

## **Why Should I Record at 24 bit?**

### **My Notes, Just to Preserve My Sanity**

A CONCLUSION AS AN INTRODUCTION? .....	2
CALCULATED ANALOGUE LEVELS .....	2
WHAT DIGITAL LEVELS? .....	2
RESOLUTION? WHAT RESOLUTION?.....	3
WHAT ARE INTER-SAMPLE PEAKS? .....	5
THE REAL CONCLUSION .....	6
ACKNOWLEDGEMENTS.....	6

## 24 Bit Recording

### A Conclusion as an Introduction?

We are going to break with tradition and begin with the conclusion.

By recording at 24 bit and making sure no level on a single track peaks above -6dbfs whilst tracking, we eliminate overs and inter-sample peaks and our multi-track recordings will sound better. It is a simple and concise conclusion to the in-depth question of why we should record at 24 bit despite the fact that all 24 bit does is give us a greater signal to noise ratio. The above short statement will solve a lot of issues, the rest of this document will now try to explain why these very simple steps are true and why everyone should be recording their multi-tracks at 24 bit and keeping their levels down.

### Calculated Analogue Levels

A fundamental mistake was made in the infancy of digital recording when the crossover from analogue to digital was just being born. To illustrate this we must go back to how levels were measured and referenced in an analogue tape and console environment and how this was incorrectly passed over to the digital environment. Our journey begins with headroom, as measured by meters, and it is worth reminding ourselves that headroom is nothing more than aiming at a sensible level lower than the maximum possible. This is, at its most basic level, the very definition of headroom.

In the days of purely analogue systems it was quite normal to set levels that were -24db or more lower than the maximum possible. This gave you greater flexibility in the analogue console to boost and cut whether with faders or EQ or some other process without the electronics going into clipping or distortion.

The meters across an analogue system were all calibrated to correspond to a deliberately and sensibly chosen operating level that was based on both science and experience. A typical analogue console would have its meters calibrated to read 0Vu at 0dBu and everything above this level was deemed to be in the red area of the meter. The maximum possible output of the console, however, was around +24dBu so the console had +24dBu of headroom. Any signals that strayed into the red, or sometimes deliberately pushed into the red, were not lost, clipped or unrecoverable because of the headroom. Later on, the 0dBu reference was increased to +4dBu in order to reduce noise and to take advantage of higher peak outputs available in later and better circuit designs. When this happened the meters were simply re-calibrated to read 0Vu when the level was +4dBu. This meant that even though the engineer was still aiming at 0Vu, the actual signal was 4dB hotter to take advantage of the greater headroom in the newer and better equipment.

For the tape machines a very similar process was used. Differing makes of tape would handle levels slightly differently so the studio engineers would rely on the tape manufactures specification and their own experience when calibrating. The optimum level was around 10dB or so below the maximum possible level you could record and playback a 1KHz test tone. To achieve the correct operating level and sufficient headroom the engineer would play a reference tape with tones recorded at the recommended levels and adjust the playback meters to read 0Vu. The engineer would then adjust the tape machines output gain to read 0Vu on the mixing desk meters. Once this was done the engineer would send a 0Vu tone from the mixing desk and adjust the tape machines record gain to read 0Vu as well.

This all meant that you could set your record levels from the mixing console because you knew that all the meters were calibrated correctly but more importantly, even when you hit 0Vu on your meters you still had plenty of headroom on your tape machines. This became pretty crucial when tape machines were sitting in their own room, away from the control room.

All you really need to take from this is that the setting of operating levels and headroom, even in a complex analogue environment, was only a question of setting gains and calibrating meters, the sole purpose of which was to get people to aim at the optimum levels for the equipment being used. These optimum levels naturally incorporated headroom of around +28dBu.

### What Digital Levels?

We are now going to skip quite mercilessly to the digital environment because in the digital domain the industry decided to set meters to read 100% at digital clipping with no allowance for headroom at all. The old school way of aiming for 0Vu on your meter, or even pushing it into the red, suddenly, and for no other reason than lack of

forethought, became a very dangerous thing to do indeed. As the years past and the technology progressed, the metering did not. A digital meter, as in your workstation, will display a maximum value of 0dBfs, which is full scale. Anything over this will produce digital clipping. At this point some of you will be, and quite rightly, protesting that whilst mixing in a DAW you can quite easily push the meters into the red and hence over 0dBfs. This is correct, you can. You shouldn't, but you can and we will cover this later in the document. Remember, we're talking about tracking at the moment.

