer preceding it are together an accumulator that accumulates the differences between the input signal X and the output signal Y. The output signal of the accumulator goes into a quantizer that produces a -1 when the accumulator output signal is negative and a +1 when it is not.

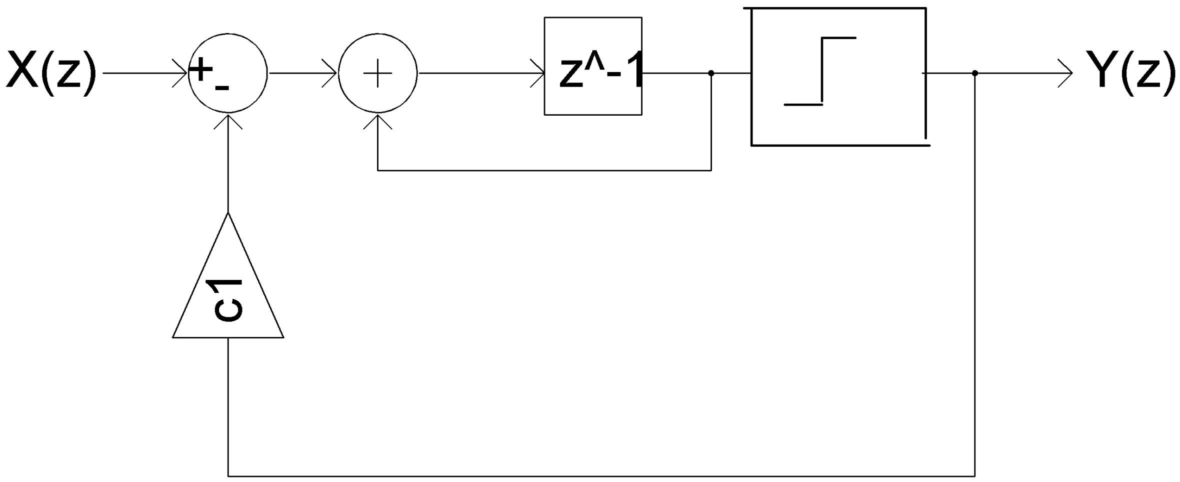

Figure 1: First-order single-bit sigma-delta (or delta-sigma) modulator

All in all, the loop produces a bit pattern that keeps the output of the accumulator as close to zero as possible. Unfortunately this can lead to regular, repeating patterns.

For example, suppose the input signal is constant at 0, coefficient c_1 is 1 and you start with a value of 0.5 at the output of the accumulator (similar results are found for any other initial value). Output *Y* will then become +1, and 1 will be subtracted from the accumulator. Hence, in the next clock cycle, the accumulator output signal will be -0.5, *Y* will become -1 and 1 will be added to the accumulator. As a result, you get a perfectly regular +1, -1, +1, -1, +1, -1 pattern coming out of the sigma-delta modulator. The accumulator output signal is +0.5, -0.5, +0.5, -0.5 and so on. Spectrally, the output signal only has peaks at half the clock frequency and its odd multiples.

If the input signal *X* is not 0 but 0.1, the accumulator output signal will drift upwards. This table has the accumulator levels in the upper row and the quantizer levels in the bottom row:

+0.5	-0.4	+0.7	-0.2	+0.9	0	-0.9	+0.2	-0.7	+0.4	-0.5	+0.6	-0.3	+0.8	-0.1	+1	+0.1	-0.8	+0.3	-0.6	+0.5	-0.4
1	-1	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1	1	-1	1	-1

The last two entries are the same as the first two, so the whole pattern repeats. You therefore get a repeating pattern with a length of 20 clock cycles having 11 ones and 9 minus ones. Spectrally, it has peaks at multiples of 1/20 of the clock frequency, with particularly strong ones at 0.45 and 0.55 times the clock frequency (the average repetition rates of the +1's and -1's), at 0.1 times the clock frequency (the average repetition rate of the extra +1's), and of course at zero frequency due to the DC offset. The component at zero frequency is the only one that is also present in the input signal, so all other components are the shaped quantization error of the sigma-delta modulator. Note that the shaped quantization error consists of discrete spectral peaks and doesn't resemble noise at all, so calling it quantization noise would be a misnomer.

Hence, an offset of 0.1 shifts the tone at half the clock frequency to 0.45 and 0.55 times the clock frequency, and produces peaks at other multiples of 0.05 times the clock frequency, particularly 0.1 times the clock frequency. The shift is proportional to the offset because the offset determines how many extra +1's (or extra -1's for negative offsets) are needed. Hence, an offset of 0.001 would produce tones at 0.4995 and 0.5005 times the clock frequency and peaks at other multiples of 0.0005 times the clock frequency, particularly 0.001 times the clock frequency. When the sigma-delta modulator runs at a rate of a couple of megahertz, this last peak will be in the audio band.

An irrational number offset (for example $\pi/1000$) would result in a pattern that never repeats and that has a continuous spectrum, so the quantization error would finally resemble noise. However, it still has peaks above and below half the clock frequency with a frequency distance that is proportional to the offset and a low-frequency peak of which the frequency is proportional to the offset.

