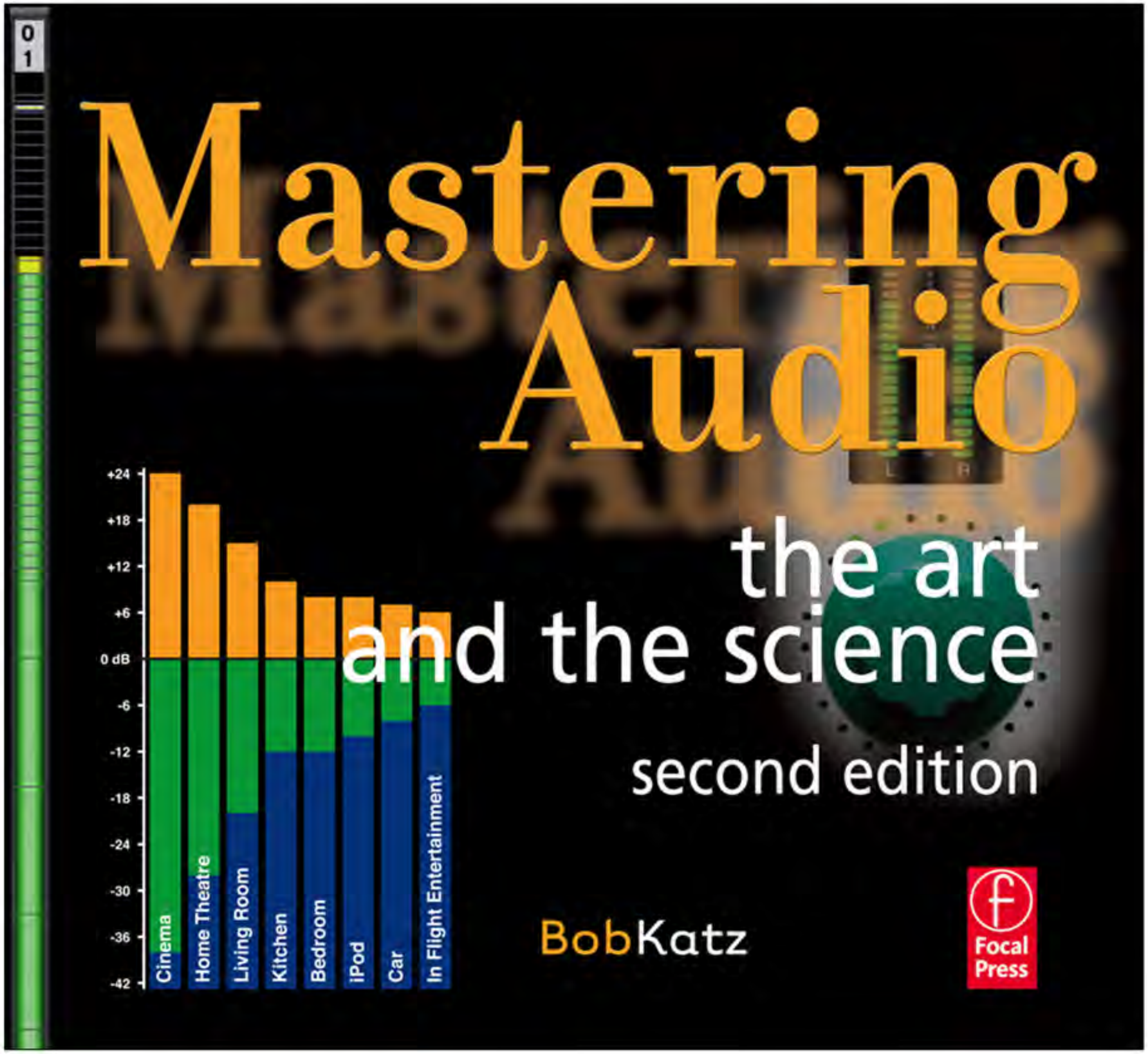

Mastering Audio
the art and the science
second edition
BobKatz
Focal Press
+24
+18
+12
+6
0 dB
-6
-12
-18
-24
-30
-36
-42
Cinema
Home Theatre
Living Room
Kitchen
Bedroom
iPod
Car
In Flight Entertainment

Mastering Audio

the art and the science
second edition

Excerpted from:
Mastering Audio: The Art and the Science, Second Edition

Most chapters in this book focus on the art of producing good-sounding audio masters. This introductory chapter is perhaps the most technical, but it covers required basic definitions and audio concepts that all audio engineers should learn. It then advances to the most up-to-date metering and leveling approaches for mixing and mastering engineers.

CHAPTER 5

Decibels: Not For Dummies

> "**Level** *is often confused with* **Gain**!"

Introduction

So many of us take our meters for granted—after all, recording is simple: *all you do is peak to 0 dB and never go over!* But things only appear that simple until you discover that with the same material, one machine says that it peaks to -1 dB, another machine shows an OVER level, and yet your workstation tells you it just reaches 0 dB! To make things worse, among the expensive digital meters available, only a handful accurately convey the information we really need to know. In this chapter we will explore the different types of meters, the concept of the digital OVER, analog and digital headroom, gain staging, loudness, signal-to-noise ratio and we will also take a fresh look at the common practices of dubbing and level calibration.

I. Stamp Out Slippery Language

Bob's Top 10 List of Slippery & Confusing Audio Terms[1]

10. Intensity... is a measure of energy flow per unit area. For practical purposes, sound intensity is the same as SPL (see below).[2]

9. Level... is a measure of intensity, but when used alone, because it can mean almost anything, it means absolutely nothing! To avoid confusion, the *level figure should always be qualified by a 'unit' term, e.g. voltage level, sound pressure level, digital level.* Examples: 40 dB SPL, -20 dBu, -25 dBFS. Each suffix defines the reference. SPL (sound pressure level) is a measure of the amplitude or energy of the physical sound present in the atmosphere. 40 dB SPL and 0.002 Pa (Pascals) are the same pressure, the first expressed in decibels relative to 0 dB SPL, the second in absolute pressure units.

