


[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## Level Practices

### (Part 1)

#### Part I: The 20th Century Dealing With Peaks

#### Overs, levels, and headroom, how to get the most from your equipment

*Digital recording is simple--all you do is peak to 0 dB and never go over!* And things remain that simple until you discover one plugin or processor telling you a signal peaks to -1 dB while another meter (e.g. in your DAW) shows an OVER level, yet your digital outboard processor tells you it just reaches 0 dB! This article will explore **concepts of the digital OVER**, machine meters, **loudness**, and take a fresh look at the common practices of **dubbing and level calibration**.

#### Section I: Digital Meters and OVER Indicators

Manufacturers often have to pack a lot in their product, therefore compromising on meter design and accuracy to cut costs. A few outboard machines' meters are driven from analog circuitry, a definite source of inaccuracy. Even manufacturers who drive their meters digitally (by the values of the sample numbers) cut costs by putting large gaps on the meter scale (avoiding costly illuminated segments), using inaccurate calculations and/or time constants or by just not translating the values right to the visible meter. As a result, there may be a -3 point and a 0 dB point, with a big no man's land in between and the values not being representative for the signals momentary peak-level. The manufacturer may feel he's doing you a favor by making the meter read 0 even if the actual level is between -1 and 0, or by setting the threshold of the OVER indicator inaccurately or too conservatively (longbefore an OVER actually occurs). Even if the meter has a segment at every decibel, on playback, the plugin or DAW may not be able to tell the difference between a level of 0 dBFS (*FS = Full Scale*) and an OVER. I would question the

OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

Password

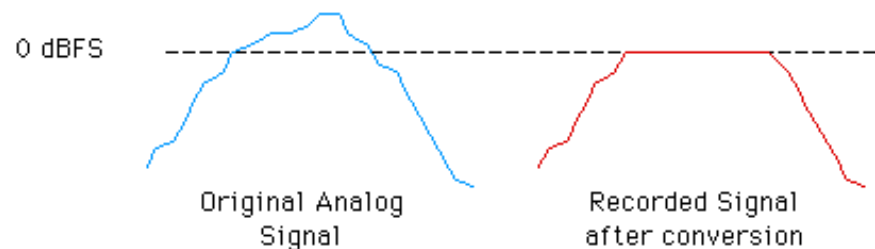
Remember me

machine's manufacturer if the OVER indicator lights on playback; it's probably a simple 0 dB detector rather than an OVER indicator.

There's only one way around this problem. Get a calibrated digital meter. Every studio should have one or two. There are lots of choices, from Dorrough, DK, Mytek, NTT, Pinguin, Sony, and others, each with unique features (including custom decay times and meter scales), but all the good meters agree on one thing: the definition of the highest measured digital audio level. A true digital audio meter reads the numeric code of the digital audio, and converts that to an accurate reading. A good digital audio meter can also distinguish between 0 dBFS and an OVER.

### The Paradox of the Digital OVER

If digital levels cannot exceed 0 dB (by definition, there's nothing higher), then how can a digital signal go OVER? One way a signal can go OVER is during recording from an analog source. Of course the digitally encoded level cannot exceed 0 dBFS, but a level sensor in an A/D converter causes the OVER indicator to illuminate if the analog level is greater than the voltage equivalent to 0 dBFS. If the recordist does not reduce the analog record level, then a maximum level of 0 dB will be recorded for the duration of the overload, producing a nicely distorted square wave. There is a simple (digital) way of detecting if an OVER had occurred, even on playback--by looking for consecutive samples at 0 dB, which is a square wave. A specialized digital meter determines an OVER by counting the number of samples in a row at 0 dB. The Sony 1630 OVER standard is three samples, because it's fair to assume that the analog audio level must have exceeded 0 dB somewhere between sample number one and three. Three samples is a very conservative standard--most authorities consider distortion lasting only 33 microseconds (three samples at 44.1 KHz) to be inaudible. Manufacturers of digital meters often provide a choice of setting the OVER threshold to 4, 5, or 6 contiguous samples, but in this case it's better to be conservative. Even 6 samples is hard to hear on many types of music, so if you stick with the 3-sample standard, you'll guarantee that virtually all audible OVERs will be nipped in the bud, or at least detected! Once you've used a good digital meter, you'll never want to go back to the built-in kind. In the diagram below, a positive-going analog signal goes OVER in the area above the dotted line.



### Using External A/D Converters or Processors

There is no standard for communicating OVERs on an AES/EBU or S/PDIF line. So if you're using an external A/D converter and feed the signal into any machine, the OVER indicator there will probably not function properly or at all. I advise ignoring the indicator if it does light up, unless

the manufacturer confirms that it's a sample counting OVER indicator. They'll probably reveal that it's an analog-driven level detector. Some external A/D converters do not have OVER indicators, so in this case, there's no substitute for an accurate external meter; without one I would advise not exceeding -1 dB on the feeded machine.

When making a digital dub through a digital processor you'll find most do not have accurate metering (be sure to read [The Secrets of Dither](#) before using any digital processor). Equalizer or processor sections can cause OVERs. Contrary to popular belief, an OVER can be generated even if a filter is set for attenuation instead of boost, because filters can ring. Digital processors can also overload internally in a fashion undetectable by a digital meter. Cascaded internal stages may "wrap around" when they overload, without transferring OVERs to the output. In those cases, a digital meter is not a foolproof OVER detector, and there's no substitute for the ear, but a good digital meter will catch most other transgressions. When you hear or detect an overload from a digital processor, try using the processor's digital input attenuator.

### Practice Safe Levels

When recording to digital from an analog source, if you have an external digital meter set to 3 samples, then trust its OVER indicator and reduce gain slightly if it illuminates during recording. If you've been watching your levels prior to generating the OVER, chances are it will be an inaudible 3 sample OVER. You won't lose any meaningful signal-to-noise ratio, and you'll end up with a cleaner recording, especially when sending it for mastering. At the mastering studio, a file which is too hot can cause a digital EQ or sample rate converter to overload. There are ways around that, but not without complicating the mastering engineer's life.

### Section II: How Loud is It?

Contrary to popular belief, the levels on a digital peak meter have (almost) nothing to do with loudness. For example, you're doing a direct to two-track recording (some engineers still work that way!) and you've found the perfect mix. Now, keep your hands off the faders, watch the levels to make sure they don't overload, and let the musicians make a perfect take. 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? If you answered "both takes are about the same loudness", you're probably right, **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, it will now sound 4 dB louder than take two, even though they both now measure the same on the peak meter.

**Do not confuse the peak-reading meters on digital recorders with VU meters.** Besides having a different scale, a VU meter has a much slower attack time than a digital peak meter. In PART II, we will discuss loudness in more detail, but let's summarize by saying that the VU meter responds more closely to the response of the ear. For loudness judgment, if all you have is a peak meter, use your ears. If you have a VU, use it as a guide, not an absolute, because the meter can be fooled (see [PART II](#)).

**Did you know that an analog and digital recording of the same source sound very different in terms of loudness?** Make an analog recording and a digital recording of the same music. Dub the analog recording to the digital domain, peaking at 0 dB. The analog dub will sound about 6 dB louder than the all-digital recording! That's a lot. This is because the typical peak-to-average ratio of an analog recording is about 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?). That's why pop producers who record digitally may have to compress or limit to compete with the loudness of their analog counterparts.

### **The Myth of "Normalization"**

Digital audio editing programs have a feature called "Normalization," a semi-automatic method of adjusting levels. The engineer selects all the segments (songs), and the computer grinds away, searching for the highest peak on the album. Then the computer adjusts the level of all the material until the highest peak reaches 0 dBFS. This is not a serious problem esthetically, 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. Since the ear responds to average levels, and normalization measures peak levels, the result can totally distort musical values. A compressed ballad will end up louder than a rock piece! In short, normalization should not be used to regulate song levels in an album. There's no substitute for the human ear.

### **Judging Loudness the Right Way**

Since the ear is the only judge of loudness, is there any objective way to get a handle on how loud your CD will sound? The first key is to use a single D/A converter to reproduce all your digital sources. That way you can compare your *CD in the making* against other CDs, in the digital domain. Judge plugins, CDs, workstations, and digital processors through this single converter. Another important tool is a calibrated monitor level control with 1 dB per step settings. In a consistent monitoring environment, you can become familiar with the level settings of the monitor control for many genres of music, and immediately know how far you are (in dB) from your nearest competitor, just by looking at the setting of the monitor knob. At *Digital Domain*, we log all monitor settings used on a given project, so we can return to the same setting for revisions. In **PART II**, we will discuss how to use our knowledge to make a better system in the 21st Century.

### **The Moving Average Goes Up and Up...**

Some of the latest digital processors permit making louder-sounding recordings than ever before. Today's mastering tools could make a nuclear bomb out of yesterday's firecrackers. But the sound becomes squashed, distorted and usually uninteresting. Visit my article on **Compression** for a more detailed description of the loudness race. While it seems the macho thing to do, you don't have to make your CD louder than the loudest current CD; try to make it sound *better*, which is much harder to do.

## **Section III: Calibrating Studio Levels**

That concludes our production discussion. This next section is intended primarily for the maintenance engineer. Let's talk about alignment of studio audio levels. Stick around for a fresh perspective on level setting in the hybrid analog-digital studio.

### Marking Tapes

dBm and dBv do not travel from house to house. These are measurements of voltages expressed in decibels. I once received a 1/4" tape in the mail marked "the level is +4 dBm." +4 dBm is a voltage (it's 1.23 volts, although the "m" stands for milliwatts). The 1/4" tape has no voltage on it, it doesn't have any idea whether it was made with a semi-pro level of 0 VU = -10 dBv or a professional level of +4. Voltages don't travel from house to house, only *nanowebers per meter* on analog tapes, and *dBFS* on digital tapes.

That doesn't diminish the importance of the analog reference level you use in-house. It's just irrelevant to the recipient of the tape. Just indicate the magnetic flux level which was used to coordinate with 0 VU. For example, *0 VU=400 nW/m at 1 KHz*. Most alignment tapes have tables of common flux levels, where you'll find that 400 nW/M is 6 dB over 200 nW/m. Engineers often abbreviate this on the tape box as *+6dB/200*.

### Deciding On an In-House Analog (voltage) Level

Just use the level provided by your console manufacturer, right? Well, maybe not. +4 dBv (reference .775 volts) may be a bad choice of reference level. Let's examine some factors you may not have considered when deciding on an in-house standard analog (voltage) level. When was the last time you checked the clipping point of your console and outboard gear? Before the advent of inexpensive 8-buss consoles, most professional consoles' clipping points were +24 dBv or higher. A frequent compromise in low-priced console design is to use internal circuits that clip around +20 dBv (7.75 volts). This can be a big impediment to clean audio, especially when cascading stages (how many of those amplifiers are between your source and your multitrack?). In my opinion, to avoid the "solid state edginess" that plagues a lot of modern equipment, the *minimum* clip level of every amplifier in your system should be 6 dB above the potential peak level of the music. The reason: Many opamps and other solid state circuits exhibit an extreme distortion increase long before they reach the actual clipping point. This means at least +30 dBv (24.5 volts RMS) if 0 VU is +4 dBv.

### How Much Headroom is Enough?

Have you noticed that solid-state equipment starts to sound pretty nasty when used near its clip point? All other things being equal, the amplifier with the higher clipping point sounds better, in my opinion. Perhaps that's why tube equipment (with their 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.

Traditionally, the difference between average level and clip point has been called the *headroom*, but in order to emphasize the need for even more than the traditional amount of headroom, I'll call the space between the peak level of the music and the amplifier clip point a *cushion*. In the days of analog tape, a 0 VU reference of +4 dBv with a clipping point of +20 dBv provided



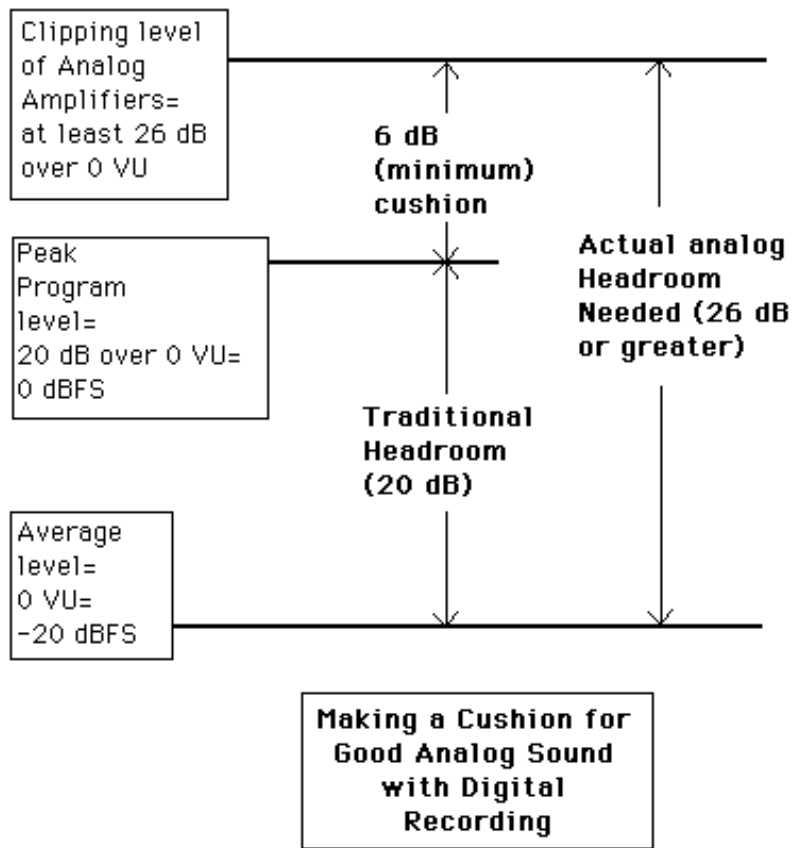
reasonable amplifier headroom, because musical peak-to-average ratios were reduced to the compression point of the tape, which maxes out at around 14 dB over 0 VU. Instead of clipping, analog tape's gradual saturation curve produces 3rd and 2nd harmonics, much gentler on the ear than the higher order distortions of solid state amplifier clipping.

But it's a different story today, where the peak-to-average ratio of raw, unprocessed digital audio tracks can be 20 dB. Adding 20 dB to a reference of +4 dBv results in +24 dBv, which is beyond the clipping point of many so-called *professional* pieces of gear, and doesn't leave any room for a *cushion*. If you adapt an active balanced output to an unbalanced input, the clipping point reduces by 6 dB, so the situation becomes proportionally worse (all those headroom specs have to be reduced by 6 dB if you unbalance an amplifier's output). Be particularly suspicious of consoles that are designed to work at either professional or semi-pro levels. To meet price goals, manufacturers often compromise on headroom in professional mode, making the so-called semi-pro mode sound cleaner! You'll be unpleasantly surprised to discover that many consoles clip at +20 dBv, meaning they should never be using a professional reference level of +4 dBv (headroom of only 16 dB and no cushion). Even if the console clips at +30 dBv (the minimum clipping point I recommend), that only leaves a 6 dB cushion when reproducing music with 20 dB peak-to-average ratio. That's why more and more high-end professional equipment have clipping points as high as +37 dBv (55 volts!). To obtain that specification, an amplifier must use very high output devices and high-voltage power supplies. Translation--better sound.

To summarize, make sure the clip point of all your analog amplifiers is at least 6 dB (preferably 12 or more dB) above the peak level of analog material that will run in the system. I call this additional headroom *the cushion*.

**How can you increase the cushion in your system, short of junking all your distribution amplifiers and consoles for new ones?** One way to solve the problem is to recalibrate all your VU meters. You will not lose significant signal-to-noise ratio if you set 0 VU= 0 dBv or even -4 dBv (not an international standard, but a decent compromise if you don't want to throw out your equipment, and you have the expertise to make this standard stick throughout your studio). Try it and let me know if things sound cleaner in your studio.

Once you've decided on a standard analog reference level, calibrate all your analog-driven VU meters to this level. Here's a diagram describing the concept of *cushion*.



### Dubbing and Copying - Translating between analog and digital points in the system

Let's discuss the interfacing of analog devices equipped with VU meters and digital devices equipped with digital (peak) meters. When you calibrate a system with sine wave tone, what translation level should you use? There are several de facto standards. Common choices have been -20 dBFS, -18 dBFS, and -14 dBFS translating to 0 VU. I'd like to see accurate calibration marks in digital recorders and DAWs at -12, -14, -18, and -20 dB, which covers most bases. Most of the external digital meters provide means to accurately calibrate at any of these levels.

**How do you decide which standard to use? Is it possible to have only one standard? What are the compromises of each?**

To make an educated decision, ask yourself: What is my system philosophy?

- Am I interested in maintaining headroom and avoiding peak clipping or do I want the highest possible signal-to-noise ratio at all times?
- Do I need to simplify dubbing practices or am I willing to require constant supervision during dubbing (operator checks levels before each dub, finds the peaks, and so on)?
- Am I adjusting levels or processing dynamics--mastering for loudness and consistency with

only secondary regard for the peak level?

Consider your typical musical sources. Are your sources totally digital (DDD)? Did they pass through extreme processing (compression) or through analog tape stages? Pure, unprocessed digital sources, particularly individual tracks on a multitrack, will have peak levels 18 to 20 dB above 0 VU. Whereas processed mixdowns will have peak-to-average ratios of up to 18 dB (rarely up to 20). Analog tapes will have peak levels up to 14 dB, almost never greater. And that's how the three most common choices of translation numbers (-18, -20, and -14) were derived.

### **Broadcast Studios**

In Broadcast, *Practicality* is our object, simplifying day-to-day operation, especially if your consoles are equipped with VU meters and your recorders are digital. In broadcast studios, it is desirable to use fixed, calibrated input and output gains on all equipment. My personal recommendation for the vast majority of studios is to standardize on reference levels of -20 dBFS ~0 VU, particularly when mixing to 2-track digital from live sources or tracking live to multitrack digital. If you're watching the console's VU meters, you will probably never clip a digital tape if you use -20 dBFS as a reference.

For a busy recording studio that does most of its mixing, recording and dubbing to harddisc, standardizing on -20 dBFS will simplify the process. Recording studios who decide on -18 dBFS ~0 VU will run into occasional digital clipping. That's why I'm against -18 dBFS as a standard for recording studios using VU meters for recording.

If you standardize on a -20 dBFS reference, the more compressed your musical material, the more signal-to-noise ratio you seem to be throwing away, but this is not true. If your source is analog tape, you might throw away 6 or more dB of signal, but this is less important than maintaining the convenience of never having to adjust dubbing levels on equipment. Furthermore, the ear judges noise level by average levels, and if the crest factor of your material is 6 dB less, it will seem just as loud as the uncompressed material peaking to 0 dBFS, you will not have to turn up your monitor, and you will not hear additional noise. Remember: analog tapes typically sound 6 dB louder than digital tapes, if peaked to the same peak level.

A -20 reference is only a potential problem when dubbing from digital source to analog tape. In many cases, you can accept the innocuous 6 dB compression. We've been enjoying that for years when we mixed from live material on VU-equipped console direct to analog tape. When making dubs to analog for archival purposes, choose a tape with more headroom, or use a custom reference point (-14 to -18 dBFS), as the goal is to preserve transients for the enjoyment of future listeners. A calibrated peak level meter on the analog machine will tell you what it's doing more than a VU meter. For archival purposes, I prefer to use the headroom of the new high-output tapes for transient clarity, rather than to jack up the flux level for a better signal-to-hiss ratio.

If working in a broadcast facility which seems no live (uncompressed) material, then for the broadcast dubbing room, -14 is a good number (dubbing between analog and digital tapes). -18 is



a safe all-around reference for all the other A/D/A converters in the broadcast complex, since most of the material will have 18 dB or lower peak-to average ratio, and occasional clipping maybe tolerated.

### Mastering Studios

Mastering studios are working more frequently in 20-bit or 24-bit. In [Part II](#), I suggest the 21st Century approach to mastering.

### Analog PPMs

Analog PPMs have a slower attack time than digital PPMs. When working with a digital recorder, a live source, and desk equipped with analog PPM, I suggest a 5 dB "lead." In other words, align the highest peak level on the analog PPM to -5 dBFS with sine wave tone.

---

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

[< Prev](#)      [Next >](#)

[\[ Back \]](#)

#### Popular

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)
- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us


[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## Level Practices (Part 2) (Includes the K-System)



### Part II: How To Make Better Recordings in the 21st Century - An Integrated Approach to Metering, Monitoring, and Leveling Practices.

(includes a description of the K-System, an integrated system of metering and monitoring)

*Updated from the article published in the September 2000 issue of the AES Journal by Bob Katz*

#### A: Two-Channel

For the last 30 years or so, film mix engineers have enjoyed the liberty and privilege of a controlled monitoring environment with a fixed (calibrated) monitor gain. The result has been a legacy of feature films, many with exciting dynamic range, consistent and natural-sounding dialogue, music and effects levels. In contrast, the broadcast and music recording disciplines have entered a runaway loudness race leading to chaos at the end of the 20th century. I propose an integrated **system** of metering and monitoring that will encourage more consistent leveling practices among the three disciplines. This system handles the issue of differing dynamic range requirements far more elegantly and ergonomically than in the past. We're on the threshold of the introduction of a new, high-resolution consumer audio format and we have a unique opportunity to implement a 21st Century approach to leveling, that integrates with the concept of **Metadata**. Let's try to make this a **worldwide standard** to leave a legacy of better recordings in the 21st Century.