2. Higher-order sigma-delta modulator

Figure 2 shows an example of a higher-order sigma-delta modulator, a third order to be precise.

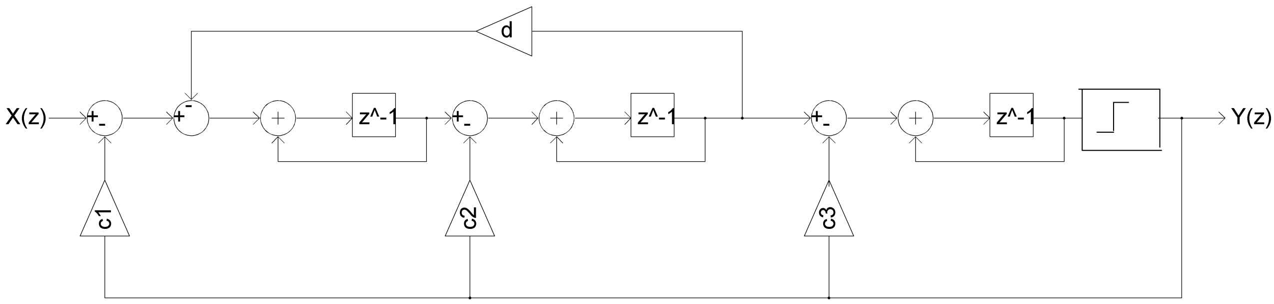

Figure 2: Third-order sigma-delta modulator

It consists of a structure that is exactly the same as figure 1 (except that coefficient c_1 has been renamed to c_3) with an extra feedback loop around it. This extra feedback loop has a large loop gain at audio frequencies, but not around half the clock frequency. As a result, any tones in the audio band are substantially reduced (the higher the order the more they are reduced), but the tones around half the clock frequency remain.

The signal coming out of the second accumulator will normally not be a constant rational number, so the patterns that come out of a higher-order sigma-delta modulator are generally much more irregular and noise-like than those coming out of a first-order sigma-delta modulator with a constant rational input signal. You still have the tones around half the clock frequency, though.

In fact, as was already noted in section 1, the tones around half the clock frequency are related to the average repetition rates of the +1's and the -1's, and those are determined by the clock frequency and the desired signal, no matter what the order of the sigma-delta modulator may be. As each clock cycle produces either a +1 or a -1, the sum of the average repetition rates of the +1's and the -1's has to be equal to the clock frequency. What percentage of the bits is +1 and what percentage -1 determines the average output level, and is therefore set by the input signal and the signal transfer of the modulator.

When the output signal goes into something that distorts, intermodulation between the tones just above and below half the clock frequency can again produce a tone in the audio band, even when the modulator itself suppresses it well. When the output signal goes into a DAC that has some second clock subharmonic (interference at half the clock frequency) at its clock or its reference, intermodulation between the idle tones and half the clock frequency can again result in an audio tone.

3. Chaos, dither and deliberate offset

There are two methods to make the quantization error more noise-like and reduce the idle tones, namely chaos and dither. I've written more about that in section 3.1 of https://linearaudio.net/sites/linearaudio.net/files/03%20Didden%20LA%20V13%20mvdg.pdf

Suffice it to say that dither can be extremely effective in multibit and quasi-multibit modulators. It is very useful but has its limitations in single-bit modulators.

One way to make the low-frequency idle tone/the intermodulation product between the idle tones just above and below half the clock frequency inaudible when playing very soft music is to just add a deliberate offset that shifts it far into the ultrasonic region. While this works well for signals that are much smaller than the deliberately added offset, it doesn't work well when you play some deep bass tone that is about the same size as the offset: the bass tone plus offset then sweep the idle tone between zero and some ultrasonic frequency, so it periodically passes through the audio band.

Note that when playing absolute silence through a modulator with no deliberately added offset, the low-frequency idle tone will be at 0 Hz and will therefore be inaudible. Very small signals will then already sweep it into the audio band, though.

4. Test with a soft 1 Hz or 0.1 Hz signal

A way to check for these idle tone issues is to put a DC level at the input and slowly sweep it. If there is such an issue, there will be DC levels that move the idle tone into the audio band. As an alternative, you can use a small deep subsonic sine wave, like 1 Hz or even 0.1 Hz. The sine wave has to be large enough to shift the tone into the audio band; this implies it has to be larger than the DC offset that the DAC manufacturer added deliberately, if any. It shouldn't be too large, though, because then the tone may sweep through the audio band too quickly to be noticed. Both with swept DC levels and with deep subsonic sine waves, it is advisable to put a high-pass between the DAC and the amplifier to ensure you don't trigger any DC protections or fry any woofers.

5. Test with a 0 dBFS 1 Hz signal

A full-scale 1 Hz sine wave can show if the noise floor of the DAC gets amplitude modulated by the signal, the noise will then sound woosh-woosh-like instead of constant. The rate of the transitions between +1 and -1 varies with the signal, so it will show up when the DAC has some imperfection that makes its noise dependent on this rate. It is again advisable to put a high-pass between the DAC and the amplifier to ensure you don't trigger any DC protections or fry any woofers.