Level is very often confused with **Gain**. For example, the sound pressure level from your monitor loudspeakers is often confused with the monitor gain. The term *monitor gain* is so slippery that I have started using a more solid term that everyone seems to understand: **MONITOR (control) POSITION.** For example, I say "the monitor control is at the 0 dB position."

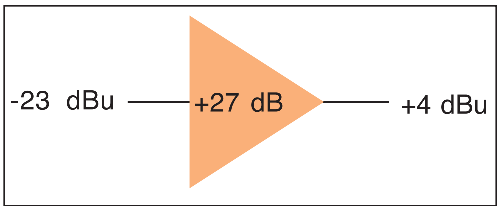
Stamp Out Monitor Controls Marked in SPL! Some manufacturers think it's sophisticated to label a monitor control in dB SPL, such as 73, 74, 75. But this is confusing. They forget that gain must always be expressed in simple decibels with no suffix. When people use those mislabeled controls, they talk in circles. Speaking of the control position, Betty says: "I always play back at 83 dB." Thinking she meant sound pressure level, Fred replies: "But I measured 90 dB at your loudspeakers." That's why a monitor control should be simply marked from 0 dB to -∞, like any other attenuator.

8. Decibel... (dB) is a relative quantity; it is always expressed as a ratio, compared to a *reference*. For example, what if every length had to be compared to one centimeter? You'd say, "this piece of string is ten times longer than one centimeter." It's the same thing with decibels, though sometimes the reference is not explicitly stated but just implied. +10 dB means "10 dB more than my reference, which I defined as 0 dB." Decibels are logarithmic ratios, so if we mean "twice as much voltage," we say "6 dB more" [20 * log (2) = 6].

dBu, dBm, dB SPL, dBFS... are ratios with predefined references, so they can be converted to absolute values in volts, power, etc. I believe the term **dBu** was coined in the 1960's by the Neve Corporation; it means *decibels unterminated, compared to a voltage reference of 0.775 volts.* **dBm** means *decibels compared to a power reference of one milliwatt.* **dBFS** means *decibels compared to full scale PCM;* 0 dBFS represents the highest digital level we can encode.

Plurals. We **do** say "two decibels", but we do **not** pluralize the abbreviation. We do say "two dee bee", but we do not say "two dee bees".

7. Gain or Amplification... is a **relative** term expressed in plain decibels with no suffix: it is the ratio of the amplifier's output level to the input level. If an amplifier receives an input of -23 dBu and puts out an output of +4 dBu, it has 27 dB gain (without any suffix). See Sidebar.



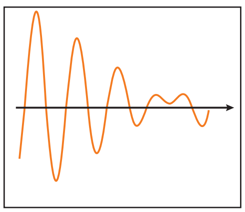
The meaning of Gain vs. Level. An amplifier with 27 dB gain is fed an input signal whose level is -23 dBu to yield an output level of +4 dBu. The decibels of gain should never need a suffix.

6. Attenuation... when expressed in dB is an optional term for negative gain, e.g. a loss. Examples: 20 dB attenuation is the same as -20 dB gain.

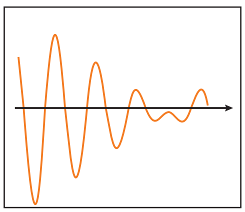
5. Sound Pressure Level (SPL)... is a measure of sound pressure in dB relative to 0.0002 dyne/cm^2 (0 dB SPL).[3] 74 dB SPL is the typical level of spoken word 12 inches away, which increases to 94 dB SPL at one inch distance. While we often see language like *95 dB SPL loud*, this is both inaccurate and ill-defined, as *loud* refers to the user's perception, and *SPL* to the physical intensity.

4. Loudness... is used specifically and precisely for *the listener's perception. Loudness* is much more difficult to represent in a metering system, in fact, it's best presented as a series of numbers rather than as one overall figure of "loudness." Two pieces of music that measure the same on a flat level meter can have drastically different *loudness.* A true loudness meter makes a complex calculation using SPL, frequency content, and duration. Exposure time also affects our perception; after a five minute rest, the music seems much louder, but then we get used to it again—another reason why it is wise to have an SPL meter around to keep us from damaging our ears.

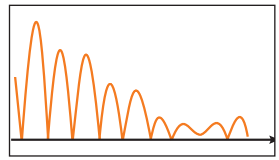
3. Intrinsic Loudness... In the first edition I invented a term "absolute loudness", but it wrongly gives the impression that we somehow have control over the consumer's volume control. So I'm replacing that with a new term **intrinsic loudness**, which I define as the loudness of a program **before the level is adjusted using the monitor control**. Since there is no SPL reference in a digital file, intrinsic loudness has no absolute units, but the term can be used in a relative way. We can compare two programs' intrinsic loudness by switching between them, adjusting the monitor control until they sound equally loud, and noting the decibel difference in the monitor positions. Then we can say that program 1 has "2 dB more intrinsic loudness than program 2" though, for brevity, I may say program 1 is 2 dB "louder", using quotation marks. When I use the term **hot CD** or **hot master**, I am referring to a recording which has a high intrinsic loudness. Our perception of the program's loudness is also affected by the behavior of the monitor DAC. For if a program is so distorted that its analog reconstructed level would cause a certain DAC to overload, this DAC may appear louder due to the high frequency distortion. **Intrinsic loudness** would not have been a meaningful term in the analog era because analog tapes and LPs do not have a consistent reference, but with digital 0 dBFS is always the same.