#### History of the VU meter

On May 1, 1999, the VU meter celebrated its 60th birthday. 60 years old, but still widely misunderstood and misused. The VU meter has a carefully-specified time-dependent response to

OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

Password

Remember me

program material which this paper refers to as "Average," or "averaging", but means the particular VU meter response. This instrument was intended to help program producers create consistent loudness amongst program elements, but was not a suitable measure of when the recording medium was being exceeded, or overloaded. Therefore the meter's designers assumed that the recording medium would have at least 10 dB Headroom over 0 VU, like the analog media then in use.

### Summary of VU Inconsistencies and Errors

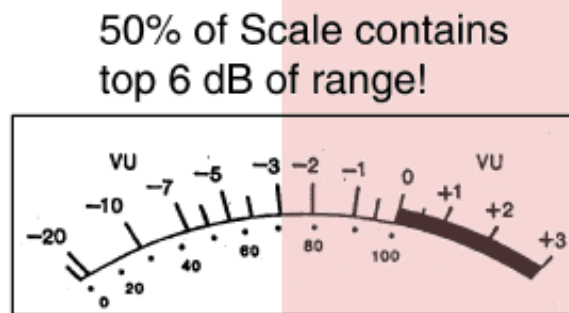
In General, the meter's ballistics, scale, and frequency response all contribute to an inaccurate indicator. The meter approximates momentary loudness changes in program material, but reports that moment-to-moment level differences are greater than the ear actually perceives.

#### Ballistics

The meter's ballistics were designed to "look good" with spoken word. Its 300 ms integration time gives it a syllabic response, which looks very "comfortable" with speech, but doesn't make it accurate. One time constant cannot sum up the complex multiple time constants required to model the loudness perception of the human listener. Skilled users soon learned that an occasional short "burst" from 0 to +3 VU would probably not cause distortion, and usually was meaningless as far as a loudness change.

#### Scale

In 1939, logarithmic amplifiers were large and cumbersome to construct, and it was desirable to use a simple passive circuit. The result is a meter where every decibel of change is not given equal merit. The top 50% of the physical scale is devoted to only the top 6 dB of dynamic range, and the meter's useable dynamic range is only about 13 dB. Not realizing this fundamental fact, inexperienced and experienced operators alike tend to push audio levels and/or compress them to stay within this visible range. With uncompressed material, the needle fluctuates far greater than the perceived loudness change and it is difficult to distinguish compressed from uncompressed material by the meter. Soft material may hardly move the meter, but be well within the acceptable limits for the medium and the intended listening environment.



#### Frequency response

The meter's relatively flat frequency response results in extreme meter deflections that are far greater than the perceived loudness change, since the ear's response is non-linear with respect to frequency. For instance, when mastering reggae music, which has a very heavy bass content, the VU meter may bounce several dB in response to the bass rhythm, but perceived loudness change

is probably less than a dB.

### Lack of conformance to standards

There are large numbers of improperly-terminated mechanical VU meters and inexpensively-constructed indicators which are labelled "VU" in current use. These disparate meters contribute to disagreements among program producers reading different instruments. A true VU meter is a rather expensive device. It's not a VU meter unless it meets the standard.

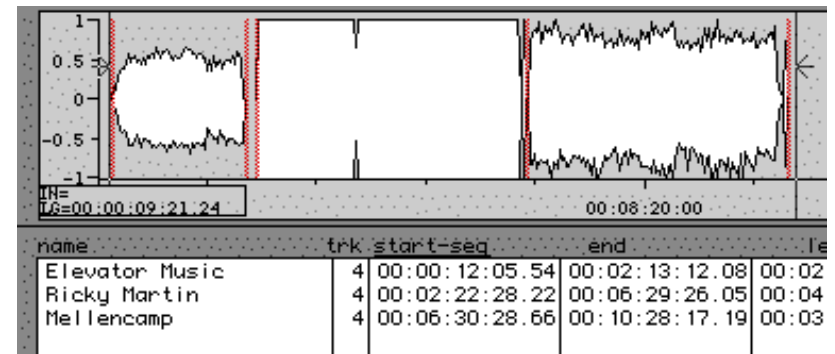
Over the past 60 years, psychoacousticians have learned how to measure perceived loudness much better than a VU. Despite all these facts, *the VU meter is a very primitive loudness meter*. In addition, current digital technology permits us to easily correct the non-linear scale, its dynamic range, ballistics, and frequency response.

### II. Current-day levelling problems

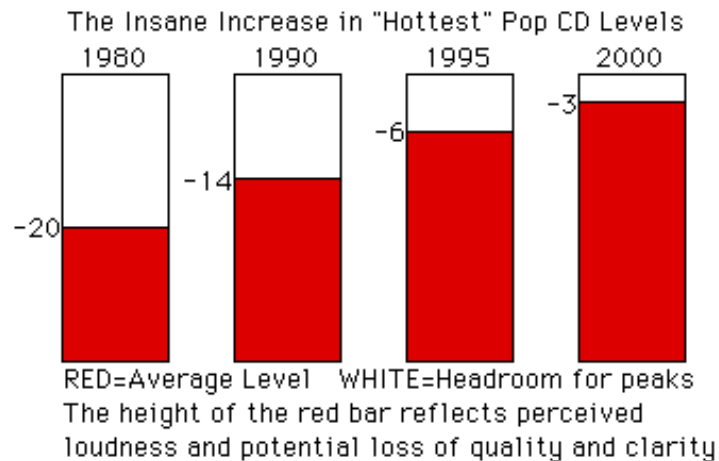
In the music and broadcast industries, chaos currently prevails. Here is a waveform taken from a digital audio workstation, showing three different styles of music recording. The time scale is about 10 minutes total, and the vertical scale is linear, +/- 1 at full digital level, 0.5 amplitude is 6 dB below full scale. The "density" of the waveform gives a rough

approximation of the music's dynamic range and Crest Factor (headroom for peaks above the average level). On the left side is a piece of heavily compressed pseudo "elevator music" I constructed for a demonstration at the 107th AES Convention. In the middle is a four-minute popular compact disc single produced in 1999, with sales in the millions. On the right is a four-minute popular rock and roll recording made in 1990 that's quite dynamic-sounding for rock and roll of that period. The perceived loudness difference between the 1990 and 1999 CDs is greater than 6 dB, though both peak to full scale. Auditioning the 1999 CD, one mastering engineer remarked "this CD is a lightbulb! The music starts, all the meterlights come on, and it stays there the whole time." To say nothing about the distortion. Are we really in the business of making square waves?

The average level of popular music compact discs continues to rise. Popular CDs with this problem are becoming increasingly prevalent, coexisting with discs that have beautiful dynamic range and impact, but whose loudness (and distortion level) is far lower. There are many technical, sociological and economic reasons for this chaos that are beyond the scope of this paper. Let's concentrate on what we can do as an engineering body to help reduce this chaos, which is a disservice to the consumer. It's also an obstacle to creating quality program material in the 21st century. What good is a 24-bit/96 kHz digital audio system if the programs we create only have 1



bit dynamic range?



Is this what will happen to the next generation carrier? (e.g. DVD-A, SACD). It will, if we don't take steps to stop it. Unlike with the LP, there is no PHYSICAL limit to the average level we can place on a digital medium. Note that there is a point of diminishing returns above about -14 dBFS. Dynamic inversion begins to occur and the program material usually stops sounding louder because it loses clarity and transient response.

### III. The Magic of "83" with Film Mixes

In the music world, everyone currently determines their own average record level, and adjusts their monitor accordingly. With no standard, subjective loudness varies from CD to CD in popular music as much as 10-12 dB, which is unacceptable by any professional standard. But in the film world, films are consistent from one to another, because the monitoring gain has been standardized. In 1983, as workshops chairman of the AES Convention, I invited Tomlinson Holman of Lucasfilm to demonstrate the sound techniques used in creating the Star Wars films. Dolby systems engineers labored for two days to calibrate the reproduction system in New York's flagship Ziegfeld theatre. Over 1000 convention attendees filled the theatre center section. At the end of the demonstration, Tom asked for a show of hands. "How many of you thought the sound was too loud?" About four hands were raised. "How many thought it was too soft?" No hands. "How many thought it was just right?" At least 996 audio engineers raised their hands.

This is an incredible testament to the effectiveness of the **83 dB SPL** reference standard proposed by Dolby's Ioan Allen in the mid-70's, originally calibrated to a level of 0 VU for use with analog magnetic film. The choice of 83 dB SPL has stood the test of time, as it permits wide dynamic range recordings with little or no perceived system noise when recording to magnetic film or 20-bit digital. Dialogue, music and effects fall into a natural perspective with an excellent signal-to-noise ratio and headroom. A good film mix engineer can work without a meter and do it all by the monitor, using the meter simply as a guide. In fact, working with a fixed monitor gain is *liberating*,

*not limiting*. When digital technology reached the large theatre, the SMPTE attached the SPL calibration to a point below full scale digital. When we converted to digital technology, the VU meter was rapidly replaced by the peak program meter.

When AC-3 and DTS became available for home theatre, many authorities recommended lowering the monitor gain by 6 dB because a typical home listening room does not accommodate high SPLs and wide dynamic range. If a DVD contains the wide range theatre mix, many home listeners complain that "this DVD is too loud", or "I lose the dialogue when I turn the volume down so that the effects don't blast." With reduced monitor gain, the soft passages become too soft. For such listeners, the dynamic range may have to be reduced by 6 dB (6 dB upward Compression, or dynamic range reduction) in order to use less monitor gain.

**Metadata** are coded data which contain information about signal dynamics and intended loudness; this will resolve the conflict between listeners who want the full theatrical experience and those who need to listen softly. But without metadata there are only two solutions: a) to compromise the audio soundtrack by compressing it, or better, b) use an optional compressor for the home system. With the later approach the source audio is uncompromised.

#### **IV. The Magic of "-6 dB" Monitor Gain for the Home**

In the 21st century, home theatre, music, and computers are becoming united. Many, if not most, consumers will eventually be auditioning music discs on the same system that plays broadcast television, home theatre (DVDs), and possibly even web-audio, e.g. MP3. Music-only discs are often used as casual or background music, but I am specifically referring to foreground music that the discerning consumer or audiophile will play at normal or full "enjoyment" loudness.

With the integration of media into a single system, it is in the direct interest of music producers to think holistically and unite with video and film producers for a more consistent consumer audio presentation. Music producers experimenting with 5.1 surround must pay more than casual attention to monitor level calibration. They have already discovered the annoyance that a typical pop CD will **blast the sound system** when inserted into a DVD player after a movie has been played. Recently a DVD and soundtrack CD were produced of the classic rock music movie **Yellow Submarine**. Reviewers complained that the CD is much louder and less dynamic than the DVD. Audio CDs should not be degraded for the sake of a "loudness competition". CDs can and should be produced to the same audio quality standard as the DVD.

New program producers with little experience in audio production are coming into the audio field from the computer, software and computer games arena. We are entering an era where the learning curve is high, engineer's experience is low, and the monitors they use to make program judgments are less than ideal. It is our responsibility to educate engineers on how to make loudness judgments. A plethora of peak-only meters on every computer, DAT machine and digital console do not provide information on program loudness. Engineers must learn that the sole purpose of the peak meter is to protect the medium and that something more like average level



affects the program's loudness. Bear in mind that the bandwidth and frequency distribution of the signal also affect program loudness.

As a music mastering engineer, I have been studying the perceived loudness of music compact discs for over 15 years. Around 1993, I installed a 1 dB/per step monitor control for repeatability. In an effort to achieve greater consistency from disc to disc, I made it a point to try to set the monitor gain first, and then master the disc to work well at that monitor gain.

In 1996, we measured that monitor gain, and found it to be 6 dB less than the film-standard for most of the pop music we were mastering. To calibrate a monitor to the film-standard, play a standardized pink noise calibration signal whose amplitude is -20 dB FS RMS, on one channel (loudspeaker) at a time. Adjust the monitor gain to yield 83 dB SPL using a meter with C-weighted, slow response. Call this gain 0 dB, the reference, and you will find the pop-music "standard" monitor gain at 6 dB below this reference.

By now, we've mastered hundreds of pop CDs working at monitor gain 6 dB below the reference, with very satisfied clients. **However, if monitor gain is further reduced, average recorded level tends to go up because the mastering engineer seeks the same loudness to the ears. Since the average program level is now closer to the maximum permissible peak level, more compression/limiting must be used to keep the system from overloading.** Increased compression/limiting is potentially damaging to the program material, resulting in a distorted, crowded, unnatural sound. Clients must be informed that they can't get something for nothing; a hotter record means lower sound quality.

### Mastering and the Loudness Race

By 1997, some music clients were complaining that their reference CDs were "not hot enough", a tragic testimony on the loudness race which is slowly destroying the industry. Each client wants his CD to be as loud as or louder than the previous "winner", but every winner is really a loser. Fueling that race are powerful digital compressors and limiters which enable mastering engineers to produce CDs whose average level is almost the same as the peak level! There is no precedent for that in over 100 years of recording. We end up mastering to the lowest common denominator, and fight desperately to avoid that situation, wasting a lot of time showing clients that the sound quality suffers as the average level goes up. The psychoacoustic problem is that when two **identical** programs are presented at slightly differing loudness, the louder of the two often appears "better" in short term listening. This explains why CD loudness levels have been creeping up until sound quality is so bad that everyone can perceive it. Remember that the loudness "race" has always been an artificial one, since the consumer adjusts their volume control according to each record anyway.

In addition, it should be more widely known **that hyper-compressed recordings do not play well on the radio.** They sound softer and seriously distorted, pointing out that the loudness race has no winners, even in radio airplay. The best way to make a "radio-ready" recording is not to

squash it, but rather produce it with the typical peak to average ratios that have worked for about a hundred years.

As the years went on, trying to "hold the fort", I gradually raised the average level of mastered CDs only when requested, which forced the monitor gain to be reduced from 1 to several dB. For every decibel of increased average level, considerably more damage is done to the sound. We often note severe processor distortion when the monitor gain falls below -6 dB. Consumers find their volume controls at the bottom of their travel, where a small control movement produces awkward level changes.

#### V. The relationship between SPL and 0 VU

In 1994, I installed a pair of Dorrough meters, in order to view the average and peak level simultaneously on the same scale. These meters use a scale with 0 "average" (a quasi-VU characteristic I'll call "AVG") placed at 14 dB below full digital scale, and full scale marked as +14 dB. Music mastering engineers often use this scale, since a typical stereo 1/2" 30 IPS analog tape has approximately 14 dB headroom above 0 VU.

The next step is to examine a simple relationship between the 0 AVG level and the sound pressure level. For typical pop productions, our monitor gain has been adjusted to -6 dB (below the standard reference, which yields 77dB SPL with -20 dBFS pink noise).



Since -20 dBFS reads -6 AVG, then 6 dB higher, or 0 AVG must be 83 dB SPL. In other words, we're really running average SPLs similar to the original theatre standard. The only difference is that headroom is 14 dB above 83 instead of 20. Running a sound pressure level meter during the mastering session confirms that the ear likes 0 AVG to end up circa 83 dB (~86 dB with both loudspeakers operating) on forte passages, even in this compressed structure. If the monitor gain is further reduced by 2 dB the mastering engineer judges the loudness to be lower, and thus raises average recorded level--and the AVG meter goes up by 2 dB. It's a linear relationship. **This leads us to the logical conclusion that we can produce programs with different amounts of dynamic range (and headroom) by designing a loudness meter with a sliding scale, where the moveable 0 point is always tied to the same calibrated monitor SPL. Regardless of the scale, production personnel would tend to place music near the 0**

point on forte passages.

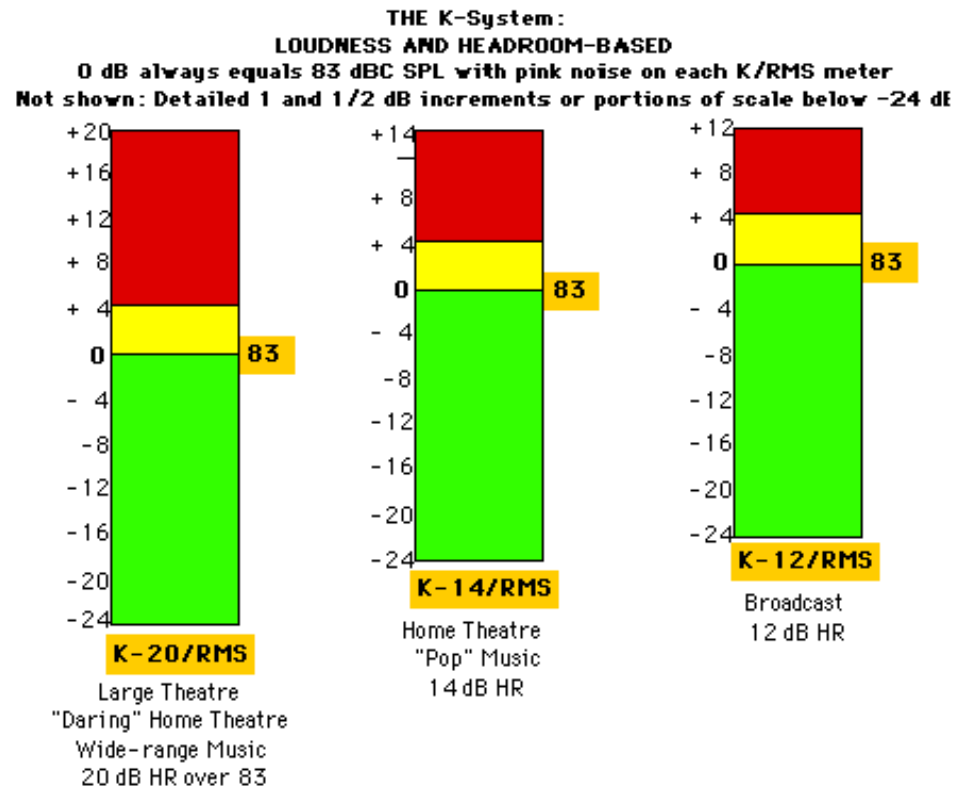
## VI. The K-System Proposal

The proposed **K-System** is a metering and monitoring standard that integrates the best concepts of the past with current psychoacoustic knowledge in order to avoid the chaos of the last 20 years.

In the 20th Century we concentrated on the *medium*. In the 21st Century, we should concentrate on the *message*. We should avoid meters which have 0 dB at the top--this discourages operators from understanding where the message really is. Instead, we move to a metering system where 0 dB is a **reference loudness**, which also determines the monitor gain. In use, programs which exceed 0 dB give some indication of the amount of processing (compression) which must have been used. There are three different K-System meter scales, with 0 dB at either 20, 14, or 12 dB below full scale, for typical headroom and SNR requirements. The dual-characteristic meter has a bar representing the average level and a moving line or dot above the bar representing the most recent highest instantaneous (1 sample) peak level.

Several accepted methods of measuring loudness exist, of varying accuracy (e.g., ISO 532, LEQ, Fletcher-Harvey-Munson, Zwicker and others, some unpublished). The extendable K-system accepts all these and future methods, plus providing a "flat" version with RMS characteristic. Users can calibrate their system's electrical levels with pink noise, without requiring an external meter. RMS also makes a reasonably-effective program meter that many users will prefer to a VU meter.

The three K-System meter scales are named K-20, K-14, and K-12. I've also nicknamed them the *papa*, *mama*, and *baby meters*. The K-20 meter is intended for wide dynamic range material, e.g., large theatre mixes, "daring home theatre" mixes, audiophile music, classical (symphonic) music, "audiophile" pop music mixed in 5.1 surround, and so on. The K-14 meter is for the vast majority of moderately-compressed high-fidelity productions intended for home listening (e.g. some home theatre, pop, folk, and rock music). And the K-12 meter is for productions to be dedicated for broadcast.



Note that full scale digital is always at the top of each K-System meter. The 83 dB SPL point slides relative to the maximum peak level. Using the term K-(N) defines simultaneously the meter's 0 dB point and the monitoring gain.

The peak and average scales are calibrated as per AES-17, so that peak and average sections are referenced to the same decibel value with a sine wave signal. In other words, +20 dB RMS with sine wave reads the same as +20 dB peak, and this parity will be true only with a sine wave. Analog voltage level is not specified in the K-system, only SPL and digital values. There is no conflict with -18 dBFS analog reference points commonly used in Europe.