For now we have a single question which is, what is the meter in our DAW measuring? Level? Volume?

The answer is neither.

One of the most essential things to grasp is that there is no sound inside your computer. There is only maths and 1's and 0's. Your audio interface, although offering lots of lovely things like inputs, outputs, MIDI etc; serves two very essential services. They are Analogue to Digital Conversion (ADC - on the way in) and Digital to Analogue Conversion (DAC - on the way out). We take this very much for granted but let's take a minute to think about what is going on here. We have our source, let's say it is an electric guitar. We plug our guitar into the guitar amp and choose a suitable mic, put the mic in an equally suitable position in front of the guitar amp speaker and run it into our chosen mic pre. All this is totally analogue with most of us striving for the best analogue front ends we can afford or get our hands on. The signal will then run into our interface and we'll assign a track in our DAW for recording. The moment our signal passes into the interface it passes through the ADC and becomes a digital signal inside the computer. It is no longer an analogue sound. It is no longer sound at all.

If this is the case, what are the meters measuring? Current digital metering systems in workstations measure sample value only. Since this is not decoded signal, the meters do not show actual signal level. One of the biggest and most widespread misapprehensions in digital is that sample values (as read on meters or seen on your editor) are signal. When in fact all that passes within the digital application is un-decoded sample value numbers, which are only turned back into signal at the very final stage when you play it back into the real world. The signal that you hear through your speakers is the reconstructed analogue waveform that your DAC constructs from the digital information.

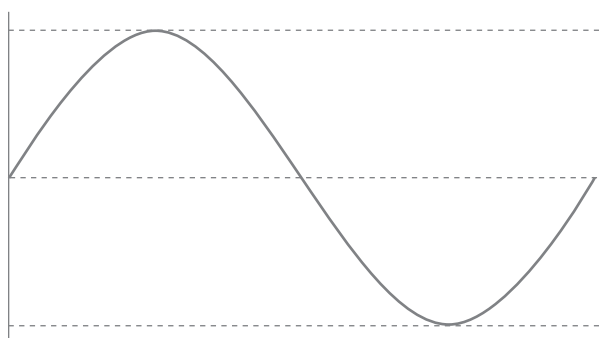
So are we saying that the meters in our DAWs are useless? Not at all, they just differ from their analogue counterparts and should be treated differently. It is the engineers responsibility to make sure there is enough headroom as there is no built in headroom and trying to hit 0dBfs whilst tracking will not result in a better recording. So what does this have to do with recording at 24bit? Where does sample rate come into it and if you don't record hot, aren't you losing resolution?

## Resolution? What Resolution?

As Paul Frindle points out, there is no such thing as resolution although I can see why people make comparisons to resolution.

Resolution is one of the biggest misconceptions that plagues our whole industry and probably the cause of more misunderstanding than any other issue. There is no such thing as resolution - it's a complete myth - please forget anyone ever used the word! The mathematical precision of the bit depth dictates the signal to noise ratio - NOT - the distortion or purity of the signal, and the bit depth has no impact on the frequency response, which is constrained by the sample rate.

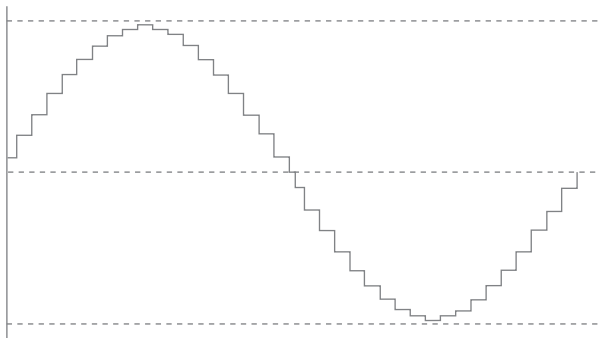
We are going to take a few paragraphs to illustrate why the above statements are true. Very basically, the sample rate controls how much frequency response is captured. Old telephone systems were the equivalent of about 3-4kHz of usable sound. It didn't really need to be any more than that. 44.1kHz will easily give us 22.05kHz of usable frequency, a range greater than most human hearing. This ratio is set by the Nyquist-Shannon sampling theorem which you can look up if you want a more thorough and mathematical explanation.