Which of these two waves is louder?



2A. Average vs. Peak Level... Which of these two waves is louder? The answer is: Both have the same loudness. Both have the same average and maximum peak level. The first wave is identical to the second except its polarity is exactly reversed.[4] Think of this picture as the graph of the movement of a drumhead seen from the front and back side. The terms average and peak are a bit misleading if we consider the up-going direction as *positive* and the down-going as *negative*, because the top wave has a maximum in the positive direction and the bottom has a maximum in the negative; however the **absolute value** of the greatest peak in each wave is the same, so we say each wave has the same maximum peak level. Over a long period of time, the average of the positive and negative numbers must be zero, which is the static pressure of the atmosphere (the position of the drumhead at rest). However, the ear generally ignores polarity when it considers loudness, so it sees this wave **rectified** (pictured), where all the negative segments have been converted to a positive, or more correctly, absolute value.



Rectified wave

The average value of this rectified wave over time is somewhere between its peak value and zero; our meters determine this average depending on the method of averaging and the time length of the average. Sometimes we set our SPL meters to read peak level, but generally we set them to read average sound pressure level because this correlates closer to how the ear hears. For proper monitor calibration, a good SPL meter should use the *RMS* averaging method, as opposed to a simple averaging which can produce as much as 2 dB measurement error. RMS ignores issues such as phase shift and tells us the true **energy level** of the recording. The word *average* in this book can refer to the true RMS level or the simple average; when it is important, I will specifically say RMS.

MYTH:

"The red light came on while I was recording, but when I played it back, there weren't any overs, so I thought it was OK."

Contributed by Lynn Fuston

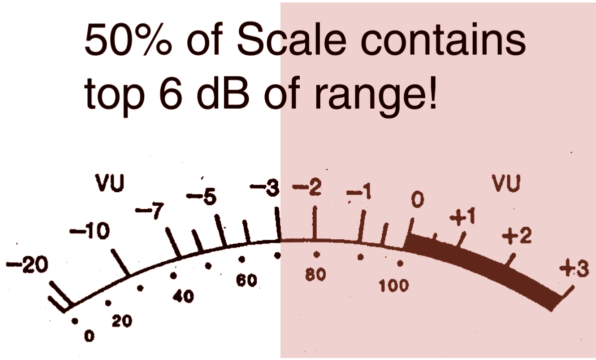
2B. Crest Factor, also known as **Peak-To-RMS Ratio**, is the difference between the RMS level of a musical passage and its instantaneous peak level. In practice, any averaging meter may be used, e.g. if a fortissimo passage measures -20 dBFS on the averaging meter and the highest momentary peak is -3 dBFS on the peak meter, it has a crest factor of 17 dB. It is extremely rare to encounter a piece of music with a crest factor greater than 20 dB, so this is the commonly cited maximum. When the dynamic range (the difference between the loudest and softest passages) of a recording has been reduced, we say the material has been **compressed**, and that a compressed recording has a lower crest factor than an uncompressed one.[5]

And the #1 most confusing audio term is...

1. Volume ... is **usually associated with an audio level control, but is an imprecise consumer term.** Volume is measured in quarts, liters and cubic meters! The words more properly used in our art are **Level** and **Loudness.** The big problem is that consumers use the term ambiguously, to mean both the loudness they perceive and the position of the "volume control"—a perfect example of confusing gain with level! So in this book I prefer to use the professional term **monitor control.** I rarely use the word *volume* except when speaking informally to clients or consumers, and occasionally succumb to saying "volume control" when referring to a consumer's system.

II. Meters Meters Meters

We Won't Get Fooled Again. Recording engineers rely heavily on their favorite meter, and this book is not intended to change people's favorite. But as practicing engineers, it is prudent to learn the defects and virtues of each meter we encounter.



VU meter operators are often fooled into treating the top and bottom halves of the scale with equal weight, but the top half has only 6 dB of the total dynamic range.

The VU Meter. Relative newcomers to the industry may have never seen a VU meter, and some of them may be using the word "VU" incorrectly to describe peak-reading digital meters. **VU** should only be applied to a true VU meter that meets a certain standard. The first thing we must learn is that *the VU meter is a dreadful liar...* It is an averaging meter, and so it cannot indicate true peaks, nor can it protect us from overload. However the VU does do one thing better than a peak meter—it comes closer to our perception of loudness, but even so, it is a very inaccurate loudness meter because its frequency response gives low frequency information equal weight, and the ear responds less to low frequencies. Another problem is that the VU meter's scale is so non-linear that inexperienced operators think that the greater part of the musical action should live between -6 and +3 VU, but this is wrong. A well-engineered music program has plenty of meaningful life down to about -20 VU, but since the needle hardly moves at that level, it scares the operator into thinking the level is too low. Only highly-processed (dynamically compressed) music can swing in such a narrow range; in other words, the VU scale encourages overcompression (we'll discuss compression of dynamic range in Chapter 9). Hence the VU meter should only be taken as a guide.[6] A much better averaging meter would have a linear-decibel scale, where each decibel has equal division and weight down to -20 dB. We'll discuss the use of averaging meters in more detail in Chapter 14.