### VII. Production Techniques with the K-System

To use the system, first choose one of the three meters based on the intended application. Wide dynamic range material probably requires K-20 and medium range material K-14. Then, calibrate the monitor gain where 0dB on the meter yields 83 dB SPL (per channel, C-Weighted, slow speed). 0dB always represents the same calibrated SPL on all three scales, unifying production practices worldwide. **The K-system is not just a meter scale, it is an integrated system tied to monitoring gain.**

A manual for a certain digital limiter reads: "For best results, start out with a threshold of -6 dB

FS". This is like saying "always put a teaspoon of salt and pepper on your food before tasting it." This kind of bad advice does not encourage proper production practice. A gain reduction meter is not an indication of loudness. Proper metering and monitoring practice is the only solution.

If console and workstation designers standardize on the K-System it will make it easier for engineers to move programs from studio to studio. Sound quality will improve by uniting the steps of pre-production (recording and mixing), post-production (mastering) and metadata (authoring) with a common "level" language. By anchoring operations to a consistent monitor reference, operators will produce more consistent output, and everyone will recognize what the meter means.

**If making an audiophile recording, then use K-20, if making "typical" pop or rock music, or audio for video, then probably choose K-14. K-12 should be reserved strictly for audio to be dedicated to broadcast; broadcast recording engineers may certainly choose K-14 if they feel it fits their program material. Pop engineers are encouraged to use K-20 when the music has useful dynamic range.**

The two prime scales, K-20 and K-14, will create a cluster near two different monitor gain positions. People who listen to both classical and popular music are already used to moving their monitor gains about 6 dB (sometimes 8 to 12 dB with the hottest pop CDs). It will become a joy to find that only two monitor positions satisfy most production chores. With care, producers can reduce program differences even further by ignoring the meter for the most part, and working solely with the calibrated monitor.

**Using the Meter's Red Zone.** This 88-90 dB+ region is used in films for explosions and special effects. In music recording, naturally-recorded (uncompressed) large symphonic ensembles and big bands reach +3 to +4 dB on the average scale on the loudest (*fortissimo*) passages. Rock and electric pop music take advantage of this "loud zone", since climaxes, loud choruses and occasional peak moments sound incorrect if they only reach 0dB (*forte*) on any K-system meter. Composers have equated *fortissimo* to 88-90+ dB since the time of Beethoven. Use this range *occasionally*, otherwise it is musically incorrect (and ear-damaging). If engineers find themselves using the red zone all the time, then either the monitor gain is not properly calibrated, the music is extremely unusual (e.g. "heavy metal"), or the engineer needs more monitor gain to correlate with his or her personal sensitivities. Otherwise the recording will end up overcompressed, with squashed transients, and its loudness quotient out of line with K-System guidelines.

### **Equal Loudness Contours**

Mastering engineers are more inclined to work with a constant monitor gain. But many music mixing engineers work at a much higher SPL, and also vary their monitor gain to check the mix at different SPLs. I recommend that mix engineers calibrate your monitor attenuators so you can always return to the recommended standard for the majority of the mix. Otherwise it is likely the mix will not translate to other venues, since the equal-loudness contours indicate a program will be bass-shy when reproduced at a lower (normal) level.

## Tracking/Mixing/Mastering

The K-System will probably not be needed for multitracking--a simple peak meter is probably sufficient. For highest sound quality, use K-20 while mixing and save K-14 for the calibrated mastering suite. If mixing to analog tape, work at K-20, and realize that the peak levels off tape will not exceed about +14. K-20 doesn't prevent the mix engineer from using compressors during mixing, but the author hopes that engineers will return towards using compression as an esthetic device rather than a "loudness-maker."

Using K-20 during mix encourages a clean-sounding mix that's advantageous to the mastering engineer. At that point, the producer and mastering engineer should discuss whether the program should be converted to K-14, or remain at K-20. *The K-System can become the lingua franca of interchange within the industry, avoiding the current problem where different mix engineers work on parts of an album to different standards of loudness and compression.*

### When the K-System is not available

Current-day analog mixing consoles equipped with VUs are far less of a problem than digital models with only peak meters. Calibrate the mixdown A/D gain to -20 dBFS at 0 VU, and mix normally with the analog console and VUs. However, mixing consoles should be retro fitted with calibrated monitor attenuators so the mix engineer can repeatably return to the same monitor setting.

**Compression** is a powerful esthetic tool. But with higher monitor gain, less compression is needed to make material sound good or "punchy." For pop music, many K-14 presentations sound better than K-20, with skillfully-applied dynamics processing by a mastering engineer working in a calibrated room. But clearly, the higher the K-number, the easier it is to make it sound "open" and clean. Use monitor systems with good headroom so that monitor compression does not contaminate the judgment of program transients.

**Adapting large theatre material to home use** may require a change of monitor gain and meter scale. Producers may choose to compress the original 6-channel theatre master, or better, remix the entire program from the multi-track stems (submixes). With care, most of the virtues and impact of the original production can be maintained in the home. Even audiophiles will find a well-mastered K-14 program to be enjoyable and dynamic. It is desirable to try to fit this reduced-range mix on the same DVD as the wide-range theatre mix.

### Multichannel to Stereo Reductions

The current legacy of loud pop CDs creates a dilemma because DVD players can also play CDs. Producers should try to create the 5.1 mix of a project at K-20. If possible, the stereo version should also be mixed and mastered at K-20. While a K-20 CD will not be as loud as many current pop CDs, it may be more dynamic and enjoyable, and there will not be a serious loudness jump compared to K-20 DVDs in the same player. If the producer insists on a "louder" CD, try to make it no louder than K-14, in which case there will only be 6 dB loudness difference between the DVD



and the audio CD. Tell the producer that the vast majority of great-sounding pop CDs have been made at K-14 and the CD will be consistent with the lot, even if it isn't as hot as the current hypercompressed "fashion." It's the hypercompressed CD that's out of line, not the K-14.

### **Full scale peaks and SNR**

It is a common myth that audible signal-to-noise ratio will deteriorate if a recording does not reach full scale digital. On the contrary, the **actual loudness** of the program determines the program's perceived signal-to-noise ratio. The position of the listener's monitor level control determines the perceived loudness of the system noise. If two similar music programs reach 0 on the K-system's average meter, even if one peaks to full scale and the other does not, both programs will have similar perceived SNR. Especially with 20-24 bit converters, the mix does not have to reach full scale (peak). Use the averaging meter and your ears as you normally would, and with K-20, even if the peaks don't hit the top, the mixdown is still considered normal and ready for mastering, with no audible loss of SNR.

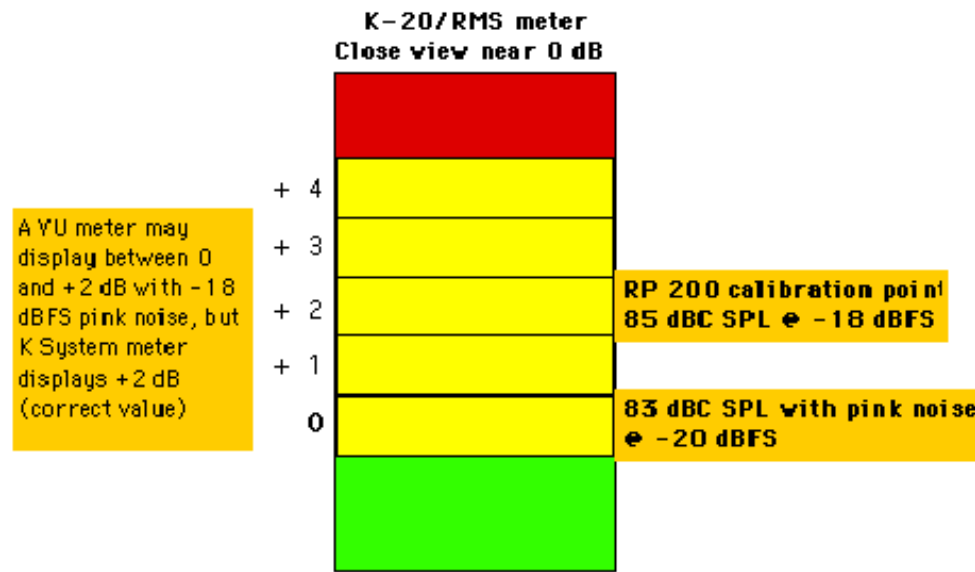
### **Multipurpose Control Rooms**

With the K-System, multipurpose production facilities will be able to work with wide-dynamic range productions (music, videos, films) one day, and mix pop music the next. A simultaneous meter scale and monitor gain change accomplishes the job. It seems intuitive to automatically change the meter scale with the monitor gain, but this makes it difficult to illustrate to engineers that K-14 really is louder than K-20.

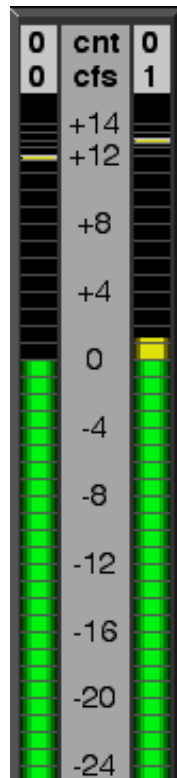
A simple 1 dB per step monitor attenuator can be constructed, and the operator must shift the meter scale manually.

Calibrate the gain of the reproduction system power amplifiers or preamplifiers with the K-20 meter, and monitor control at the "83" or 0 dB mark. Operators should be trained to change the monitor gain according to the K-System meter in use.

Here is the K-20/RMS meter in close detail, with the calibration points.



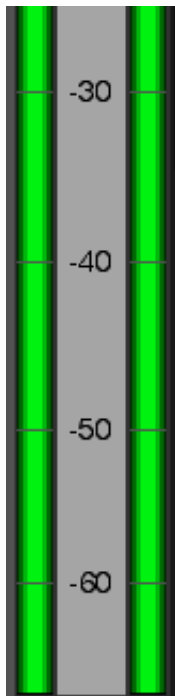
Individuals who decide to use a different monitor gain should log it on the tape (file) box, and try to use this point consistently. Even with slight deviations from the recommended K(N) practice, the music world will be far more consistent than the current chaos. Everyone should know the monitor gain they like to use.



At left is a picture of an actual K-14/RMS Meter in operation at the Digital Domain studio, as implemented by Metric Halo labs in the program *Spectrafoo* for the Macintosh. Spectrafoo versions 3f17 and above include full K-System support and a calibrated RMS pink noise generator. Other meters that conform exactly with K-System guidelines have been implemented by *Penguin* for PC, *RME* in their Digicheck software, and *Roger Nichols Digital* (formerly Elemental audio) Inspector XL. The *Dorrough* and *DK* meters nearly meet K-System guidelines but an external RMS meter must be used for pink noise calibration since they use a different type of averaging. In practice with program material, the difference between RMS and other averaging methods is insignificant, especially when you consider that neither method is close enough to a true loudness meter. As of this date, 12/05/07, we are still awaiting a company that will implement the K-System with a loudness characteristic, such as Zwicker.

#### Audio Cassette Duplication

Cassette duplication has been practiced more as an art than a science, but it should be possible to do better. The K-System may finally put us all on the same page (just in time for obsolescence of the cassette format). It's been difficult for mastering engineers to communicate with audio cassette duplicators, finding a reference level we all can understand. A knowledgeable duplicator once explained that the tape most



commonly used cannot tolerate average levels greater than +3 over 185 nW/m (especially at low frequencies) and high frequency peaks greater than about +5-6 are bound to be distorted and/or attenuated. Displaying crest factor makes it easy to identify potential problems; also an engineer can apply cassette high-frequency preemphasis to the meter. Armed with that information, an engineer can make a good cassette master by using a "predistortion" filter with gentle high-frequency compression and equalization. Meter with K-14 or K-20, and put test tone at the K-System reference 0 on the digital master. Peaks must not reach full scale or the cassette will distort. Apparent loudness will be less than the K-standard, but this is a special case.

### Classical music

It's hard to get out of the habit of peaking our recordings to the highest permissible level, even though 24-bit systems have a theoretically 48 dB better signal-to-dither-ratio than 16-bit. It is much better for the consumer to have a consistent monitor gain than to peak every recording to full scale digital. I believe that attentive listeners prefer auditioning at or near the natural sound pressure of the original classical ensemble ([see Footnote](#)). The dilemma is that string quartets and Renaissance music, among other forms, have low crest factors as well as low natural loudness.

Consequently, the string quartet will sound (unnaturally) much louder than the symphony if both are peaked to full scale digital.

I recommend that classical engineers mix by the calibrated monitor, and use the average section of the K-meter only as a guide. It's best to fix the monitor gain at 83 dB and always use the K-20 meter even if the peak level does not reach full scale. There will be less monitoring chaos and more satisfied listeners. However, some classical producers are concerned about loss of resolution in the 16-bit medium and may wish to peak all recordings to full scale. I hope you will reconsider this thought with 24 bit media or SACD.

### Narrow Dynamic Range Pop Music

We can avoid a new loudness race and consequent quality reduction if we unite behind the K-System before we start fresh with high-resolution audio media such as DVD-A and SACD. Similar to the above classical music example, pop music with a crest factor much less than 14 dB should not be mastered to peak to full scale, as it will sound too loud.

Recommended:

- 1: Author with metadata to benefit consumers using equipment that supports metadata
- 2: If possible, master such discs at K-14
- 3: Legacy music, remasters from often overcompressed CD material should be reexamined for its loudness character. If possible, reduce the gain during remastering so the average level falls within K-14 guidelines. Even better, remaster the music from unprocessed mixes to undo some of the unnecessary damage incurred during the years of chaos. Some mastering engineers already

have made archives without severe processing.

### VIII. An Extendable System

Since the K-System is extendable to future methods of measuring loudness, program producers should mark their tape boxes or digital files with an indication which K-meter and monitor calibration was used. For example, "K-14/RMS," or "K-20/Zwicker." I hope that these labels will someday become as common as listings of nanowebers per meter and test tones for analog tapes. If a non-standard monitor gain was used, note that fact on the tape box to aid in post-production authoring and insertion of metadata.

### IX. Metadata and the K-System

Dolby AC-3, MPEG2, AAC, and hopefully MLP will take advantage of metadata control words. Pre-production with the K-System will speed the authoring of metadata for broadcast and digital media. Music producers must familiarize themselves with how metadata affects the listening experience. First we'll summarize how the control word *Dialnorm* is used in digital television. Then we will examine how to take advantage of *Dialnorm* and *MixLevel* for music-only productions.

#### Dialnorm

*Dialogue normalization*, is used in digital television and radio as "ecumenical gain-riding". Program level is controlled at the decoder, producing a consistent average loudness from program to program; with the amount of attenuation individually calculated for each program. The receiver decodes the dialnorm control word and attenuates the level by the calculated amount, resulting in the "table radio in the kitchen" effect. In an unnatural manner, average levels of sports broadcasts, rock and roll, newscasts, commercials, quiet dramas, soap operas, and classical music all end up at the loudness of average spoken dialogue.

With *Dialnorm*, the average loudness of all material is reduced to a value of -31 dB FS (LEQ-A). Theatrical films with dialogue at around -27 dB FS will be reduced 4 dB. -31 corresponds not with musical *forte*, but rather *mezzo-piano*. For example, a piece of rock and roll, normally meant to be reproduced *forte*, may be reduced 10 or more dB, while a string quartet may only be reduced 4-5 dB at the decoder. The *dialnorm* value for a symphony should probably be determined during the second or third movement, or the results will be seriously skewed. We do want the forte passages to be louder than the spoken word! Rock and roll, with its more limited dynamic range, will be attenuated farther from "real life" than the symphony. However, unlike the analog approach, the listener can turn up his receiver gain and experience the original program loudness--without the noise modulation and squashing of current analog broadcast techniques. Or, the listener can choose to turn off *dialnorm* (on some receivers) and experience a large loudness variance from program to program.

Each program is transmitted with its full intended dynamic range, without any of the compression used in analog broadcasting--the listener will hear the full range of the studio mix. For example, in variety shows, the music group will sound pleasingly louder than the presenter. Crowd noises in

sports broadcasts will be excitingly loud, and the announcer's mike will no longer "step on" the effects, because the bus compressor will be banished from the broadcast chain.

### **Mixlev**

*Dialnorm* does not reproduce the dynamic range of real life from program to program. This is where the optional control word *mixlev* (mix level) enters the picture. The *dialnorm* control word is designed for casual listeners, and *mixlev* for audiophiles or producers. Very simply, *mixlev* sets the listener's monitor gain to reproduce the SPL used by the original music producer. Only certain critical listeners will be interested in *mixlev*. If the K-system was used to produce the program, then K-14 material will require a 6 dB reduction in monitor gain compared to K-20, and so on. *Mixlev* will permit this change to happen automatically and unattended. Attentive listeners using *mixlev* will no longer have to turn down monitor gains for string quartets, or up for the symphony or (some) rock and roll.

The use of *dialnorm* and *mixlev* can be extended to other encoded media, such as DVD-A. Proper application of *dialnorm* and *mixlev*, in conjunction with the K-System for pre-production practice--will result in a far more enjoyable and musical experience than we currently have at the end of the 20th century of audio.

### **X. In Conclusion**

Let's bring audio into the 21st century. The K-system is the first integrated approach to monitoring, levelling practices, metering and metadata.

### **B: Multichannel**

There's good news for audio quality: 5.1 surround sound. Current mixes of popular music that I have listened to in 5.1 sound open, clear, beautiful, yet also impacting. I've done meter measurements and listening to a few excellent 20 and 24 bit 5.1 mixes, and they all fall perfectly into the K-20 Standard. Monitor gain ran from 0 dB to -3 dB, mostly depending on taste, as it was perfectly comfortable to listen to all of these particular recordings at 0 dB (reference RP 200).

What became clear while watching the K-20 meter is that the best engineers are using the peak capability of the 5.1 system strictly for headroom. It is possible that I didn't see a single peak to full scale (+20 on the K-20 Meter) on any of these mixes. The averaging portion of the meter operated just as in my recommendations, with occasional peaks to +4 on some of the channels.

Monitor calibration made on an individual speaker basis worked extremely well, with the headroom in each individual channel tending to go up as the number of channels increases. This is simply not a problem with 24 bit (or even 20 bit) recording. System hiss is not evident at RP 200 monitor gains with long-wordlength recording, good D/A converters, modern preamps and power amplifiers.

Another question is: Should we have an overall meter calibrated to a total SPL? If so, what should that SPL be? My initial reactions are that an overall meter is not necessary, at least in mix situations where mix engineers use calibrated monitoring and monitors with good headroom.

Another positive thought. I've been giving 5.1 seminars sponsored by TC, Dynaudio, and DK Meters. To begin the show, I played two stereo masters that I had mastered, and demonstrated some very sophisticated techniques to bump them up (transparently) to 5.1. This is a growing field, and you'll see increasing techniques for doing this, especially when the record company wants a DVD or DVD-A remaster without (horrors) having to pay for a remix.

The good news is I found that the true 5.1 mixes by George Massenburg and others that I was demonstrating sounded so OPEN and clear and beautiful that even I was embarrassed to start from a 24-bit version of my own two masters. I had to remaster the two pieces with about 2 to 4 dB LESS LIMITING in order to make them COMPETE SONICALLY with the 5.1 stuff!!! "Louder is better" just doesn't work when you're in the presence of great masters.

That's right, I predict that the critical mastering engineers of the future will be so embarrassed by the sound quality of the good 5.1 stuff that they won't be able to get away with smashing 5.1 masters. And, hopefully, the two-track reductions that they also remaster (the CD versions) especially if there is a CD layer on the same disc, will be mastered to work at the same LOUDNESS.

In fact, if you tried to turn 5.1 Lyle Lovett, Michael Jackson, Aaron Neville, or Sting into a K-14, they just would sound horrid, on any reasonable 5.1 playback system!

The DK meters, set to K-20 demonstrated clearly that K-20 rules in 5.1. In fact, after a while I simply turned off the peak portion of the meter as it was distracting. So we could watch the VU-style levels and see the techniques used by each of the mix engineers. At K-20 and with 6 speakers running, you have so much headroom that it is hardly necessary to watch the peak meters at all. Furthermore, at 24 bits, there is absolutely no necessity to hit 0 dBFS ANYMORE AT ALL.

The proof is in the pudding, when you try your first 5.1 master you will see clearly what I mean. K-20-style metering and calibrated monitoring becomes a MUST in 5.1.

**If you are interested in discussing the ramifications of these topics, please contact the author, [Bob Katz](#).**

### Credits

Many thanks to: Ralph Kessler of Pinguin for reviewing the manuscript and suggesting valuable corrections and additions.

-----

### Appendix 1: Definition of Terms

**Average** - "Integrated" level of program, as distinguished from its momentary peak levels.

**Average level** - Area under the rough waveform curve, ignoring momentary peaks.

**Averaging method** - (such as *arithmetic mean*, or *root-mean-square*) must be specified in order to determine area under curve.



**Compression** - "dynamic range reduction". Not to be confused with the recent use of the word to describe digital audio *coding* systems such as **AC-3, MPEG, DTS and MLP**. To avoid ambiguity, refer to the latter as coding systems, or more exactly, data-rate-reduction systems.