If we play a sine wave into an oscilloscope, we will see an image that looks very much like the image on the left.

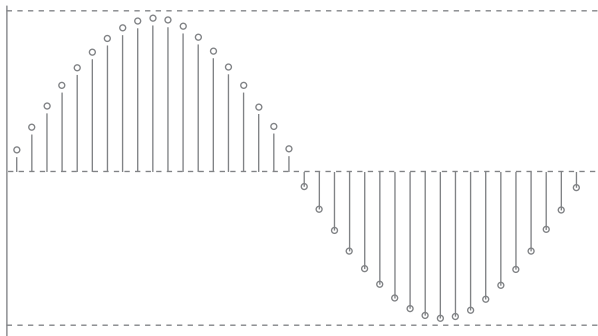
We like to describe this as an analogue image of a sine wave, perfectly smooth.

When we sample an analogue signal, the analog to digital converter will sample the sound at discrete points. These points are defined in time by the sampling frequency. So at 44.1 kHz we will get 44100 samples per second and at 48kHz we get 48000 samples per second.



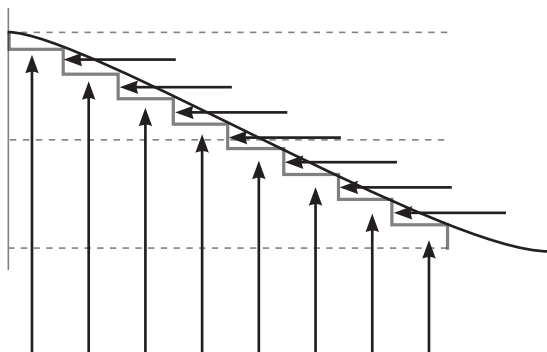
We often see a graph of a sampled sine wave as a stepped graph as pictured on the left. This is actually a misconception of what is really happening. It's not exactly wrong, it is just incomplete and it's the way scientists like to draw graphs.

The steps are not actually there at all. Drawing a graph like this is called Zero-Order Hold and it is incomplete because it is trying to depict the samples before reconstruction. We, ourselves, actually draw in the steps. There are no steps.

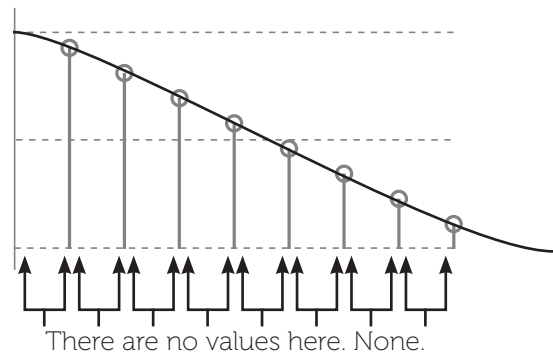


A much better way to represent the digitally sampled graph is as a lollipop chart as depicted on the left. The stepped graph above is indicating that there are values at every point along the stepped line. This is simply not happening. There is only a value where there is a sample and in between the samples there is no value at all. When the samples present themselves at the digital to analogue converter, the DAC reconstructs the signal by drawing in the analogue sine wave. What we get out is exactly the same as the input signal because there is only one way to join the dots.

The audio that is reconstructed after the DAC will again be a completely smooth analogue line if displayed on an oscilloscope when it comes out of your audio interface. There is no resolution. If you had a greater sample rate you would get more 'steps' on the Zero-Order Hold graph but we have shown that more samples, or increased sample rate, gives us greater frequency response only. It will not mean that the reconstructed curve is more or less accurate only what frequencies it can faithfully reproduce. The reconstructed output will have a pure, smooth analogue curve.