Digital Peak Meters

Digital Peak meters come in three varieties:

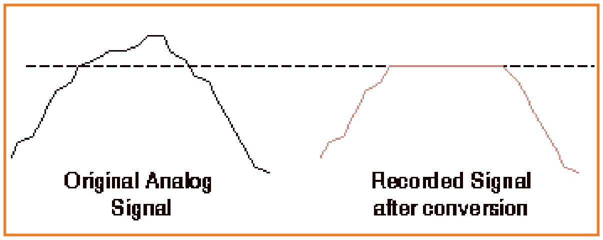
1. Cheap and dirty
2. Sample-accurate and sample-counting (but misleading)
3. Reconstruction (oversampling)

Cheap and Dirty Peak Meters. Recorder manufacturers pack a lot in a little box, often compromising on meter design to cut production costs. A few machines even have meters which are driven from analog circuitry—a definite source of inaccuracy. Some manufacturers who drive their meters digitally (by the values of the sample numbers) cut costs by putting large gaps on the meter scale (avoiding expensive illuminated segments). The result is that there may be a -3 and a 0 dB point, with a large unhelpful no man's land in between. When recording with a meter that has a wide gap between -3 and 0, it is best practice to stay well below full scale.

Sample-Accurate and Sample-Counting Meters.

Several manufacturers have produced sample-accurate meters with 1 dB (or smaller) steps, that convert the numeric value of the samples to a representation of the sample value, expressed in dBFS.[7]

The Paradox of the Digital OVER. When it comes to playback, a meter cannot tell the difference between a level of 0 dBFS (*FS = Full Scale)* and an OVER. That's because once the digital signal has been recorded, the sample level cannot exceed full scale, as in this figure.



While an original analog signal can exceed the amplitude of 0 dB, after conversion there will be no level above 0, yielding a distorted square wave. This diagram shows a positive-going signal, but the same is true on the negative-going end.

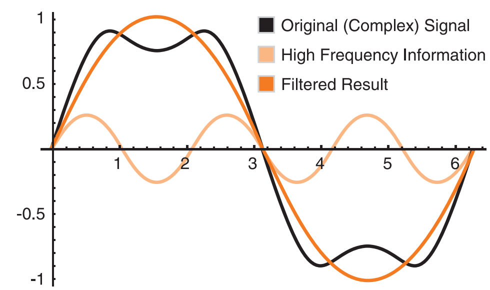
We need a means of knowing if the ADC is being overloaded during recording. So we can use an early-warning indicator—an **analog** level sensor prior to A/D conversion—which causes the OVER indicator to illuminate if the analog level is greater than the voltage equivalent to 0 dBFS. If the analog record level is not reduced, then a maximum level of 0 dB will be recorded for the duration of the overload, producing a distorted square wave.

After the signal has been recorded, a standard sample-accurate meter cannot distinguish between full scale and any part of the signal that had gone over during recording, it shows the highest level as 0 dBFS. However, a **sample-counting meter** can analyze a recording to see if the ADC had been overdriven. This meter counts contiguous samples and can actually distinguish between 0 dBFS and an OVER after the recording has been made! The sample-counting digital meter determines an OVER by counting the number of samples in a row at 0 dB. If 3 contiguous samples equal 0 dBFS, the meter signals an OVER, because it's fair to assume that the incoming analog audio level must have exceeded 0 dBFS somewhere between sample number one and three.[8] Three samples at 44.1 kHz is a very conservative standard; on that basis, the recorded distortion would last only 33 **microseconds** and would probably be inaudible. While this type of meter was sophisticated in its day, current thinking is that the sample-counting meter is only suitable for evaluating whether an ADC has overloaded. Authorities now feel that meters which display the digital value of the samples and which count samples to determine an OVER are no longer sufficient for mastering purposes and should be used with caution during mixing. Their place is taken by…

The Reconstruction Meter: Even More Sophisticated

As long as a signal remains in the digital domain, the sample level of the digital stream is sufficient to tell us if we have an OVER. However, **signals which migrate between domains can exceed 0 dBFS and cause distortion.** This includes any signal that passes through a DAC, a sample rate converter, or is converted through a codec such as mp3 or AC3. During the conversion from PCM digital to analog or mp3, filtering within the converter yields occasional peaks **between the samples** that are higher than the digital stream's measured level, which can be higher than full scale. This next figure shows that contrary to what we might assume, **filtering or dips in an equalizer which we'd imagine would produce a** ***lower*** **output can actually produce a higher output level than the source signal**. B.J. Buchalter explains that

> the third harmonic is out of phase with the fundamental at the peak values of the fundamental, so it serves to reduce the overall amplitude of the composite signal. By introducing the filter, you have removed this canceling effect between the two harmonics, and as a result the signal amplitude increases. Another reason for the phenomenon is that all filters resonate, and generally speaking, the sharper the filter, the greater the resonance.[9]



In black is a complex wave. When the high frequency information (light orange) is filtered out, the result is a signal (orange) that is higher in amplitude than the original!

Equipment designers have known for years that because of filtering, the analog output level of

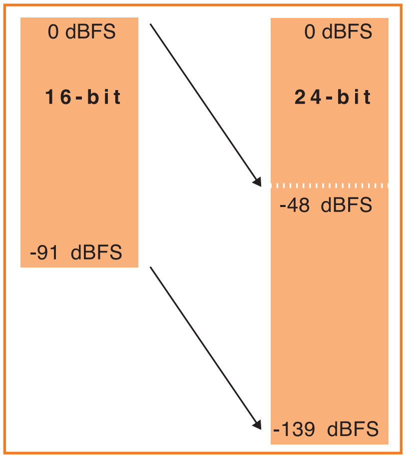
complex audio from a DAC can exceed the sinewave value of 0 dBFS but very few have taken this into account in the design. TC Electronic has performed tests on typical consumer DACs,[10] showing that many of them distort severely since their digital filters and analog output stages do not have the headroom to accommodate output levels which exceed 0 dBFS! While typical 0 dBFS+ peaks do not exceed +0.3 dBFS, some very rare 0 dBFS+ peaks may exceed full scale by as much as 4 or 5 dB with certain types of signals—especially mastered material which has been highly processed, **clipped** (turned into a square wave on top and bottom), and/or brightly equalized.[11] By oversampling the signal, we can measure peaks that would occur after filtering. An oversampling meter (or reconstruction meter) calculates what these peaks would be, but these meters are still rare. Products from TC Electronic (System 6000) and Sony (Oxford) have an oversampling limiter and reconstruction peak meter. RME's Digicheck software includes an oversampling meter.[12]