**Crest Factor** - ratio between peak and average program levels, or ratio of level of instantaneous highest peak to average level of program. There is no standard for the averaging method to be used in determining crest factor. I've used a VU characteristic for purposes of illustration. Unprocessed music exhibits a high crest factor, and a low crest factor can only be obtained using dynamic-range compression.

**Headroom** - ratio between peak capability of medium and average level of program. There is no standard averaging method for determining headroom. I've used a VU characteristic for purposes of discussion.

**Metadata** - "data about data" Coding systems such as **AC-3, DTS, and MLP** can insert control words in the data stream which describe the data, the audio levels, and ways in which the audio can be manipulated. Metadata permits the insertion of an optional dynamic-range compressor located in the listener's decoder, bringing up soft passages to permit listening at reduced average loudness. The control word *dynrng* controls the parameters of this compressor in the AC-3 system and hopefully will also be used in **MLP**. The advantage of this approach is that the source audio remains uncompromised. Other important control words include *dialnorm* and *mixlev*.

**MLP** - (Meridian lossless packing). The lossless coding system specified for the DVD-Audio disc.

**VU meter** - According to **A New Standard Volume Indicator and Reference Level**, Proceedings of the I.R.E., January, 1940, the mechanical VU meter used a copper-oxide full-wave rectifier which, combined with electrical damping, had a defined averaging response according to the formula  $i = k * e$  to the  $p$  equivalent to the actual performance of the instrument for normal deflections. (In the equation  $i$  is the instantaneous current in the instrument coil and  $e$  is the instantaneous potential applied to the volume indicator)...a number of the new volume indicators were found to have exponents of about 1.2. Therefore, their characteristics are intermediate between linear ( $p = 1$ ) and square-law or root-mean-square ( $p=2$ ) characteristic."

## Appendix 2: SMPTE Practice

All quoted monitor SPL calibration figures in this paper are referenced to -20 dB FS. The "theatre standard", **Proposed SMPTE Recommended Practice: Relative and Absolute Sound Pressure Levels for Motion-Picture Multichannel Sound Systems**, SMPTE Document RP 200, defines the calibration method in detail. In the 1970's the value was quoted as "85 at 0 VU" but as the measurement methods became more sophisticated, this value proved to be in error. It has now become "85 at -18 dB FS" with 0 VU remaining at -20 dBFS (sine wave). The history of this metamorphosis is interesting. A VU meter was originally used to do the calibration, and with the advent of digital audio, the VU meter was calibrated with a sine wave to -20 dB FS. However, it was forgotten that a VU meter does not average by the RMS method, which results in an error between the RMS electrical value of the pink noise and the sine wave level. While 1 dB is the theoretical difference, the author has seen as much as a 2 dB discrepancy between certain VU

meters and the true RMS pink noise level.

The other problem is the measurement bandwidth, since a widerange voltmeter will show attenuation of the source pink noise signal on a long distance analog cable due to capacitive losses. The solution is to define a specific measurement bandwidth (20 kHz). By the time all these errors were tracked down, it was discovered that the historical calibration was in error by 2dB. Using pink noise at an RMS level of -20 dBFS RMS must correctly result in an SPL level of only 83 dB. In order to retain the magic "85" number, the SMPTE raised the specified level of the calibrating pink noise to -18dB FS RMS, but the result is the identical monitor gain. One channel is measured at a time, the SPL meter set to C weighting, slow. The K-System is consistent with RP 200 only at K-20. I feel it will be simpler in the long run to calibrate to 83 dB SPL at the K-System meter's 0 dB rather than confuse future users with a non-standard +2 dB calibration point. It is critical that the thousands of studios with legacy systems that incorporate VU meters should adjust the electrical relationship of the VU meter and digital level via a sine wave test tone, then **ignore the VU meter** and align the SPL with an RMS-calibrated digital pink noise source.

### **Improved measurement accuracy if narrow-band pink noise is used**

There are many sources of inaccuracy when determining monitor gain when using pink noise. Using wideband (20-20 kHz) pink noise and a simple RMS meter can result in low frequency errors due to standing waves in the room, high frequency errors due to off-axis response of the microphone, and variations in filter characteristics of inexpensive sound level meters. For the most accurate measurement, use narrow-band pink noise limited 500-2kHz, whose RMS level is -20 dBFS. This noise will read the same level on SPL meters with flat response, A weighting, or C weighting, eliminating several variables.

For even more accuracy, a spectrum analyzer can be used to make the critical 1/3 octave bands equal and reading ~68 dB SPL, yet totalling the specified 83 dB SPL.

### **Appendix 3: Detailed Specifications of the K-System Meters**

**General:** All meters have three switchable scales: K-20 with 20 dB headroom above 0 dB, K-14 with 14 dB, and K-12 with 12 dB. The **K/RMS meter version (flat response)** is the only required meter--to allow RMS noise measurements, system calibration, and program measurement with an averaging meter that closely resembles a "slow" VU meter. The other K-System versions measure loudness by various known psychoacoustic methods (e.g., LEQ and Zwicker).

**Scales and frequency response:** A tri-color scale has green below 0 dB, amber to +4 dB, and red above that to the top of scale. The peak section of the meters always has a flat frequency response, while the averaging section varies depending on version which is loaded. For example: Regardless of the sampling rate, meter version **K-20/RMS** is band-limited as per SMPTE RP 200, with a flat frequency response from 20-20 kHz +/- 0.1 dB, the average section uses an RMS detector, and 0 dB is 20 dB below full scale. To maintain pink noise calibration compatibility with SMPTE proposal RP 200, the meter's bandpass will be 22 kHz maximum regardless of sample rate. Other loudness-determining methods are optional. The suggested average section of Meter **K-20/**

**LEQA** has a non-flat (A-weighted) frequency response, and response time with an equal-weighted time average of 3 seconds. The average section of Meter **K-20/Zwicker** corresponds with Zwicker's recommendations for loudness measurement. Regardless of the frequency response or methodology of the loudness method, reference 0 dB of all meters is calibrated such that 20-20 kHz pink noise at 0 dB reads 83 dB SPL, C weighted, slow. Psychoacousticians designing loudness algorithms recognize that the two measurements, SPL and loudness are *not* interchangeable and take the appropriate steps to calibrate the K-system loudness meter 0 dB so that it equates with a standard SPL meter at that one critical point with the standard pink noise signal.

**Scale gradations:** The scale is linear-decibel from the top of scale to at least -24 dB, with marks at 1 dB increments except the top 2 decibels have additional marks at 1/2 dB intervals. Below -24 dB, the scale is non-linear to accommodate required marks at -30, -40, -50, -60. Optional additional marks through -70 and below. Both the peak and averaging sections are calibrated with sine wave to ride on the same numeric scale. **Optional (recommended):** A "10X" expanded scale mode, 0.1 dB per step, for calibration with test tone.

**Peak section of the meter:** The peak section is always a flat response, representing the true (1 sample) peak level, regardless of which averaging meter is used. An additional pointer above the moving peak represents the highest peak in the previous 10 seconds. A peak hold/release button on the meter changes this pointer to an infinite high peak hold until released. The meter has a fast rise time (aka *integration time*) of one digital sample, and a slow fall time, ~3 seconds to fall 26 dB. An adjustable and resettable OVER counter is highly recommended, counting the number of contiguous samples that reach full scale.

#### FOOTNOTE

The late Gabe Wiener produced a series of classical recordings noting in the liner notes the SPL of a short (test) passage. He encouraged listeners to adjust their monitor gains to reproduce the "natural" SPL which arrived at the recording microphone. The author used to second-guess Wiener by first adjusting monitor gain by ear, and then measuring the SPL with Wiener's test passage. Each time, the author's monitor was within 1 dB of Wiener's recommendation. Thus demonstrating that for classical music, the natural SPL is desirable for attentive, foreground listeners.

---

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

<                      Next  
Prev                      >

[ Back ]

**Popular**

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)

- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us

© 2007 Digital Domain - All Rights Reserved


[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## Depth and Dimension



**Master the 2-channel art first** Everyone is talking about multichannel sound. I have no doubt that well-engineered multi-channel recordings will produce a more natural soundfield than we've been able to achieve in our 2-channel recordings, but it amazes me how few engineers really know how to take advantage of good ol' fashioned 2-channel stereo. I've been making "naturalistic" 2-channel recordings for many years, and there are others working in the pop field who produce 2-channel (pop, jazz, even rock) recordings with beautiful depth and space. I'm rather disappointed in the sound of 2-channel recordings made by simple "pan-potted mono", the typical sound of a rock mix. But it doesn't have to be, if you study the works of the masters.

I wonder if the recording engineers who are disappointed in 2-channel recording may simply be using the wrong techniques. Pan-potted mono techniques, coupled by artificial reverberation--tend to produce a vague, undefined image, and I can understand why many engineers complain about how difficult it is to get definition working in only two channels. They say that when they move to multichannel mixing (e.g., 5.1) that they have a much easier time of it. Granted, though I suggest that first they study how to make a good 2-channel mixdown with depth, space, clarity, and definition. It's possible if you know the tricks. Most of those tricks involve the use of the Haas effect, phase delays, more natural reverbs and unmasking techniques. If engineers don't study the art of creating good 2-channel recordings, when we move to 5.1, ultimately we will end up with more humdrum mixes, more "pan-potted mono", only with more speakers. This article describes techniques that will help you with 2-channel and multichannel recordings. Furthermore, well-engineered 2-channel recordings have encoded ambience information which can be extracted to multichannel, and it pays to learn about these techniques.

### The Perception of Depth

OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

Password

Remember me

At first thought, it may seem that depth in a recording is achieved by increasing the ratio of reverberant to direct sound. But it is a much more involved process. Our binaural hearing apparatus is largely responsible for the perception of depth. But recording engineers were concerned with achieving depth even in the days of monophonic sound. In the monophonic days, many halls for orchestral recording were deader than those of today. Why do monophonic recording and dead rooms seem to go well together? The answer is involved in two principles that work hand in hand: 1) The masking principle and 2) The Haas effect.

### **The Masking Principle**

The masking principle says that a louder sound will tend to cover (mask) a softer sound, especially if the two sounds lie in the same frequency range. If these two sounds happen to be the direct sound from a musical instrument and the reverberation from that instrument, then the initial reverberation can appear to be covered by the direct sound. When the direct sound ceases, the reverberant hangover is finally perceived.

In concert halls, our two ears sense reverberation as coming diffusely from all around us, and the direct sound as having a distinct single location. Thus, in halls, the masking effect is somewhat reduced by the ears' ability to sense direction.

In monophonic recording, the reverberation is reproduced from the same source speaker as the direct sound, and so we may perceive the room as deader than it really is, because of directional masking. Furthermore, if we choose a recording hall that is very live, then the reverberation will tend to intrude on our perception of the direct sound, since both will be reproduced from the same location--the single speaker. So there is a limit to how much reverberation can be used in mono.

This is one explanation for the incompatibility of many stereophonic recordings with monophonic reproduction. The larger amount of reverberation tolerable in stereo becomes less acceptable in mono due to directional masking. As we extend our recording techniques to 2-channel (and eventually multichannel) we can overcome directional masking by spreading reverberation spatially away from the direct source, achieving both a clear (intelligible) and warm recording at the same time.

### **The Haas Effect**

The Haas effect can be used to overcome directional masking. Haas says that, in general, echoes occurring within approximately 40ms of the direct sound become fused with the direct sound. We say that the echo becomes "one" with the direct sound, and only a loudness enhancement occurs. A very important corollary to the Haas effect says that fusion (and loudness enhancement) will occur even if the closely-timed echo comes from a different direction than the original source. However, the brain will continue to recognize (binaurally) the location of the original sound as the proper direction of the source. The Haas effect allows nearby echoes (up to approximately 40ms delay, typically 30ms) to enhance an original sound without confusing its directionality. We can take advantage of the Haas effect to naturally and effectively convert an existing 2-channel recording to a 4-channel or surround medium. When remixing, place a discrete delay in the



surround speakers to enhance and extract the original ambience from a previously recorded source! No artificial reverberator is needed if there is sufficient reverberation in the original source. Here's how it works:

Because of the Haas effect, the ear fuses the delayed with the original sound, and still perceives the direct sound as coming from the front speakers. But this does not apply to ambience-- ambience will be spread, diffused between the location of the original sound and the delay (in the surround speakers). Thus, the Haas effect only works for correlated material; uncorrelated material (such as natural reverberation) is extracted, enhanced, and spread directionally. Dolby laboratories calls this effect "the magic surround," for they discovered that natural reverberation was extracted to the rear speakers when a delay was applied to them. Dolby also uses an L minus R matrix to further enhance the separation. The wider the bandwidth of the surround system and the more diffuse its character, the more effective the psychoacoustic extraction of ambience to the surround speakers.

There's more to Haas than this simple explanation. To become proficient in using Haas in mixing, study the original papers on the various fusion effects at different delay and amplitude ratios.

### **Haas' Relationship to Natural Environments**

We may say that the shorter echoes which occur in a natural environment (from nearby wall and floor) are correlated with the original sound, as they have a direct relationship. The longer reverberation is uncorrelated; it is what we call the ambience of a room. Most dead recording studios have little or no ambient field, and the deadest studios have only a few perceptible early reflections to support and enhance the original sound.

In a good stereo recording, the early correlated room reflections are captured with their correct placement; they support the original sound, help us locate the sound source as to distance and do not interfere with left-right orientation. The later uncorrelated reflections, which we call reverberation, naturally contribute to the perception of distance, but because they are uncorrelated with the original source the reverberation does not help us locate the original source in space. This fact explains why the multitrack mixing engineer discovers that adding artificial reverberation to a dry, single-miked instrument may deteriorate the sense of location of that instrument. If the recording engineer uses stereophonic miking techniques and a live room instead, capturing early reflections on two tracks of the multitrack, the remix engineer will need less artificial reverberation and what little he adds can be done convincingly.

### **Using Frequency Response to Simulate Depth**

Another contributor to the sense of distance in a natural acoustic environment is the absorption qualities of air. As the distance from a sound source increases, the apparent high frequency response is reduced. This provides another tool which the recording engineer can use to simulate distance, as our ears have been trained to associate distance with high-frequency rolloff. An interesting experiment is to alter a treble control while playing back a good orchestral recording. Notice how the apparent front-to-back depth of the orchestra changes considerably as you

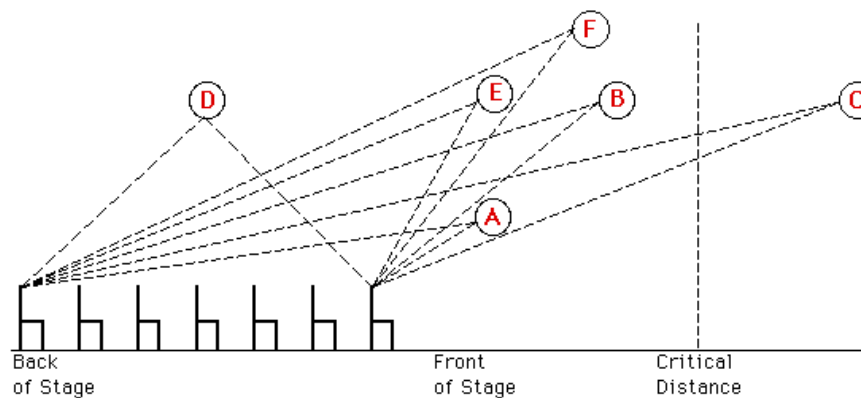
manipulate the high frequencies.

## Recording Techniques to Achieve Front-to-Back Depth Minimalist Techniques

### Balancing the Orchestra

A musical group is shown in a hall cross section. Various microphone positions are indicated by letters A-F.

Microphones **A** are located very close to the front of the orchestra. As a result, the ratio of **A**'s distance from the back compared to the front is very large. Consequently, the front of the orchestra will be much louder in comparison to the rear. Front-to-back balance will be exaggerated. However, there is much to be said in favor of mike position **A**, since the conductor usually stands there, and he purposely places the softer instruments (strings) in the front, and the louder (brass and percussion) in the back, somewhat compensating for the level discrepancy due to location. Also, the radiation characteristics of the horns of trumpets and trombones help them to overcome distance. These instruments frequently sound closer than other instruments located at the same physical distance because the focus of the horn increases direct to reflected ratio. Notice that orchestral brass often seem much closer than the percussion, though they are placed at similar distances. You should take these factors into account when arranging an ensemble for recording. Clearly, we also perceive depth by the larger ratio of reflected to direct sound for the back instruments.



The further back we move in the hall, the smaller the ratio of back-to-front distance, and the front instruments have less advantage over the rear. At position **B**, the brass and percussion are only two times the distance from the mikes as the strings. This (according to theory) makes the back of the orchestra 6 dB down compared to the front, but much less than 6 dB in a reverberant hall, because level

changes less with distance.

For example, in position **C**, the microphones are beyond the critical distance--the point where direct and reverberant sound are equal. If the front of the orchestra seems too loud at **B**, position **C** will not solve the problem; it will have similar front-back balance but be more buried in reverberation.

### Using Microphone Height To Control Depth And Reverberation

Changing the microphone's height allows us to alter the front-to-back perspective independently of reverberation. Position **D** has no front-to-back depth, since the mikes are directly over the center of the orchestra. Position **E** is the same distance from the orchestra as **A**, but being much

higher, the relative back-to-front ratio is much less. At **E** we may find the ideal depth perspective and a good level balance between the front and rear instruments. If even less front-to-back depth is desired, then **F** may be the solution, although with more overall reverberation and at a greater distance. Or we can try a position higher than **E**, with less reverb than **F**.

### **Directivity of Musical Instruments**

Frequently, the higher up we move, the more high frequencies we perceive, especially from the strings. This is because the high frequencies of many instruments (particularly violins and violas) radiate upward rather than forward. The high frequency factor adds more complexity to the problem, since it has been noted that treble response affects the apparent distance of a source. Note that when the mike moves past the critical distance in the hall, we may not hear significant changes in high frequency response when height is changed.

The recording engineer should be aware of how all the above factors affect the depth picture so he can make an intelligent decision on the mike position to try next. The difference between a B+ recording and an A+ recording can be a matter of inches. Hopefully you will recognize the right position when you've found it.

### **Beyond Minimalist Recording**

The engineer/producer often desires additional warmth, ambience, or distance after finding the mike position that achieves the perfect instrumental balance. In this case, moving the mikes back into the reverberant field cannot be the solution. Another call for increased ambience is when the hall is a bit dry. In either case, trucking the entire ensemble to another hall may be tempting, but is not always the most practical solution.

The minimalist approach is to change the microphone pattern(s) to less directional (e.g., omni or figure-8). But this can get complex, as each pattern demands its own spacing and angle. Simplistically speaking, with a constant distance, changing the microphone pattern affects direct to reverberant ratio.

Perhaps the easiest solution is to add ambience mikes. If you know the principles of acoustic phase cancellation, adding more mikes is theoretically a sin. However, acoustic phase cancellation does not occur when the extra mikes are placed purely in the reverberant field, for the reverberant field is uncorrelated with the direct sound. The problem, of course, is knowing when the mikes are deep enough in the reverberant field. Proper application of the **3 to 1 rule** will minimize acoustic phase cancellation. So will careful listening. The ambience mikes should be back far enough in the hall, and the hall must be sufficiently reverberant so that when these mikes are mixed into the program, no deterioration in the direct frequency response is heard, just an added warmth and increased reverberation. Sometimes halls are so dry that there is distinct, correlated sound even at the back, and ambience mikes would cause a comb filter effect.

Assuming the added ambience consists of uncorrelated reverberation, then theoretically an artificial reverberation chamber should accomplish similar results to those obtained with ambience microphones. The answer is a qualified yes, assuming the artificial reverberation chamber sounds

very good and consonant with the sound of the original recording hall.

What happens to the depth and distance picture of the orchestra as the ambience is added? In general, the front-to-back depth of the orchestra remains the same or increases minimally, but the apparent overall distance increases as more reverberation is mixed in. The change in depth may not be linear for the whole orchestra since the instruments with more dominant high frequencies may seem to remain closer even with added reverberation.

### **The Influence of Hall Characteristics on Recorded Front-to-Back Depth in Live Halls**

In general, the more reverberant the hall, the further back the rear of the orchestra will seem, given a fixed microphone distance. In one problem hall the reverberation is much greater in the upper bass frequency region, particularly around 150 to 300 Hz.

A string quartet usually places the cello in the back. Since that instrument is very rich in the upper bass region, in this problem hall the cello always sounds further away from the mikes than the second violin, which is located at his right. Strangely enough, a concertgoer in this hall does not notice the extra sonic distance because his strong visual sense locates the cello easily and does not allow him to notice an incongruity. When he closes his eyes, however, the astute listener notices that, yes, the cello sounds further back than it looks!

It is therefore rather difficult to get a proper depth picture with a pair of microphones in this problem hall. Depth seems to increase almost exponentially when low frequency instruments are placed only a few feet away. It is especially difficult to record a piano quintet in this hall because the low end of the piano excites the room and seems hard to locate spatially. The problem is aggravated when the piano is on half-stick, cutting down the high frequency definition of the instrument.