The stepped Zero-Order Hold graph, if read literally, would give us values on all parts of the steps, where the arrows are pointing. There are no values here at all. There are only the samples on the very corners of each step. When the DAC reconstructs the analogue waveform it will be a proper curve.



The lollipop graph has sample points in identical positions as the stepped graph on the left. It is the same chart but drawn so that it reflects way more accurately what is actually going on. The reconstructed curve on both graphs is exactly the same. No matter how many steps you have, there is no resolution.

Now we are going to begin to link our sample rate and our bit depth together. When an analogue sound source is sampled it is sampled in both time and frequency. Seeing as the analogue signal can be anywhere in nature and a digital capture is fixed to a grid, the converted signal is quantised to the nearest digital value if need be. This quantisation adds noise and this noise will turn up as harmonics in the signal. These harmonics are unwanted frequencies and were not there in the original signal. In fact, quantisation is the only thing that adds noise to a digital signal. This does not eliminate a digital recording capturing the noise from a noisy piece of gear, of course. Using 16 bits will give us 65536 values and using 24 bits will give us over 16.7 million values.

These unwanted harmonics are caused by our regular discrete measurements as set by our bit depth. This is also partly to blame for the term 'resolution'. You would be right in thinking the more values you had (higher bit depth) the greater precision you would have and the more accurately you could describe an event. Although this is true, no amount of counting, short of infinity, will give you a completely accurate representation of a natural event. Increasing the steps to try and eliminate distortion is a lost cause as some will remain whatever you do and however much data you waste on it.

The way we avoid this harmonic error is to turn it into random noise by adding statistically random values to the signal so that the quantisation value steps are no longer the same every time. The steps are blurred out and this is what we call dithering.

We are left with a signal that has no harmonic error, just the signal, with some added random noise due to the dither blurring the measurement steps. This becomes the equivalent of an analogue signal passed through an unquantised system that has a finite signal to noise ratio - i.e. just like the real world around us.

So the real perception of resolution comes from our graphs on the previous page. People are looking, all the time, at editing screens that show discrete sample values purely because they have to, because this is all there is to see within data on our systems. This leads us, incorrectly, to the conclusion that the signal will end up sounding like it looks on the screen. It won't. These sample values are not signal until they are decoded and reconstructed by your DAC. When they are decoded and reconstructed they are pure analogue. The term 'High Resolution Audio' is nothing but a marketing term. It means nothing. Your DAC is a Digital to Analogue converter.

I am trying to keep maths to a minimum in this document but it is worth doing a little now to show us how bit depth gives us a greater signal to noise ratio. This is all bit depth gives us.

24 bit is (2 to the power of 24) = 16777216 steps. This means that each step is 1/16777216 of the total.

If the total is called 0dBfs then the error of the steps would be  $20 \cdot \log(1/16777216) = 144.49\text{dB}$

Smoothing out the steps with noise will cost another 3dB so the total SNR would be 141.49dB.

For us, this means that after we dither we have a perfect signal without any steps or distortion with some random noise around -141.49dB below 0dBfs.

16 bit gives us 65536 steps. After the maths, it gives us -93dB.

So, the mathematical precision dictates the signal to noise ratio - NOT - the distortion or purity of the signal.

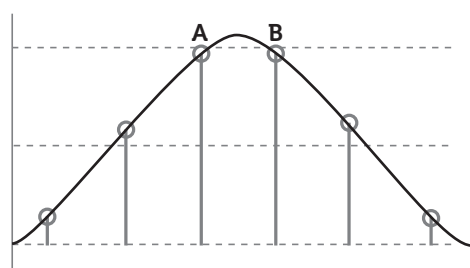
## What are Inter-Sample Peaks?

Why should I keep my levels low and aim for -6dBfs, max, whilst tracking? Why does multi-track audio in the box work and sound better at lower levels if we record at 24 bit? By this point we know that it doesn't because there is no audio in the box, there is only maths. So our real question is, why does the maths work better?

Earlier, we briefly touched on mixing in our workstations and how we can, theoretically, hit the red and go over 0dBfs. Most workstations have an internal audio engine that works at 32 bit float. To cut a long story short this gives us phenomenal headroom, but there's a price. There's always a price.