"Signals which cross domains can exceed 0 dBFS"

Reconstruction meters tell us not only how our DAC will react, but what may happen to the signal after it is converted to mp3 or sent to broadcast, both of which employ many filters and post-processes. Many DSP-based consumer players cannot handle the high levels at all and exhibit severe distortion with 0 dBFS+ signals. Armed with this knowledge, no mastering engineer should produce a master that may sound acceptable in the control room but which she knows will likely produce severe distortion when post-processed or auditioned in the real world. If the reconstruction meter is not enough to convince the client, she should also demonstrate that this "loud" signal becomes distorted, ugly, and soft when it is converted to low bit rate mp3. All the harmonics which made the signal seem loud in the control room have been converted to additional distortion. For example, ***Figure C5-1*** in the color plates is a spectragram of the remnant distortion from an mp3 conversion of two different CDs. Time moves from left to right, levels vary from highest in red to lowest in blue, and high frequencies are at the top. At left, a clipped CD master; at right, a "loud" master made without clipping, then brickwall limiting. Note how the high frequency distortion in CD #1 cuts off sharply below about 12 kHz, indicating some kind of aliasing distortion in the mp3 converter.[13]

Practice Safe Levels

What this means is that if you are mixing with a standard digital meter, keep peaks below -3 dBFS, especially if you are using aggressive bus processing.[14] The more severely processed, equalized or compressed a master, the more problems it can cause when it leaves the mastering studio. We didn't start hearing about this problem, or at least the severity of it, before the loudness race and the invention of digital processing which could be egregiously abused. Maximizing engineers should try to use a



A 24-bit recording would have to be lowered in level by 48 dB in order to reduce it to the SNR of 16-bit. The noise floors shown are with flat dither.

reconstruction meter and/or an oversampled brickwall limiter. If these are not available, use a standard peak limiter whose ceiling is set to -0.3 dB (see Chapter 10) and exercise caution. But even the oversampled brickwall limiter is not foolproof; I've discovered that such limiters do not protect from very severe processing and can still make a consumer DAC overload unpleasantly. The best solution is to be conservative on levels. Clipping of any type is to be avoided, as demonstrated in Appendix 1.

The Myth of the Magic Clip Removal

If the level is turned down by as little as 0.1 dB, then a recording which may be full of OVERs will no longer measure any overs. But this does not get rid of the clipping or the distortion, it merely prevents it from triggering the meter. Some mastering engineers deliberately clip the signal severely, and then drop the level slightly, so that the meters will not show any OVERs. This practice, known as **SHRED**, produces very fatiguing (and potentially boringly similar) recordings.[15]

> *"You would have to lower the peak level of a 24-bit recording by 48 dB to yield an effective 16-bit recording!"*

Peak Level Practice for Good 24-bit Recording

Even though 24-bit recording is now the norm, some engineers retain the habit of trying to hit the top of the meters, which is totally unnecessary as illustrated at left. Note that a 16-bit recording fits entirely in the bottom 91 dB of the 24-bit. You would have to lower the peak level of a 24-bit recording by 48 dB to yield an effective 16-bit recording! There is a lot of room at the bottom, so you won't lose any dynamic range if you peak to -3 dBFS or even as low as -10 dBFS, and you'll end up with a cleaner recording. Since distortion accumulates, if a "hot" recording arrives for mastering, the mastering engineer doing analog processing may have to attenuate the level to prevent the processing DAC from overloading. A digital mix that peaks to -3 dBFS or lower makes it easier to equalize and otherwise process without needing an extra stage of attenuation in the mastering.

A number of 24-bit ADCs are advertised as having *additional headroom*, achieved by employing a built-in compressor at the top of the scale, claiming that the compressor can also protect the ADC from accidental overloads. But this is specious advertising. Level accidents don't occur in a mix studio; engineers have control over their levels and when tracking live musicians, it is better to turn off the ADC's compressor, drop the level and leave plenty of headroom for peaks. The only possible use of this function of a compressor is if you like its *sonic qualities* and are trying to emulate the sound of tracking to analog tape. But since tracking decisions are not reversible, I suggest postponing "analog simulation" to the mixing stage. It's easier to add warmth later than try to take away some mushiness due to an overdriven compressor. As we have just seen, there is no audible improvement in SNR by maximizing a 24-bit recording and no SNR advantage to compressing levels with a good 24-bit ADC.

How Loud is It?

Contrary to popular belief, the levels on a digital peak meter have (almost) nothing to do with loudness. Here is an illustration. Suppose you are doing a direct to two-track recording (some engineers do still work that way!) and you've found the perfect mix. Leaving the faders alone, you let the musicians do a couple of great takes. During take one, the performance reached -4 dB on the meter; and in take two, it reached 0 dB for a brief moment during a snare drum hit. Does that mean that take two is louder? No: **because in general, the ear responds to average levels, not peak levels when judging loudness**. If you raise the master gain of take one by 4 dB so that it too reaches 0 dBFS peak, it will sound 4 dB louder than take two, even though they both now measure **the same** on the peak meter.