The miking solution I choose for this problem is a compromise; close mike the piano, and mix this with a panning position identical to the piano's virtual image arriving from the main mike pair. I can only add a small portion of this close mike before the apparent level of the piano is taken above the balance a listener would hear in the hall. The close mike helps solidify the image and locate the piano. It gives the listener a little more direct sound on which to focus.

### **Very Dead Rooms**

Can minimalist techniques work in a dead studio? Not very well. My observations are that simple miking has no advantage over multiple miking in a deadroom. I once recorded a horn overdub in a dead room, with six tracks of close mikes and two for a more distant stereo pair. In this dead room there were no significant differences between the sound of this "minimalist" pair and six multiple mono close up mikes! The close mikes were, of course, carefully equalized, leveled and panned from left to right. This was a surprising discovery, and it points out the importance of good hall acoustics on a musical sound. In other words, when there are no significant early reflections, you might as well choose multiple miking, with its attendant post-production balance advantages.

### **Miking Techniques and the Depth Picture**

The various simple miking techniques reveal depth to greater or lesser degree. Microphone

patterns which have out of phase lobes (e.g., hypercardioid and figure-8) can produce an uncanny holographic quality when used in properly angled pairs. Even tightly-spaced (coincident) figure-8's can give as much of a depth picture as spaced omnis. But coincident miking reduces time ambiguity between left and right channels, and sometimes we seek that very ambiguity. Thus, there is no single ideal minimalist technique for good depth, and you should become familiar with the relative effects on depth caused by changing mike spacing, patterns, and angles. For example, with any given mike pattern, the farther apart the microphones of a pair, the wider the stereo image of the ensemble. Instruments near the sides tend to pull more left or right. Center instruments tend to get wider and more diffuse in their image picture, harder to locate or focus spatially.

The technical reasons for this are tied in to the Haas effect for delays of under approximately 5ms. vs. significantly longer delays. With very short delays between two spatially located sources, the image location becomes ambiguous. A listener can experiment with this effect by mistuning the azimuth on an analog two-track machine and playing a mono tape over a well-focused stereo speaker system. When the azimuth is correct, the center image will be tight and defined. When the azimuth is mistuned, the center image will get wider and acoustically out of focus. Similar problems can (and do) occur with the mike-to-mike time delays always present in spaced-pair techniques.

### **The Front-to-back Picture with Spaced Microphones**

I have found that when spaced mike pairs are used, the depth picture also appears to increase, especially in the center. For example, the front line of a chorus will no longer seem straight. Instead, it appears to be on an arc bowing away from the listener in the middle. If soloists are placed at the left and right sides of this chorus instead of in the middle, a rather pleasant and workable artificial depth effect will occur. Therefore, do not overrule the use of spaced-pair techniques. Adding a third omnidirectional mike in the center of two other omni's can stabilize the center image, and proportionally reduces center depth.

### **Multiple Miking Techniques**

I have described how multiple close mikes destroy the depth picture; in general I stand behind that statement. But soloists do exist in orchestras, and for many reasons, they are not always positioned in front of the group. When looking for a natural depth picture, try to move the soloists closer instead of adding additional mikes, which can cause acoustic phase cancellation. But when the soloist cannot be moved, plays too softly, or when hall acoustics make him sound too far back, then a close mike or mikes (known as spotmikes) must be added. When the close solo mikes are a properly placed stereo pair and the hall is not too dead, the depth image will seem more natural than one obtained with a single solo mike.

Apply the **3 to 1 rule**. Also, listen closely for frequency response problems when the close mike is mixed in. As noted, the live hall is more forgiving. The close mike (not surprisingly) will appear to bring the solo instrument closer to the listener. If this practice is not overdone, the effect is not a

problem as long as musical balance is maintained, and the close mike levels are not changed during the performance. We've all heard recordings made with this disconcerting practice. Trumpets on roller skates?

### **Delay Mixing**

At first thought, adding a delay to the close mike seems attractive. While this delay will synchronize the direct sound of that instrument with the direct sound of that instrument arriving at the front mikes, the single delay line cannot effectively simulate the other delays of the multiple early room reflections surrounding the soloist. The multiple early reflections arrive at the distant mikes and contribute to direction and depth. They do not arrive at the close mike with significant amplitude compared to the direct sound entering the close mike. Therefore, while delay mixing may help, it is not a panacea.

### **Influence of the Control Room Environment on Perceived Depth**

At this point, many engineers may say, "I've never noticed depth in my control room!" The widespread practice of placing near-field monitors on the meter bridges of consoles kills almost all sense of depth. Comb-filtering and sympathetic vibrations from nearby surfaces destroy the perception of delicate time and spatial cues. The recent advent of smaller virtual control surfaces has helped reduce the size of consoles, but seek advice from an expert acoustician if you want to appreciate and manipulate depth in your recordings. We should all do this before we expand to multi-channel, for we still have a lot to learn about taking advantage of the hidden depth in 2-channel recordings.

### **Musical Examples to Check Out**

Check out the **CD Honor Roll** for examples of fantastic recordings. Standard multitrack music recording techniques make it difficult for engineers to achieve depth in their recordings. Mixdown tricks with reverb and delay may help, but good engineers realize that the best trick is no trick: Record as much as you can by using stereo pairs in a live room. Here are some examples of audiophile records I've recorded that purposely take advantage of depth and space, both foreground and background, on Chesky Records. Sara K. Hobo, Chesky JD155. Check out the percussion on track 3, *BrickHouse...* Johnny Frigo, Debut of a Legend, Chesky JD119. Check out the sound of the drums and the sax on track 9, *I Love Paris*. Ana Caram, The Other Side of Jobim, Chesky JD73. Check out the percussion, cello and sax on *Correnteza*. Carlos Heredia, Gypsy Flamenco, Chesky WO126. Play it loud! And listen to track 1 for the sound of the background singers and handclaps. Phil Woods, Astor and Elis, Chesky JD146, for the natural-sounding combination of intimacy and depth of the jazz ensemble.

### **Technological Impediments to Capturing Recorded Depth**

Depth is the first thing to suffer when low-resolution technology is used. Here is a list of some of the technical practices that when misused, or accumulated, can contribute to a boringly flat, depthless recorded picture: Multitrack and multimike techniques, small/dead recording studios, low resolution recording media, amplitude compression, improper use of dithering, cumulative digital processing, and low-resolution digital processing (e.g., using single-precision as opposed to



double or higher-precision equalizers). When recording, mixing and mastering-use the best miking techniques, room acoustics, and highest resolution technology, and you'll resurrect the missing depth in your recordings.

**Thanks to:**

My assistant, David Holzmann, for transcribing my original 1981 article, which I have herein revised and updated.

Lou Burroughs, whose 1974 book **Microphones: Design and Application**, now out of print, is still one of the prime references on this subject and covers the topic of acoustic phase cancellation. **Burroughs invented the 3-to-1 rule, expressed simply:** *When a sound source is picked up by one microphone and also "leaking" into another microphone that is mixed to the same channel---make sure the distance between microphones is at least three times the distance of the first mike from the source. This general rule does not account for the intensities of all instruments or all room acoustics, but you should listen critically when your microphone distances decrease. And remember, this applies to mixing to one channel, for non-coincident stereo microphone techniques can break the 3-t-1 rule, and if so, be sure to check the sound in mono for phase cancellation.*

E. Roerback Madsen, whose article "Extraction of Ambiance Information from Ordinary Recordings" can be found in the 1970 October issue of the **Journal of the Audio Engineering Society**. Covers the Haas effect and its correlary.

Don Davis, who first defined "critical distance" and many other acoustic terms.

---

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

<  
Prev

Next  
>

[ [Back](#) ]

**Popular**

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)

- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us

© 2007 Digital Domain - All Rights Reserved


[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## Mixing Tips and Tricks



**First, read all the articles on this web site!**

**Here's one of the secrets of the mixing engineers**

To avoid squashing, if it doesn't sound loud enough to your ears, turn up the monitor! If you find that you've been forced to apply limiting or compression just to keep the meters from overloading, then you've been going about this backwards. Instead, turn down your individual mix levels several dB, then get rid of any compression you were using to "protect" the 2-mix. Now your mix is at a lowered meter level, so turn up your monitor gain to arrive at the same loudness--only this time it won't sound squashed. Leave the monitor at that position as you continue to mix (mark it so you can get back to it).

In 24-bit recording you can make a perfectly good mix that peaks between -3 and -10 dBFS with no loss of quality, in fact, with improved quality. So if the mix gets too loud by your ears, then turn down the elements that are too hot in the mix instead of turning down the monitor again, with no fear of mixing "too low". In other words, a high monitor gain gives you less temptation to overcompress. High monitor gain does not necessarily mean high monitor output from the speakers--it means that the mix level had to be lower. For example, visit the [CD Honor Roll](#) and check out the great-sounding Lyle Lovett selection, which is close to the dynamics of a raw mix. Notice that in order to listen to it, you have to turn up your monitor gain. That's approximately where your monitor control for a dynamic raw mix should be sitting (within 4-6 dB) before mastering. Obviously, a lot of today's hypercompressed masters would require turning down the monitor, but we're trying to show you how not to ruin the record in the mix stage (and hopefully not in the mastering, either!).

**Know thy monitors**

OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

Password

Remember me

But even when you do, never be fooled. Take your mixes around and listen to them on several other systems that you know; then go back into the control room and if they do not translate, try to adjust your mixes in the areas where they do not translate. HOWEVER, be aware of the extremes. If it sounds reasonably good in a car, for example, don't be tempted to turn up the highs for the car or it will screech (horribly) EVERYWHERE else. First of all, in the mastering we have much more experience in knowing how far to go and make sure that a recording is not made bass-shy just because it sounds boomy in a naturally-boomy car, for example.

### **Always mix to the highest possible wordlength**

Even if the source tracks are 16-bit! Do **not** sample rate convert. When you're ready to bounce or prepare files, please see our [guidelines](#) page for suggestions on making file names and file types.

### **Track important instruments in stereo**

In the days of 8 track you had to be very careful about allocating tracks. But those days are gone. You have enough tracks to splurge now! So there's no reason to conserve on tracks during the tracking stage. The stereo image and depth of your final product will be determined by your skill in mixdown at using delays, reverberation, effects, and your skills in tracking, how you tracked your instruments. Try to make a plan beforehand of how your soundstage might look, where the instruments might be placed. Realize that it probably will not hurt, and probably will help to record your important instruments in stereo.

For example, even a pair of bongos that are destined to be on the right side of the soundstage will sound better if one bongo mike is panned full right and the other somewhat right of center. This is because the ear decodes the natural space and delays picked up by those microphones, actually enhancing their definition in the mix (if the room acoustics are good).

Another example: Electric guitar. Capture the direct to one track. Capture the output of the loudspeaker with a close mike to another track. Capture the medium distant sound of the speaker bouncing from the walls of the room with another mike. Listen to the combination of these sources panned to different places, and also listen in mono to make sure you have not created phase cancellations. By using stereo miking and natural room acoustics in the tracking, and possibly artificial delays and good stereophonic reverberation in the mixing, your mix will sound richer and deeper. Not everything should be tracked in stereo, but don't skimp on elements that will increase the depth and space of your recording. Of course you will need a foreground, middleground, and background in the mix, but it's a lot harder to create a location and space for an instrument if you had only recorded it in mono.

In the mixing, use artificial reverberators that enhance depth and space and do not sound flat, plastic or "cheesy." Use artificial delays to locate instruments in space, not just simple panners.

### **Levels**

Try to not exceed -3 dBFS peak on a peak meter on the highest peak of the mix. Low levels are perfectly acceptable in a 24 bit system. Once you see that the highest peak is in the range of, say, -10 dBFS to -3 dBFS, then from that point on, if you can hear it, the low level passages are ok.

Preserve dynamic range! Assume that if anyone is going to ruin the master, let it be me (the mastering engineer). If the mix sounds good, then soft passages automatically are NOT too soft. Of course, if you think a soft passage sounds too soft in the mixing, then of course try to fix that during the mixing. But these can easily be dealt with and often more efficiently in mastering, as we have the context of the album in mind.

If you have a VU meter, use it. With a sine wave, adjust it so 1 kHz, 0 VU is equal to -20 dBFS on the peak meter. Use the VU, ignore the peak, and you'll start making better mixes.

### **Vocal levels**

Do make a lead vocal up (1/2 or 1 dB, you be the best judge) version. Do it NOW before you forget. It's a lot easier to do it NOW than to discover in the mastering that you should have. Occasionally do a vocal down (1/2 or 1 dB) version if you think it may be useful; then again, it only takes 3 minutes to do a vocal alt version when you're in the heat of mixing, but it takes forever to try to fix it in the mastering if you forgot.

### **Original sources, please**

If at all possible, deliver a generation that is as close to the original as possible. If it's on CD ROM, then cut a CDR directly from your hard disc files. Speed of cutting? Try to use Taiyo Yuden or other reputable blanks, and cut at 4X to 8X speed. These will PROBABLY produce the best results. Murphy's Law: Allow for Murphy. Do not ASSUME that all the files will transfer successfully over here and that the CD-ROMs you have cut are perfect. Allow for the possibility that on the very last minute of the very last hour of the very last day, we may have to go to a backup CD-ROM, or you may have to cut another, because of some error or other problem in the transfer. Do not paint yourself into a corner. Make backups. Do not destroy or erase any source hard disc at the origination studio until the mastering has been completed.

### **When, Why, and How to Make Stems**

I've definitely reached the conclusion that the less compromise you can make in the mastering process, the better the result. Let's say you have an otherwise great mix, but which has too little bass instrument, too much kick, and the lower midrange is a little bit muddy. This is a potentially bad (not lethal) combination for mastering and if the client has time, I recommend a remix.

However, in situations like the afore mentioned, when time is tight, I have also asked the client for stems, and the results have ALWAYS been better than if I had mastered from the combined two-track. Next the question comes of whether to remix the stems without mastering processing in line or to try to mix/master in the same path. If it were a 40 track mix, I'd definitely mix first, then master, but with 3 to 6 stereo stems, I find that I can get the best results mixing and mastering at the same time; the result produces the best results and the least compromise. For example, the mastering processing is going to affect the clarity of the midrange and through "slop" will probably leak down into the bass region, hopefully for the better. But in the case of this lopsided mix I just cited, the mastering processing could easily make one range better while making the other worse.

So, if mixing without the mastering processing, I may even try to take that into account, but if

mixing with the mastering processing in place, I have it all in context at one time in the ideal acoustic of the mastering room.

Is this heresy? It's certainly a dangerous technique if placed in the wrong hands. You can end up with a less than ideal mix or less than ideal master if the mastering engineer does not think holistically. But if placed in the hands of an experienced mastering engineer, I think mixing from stems while mastering can produce the very best product. Separating out the bass instrument into a stereo stem with otherwise a mix minus, and sometimes separating out the vocal the same way can reduce the number of calls for a remix, I am convinced.

In other words, the lines between mixing and mastering have never been black and white. There has always been a gray area, and this method of mastering from stems grays it out even more!

The main purpose in this discussion was in the context of suggesting having a separate mastering engineer do the mastering from the stems, not in having a complete mixdown/mastering in one step.

If I am asked to both mix and master a project; if I am fortunate enough to do the mix from scratch in the mastering environment, then I probably would mix direct to 2 track without stems. I might run stems as a safety only; an ounce of prevention is worth a pound of cure. On that note, I note that with digital technology, a single 10 second mistake can cost a whole day of makeup! We don't need no stinkin' backups :-)

But I digress...

So, if I were mixing in the mastering environment I would probably just mix to 2 track WITHOUT MASTERING PROCESSING. But if I were mixing in a typical mixing environment, I would try to mix to stems if possible WITHOUT MASTERING PROCESSING.

What I'm saying is that although there is a gray area between mixing and mastering I don't advocate trying to combine the two processes when mixing completely from scratch. I only say that it is possible to do a good (better) job if you are the independent mastering engineer on the project and you receive stems instead of full mixes.

Click here to see what other engineers had to say on the subject of [Compression in Mixing](#).

---

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

<  
Prev

Next  
>

[ [Back](#) ]



**Popular**

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)

- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us

© 2007 Digital Domain - All Rights Reserved


[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## Digital Detective



### Use or Build your own Bitscope!

Attention, Sherlock Holmes's of the audio world. You can use an ordinary oscilloscope (20 MHz or better) to see the bit activity of your digital processors, consoles and workstations. Steinberg's [Wavelab](#) and many other software programs have a bitscope meter as well. Once you install it, you'll find the bitscope is as essential in the modern-day digital studio as a phase meter. To learn more about how the bitscope has saved the day in studios, read our article, [More Bits Please](#).

These photos illustrate some typical bitscope displays.

### How to Build a Bitscope

Every digital audio recorder, processor or console extracts serial DATA and WORDCLOCK from the AES/EBU or S/PDIF line. Pick a "neutral" machine or processor that you can patch into your digital audio system at the end of your processing or monitoring chain, so you can analyze what all your processors are doing to the signal. All you have to do is connect the vertical input of your oscilloscope to DATA, and its trigger or timebase to WORDCLOCK (44.1 or 48 KHz), to see which bits and how many bits are being used at all times. If you're not used to digging into audio equipment, then give the job to someone who is. Opening any manufacturer's gear may void the warranty.

Crystal Semiconductor's ubiquitous CS8412 digital receiver IC is used in many processors. You'll find DATA on pin 26, and WORDCLOCK on pin 11 of this 28-pin chip. Attach the shield of the scope lines to ground. I suggest soldering a 75 ohm build-out (isolation) resistor from the chip's pins to the scope lines, to protect the signals from accidental shorts. Use good, short coax cables (I've used three feet with no problems).

You can still be a digital detective even if you're not the do-it-yourself type. *Digital Domain* will add scope outputs to its *FCN-1 Format Converter* or *VSP/P Digital Audio Control Center* for a small

OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

Password

Remember me

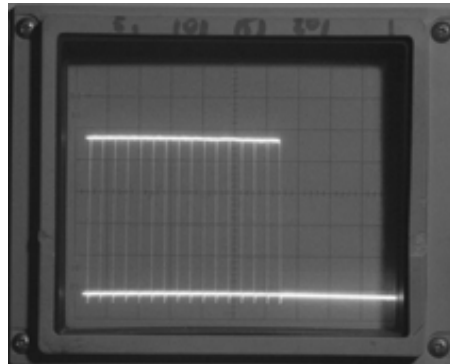
fee. For further information, contact *Digital Domain* **1-800-DIGIDO-1** or [email us](#).

### Interpreting the Display

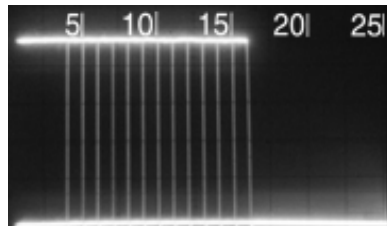
The bitscope will tell you when certain things are *wrong* (e.g., missing bits, or extra bits), but it can't guarantee that everything is *right* (e.g., harmonic distortion will not show on the bitscope). Use the bitscope as a visual aid, a first line of defense against digital audio problems. Your ears and your knowledge must do the rest.

The 8412 chip can be configured for many modes. The most common mode presents one channel's worth of data on wordclock "up," and the other channel on wordclock "down." Crystal uses a 64-bit "slot," so you'll see up to 24 bits worth of one channel, followed by 8 bits of "silence," then the other channel (another 32-bit half-slot). Counting bits is easy if you adjust your scope's timebase to show one audio channel, and 2-1/2 bits per division, which gives a convenient count of 5 bits every two divisions, and spreads 24 bits across the whole screen. The format is 2's complement, with the MSB at left, and LSB at right. When the MSB is low, the audio signal is positive, when high, it's negative, so the MSB will be toggling all the time, unless the signal is pure DC. A toggling bit will appear to have both high and low values, this just means that the eye's persistence of vision is showing both values.

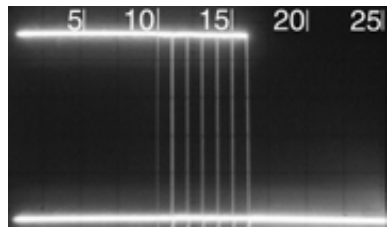
These scope pictures are a little over-exposed, so the top vertical line is fatter and brighter than the actual scope display.



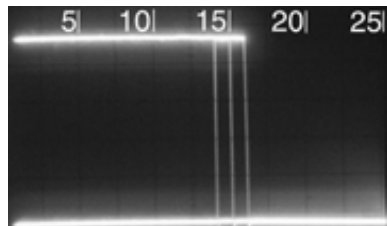
This is the bitscope, showing one channel, full scale 16-bit sine wave. Note the handwritten scale on the top of the chassis.



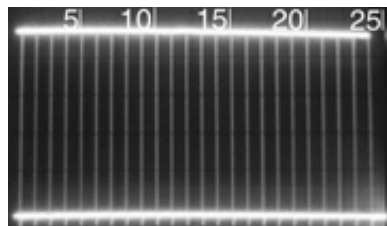
16-bit sinewave at -20 dBFS. I have added a computer-driven counter scale to these images to make it easy to identify the bits.