You only get that headroom whilst everything is inside your computer. At some point it has to come out into the real world and so has to come back to a fixed point bit depth. We must also remember that our workstations have many processes that we run our audio through. This can be as simple as nudging a fader or as complex as a string of sophisticated plug-ins.

Every signal we create is a new signal with the same requirements, just like in analogue. The fact that digital is a mathematical representation and does not have intrinsic natural uncertainty, does not let it off the hook. This means that even when we turn a fader up a bit we are creating a new signal. Every time we pass our signal through a plug-in we are creating a new signal. We keep our levels relatively low to protect us from this. Whilst in the digital domain our meters can lie to us because, as we have learned, they tell us sample value and not reconstructed analogue levels. This means that we can get overs (intersample peaks) that will clip our



DAC without reflecting it on the meters in our workstation because those meters come before the DAC, before the analogue signal is reconstructed.

The two samples labelled 'A' and 'B' in the left-hand diagram are just below 0dBfs. Our meters would not peak. When the DAC reconstructs the actual analogue signal, depicted in the diagram by the black curve, the resulting analogue signal between 'A' and 'B' would actually be above

0dBfs. It would be clipping. This is an intersample peak.

We also have to take note that some plug-ins can clip internally and if this is severe enough we can hear it as distortion. Care must be taken in pushing levels inside the plug-ins themselves as the output from a plug-in is then sent to the next process. The output of a compressor can quite easily be cranked up to give us more volume on a track instead of just being used for make-up gain. This can lead to the signal clipping if the output is pushed excessively instead of using another process, like pushing the fader up.

Generally, plug-ins running on a floating point system can stand to be overdriven providing the signals are not sent out to fixed point systems. We must always bare in mind that eventually, all signals will be sent to a fixed point system. Just be careful.

Plug-ins running in a fixed point system may stand to be overdriven if the inputs themselves are not overdriven, they have internal headroom and they have output level controls to reduce the level before it is sent to the next process.

Some plug-ins need to have internal references to real world output levels to operate properly. These would be dynamics, limiters and character plugs with distortion or saturation effects. Because they need to have this internal reference they may produce different results depending on absolute level whether they are in a fixed or floating point system (i.e. the float does not take away the need for real level references). Here's the catch, isn't there always one?

Since all of the plug-ins that need internal reference levels are designed to match full sample value level (the original default 'operating level' 0dBfs which has caused the whole problem in the first place!), they may need to be modulated to full level internally to get the intended results.

So basically, a plug-in can cause an intersample peak or just plain overloads if we're running our signals too hot. The answer? Yes, you already know it... headroom, headroom, headroom.

## The Real Conclusion

The above sections raise many issues and touch on them in language that I hope is generally understood. It seems quite amazing that keeping our levels to -6dBfs max whilst tracking and watching how we push our levels whilst mixing can fix 99% of any technical or technology based problems we may come across in our attempts to get decent signals recorded and to mix fine sounding records.

Some of the points put across in these notes are controversial, especially in an industry that seems to be lead more by marketing than it does anything else. We must not forget that the production of music is purely an artistic thing but the methods and methodology for capturing and manipulating our audio lies on a bed of science, maths and physics. It always has done, right back to the days of early analogue. But as our technology has changed, so have the parameters changed to how best to handle the capturing and replaying of audio. That does not mean, even for a second, that we have to be mathematicians or physicists to record music. We don't.

This document is really just my personal notes so that I can try to understand processes better and to try and figure out what is actually going on when I record and mix digitally. It is an ever-changing subject with ever-moving goalposts and will change over time. Additions will be made, corrections will be made and opinions will change as will the technology, which is ever improving and evolving.

The very bottom (and very basic) line is that by recording at 24 bit we get a way better signal to noise ratio, and that is all it does. With a better signal to noise ratio we can easily afford to keep the levels down and by keeping the levels down we diminish the possibility of mathematical errors inside our workstations.

## Acknowledgements

This document is made up of some of my own findings. I have read a great deal from people such as Paul Frindle, Bob Katz, Bob Olhsson and Chris Montgomery, over at xiph.org, plus many others. Some passages are taken from forum discussions that went on over a period of several years with the relevant parts collated and redistributed so that I could better understand them.