An analog tape and digital recording of the same source peaked to full scale sound very different in terms of loudness. If we make an analog tape recording and a digital recording of the same music, and then dub the analog recording to digital, peaking at the same peak level as the digital recording, the analog dub will have about 6 dB more intrinsic loudness than the all-digital recording. Quite a difference! This is because the peak-to-average ratio of an analog recording can be as much as 12-14 dB, compared with as much as 20 dB for an uncompressed digital recording. Analog tape's built-in compressor is a means of getting recordings to sound louder (oops, did I just reveal a secret?).[16] That's why pop producers who record digitally may have to compress or limit to compete with the loudness of their analog counterparts.

The Myths of Normalization

The Esthetic Myth. Digital audio editing programs have a feature called Peak Normalization, a semi-automatic method of adjusting levels. The engineer selects all the songs on the album, and the computer grinds away, searching for the highest peak level on the album and then automatically adjusts the level of all the material until the highest peak reaches 0 dBFS. If all the material is group-normalized at once, this is not a serious esthetic problem, as long as all the songs have been raised or lowered by the same amount. But it is also possible to select each song and normalize it individually, but this is a big mistake; since the ear responds to average levels, and normalization measures peak levels, the result can totally distort musical values. A ballad with little crest factor will be disproportionately increased and so will end up louder than a rock piece with lots of percussion!

The Technical Myth. It's also a myth that normalization improves the sound quality of a recording; it can only degrade it. Technically speaking, normalization adds one more degrading calculation and level of quantization distortion. And since the material has already been mixed, it has already been quantized, which predetermines its signal-to-noise ratio—which cannot then be further improved by raising it. Let me repeat: raising the level of the material will not alter its inherent signal-to-noise ratio but will add more quantization distortion. **Of course material to be mastered does not need normalizing** since the mastering engineer will be performing further

MYTH:
Peak Normalization Makes the Song Levels Correct

MYTH OF THE MAGIC CLIP REMOVAL: *Turn it down after clipping and the clip will go away.*

processing anyway.[17] Clients often ask: "do you normalize?" I reply that I never use the computer's automatic method, but rather songs are leveled by ear.

Average Normalization

This is another form of normalization, an attempt to create an intelligent loudness algorithm based on the average level of the music, as opposed to the peak. But when making an album, neither peak nor average normalization nor any intelligent loudness algorithm can do the right job, because the computer does not know that the ballad is supposed to sound soft. There's no substitute for the human ear. However, average normalization or better, a true intelligent loudness algorithm can help in situations where every program needs the same loudness, even if that doesn't sound natural, such as radio broadcast, ceiling loudspeakers in a store, a party or background listening.

Judging Loudness the Right Way

Since the ear is the only judge of loudness, is there any objective way to determine how loud your CD will sound? The first key is to use a single DAC to reproduce all your digital sources and maintain a fixed setting on your monitor gain. That way you can compare your *CD in the making* against other CDs, in the digital domain. Judge DVDs, CDs, workstations, and digital processors through this single converter.

III. Analog Studio Levels, Headroom and Cushion

Protecting the Mix from Clipping the ADC. Professional mixing studios with analog consoles are still using VU meters to measure average program level and feed the console output to an ADC. I use the term *nominal* to mean the voltage level with a sine wave that corresponds with 0 VU, typically 20 dB below full scale digital (0 dBFS). To set up the system, feed a sinewave through the analog system at 0 VU and adjust the gain of the ADC to produce -20 dBFS, measured with an accurate digital meter. This protects the mix from clipping the ADC since the peak to average ratio of typical music is no more than 20 dB, more typically 12 to 14 dB.

Headroom of the analog gear. Protecting your ADC and mix from clipping does no good if your analog console is distorting in front of the ADC! In the mastering suite we usually chain multiple pieces of analog gear, so it's important to learn about the analog levels, distortion and noise in the analog signal chain in front of your ADC.

Not all analog gear is created equal, and the standard nominal +4 dBu[18] may be too high for two reasons:

The first reason is the clipping point of cheaper analog gear has gone down over the years, to save money on parts. Before the advent of inexpensive 8-buss consoles, most professional consoles' clipping points were +24 dBu or higher. But frequently, low-priced console design uses circuits that clip at a lower level, around +20 dBu (7.75 volts). This can be a big impediment to clean audio, especially when cascading amplifiers.

The second reason is that in my opinion, some solid state circuits exhibit an extreme distortion increase **long before they reach the actual clipping point**.[19] So the music peak level must stay below the distortion region, not just the clipping point. To avoid the *solid-state edginess* that plagues a lot of solid state equipment, we should use amplifiers that clip at least 6 dB above the potential peak level of the music, or else lower the level of the signal to suit the amplifier. This means that for a 0 VU level of +4 dBu, the clipping point should be at least +30 dBu (24.5 volts RMS)! That's why more and more high-end professional equipment have clipping points as high as +37 dBu (55 volts!). To be that robust, an amplifier must use very high output devices and high-voltage power supplies. The effects of this are higher cost due to the need for more robust parts but, all other things being equal, the amplifier with the higher clipping point will sound better. Perhaps that's why tube equipment (with its 300 volt B+ supplies and headroom 30 dB or greater) often has a *good* name and solid state equipment with inadequate power supplies or headroom has a *bad* name (see Figure at right).

Cushion. Traditionally, the difference between **average level** and clip point has been called the *headroom*, but in order to deal with the increased distortion as the amplifier approaches clipping, I'll call the part of the headroom between the **peak level** of the music and the amplifier clip point a *cushion*. If an active balanced output feeds an unbalanced input, the clipping point reduces by 6 dB, so the situation becomes proportionally worse.[20] Dual-output consoles that are designed to work at either professional or semi-pro levels can be particularly problematic. Ironically, the lower output level in semi-pro mode may sound cleaner.