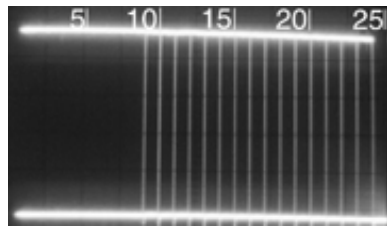
16-bit sine wave at -60 dBFS.



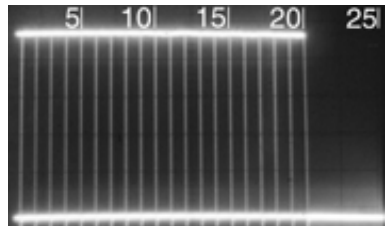
16-bit sine wave at -80 dBFS.



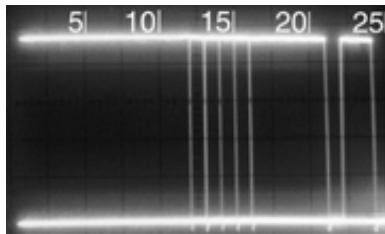
24-bit full scale sine wave.



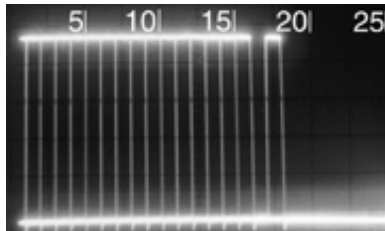
24-bit sine wave at -50 dBFS.



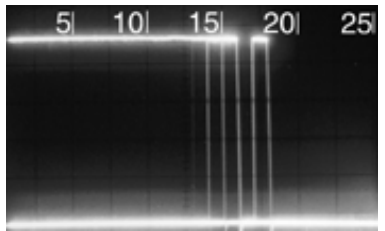
20-bit full scale sine wave.



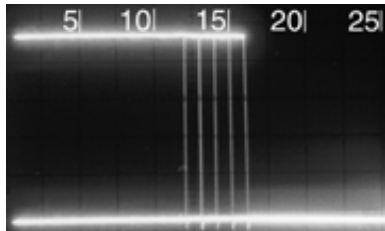
"Defective" digital processor in BYPASS. Source is a 16-bit sine wave at -70 dBFS. The additional bits could be DC offset or what?



Defective dithering processor set for 16-bit output. Source is a 16-bit full-scale sine wave. Note the missing bit in the 17th position and an extra 18th bit is toggling.



The same defective dithering processor idling (with no input signal). Note the faint line showing the 14th bit is toggling, along with the 15 and 16th bits, plus the same missing bit at the 17th position, and the toggling 18th bit.



Dithering processor in idle (no input signal), showing 4 bits toggling (high order dither with noise-shaping).

---

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

<  
Prev

Next  
>

[ Back ]

**Popular**

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)

- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us

© 2007 Digital Domain - All Rights Reserved




[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## Dither



### Keeping Your Digital Audio Pure from First Recording to Final Master

#### Part I

Dither is not the most important technical detail to learn about, but if you want to get your digital audio done just right, then you should learn about dither. Especially if you want to learn why your digital reverb has been leaching the ambience out of your music, when it's supposed to be *adding* ambience. Or why your CDs don't sound as spacious as your 24 bit sources and want to avoid that veiled, dry, and lifeless feeling!

#### Follow that Sample

Let's start with a little lesson in DSP (Digital Signal Processors). Many workstation and processor manufacturers ignore the critical issue of **wordlength**. Let's examine what happens to digital audio when you change gain (or mix, equalize, compress, sample rate convert, or perform any type of calculation) in a digital audio workstation. It's all arithmetic, isn't it? Yes, but the accuracy of that arithmetic, and how you (or the workstation) deal with the arithmetic product, can make the difference between pure-sounding digital audio or digital sand paper.

All DSPs deal with digital audio on a sample by sample basis. At 44.1kHz, there are 44,100 samples in a second (88,200 stereo samples). When changing gain, the DSP looks at the first sample, performs a multiplication, spits out a new number, and then moves on to the next sample. It's that simple.

Instead of losing you with esoteric concepts like two's complement notation, fixed vs. floating point, and other digital details, I'm going to talk about *digital dollars*. Suppose that the value of your first digital audio sample was expressed in dollars instead of volts, for example, one dollar

OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

Password

Remember me

and fifty one cents--\$1.51. And suppose you wanted to take it down (attenuate it) by 6 dB. If you do this wrong, you'll lose more than money, by the way. 6 dB is half the original value (it has to do with logarithms; don't worry about it). So, to attenuate our \$1.51 sample, we divide it by 2.

Oops! \$1.51 divided by 2 equals 75-1/2 cents, or .755. So, we've just gained an extra decimal place. What should we do with it, anyway? It turns out that dealing with extra places is what good digital audio is all about. If we just drop the extra five, we've theoretically only lost half a penny--but you have to realize that half a penny contains a great deal of the natural ambience, reverberation, decay, warmth, and stereo separation that was present in the original \$1.51 sample! Lose the half penny, and there goes your sound. The dilemma of digital audio is that most calculations result in a longer wordlength than you started with. Getting more decimal places in our digital dollars is analogous to having more bits in our digital words. When a gain calculation is performed, the wordlength can increase infinitely, depending on the precision we use in the calculation. A 1 dB gain boost involves multiplying by 1.122018454 (to 9 place accuracy). Multiply \$1.51 by 1.122018454, and you get \$1.694247866 (try it on your calculator). Every extra decimal place may seem insignificant to you, until you realize that DSPs require repeated calculations to perform filtering, equalization, and compression. 1 dB up here, 1 dB down here, up and down a few times, and the end number may not resemble the right product at all, *unless adequate precision is maintained*. Remember, the more precision, the cleaner your digital audio will sound in the end (up to a reasonable limit).

### **The First Secret of Digital Audio**

Now you know the first critical secret of digital audio: *wordlengths expand*. If this concept is so simple, why is it disregarded by some manufacturers? The answer is in your wallet. While DSPs are capable of performing double and triple precision arithmetic (all you have to do is store intermediate products in temporary storage registers), it slows them down, and complicates the whole process. It's a hard choice, entirely up to the DSP programmer/processor designer, who's been put under the gun by management to fit more program features into less space, for less money. Questions of sound quality and quantization distortion can become moot compared to the selling price.

Inside a digital mixing console (or workstation), the mix buss must be much longer than 16 bits, because adding two (or more) 16-bit samples together and multiplying by a coefficient (the level of the master fader is one such coefficient) can result in a 32-bit (or larger) sample, with every little bit significant. Since the AES/EBU standard can carry up to 24-bits, it is practical to take the 32-bit word, round it down to 24 bits, then send the result to the outside world, which could be a 24-bit storage device (or another processor). The next processor in line may have an internal wordlength of 32 or more bits, but before output it must round the precision back to 24 bits. The result is a slowly cumulating error in the least significant bit(s) from process to process. Fortunately, the least significant bit of a 24-bit word is 144 dB down, and most sane people

recognize that degree of error to be inaudible.

### Something For Nothing?

But suppose you want to record the digital console's output to a 16 bit medium, like the CD. Frankly, it's a serious compromise to take your console's 24-bit output word and truncate it to 16 bits. After processing, the mastering engineer uses a technique called *dithering* to take long wordlengths, and cleanly turn them to 16-bit for the CD. First, must ensure that our DAW is high resolution (has very low distortion at low levels) and can be bit-transparent when called upon. **Bit-transparent** means that the output is identical to the source, from the most significant to the least significant bit, that the DAW does not increase or decrease the source wordlength.

### Good Advice

Once you've verified your workstation is bit-transparent, then proceed with editing, with the goal of maintaining the integrity of your original source. **Do not change gain** unless you need to align the gains of two pieces you are editing together. **Do not normalize** (normalization is just changing gain). **Do not equalize. Do not fade in or fade out.** Just edit. This is to avoid additional DSP or degradation when the mix gets to the mastering studio .Leave the segues, fadeouts and gain changes for the mastering house, where they can properly handle the long wordlengths necessary for smooth fades (so that's why your last fadeout sounded like it dropped off a cliff!). Follow these simple guidelines and your digital audio will immediately start sounding better.

## Part II

### Dither

**How to Dither** Let's look at that long sample word. Whether it's 24 bits or 32 bits, we have to find some way to move the important information contained in the lower (least significant) bits into the upper 16 bits for recording to the CD standard. Truncation is very bad. What about rounding? In our digital dollar example, we ended up with an extra 1/2 cent. In grammar school, they taught us to round the numbers up or down according to a rule (we learned "even numbers... roundup, odd...round down"). But when we're dealing with more numerical precision and small numbers that are significant, it gets a little more complicated.

It turns out the best solution for maintaining the resolution of digital audio is to calculate random numbers and add a different random number to every sample. Then, cut it off at 16 bits. The random numbers must also be different for left and right samples, or else stereo separation will be compromised.

For example:

Starting with a 24-bit word (each bit is either a 1 or a 0 in binary notation):

Upper 16 bits                      Lower 8

Original 24-bit Word MXXX XXXX XXXX XXXW YYYY YYYY

Add random number

ZZZZ ZZZZ

The result of the addition of the Z's with the Y's gets carried over into the new least significant bit of the 16-bit word (LSB, letter W above), and possibly higher bits if you have to carry. In essence, the random number sequence combines with the original lower bit information, *modulating* the LSB. Therefore, the LSB, from moment to moment, turns on and off at the rate of the original low level musical information. The random number is called *dither*; the process is called *redithering*, to distinguish from the original *dithering* process used to during the original recording.

Random numbers such as these translate to random noise (hiss) when converted to analog. The amplitude of this noise is around 1 LSB, which for 16 bit lies at about 96 dB below full scale. By using dither, ambience and decay in a musical recording can be heard down to about -115 dB, even with a 16-bit wordlength. Thus, although the quantization steps of a 16-bit word can only theoretically encode 96 dB of range, with dither, there is an audible dynamic range of up to 115 dB! The maximum *signal-to-noise* ratio of a dithered 16-bit recording is about 96 dB. But the *dynamic range* is far greater, as much as 115 dB, because we can hear music *below the noise*. Usually, manufacturer's spec sheets don't reflect these important specifications, often mixing up dynamic range and signal-to-noise ratio. *Signal-to-noise ratio* (of a linear PCM system) is the RMS level of the noise with no signal applied expressed in dB below maximum level (without getting into fancy details such as noise modulation). It should be, ideally, the level of the dither noise. *Dynamic range* is a subjective judgment more than a measurement--you can compare the dynamic range of two systems empirically with identical listening tests. Apply a 1 kHz tone, and see how low you can make it before it is undetectable. You can actually measure the dynamic range of an A/D converter without an FFT analyzer. All you need is an accurate test tone generator and your ears, and a low-noise headphone amplifier with sufficient gain. Listen to the analog output and see when it disappears (use a real good 16 bit D/A for this test). Another important test is to attenuate music in your workstation (about 40 dB) and listen to the output of the system with headphones. Listen for ambience and reverberation; a good system will still reveal ambience, even at that low level. Also listen to the character of the noise--it's a very educating experience.

### Some Tests for Linearity

You can verify whether your digital audio workstation truncates digital words or does other nasty things, without any measurement instruments except your ears. Obtain the disc *Best of Chesky Classics and Jazz and Audiophile Test Disc, Vol. III*, Chesky JD111.\* Track 42 is a fade to noise without dither, demonstrating quantization distortion and loss of resolution. Track 43 is a fade to noise with white noise dither, and track 44 uses noise-shaped dither (to be explained). Use Track 43 as your test source; you should be able to hear smooth and distortion-free signal down to about -115 dB. Then listen to track 44 to see how much better it can sound. Try processing track 43 with digital equalization or level changes (both gain and attenuation, with and without dither, if it's available in your workstation) to see what they do to the sound. If your workstation is not up to par, you'll be shocked. Use a quiet, high-gain headphone amplifier to help reveal the low level

problems.

\*available at major record chains or through Chesky Records, Box 1268, Radio City Station, New York, NY 10101; 212-586-7799. The hard-to-find CBS CD-1, track 20, also contains a fade to noise test.

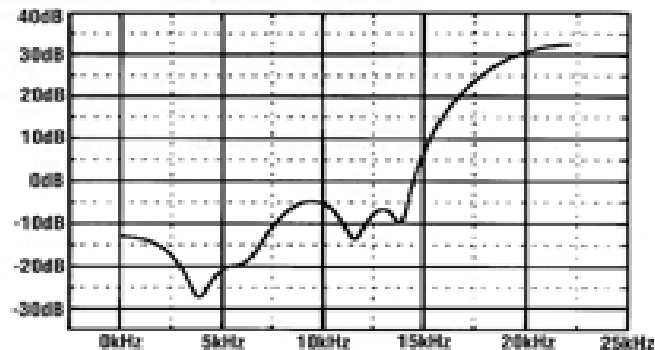
### So Little Noise, So Much Effect

-96 dB seems like so little noise. But strangely, engineers have been able to hear the effect of the dither noise, even at normal listening levels. Dither noise helps us recover ambience, but conversely it also obscures the same ambience we've been trying to recover! Dither noise adds a slight veil to the sound. That's why I say, *dither, you can't live with it, and you can't live without it.*

### Improved Dithering Techniques

Where there's a will, there's a way. Although the required amplitude of the dither is about -96 dB, it's possible to shape (equalize) the dither to minimize its audibility. Noise-shaping techniques re-equalize the spectrum of the dither while retaining its average power, moving the noise away from the areas where the ear is most sensitive (circa 3 KHz), and into the high frequency region (10-22 KHz).

Here is a picture of one of the most successful noise-shaping curves (courtesy of Meridian Audio, Ltd).



As you can see, it is a very high-order filter, requiring considerable calculation, with several dips where human hearing is most sensitive. The sonic result is an incredibly silent background, even on a 16-bit CD. The 0 dB line is around -96 dBFS in this diagram.

There are numerous noise-shaping redithering devices on the market. Very high precision (56 to 72 bit) arithmetic is required to calculate these random numbers. One box uses the resources of an entire DSP chip just to calculate dither. The sonic results of these noise-shaping techniques range from very good to marvelous. The best techniques are virtually inaudible to the ear. With 72-bit arithmetic, all the dither noise has been pushed into the high frequency region, which at -

60 or -70 dB is still inaudible. Critical listeners were complaining that the high frequency rise of the early noise-shaping curves changed the tonality of the sound, adding a bit of brightness. But it turns out that it is the shape of the curve in the midband that affects the tonality, due to masking. Two or three of the latest and best of these noise-shaping dithers are tonally neutral, to my ears. It took a long time to get there (about 10 years of development), but now we can say that the best of these processors yield 19-20 bit performance on a 16-bit CD, with virtually no tonal alteration or loss of ambience from the 24-bit source.

Noise-shapers on the market include: db Technologies model 3000 Digital Optimizer, Meridian Model 618, Sony Super Bit Mapping, Waves L1 and L2 Ultramaximizers, Prism, POW-R, and several others. When using dithering plugins, be sure to use them with the right version of workstation software to retain a 24-bit wordlength until the final mastering step.

Apogee Electronics produced the UV-22 system, in response to complaints about the sound of earlier noise-shaping systems, declaring that 16-bit performance is just fine. They do not use the word "dither" (because their noise is periodic, they prefer to call it a "signal"), but it smells like dither to me. Instead of noise-shaping, UV-22 adds a carefully calculated noise at around 22 KHz, without altering the noise in the midband.

To effectively compare the sound and resolution of these redithering techniques, perform the low level test described above. Feed low level 24-bit music (around -40 dB) into the processor, and listen to the output at high gain in a pair of headphones with a good quality D/A converter. You will be shocked to hear the sonic differences between the systems. Some will be grainy, some noisy, and some distorted, indicating improper dithering or poor calculation. The winner of this test should be your choice of dithering processor, although at high gains you are exaggerating the effect of the extra high frequencies in the noise, which would not be noticed at normal gains..

### **Damage, Destruction, or just Deterioration?**

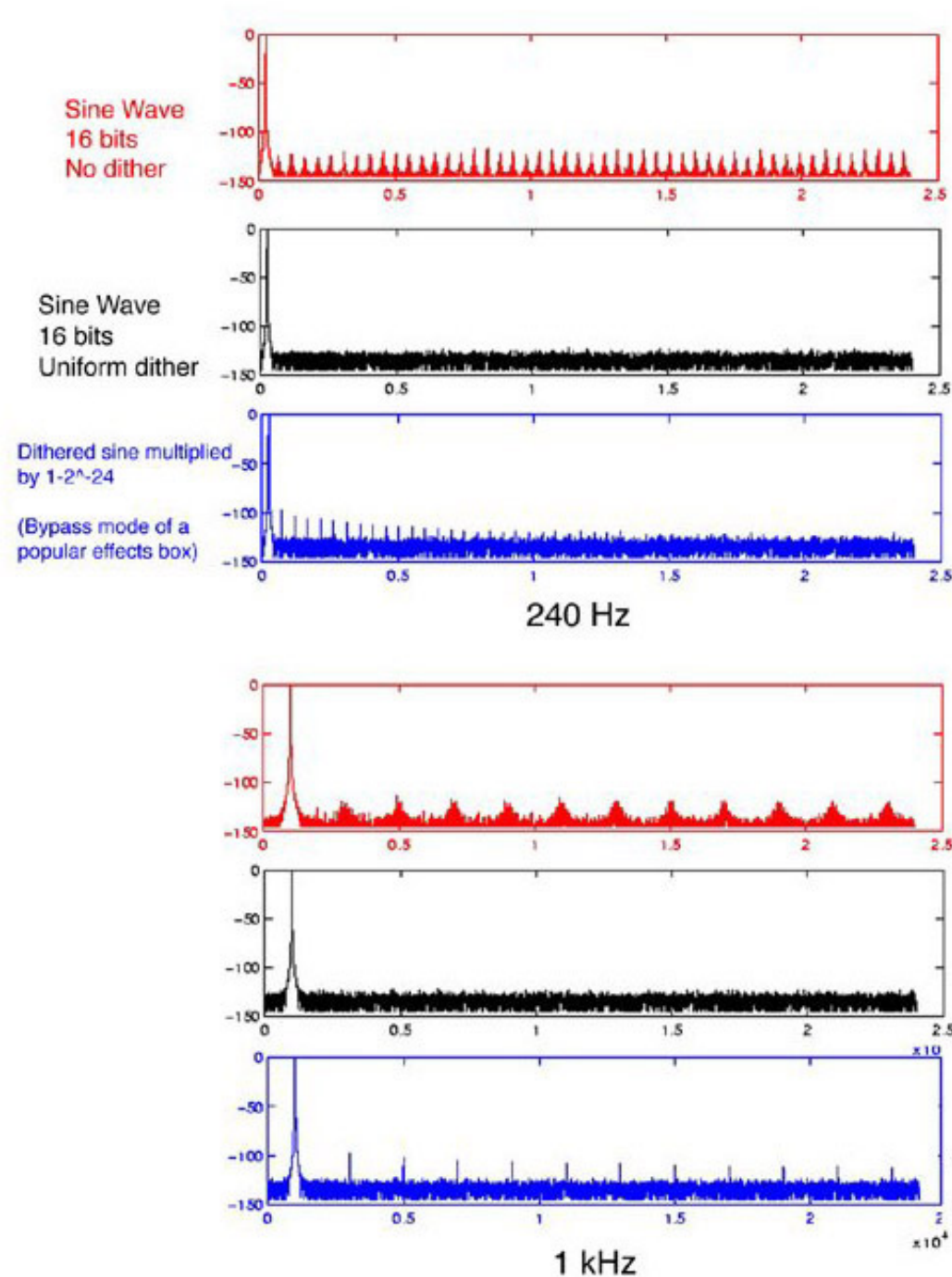
Before digital recording and editing, every edit was destructive. Every equalization or gain change involved an analog copy, with attendant noise, or remixing the multitrack, which "destroys" or replaces the previous mixdown. After DAWs were invented, people started talking about "non-destructive"-editing, and keeping your sound in the digital domain until the end. But as we have seen, even "non-destructive" may be damaging if word lengths aren't maintained.

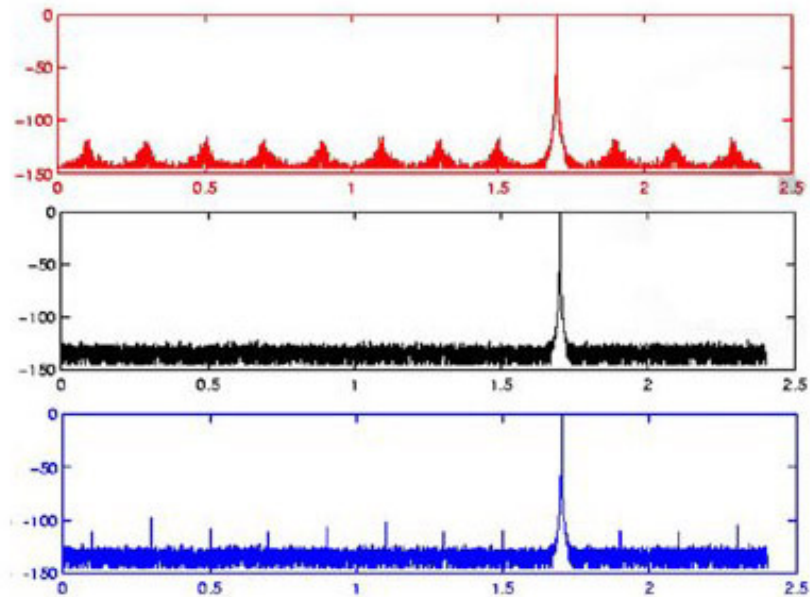
### **The Best Approach**

To maintain the quality of your digital audio, always store the full output wordlength of your digital processors. Also, be sure to *question authority*. Never take a digital processor for granted. Don't even trust BYPASS mode, unless you're sure the processor produces true clones in bypass. The following illustration (courtesy of Jim Johnston, AT&T research), shows a series of FFT plots of a sine wave. The top row is an undithered 16 bit sine wave. Note the distortion products (vertical spikes at regular intervals). The second row is that sine wave with uniform dither. Note how the distortion products are now gone. The bottom row is the dithered sine wave, going through a



popular model of digital processor set for BYPASS and truncated to 16 bits. This is what would happen if you took your source, fed it through this processor in BYPASS mode, and recorded it again!





17 kHz

Plots calculated by and courtesy of Jim Johnston.  
 Copyright jj@research.att.com 1998, all rights reserved, except transmission  
 by USENET and like facilities granted. This notice must be included. Any use by a  
 provider charging in any way for the IP represented in and by this article and any  
 inclusion in print or other media are specifically prohibited.  
 Converted from EPS by IMarc Lindahl <marc@sonorus.com>  
 President Sonorus, Inc. | <http://sonorus.com>  
 Layout by Bob Katz, Digital Domain

Disarming, isn't it? That's why you should arm yourself with a **bitscope** or test every processor you own for bit transparency before attempting to make master-quality work with those processors patched in your signal chain.

### The Cost of Cumulative Dithering

When feeding processors, DAWs or digital mixers to your recording unit, dither the output of the processor to a 24-bit word. Dithering always sounds better than truncation without dither. But to avoid adding a veil to the sound, avoid cumulative dithering, in other words, multiple generations of any dither. Make sure that redithering to 24- or 16-bit is the **one-time, final process in your project**. For related information visit my article **More Bits, Please**. When performed properly, dithering will help your music to retain its depth and purity of tone.

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

<  
Prev

Next  
>

[ Back ]

Popular

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)

- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us

© 2007 Digital Domain - All Rights Reserved


[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## Jitter



December 2007

I hesitate to remove this older article from our website, as it is still informative, but I highly recommend that those interested in the latest word on this subject please read the chapter on jitter in my new book. Some questions that this previous article has raised have been clarified in our [letters](#) section, and of course are covered much better in the book. -BK

-----  
 Jitter is so misunderstood among recording engineers and audiophiles that we have decided to devote a Section to the topic. All digital devices that have an input and an output can add jitter to the signal path. For example, Digital Domain's *FCN-1 Format Converter* adds a small amount of jitter (around 200 ps RMS) to the digital audio signal path. *Is this good? Is it bad? What sonic difference does it make?* We will attempt to answer these--and other important--questions in this Section.

### What is Jitter?

Jitter is time-base error. It is caused by varying time delays in the circuit paths from component to component in the signal path. The two most common causes of jitter are poorly-designed Phase Locked Loops (PLL's) and waveform distortion due to mismatched impedances and/or reflections in the signal path.

Here is how waveform distortion can cause time-base distortion:

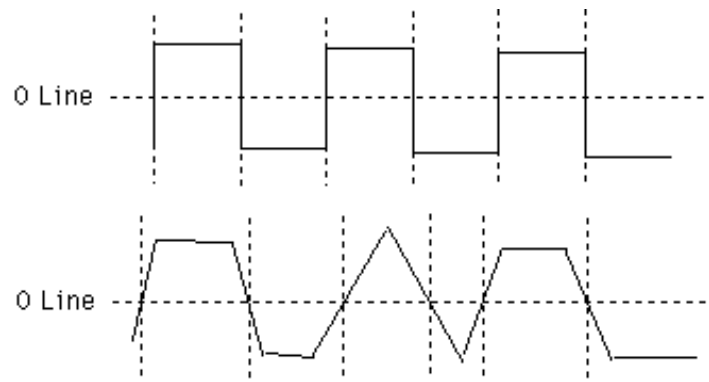
OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

Password

Remember me



The top waveform represents a theoretically perfect digital signal. Its value is 101010, occurring at equal slices of time, represented by the equally-spaced dashed vertical lines. When the first waveform passes through long cables of incorrect impedance, or when a source impedance is incorrectly matched at the load, the square wave can become rounded, fast risetimes become slow, also reflections in the cable can cause misinterpretation of the actual zero crossing point of the waveform. The second waveform shows some of the ways the first might change; depending on the severity of the mismatch you might see a triangle wave, a squarewave with ringing, or simply rounded edges. Note that the new transitions (measured at the Zero Line) in the second waveform occur at unequal slices of time. Even so, the numeric interpretation of the second waveform is still 101010! There would have to be very severe waveform distortion for the value of the new waveform to be misinterpreted, which usually shows up as audible errors--clicks or tics in the sound. If you hear tics, then you really have something to worry about.

If the numeric value of the waveform is unchanged, why should we be concerned? Let's rephrase the question: "*when* (not *why*) should we become concerned?" The answer is "hardly ever." *The only effect of timebase distortion is in the listening; as far as it can be proved, it has no effect on the dubbing of tapes or any digital to digital transfer (as long as the jitter is low enough to permit the data to be read. High jitter may result in clicks or glitches as the circuit cuts in and out).* A typical D to A converter derives its system clock (the clock that controls the sample and hold circuit) from the incoming digital signal. If that clock is not stable, then the conversions from digital to analog will not occur at the correct moments in time. The audible effect of this jitter is a possible loss of low level resolution caused by added noise, spurious (phantom) tones, or distortion added to the signal.

A **properly dithered** 16-bit recording can have over 120 dB of dynamic range; a D to A converter with a jittery clock can deteriorate the audible dynamic range to 100 dB or less, depending on the severity of the jitter. I have performed listening experiments on purist, audiophile-quality musical source material recorded with a 20-bit accurate A/D converter (dithered to 16 bits within the A/D). The sonic results of passing this signal through processors that truncate the signal at -110, -105,

or -96 dB are: increased "grain" in the image, instruments losing their sharp edges and focus; reduced soundstage width; apparent loss of level causing the listener to want to turn up the monitor level, even though high level signals are reproduced at unity gain. Contrary to intuition, you can hear these effects without having to turn up the listening volume beyond normal (illustrating that low-level ambience cues are very important to the quality of reproduction). Similar degradation has been observed when jitter is present. Nevertheless, the loss due to jitter is subtle, and primarily audible with the highest-grade audiophile D/A converters.

### **Jitter And the AES/EBU Interface**

The AES/EBU (and S/PDIF) interface carries an embedded clock signal. The designers of the interface did not anticipate that it could cause a subtle amount of jitter due to the nature of the preamble in the AES/EBU signal. The result is a small amount of program-dependent jitter which often sounds like an intermodulation, a high-frequency edge added to the music. To minimize this effect in the listening, use a D/A converter with a high degree of internal jitter reduction. An external jitter reduction device that removes the subcode signal (containing time of day, start IDs, etc.) also helps.

The SDIF-2 (Sony Digital Interface-2) uses a separate cable for the clock signal, and thus is not susceptible to program-dependent jitter. However, the quality of the PLL used to detect an SDIF-2 wordclock is still important to low jitter. It is much easier to build a low-jitter PLL for a wordclock signal than for an AES/EBU signal.

### **Is Jitter Cumulative? What About My Dubs?**

Consider a recording chain consisting of an A to D Converter, followed by the FCN-1, feeding a DAW, and finally a D to A Converter. During the recording, the jitter you will hear is dependent on the ability of the last PLL in the chain (in the D to A) to reduce the cumulative jitter of the preceding elements in the chain. The time-base error in the D to A is a complex aggregate of the timebase errors of all the preceding devices, including their ability to reject incoming jitter, plus the D to A's ability to reject any jitter coming into it. During the recording, there are 3 Phase Locked Loops in the chain: in the FCN-1, the recorder, and the D to A converter. Each PLL has its own characteristics; many good PLLs actually reduce incoming jitter; others have a high residual jitter. It is likely that during playback, you will hear far less jitter (better low level resolution, clearer highs) because there is only one PLL in the digital chain, between the playback deck and the D to A. In other words, *the playback will sound better than the sound monitored while recording!*

### **Jitter and A to D Converters**

The A to D Converter is one of the most critical digital audio components susceptible to jitter, particularly converters putting out long word lengths (e.g. 24-bits). The master clock that drives an A/D converter must be very stable. A jittery master clock in an A/D converter can cause irrevocable distortion and/or noise which cannot be cancelled out or eliminated at further stages in the chain. A/D's can run on internal or external sync. On internal sync, the A/D is running from a



master crystal oscillator. On external sync, the A/D's master clock is driven by a PLL, which is likely to have higher remnant jitter than the crystal clock. That is why I recommend running an A/D converter on internal clock wherever possible, unless you are synchronizing an A/D to video or to another A/D (in a multichannel setup). If you must use external sync, use the most stable external source possible (preferably video or wordclock over AES/EBU), and try to ensure that the A/D's designer used an ultra-stable PLL.

### **Jitter and DSP-based Processors**

Most DSP-based software acts as a "state machine." In other words, the output result on a sample by sample basis is entirely predictable based on a table of values of the incoming samples. The regularity (or irregularity) of the incoming clock has no effect on the output data.

Exceptions to "state-based" DSP processes include Asynchronous Sample Rate Converters, which are able to follow variations in incoming sample rate, and produce a new outgoing sample rate. Such devices are not "state-machines", and jitter on the input may affect the value of the data on the output. I can imagine other DSP processes that use "time" as a variable, but these are so rare that most normal DSP processes (gain changing, equalization, limiting, compression, etcetera) can be considered entirely to be state machines.

Therefore, as far as the integrity of the data is concerned, I have no problems using a chain of jittery (or non-jittery) digital devices to process digital audio, as long as the digital device has a high integrity of DSP coding (passes the "audio transparency" test).

### **Why are plug-in computer cards so jittery? Does this affect my work with the cards?**

Many computer-based digital audio cards have quite high jitter, which makes listening through them a variable experience. It is very difficult to design a computer-based card with a clean clock-- due to ground and power contamination and the proximity of other clocks on the computer's motherboard. The listener may leap to a conclusion that a certain DSP-based processor reduces soundstage width and depth, low level resolution, and other symptoms, when in reality the problem is related to a jittery phase-locked loop in the processor input, not to the DSP process itself. Therefore, always make delicate sonic judgments of DSP processors under low jitter conditions, which means placing high-quality jitter reduction units throughout the signal chain, particularly in front of (and within) the D/A converter. Sonic Solutions's USP system has very low jitter because its clocks are created in isolated and well-designed external I/O boxes.

### **Jitter and Digital Copies**

#### **The key is in the playback, not in the transfer**

Many well-known devices have high jitter on their outputs, especially DAT machines. *However*, for most digital to digital transfers, jitter is most likely irrelevant to the final result. I said "most likely" because a good scientist always leaves a little room for doubt in the face of empirical (listening) evidence, and I have discovered certain audible exceptions (see below). Until we are able to

measure jitter with widely-available high-resolution measuring equipment, and until we can correlate jitter measurements adequately against sonic results, I will leave some room for doubt.

*Playback from a DAT recorder usually sounds better than the recording, because there is less jitter. Remember, a DAT machine on playback puts out numbers from an internal RAM buffer memory, locked to its internal crystal clock. A DAT machine that is recording (from its digital input) is locked to the source via its (relatively jittery) Phase Locked Loop. As the figure above illustrates, the numbers still get recorded correctly on tape, although their timebase was jittery while going in. Nevertheless, on playback, that time base error becomes irrelevant, for the numbers are relocked by the DAT machine! I have not seen evidence that jitter is cumulative on multiple digital dubs. In fact, a Compact Disc made from a DAT master usually sounds better than the DAT... because a CD usually plays back more stably than a DAT machine. The fact that a dub can sound better than the original is certainly a tough concept to believe, but it is one key to understanding the strange phenomenon called Digital Audio.*

It's unnerving to hear a dub that sounds different from the original, so I've performed some tests to try to see if jitter is accumulated. I think I've proved with reasonable satisfaction, that under most conditions jitter is **not** accumulated on multiple dubs, and that passing jittery sources through a storage medium (such as hard disk) results in a very non-jittery result (e.g., recorded CDR).

Here are two tests I have made (this is far from a complete list):

### **Test #1**

I produced a 99th-generation versus 1st-generation audio test on Chesky Records' first Test CD. If jitter were accumulated on subsequent dubs, then the 99th generation would sound pretty bad, right? Well, most people listening to this CD can't tell the difference and there is room for doubt that there is a difference. It's pretty hard to refute a 99th generation listening test!

### **Test #2**

I built a custom clock generator and put it in a DAT machine. On purpose, I increased the jitter of that clock generator to the point that a dubbing DAT machine almost could not lock to the signal from the jittery source DAT. The sound coming out of the D/A converter of the dubbing DAT was entirely distorted, completely unlistenable. However, when played back, the dub had no audible distortion at all!

These are two scientifically-created proofs of an already well-understood digital "axiom," that the process of loading and storing digital data onto a storage medium effectively (or virtually) cancels the audible jitter coming in.

**Does copying to hard disk deteriorate the sound of the source?**

If you copy from a jittery source to a hard disk-recorder and later create a CDR from that hard disk, will this result in a jittery CDR? I cannot reach this conclusion based on personal listening experience. In most cases, the final CDR sounds better than the source, as auditioned direct off the hard disk! I must admit it is frustrating to listen to "degraded" sources and not really know how it is going to sound until you play back the final CDR.

Please note that I perform all my listening tests at Digital Domain through the same D/A converter, and that converter is preceded by an extremely powerful jitter-reduction device. Surprisingly, I can still hear some variation in source quality, depending on whether I am listening to hard disk, CDR, 20-bit tape, or DAT. The ear is an incredibly powerful "jitter detector"!

### Quiz

Is it all right to make a digital chain of two or more DAT machines in record? The answer: During record you may hear a subtle loss of resolution due to increased jitter. However, the cumulative jitter in the chain will be reduced on playback. But we advise against chaining machines; it is safer to use a distribution amplifier (like the FCN-1) to feed multiple machines, because if one machine or a cable fails, the failure will not be passed on to another machine in line.

### Can Compact Discs contain jitter?

When I started in this business, I was skeptical that there could be sonic differences between CDs that demonstrably contained the same data. But over time, I have learned to hear the subtle (but important) sonic differences between jittery (and less jittery) CDs. What started me on this quest was that CD pressings often sounded deteriorated (soundstage width, depth, resolution, purity of tone, other symptoms) compared to the CDR master from which they were made. Clients were coming to me, musicians with systems ranging from \$1000 to \$50,000, complaining about sonic differences that by traditional scientific theory should not exist. But the closer you look at the phenomenon of jitter, the more you realize that even minute amounts of jitter are audible, even through the FIFO (First in, First Out) buffer built into every CD player.

CDRs recorded on different types of machines sound different to my ears. An AES-EBU (stand-alone) CD recorder produces inferior-sounding CDs compared to a SCSI-based (computer) CD recorder. This is understandable when you realize that a SCSI-based recorder uses a crystal oscillator master clock. Whenever its buffer gets low, this type of recorder requests data on the SCSI buss from the source computer and thus is not dependent on the stability of the computer's clock. In contrast, a stand-alone CD recorder works exactly like a DAT machine; it slaves its master clock to the jittery incoming clock imbedded in the AES/EBU signal. No matter how effective the recorder's PLL at removing incoming jitter, it can never be as effective as a well-designed crystal clock.

I've also observed that a 4X-speed SCSI-based CDR copy sounds inferior to a double-speed copy and yet again inferior to a 1X speed copy.

Does a CD copy made from a jittery source sound inferior to one made from a clean source? I don't think so; I think the quality of the copy is solely dependent on clocking and mechanics involved during the transfer. Further research should be done on this question.

David Smith (of Sony Music) was the first to point out to me that power supply design is very important to jitter in a CD player, a CD recorder, or a glass mastering machine. Although the FIFO is supposed to eliminate all the jitter coming in, it doesn't seem to be doing an adequate job. One theory put forth by David is that the crystal oscillator at the output of the FIFO is powered by the same power supply that powers the input of the FIFO. Thus, the variations in loading at the input to the FIFO are microcosmically transmitted to the output of the FIFO through the power supply. Considering the minute amounts of jitter that are detectable by the ear, it is very difficult to design a power supply/grounding system that effectively blocks jitter from critical components. Crystal oscillators and phase locked loops should be powered from independent supplies, perhaps even battery supplies. A lot of research is left to be done; one of the difficulties is finding measurement instruments capable of quantifying very low amounts of jitter. Until we are able to correlate jitter measurements against audibility, the ear remains the final judge. Yet another obstacle to good "anti-jitter" engineering design is engineers who don't (or won't) listen. The proof is there before your ears!

David Smith also discovered that inserting a reclocking device during glass mastering definitely improves the sound of the CD pressing. Correlary question: If you use a good reclocking device on the final transfer to **Glass Master**, does this cancel out any jitter of previous source or source(s) that were used in the pre-production of the premaster? Answer: We're not sure yet!

### Listening tests

I have participated in a number of blind (and double-blind) listening tests that clearly indicate that a CD which is pressed from a "jittery" source sounds worse than one made from a less jittery source. In one test, a CD plant pressed a number of test CDs, simply marked "A" or "B". No one outside of the plant knew which was "A" and which "B." All listeners preferred the pressing marked "A," as closer to the master, and sonically superior to "B." Not to prolong the suspense, disc "A" was glass mastered from PCM-1630, disc "B" from a CDR.

### Attention CD Plants--a New Solution to the Jitter Problem from Sony

In response to pressure from its musical clients, and recognizing that jitter really is a problem, Sony Corporation has decided to improve on the quality of glass mastering. The result is a new system called (appropriately) *The Ultimate Cutter*. The system can be retrofitted to any CD plant's Glass Mastering system for approximately \$100,000. *The Ultimate Cutter* contains 2 gigabytes of flash RAM, and a very stable clock. It is designed to eliminate the multiple interfering clocks and mechanical irregularities of traditional systems using 1630, Exabyte, or CD ROM sources. First the data is transferred to the cutter's RAM from the CD Master; then all interfering sources may be

shut down, and a glass master cut with the stable clock directly from RAM. This system is currently under test, and I look forward to hearing the sonic results.

### Can Jitter in a Chain be Erased or Reduced?

The answer, thankfully, is "yes.". Several of the advanced D to A converters now available to consumers contain jitter reduction circuits. Some of them use a frequency-controlled crystal oscillator to average the moment to moment variations in the source. In essence, the clock driving the D/A becomes a stable crystal, immune to the pico- or nano-second time-base variations of jittery sources. This is especially important to professionals, who have to evaluate the digital audio during recording, perhaps at the end of a chain of several Phase Locked Loops. Someday all D to A converters will incorporate very effective jitter-reduction circuits.

### Good Jitter vs. Bad Jitter

The amount of jitter is defined by how far the time is drifting. Original estimates of acceptable jitter in A/D and D/A converters were around 100 to 200 picoseconds (pS). However, research into oversampling converters revealed that jitter below 10 pS is highly desirable. For D/A converters, the amount of jitter is actually less important than the type of jitter, for some types of jitter are audibly more benign than others (I repeat: **jitter does not affect D-D dubs, it only affects the D to A converter in the listening chain**).

There are three different "types" of jitter:

1. The variations in the time base which are defined as jitter are regular and periodic (possibly sinusoidal)
2. The variations are random (incoherent, white noise)
3. The variations are related to the digital audio signal

Jitter can also be a combination of the above three.

Periodic fluctuations in the time base (#1 above) can cause spurious tones to appear at low levels, blocking our ability to hear critical ambient decay and thus truncating the dynamic range of the reproduction. Often this type of jitter is caused by clock leakage. It is analogous to scrape flutter in analog recorders.

On the other hand, *Gaussian*, or random jitter (#2 above, usually caused by a well-behaved Phase Locked Loop wandering randomly around the nominal clock frequency) is the least audible type. In addition to adding some additional noise at high frequencies, gaussian jitter adds a small perfume of hiss at the lowest levels, which may or may not be audible, and may or may not mask low level musical material. Sometimes, this type of jitter puts a "veil" on the sound. This veiling is not permanent (unlike the effects of dither, which are generally permanent), and will go away with a proper relocking circuit into the D/A converter.