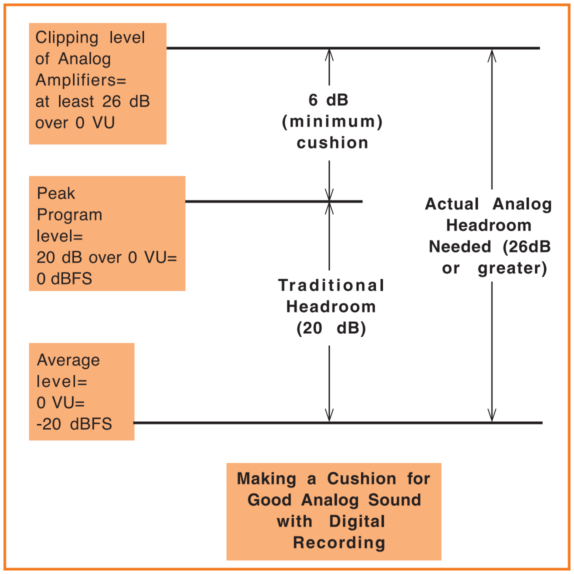
We should raise the question of whether the professional standard of +4 dBu is still appropriate because not every mastering or mixing studio has gear with extremely high clip points. Short of replacing all the gear, the easiest solution is to lower the analog level that represents nominal level or 0 VU. **I recommend a studio standard nominal analog level of 0 dBu, or 0.775 volts**. Just that simple 4 dB decrease can help produce a cleaner analog chain. Many European studios have been using a 0 dBu standard for decades for this very reason.

Internal clipping point in DAC. One of the most common mistakes made by digital equipment manufacturers is to assume that, if the digital signal *clips* at 0 dBFS, then it's OK to install a (cheap) analog output stage that would clip at a voltage equivalent to, say, 1 dB higher. This almost guarantees a nasty-sounding converter or recorder, because of the lack of cushion in its analog output section and the potential for 0 dBFS+ levels.

MYTH:
+4 dBu is always the best level to use for 0 VU with balanced analog electronics.

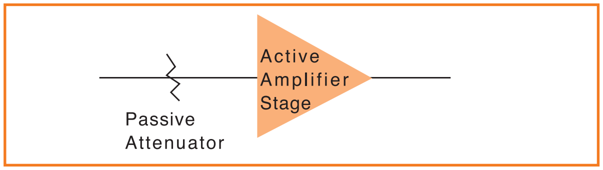
IV. Gain Staging—Analog Chains

Now that we know how to choose an analog level, it's time to chain our equipment together. To really get a handle on our equipment, we should determine its internal structure. The figures on the next page represent two possible internal structures. All complex equipment structures are variations on these themes.

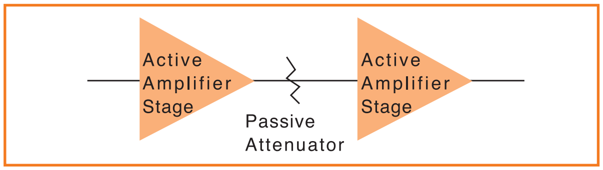


Making a Cushion for Good Analog Sound with Digital Recording

To properly test analog devices and determine their internal makeup, use a good clean monitor system, an oscilloscope, a digital voltmeter and a sine wave generator that can deliver a clean +24 dBu or higher (a tough requirement in itself). There are two different types of devices. The first type has a passive attenuator on its input, which means that we can feed it any reasonable source signal without fear of overload. We can ascertain whether there is a passive attenuator on the input side by turning the generator up and the attenuator down and observing whether or not the output clips. We can also disconnect the generator and listen to the output of the device as we raise and lower the attenuator. There should be no change in noise or hiss, and the output noise should be well below –70 dBu unweighted, preferably below –90 dBu A-Weighted. This is another indication that the device has a passive attenuator on its input. If the output noise changes significantly at intermediate positions of the attenuator, then the internal impedances of the circuit may not be optimal, or there may be some DC offset. The output noise of this device will be limited by the noise floor of its output amplifier. We determine the best **nominal operating level** of this device by taking the output clip point and subtract at least 26 dB for headroom and cushion.



In the top device, signal enters a passive attenuator and exits through an active amplifier stage. This circuit effectively has infinite input headroom. The bottom device's input headroom is determined by the headroom of the input amplifier.



The second type of device has an active amplifier stage on its input, the design of which is much more critical. It is very rare to find a solid state device built this way which won't clip with >+24 dBu input. While raising the signal generator, turn down the attenuator to keep the output from overloading. If we hear clipping prior to the generator reaching +24 dBu, then the device has a weak internal signal path. The clip point determines the nominal analog input level, which should be at least 26 dB below this clip point. Then, we should test to see that the output stage clips at a level no lower than the input stage.

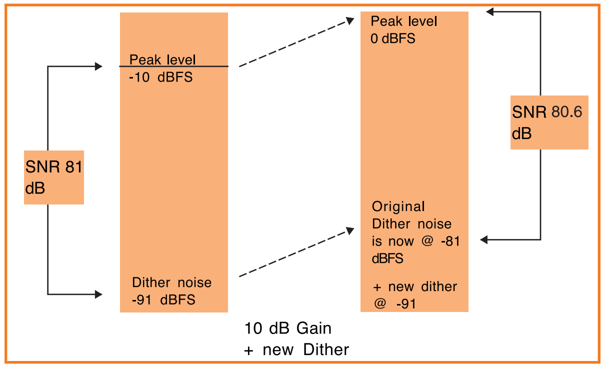
System noise. When cascading analog gear, the noise of the system is determined by the weakest link. Set your monitor gain with typical music, then listen closely to the noise floor at the last device in the chain; if the output of the chain sounds good and reasonably quiet, then don't worry about tweaking the chain. In an analog signal chain, raising the music signal level—as high as practical—as early as possible in the chain will improve the signal-to-noise ratio of the entire chain. In general, tube gear has a higher noise floor, so if gain has to be turned up, it should be in front of noisy tube gear, not after it.

V. Gain Staging—Digital Chains

There is no loss or gain in a digital interconnection such as AES/EBU or S/PDIF but we still have to be concerned about overloads. As we mentioned, equalizers can increase level even when dipping. Many outboard digital processors do not have accurate metering so I recommend patching an external digital meter to their output. If the processor overloads, try attenuating at either the input or the output.

Headroom of the Processor. We can test digital systems for headroom, clipping, and noise using digitally-generated test tones and an FFT analyzer. Suppose we have a digital equalizer with several gain controls and equalization; we feed it a 1 kHz sine wave test tone at about -6 dBFS and turn up the 1 kHz equalization by 10 dB, observing that the output clips. Then we turn down the output gain control until the output is below 0 dBFS and verify by listening or FFT measurements that the internal clipping stops. If the clipping does not stop, this indicates that the internal gain structure of the equalizer does not have enough headroom to handle wide range inputs. We may be able to get away with turning down the signal in front of the equalizer, or the EQ's input attenuator if it has one, but the early clipping indicates that this equalizer is not state-of-the-art. Modern-day digital processors should have enough internal headroom to sustain considerable boost in early stages without needing an input attenuator, and clipping can be removed solely by turning down the output attenuator.

Noise of the Digital Chain. With a digital chain, we no longer have to maximize the audio signal level in each piece of gear; a low level signal in a 24-bit digital signal chain does not hurt the SNR, considering the inaudible (approximately -139 dBFS) noise of the chain.[21] Instead of getting hung up on the signal level, we should consider every calculation stage as a source of quantization distortion. What matters most in a digital processing chain is to reduce the number of total calculations and use high-quality calculations, e.g. give the job of gain changes and other processing to the components with the highest internal resolution (those which would introduce the least quantization distortion or *grunge*). In fact, we should avoid raising the signal until it reaches a device which has the cleanest-sounding gain control, even if the source audio level is very low. For example, if the workstation has lower resolution than the outboard gear (shorter internal wordlength), we try to hold everything at unity gain in the DAW and reserve the gain changes or EQ for higher-precision devices later in the signal chain. Regardless of the level, pass a perfect clone (bit-transparent copy) of the source from the DAW onto the next device in line to do processing.



A 16-bit recording with peak level low at –10 dBFS. When gain is raised 10 dB and redither is added, the original 81 dB peak signal to RMS noise ratio is reduced by about 0.4 dB.

The significant noise floors in a 24-bit chain are not from the digital chain but from the original sources, including mike preamp noise, and our primary concern is with the impact of these higher level noises. We should be aware that noise floors sum in digital in the same way they do in analog, including dither noise. Let's take an example of a 16-bit recording whose peak level is 10 dB low, as in the above figure. In mastering we may choose to raise its level by 10 dB and so must add 16-bit dither before turning it into a 16-bit master. Disregarding the mike preamp and room noise, the original 16-bit recording's dither noise is at -91 dBFS and thus it has a peak signal to RMS noise ratio of 81 dB (-10 - -91).[22] When we raise the signal by 10 dB, both the original signal and the original noise are raised equally, so the original signal to noise ratio is unchanged. However, the new noise floor is the RMS sum of the original dither which has simply become noise at -81 dBFS and the added (new) dither which is at -91 dBFS. We ignore the insignificant noise of the gain processing, well below -130 dBFS, so the new dither raises the noise to -80.6 dBFS, less than 0.5 dB worse, so by raising the gain of the source, the new dither is an insignificant contributor to the total noise. However, if we add two 16-bit dithers without changing the gain of the source, the noise floor goes up by 3 dB—and this may cause a sonic veiling. Despite the potential for degradation, many times we still receive 16-bit sources; and we are forced to adjust the level according to the esthetics of the album. Fortunately I've had considerable success reducing cumulative sonic veiling by using noise-shaped dither.[23]

The manufacturers of the Waves L2 are concerned about losing resolution when converting from 24 to 16 bit. They feel that peak limiting allows raising level enough for it to be significantly above the dither noise, thus increasing the resolution. But exercise caution, because to my ears the apparent noise improvement is more than offset by the loss of transient clarity when peak limiting excessively. Is it worth sacrificing transient clarity just to gain a couple of dB increased signal-to-dither ratio?

If we could avoid 16-bit dither, by producing an output at 24-bit that the consumer could use, then mastering processing and gain-changing could be performed with no significant penalty, with a noise floor 48 dB below the noise of 16-bit. This is the promise of delivering higher wordlengths to the consumer and another reason to record in 24-bit in the first place.

The Imaginary World of Floating Point

A **fixed point processor** has a fixed maximum peak level of 0 dBFS and a fixed noise floor as per its wordlength, which for dithered 24 bit is approximately -139 dBFS. But a **floating point processor** is capable of doing tricks that do not relate to the real world. It is practically impossible to clip a floating point processor, you can raise the gain by hundreds of dB without clipping. Probably 95% of current native (CPU-based) plugins use floating point processing. Probably 80% of current outboard digital processors use floating point processing. However, **all converters use fixed point**, so wherever a floating point processor meets "the real fixed-point world", the signal must be regulated to a normal level. In other words, you can construct a floating point signal chain only internally within a DAW.[24]

In a floating point system, you can break all the rules: floating point can literally ignore the individual

From:

Mastering Audio: The Art and the Science, Second Edition.

The balance of this chapter explains how to take advantage of floating point processors in a DAW and avoid their weaknesses.