Finally, timing variations related to the digital audio signal (#3 above) add a kind of intermodulation distortion that can sound quite ugly.

### More to Come

Jitter bibliography and credits. Clarifications of some apparent contradictions in the above essay. While you're waiting for "The Jitter Bible," I urge you to listen, listen, listen, and see if you hear the problems of jitter in your audio systems, where and when they seem to occur.

---

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

[< Pre](#)      [Next >](#)

[\[ Back \]](#)

### Popular

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)

- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us




[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## Subwoofers



### Accurately Set Up a Subwoofer With (Almost) No Test Instruments

Bass frequencies are extremely important to sound reproduction. Everyone is interested in getting their bass right, but most people haven't a clue how to proceed. This article will help to settle the process of integrating an active subwoofer with an existing "satellite" system. If your room and loudspeakers are good, you'll only need two test CDs and your ears to adjust your subwoofer. If your room is not so good, or you want to refine the sound even further, then we'll discuss the best way to integrate test equipment measurements with your hearing. The simple listening test will also reveal if your room has problems and if it's time to hire an acoustician.

Let's review the basic requirements for smooth, extended bass response.

### Conquering the Room

Many people are proud of the "ideal dimensions" of their listening room. In general, the larger the room, the fewer audible problems with low frequency standing waves (nodes and antinodes). To get smooth and even bass requires ceilings taller than 10 feet, width greater than 12 feet, and length greater than 25 feet (30 or more for deep bass). Dimensions (including diagonals) should not be multiples of identical wavelengths, to avoid buildup at octave resonances. Of course, larger rooms may need absorption to keep the reverberation time down, but standing waves don't tend to build up awkwardly in larger rooms. It's also important to use absorption so that the decay time at low frequencies is roughly similar to that at mid and high frequencies. This is called a "neutral room."

Lightweight, flexible walls act as *diaphragmatic absorbers*, where some bass frequencies will escape out the walls, never to return. In my opinion, the ideal is a solid concrete (block) wall, but proper construction with plaster lathe, wood, and/or double sheet rock can accomplish similar results. But solid walls create problems of their own; a world-class room usually requires some absorption and/

OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

Password

Remember me

or diffusion to deal with resonances and echos. Watch out for cavities within the walls, which can cause resonances. Creating a large room with good bass response, interior acoustics, and outside isolation, is the role of a professional acoustician. This article will share some secrets in the fine tweaking of systems in good rooms; don't dream of building a room from scratch without hiring an acoustician.

### Speaker Mounting - Spikes or Isolators?

Soffit-mounting involves recessing loudspeakers into a cavity in the wall, with the edge of the loudspeaker flush to the wall. Soffit-mounting requires the expertise of an experienced acoustician, and is beyond the scope of this article. The main loudspeakers must be decoupled from the floor. Heavy, rigid stands should have a top no larger than the bottom of the loudspeaker to avoid diffraction (a form of comb filtering). I've had great success spiking speaker stands (using spikes, or "tiptoes") through holes in the carpet. Some authorities recommend a damping pad underneath a heavy, full range speaker instead of spikes. Whichever mounting method, the goal is to reduce sympathetic vibrations or traveling waves in cabinets, floor and walls. The resonant frequency of the box and stand should be extremely low. Hit the box with your fist and confirm it does not have a resonant character; sweep a sine wave through the system and listen for vibrations. I've had great success with a very **thin isolator** (Dr. Scholl's) between the speaker and the stand which compresses almost completely under the speaker's weight.

### Listener position

If you're sitting in an antinode, there's always going to be a dip at that frequency, and no amount of equalization will correct the acoustic problem.

### Speaker position

Ironically, solid walls aggravate the interaction of loudspeaker position and frequency response. The closer the loudspeaker to walls and especially corners, the greater the bass level. You may have the "smoothest," most accurate satellite (main) speakers in the world, but they must be positioned to avoid side wall reflections and must be far enough from all walls to reduce resonances.

### Near Field Monitoring?

I wouldn't master with near-field monitors, but I will mix with them. Near-field monitoring was devised to reduce the effects of adverse room acoustics, but if your room acoustics are good, then "Mid-field" or "Far-field" will provide a more accurate depth and spatial picture. There must be an obstruction-free path between the monitors and the listener. **What is the biggest impediment to good sound reproduction in a recording studio? The console.** No matter how you position the monitors, the console's surface reflects sound back to your ears, which causes *comb filtering*, the same tunnel effect you get if you put your hand in front of your face and talk into it. Or if you wear a wide-brimmed hat, which produces an irregular dip around 2 kHz. It amazes me that some engineers aren't aware of the deterioration caused by a simple hat brim! Similarly, I shudder when I see a professional loudspeaker sitting on a shelf inches back from the edge, which compromises

the reproduction. The acoustic compromise of the console can only be minimized, not eliminated, by positioning the loudspeakers and console to increase the ratio of the direct to reflected path. Lou Burroughs' 3 to 1 rule can be applied to acoustic reflections as well as microphones, meaning that the reflected path to the ear should ideally be at least 3 times the distance of the direct path.

### **What about measurements?**

Can't we just measure, adjust the crossovers and speaker position for flattest response, then sit down and enjoy? Well, since no room or loudspeaker is perfect, measurements are open to interpretation, and frequency response measurements will always be full of peaks and dips, some of which are more important to the ear than others. Which of those many peaks and dips in the display are important and which ones should we ignore?

I've found the ear to be the best judge of what's important, especially in the bass region. The ear will detect there's a bass problem faster than any measurement instrument. The measurement instrument will help to pinpoint the specific problem frequencies, whether they're peaks or dips, and by supplying numbers, aid in making changes. The whole process is very frustrating, and it's inspired my search for setup and test methods that use the ear. A perfect setup still requires a multistep process: listen, measure, adjust, listen again, and repeat until satisfied, but it's possible to streamline that process. Here's a listening test for adjusting subwoofer crossovers that uses simple, readily obtainable and cheap test materials, and that's generally as precise as most more formal measurement techniques! If you're setting up a permanent system, dedicate a day to the process; even the easy doesn't come easy. Some brands of subwoofer amplifiers have all the controls or connectors you need; you may have to adapt the process described below to your particular woofer system.

### **Polarity is not Phase**

This is still a confusing topic, perhaps because people are too timid to say *polarity* when they mean it. The *polarity* of a loudspeaker refers to whether the driver moves outward or inward with positive-going signal, and can be corrected by a simple wire reversal. Remember that *phase* means *relative time*; phase shift is actually a time delay. The so-called *phase* switches on consoles are actually *polarity* switches, they have no effect on the time of the signal! Sometimes this is referred to as *absolute phase*, but I recommend avoiding the use of the term *phase* when you really mean *polarity*. If two loudspeakers are working together, their polarity must be the same. If they are separated by space, or if a crossover is involved, there may be a *phase* difference between them, measured in time or degrees (at a specific frequency). I have a pair of *Genesis* subwoofers with separate servo amplifiers. There are three controls on the crossover/amplifier: volume (gain), phase (from 0 to 180 degrees), and low pass crossover frequency (from 35 Hz to over 200 Hz). Notice there is no high pass adjustment. The *natural* approach to subwoofer nirvana assumes that your (small) *satellite* loudspeakers have clean, smooth response down to some bass frequency, and gradually roll off below that. It's logical to use the natural bass rolloff of the satellites as the high pass portion of the system and to avoid adding additional electronics that will affect the delicate midrange frequencies. So we use a combination of lowpass crossover

adjustment and subwoofer positioning to fine-tune the system.

A good subwoofer crossover/amplifier usually provides more than one method of interconnection with the satellite system. The best is the one which has the least effect on the sound of the critical main system. I prefer not to interfere with the line level connections to the (main) power amp feeding the satellites. If your preamplifier does not have a spare pair of buffered outputs, I recommend using the speaker-level outputs of the main power amp. The *Genesis* provides high-impedance transformer-coupled balanced inputs on banana connectors designed to accept speaker-level signals. Connect the main power amp's output to the sub amp's input with simple zip cord with bananas on each end. No real current is being drawn, so wire gauge does not have to be heavy. Double-bananas make it easy to reverse the polarity of the subwoofer, a critical part of the test procedure. Some subwoofers use a 12 dB/octave crossover, others 18 or more. Interestingly, for reasons we will not discuss here, a 12 dB crossover slope requires woofers that are wired out of polarity with the main system. My new sub crossover uses a 24 dB lowpass slope, which also requires polarity reversal, but to make it easy on the mindless, the internal connections are reversed, and you're supposed to connect "hot to hot" between the main power amplifier and woofer amplifier. Leave nothing to doubt—we must confirm the correct polarity. Steep slopes like 18 and 24 dB are good choices to get the subwoofer to roll off before it interferes with the midrange response.

You have to sit in the "sweet spot" for the listening evaluation. If your subwoofers have an integrated amplifier, you'll need a cooperative friend to make adjustments. Since the *Genesis* amplifier is physically separate, I was able to move the subwoofer amplifiers to the floor in front of the sweet spot, and make my own adjustments. Here are the two test CDs:

1. The Mix Reference Disc, Deluxe Edition, MRD2A. Since this disc is now out of print, Mix Editor George Petersen has kindly given me permission to put the test tones up on our site so you can make your own custom **Subwoofer test CD**. You can also print a traycard from the PDF file I've provided. You can also use any source of 1/3 octave filtered pink noise.
2. Rebecca Pidgeon, *The Raven*, Chesky JD115, available at record stores, high-end stereo stores, or from **Chesky Records**.

I recorded Rebecca's disc in 1994. Track 12 is *Spanish Harlem*, which has a slow, deliberate acoustic bass part that makes it easy to identify notes that "stick out" too far and covers the major portion of the bass spectrum. This record has never failed to reveal the anomalies of different rooms and loudspeakers in several years of use as a musical reference. The ear is better with instant comparisons than absolute judgments, and this test relies on our ear's ability to make comparisons. All musical instruments and transducers produce harmonics as well as fundamentals. To the best of our ability to discriminate, we will be concentrating on the fundamental tones in this piece of music. If your loudspeakers have significant harmonic distortion, they can complicate or confuse the test. Many studio loudspeakers are designed for high power handling at the expense

of tonal accuracy or distortion. This test is not for them. If you want accurate bass, it's time to replace the loudspeakers and probably hire an acoustician with a distortion analyzer.

Start by evaluating the satellite system with the subs turned off. Listen to the bass at a moderate level equal to or slightly louder than the natural level of an acoustic bass. Listen for harmonic distortion: if it doesn't sound like a "transparent" acoustic bass, fix the problem with the satellites, first. Listen for uneven notes. If the lower note(s) of the scale are successively softer in level than the higher notes, then you have a perfect candidate for a subwoofer. If intermediate bass notes are weak or strong (uneven bass), the satellite loudspeakers may be too close to the corners, in a node or antinode, the listening position may be in a standing wave, or the satellites themselves poorly designed. It may be time to bring in an acoustician. But if the satellite bass is even, you can move on to the next step, adjusting the subwoofers.

*Spanish Harlem*, in the key of G, uses the classic 1, 4, 5 progression. Here are the frequencies of the fundamental notes of the bass. If your loudspeaker has sufficiently low harmonic distortion, it will not affect your judgment of the power of the bass notes, which are already affected by the natural harmonics of the instrument.

					49	62	73	85
82	98		73	93	110			

As you can see, this covers most of the critical bass range. If the lowest note(s) is weaker than the rest, then you are a candidate for a subwoofer. My satellites behave in the classic manner, with the lowest note (G, 49Hz) slightly low in level, but the rest fall in a balanced line. I've been in small rooms where one or more of the intermediate notes are emphasized or weak, which suggests standing wave problems. Repositioning the satellites may help. Avoid equalization, which is a nasty band aid...proper acoustic room treatment is the cure. You could conceivably add a subwoofer out of phase at the frequencies in question, but that's a technique that should remain confidential between you and your analyst. Fix the acoustic problems first and you'll be happier.

If your satellite system passed the initial examination, next step is to decide on a starting (approximate) subwoofer location. A satellite-subwoofer system has tremendous flexibility, offering in theory the best of two worlds. The satellites can be placed on rigid stands at ear level, far from corners and side walls, reducing floor and wall reflections and comb-filtering in the midband. And the subs can be placed on the floor, in the position that gives the most satisfactory bass response, integrated with the satellites. If you only have one (mono) subwoofer, start by placing it in the middle between the stereo speakers. Contrary to popular belief, stereo subwoofers are important, they can improve the sense of "envelopment", the concert hall realism that bass waves are passing by you. Authorities are split on the issue whether a mono or stereo subwoofer setup is more forgiving of room modes. I prefer the sound of stereo subwoofers. A complete discussion of how to place the satellites would require another article, but let's start by saying that



you may have to deal with reflections from the side walls by placing absorbers in critical locations. Consider consulting a competent acoustician.

Assuming your satellite system passes the listening test, it's time to find the right crossover frequency, phase and woofer amplitude that will just supplement the lower notes of the scale. Start by placing the subwoofers next to and slightly in front of the satellites. First we must determine the proper polarity for the subwoofers. If your system uses XLR input connectors, build a polarity reversing adapter for this part of the test. This is easier with only one channel playing. Put on the **MRC**D with full bandwidth pink noise, at a moderate level (70-80 dB SPL). Adjust the crossover to its highest frequency, the phase to 0, and turn up the subwoofer gain until you're sure you can hear the woofer's contribution. Reverse the polarity of the sub. The polarity which produces the loudest bass is the *correct polarity*. Mark it on the plugs, and don't forget it!

Next comes an iterative process ("lather, rinse, repeat until clean"). Here's a summary of the four-steps: (1, 2, & 3) Using filtered pink noise, we'll determine the precise phase, amplitude and crossover dial position for *any one crossover frequency*. (4) Then we'll put Rebecca back on and see if all the bass notes now sound equally loud. If not equally loud, then we'll go back to the filtered pink noise and try a different crossover frequency. We keep repeating this test sequence until the bottom note(s) has been made "even" without affecting the others. With practice you can do this in less than half an hour. Adjust each subwoofer individually, playing one channel at a time. And now in detail:

### 1) Crossover frequency (lowpass)

Play filtered pink noise (or the Mix CD's multifrequencies) at your best guess of crossover frequency, say 63 or 80 Hz. Notice that the signal has a *pitch center, or dominant pitch quality*. If the subwoofer is misadjusted, adding the sub to the satellites will slide the pitch center of the satellite's signal. Reverse the sub's polarity (set it to *incorrect polarity*). With the sub gain at a medium level, start at the lowest frequency, and raise the frequency until you hear the dominant pitch begin to rise (literally, the center "note" of the pink noise appears to go sharp, to use musical terms). Back it off slightly (to a point just below where the pitch is affected), and you have correctly set the crossover to this frequency. Recheck your setting. That's it.

### 2) Phase

The sub should always be on a line with or slightly in front of the satellite. With the woofer a moderate amount in front of the satellites, the phase will generally need to be set something greater than 0 degrees. Return the sub(s) to *the correct polarity*. Play the same frequency of filtered noise and increase the amount of "phase" until you hear the dominant pitch rise. Back it off slightly, recheck your setting, and that's it.

### 3) Amplitude

The subwoofer's settings are exactly correct when its amplitude is identical to the satellite's at the crossover frequency. The subwoofer gain is the easiest to get right because there will be a clear center point, just like focusing a camera. Play the filtered noise, and discover that the pitch is only



correct at a certain gain, above which the pitch goes up (sharp), and slightly below which it goes down (flat). "Focus" the gain for the center pitch, which will match the pitch of the satellites without the sub. Recheck your work by disconnecting and reconnecting the sub. The pitch should not change when you reconnect the sub, otherwise the gain is wrong. To be extremely precise, increase the gain in tiny increments until you find the point where the pitch rises when the sub is connected, then back the gain off by the last increment. This process is extremely sensitive.

#### **4) Rebecca**

Play *Spanish Harlem* again. If all the levels of the bass notes are even, you're finished with steps 1-4. If you hear a rise in level below some low note, then the crossover frequency is too high and vice versa. Do not attempt to fix the problem with the subwoofer gain, because that has been calibrated by this procedure, which leaves nothing in doubt except the choice of crossover frequency. Go back to step one and try again. Once all the notes are even, your crossover is perfectly adjusted. Write that frequency down. Then, for complete confidence, check the nearest frequency above and below (go back through steps 1-4), proving you made the right choice. This piece of test music is sufficiently useful that there will be a clear difference between each 1/3 octave frequency choice and it will be comparatively easy to determine the winner. The trick is not to rely on our faulty acoustic memory, but on the ear's ability to make relative comparisons.

#### **More Refinement**

##### **Fine tuning the stereo separation (space between the woofers)**

If you have stereo subwoofers, their left-right separation must be adjusted. Play *Spanish Harlem*. Listen to the sound of the bass with the subs off. It should be perfectly centered as a phantom image and its apparent distance from the listener should subtend a line between the satellites. If it is not perfectly centered or its image is vague, the satellites are too far apart. Now add the subwoofers. The bass should not move forward or backward, and its image should not get wider or vaguer. Adjust the physical separation of the subwoofers until the bass image width is not disturbed when they are turned on. This "integrates" the system. Go back to step one, recheck the amplitude and phase settings for the new woofer position. Everything is now spot on.

Congratulations, you've just aligned a world-class reproduction system! A subwoofer should not call attention to itself, either by location or amplitude. When you play music, the combination of the sub and mains will sound like a single, seamless source.

Now, after logging your settings, sit back, listen and enjoy. You've earned the time off. Don't let anyone touch those hard-earned adjustments, for you can be confident that they are about as good as they're going to get. Play several of your favorite recordings, and listen to the bass. The bass on the best recordings will be acceptable on your reference system; the worst recordings will have too much or too little bass. Now you can be reasonably sure the problem is in the recording, not your room or woofers. What a nice feeling!

#### **How The Pitch Detection Method Works**

The 1/3 octave pink noise signal (or the multitone test signal) contains a narrow band of

frequencies, whose dominant level is at the center of the band. Thus, you perceive a "pitch" to the signal. When you add a second loudspeaker driver (the subwoofer) driven by the same signal, if the woofer's output does not exactly match the level and distribution of frequencies produced by the main loudspeaker, there will be a shift in the dominance of the multifrequencies, either towards the high end of the band or the low end, perceived as a pitch shift. When the two signals are well-matched in level, frequency distribution and phase, you will hear a delicate increase in level, but no change in pitch. By simple comparative listening, taking the woofer in and out of the circuit, you have confirmed that your drivers are matched at the crossover frequency, and that the wavefronts of your main speakers and subs are aligned at the critical crossover frequency.

Of course, we're making certain assumptions...that:

- your satellite system is well designed, linear and rolls off below some defined frequency.
- your subwoofer system is linear and rolls off above some defined frequency.
- the slopes of the two rolloffs are compatible and will integrate.

Your degree of success depends on how closely the two systems meet those requirements.

### **What To Do When the Results are Less Than Perfect**

When interpreting *Spanish Harlem*, don't get too hung up on little "dips" in level. Dips are less objectionable to the ear than peaks. First, attack problems with resonant notes and then look at the dips. Everything may not be rosy the first time around. Supposing that the subwoofer helped the bottom note(s), which means the crossover is at the right frequency, but some upper note in the progression has been affected. This means the subwoofer position is not optimized, or the subwoofer has some frequency response anomaly. As the sub is moved towards the room corners, the low bass response goes up, previous dips become peaks. There's cancellation/reinforcement between the subs and the satellites, which changes complexly as the sub is moved. Thus, adjusting the subwoofer position is a powerful method to even out the bass, but this type of trial and error is too complicated without test equipment. You could slide the woofer slightly, adjust the crossover as above, listen, move it again, readjust, and listen, but our acoustic memory is too short to tell when we've hit the perfect spot.

### **Advanced Techniques**

#### **Integrating the Instruments with the Ears**

Here's where it gets complicated. If you are having problems with uneven bass, we can no longer rely strictly on our ears. If you're comfortable with measurement instruments, then let's proceed. First, listen to Rebecca and mark down the problem frequency or frequencies, either peaks or dips. You'll use that knowledge when you bring in the big guns, the 1/3 octave analyser. The good thing is that Rebecca has already told you where the problems are, so you'll know how to separate the forest from the trees in the 1/3 octave display. I used *Spectrafoo* (an excellent analysis program for the Mac) in transfer function mode with wide band pink noise into both satellites and subs (one

channel at a time). *Spectrafoo* time aligns the stimulus and response, which helps to separate direct from reflected sound, more accurately representing what the ear hears. *Spectrafoo* revealed a rising response in my room below 40 Hz, and more important, a little dip in the combined response circa 63 Hz which corresponded with my perception that note was perhaps a little weak. By moving the sub around very slightly and watching the display, I was able to exchange the weakness against the surplus without aggravating any other peaks.

The strength of this method is we're continuously integrating our powerful (almost objective) listening judgments with the "over-powerful" analysis tool. We're using the analyser for general trends, not absolute amplitudes; that's what I mean by separating the forest from the trees. The position of the test microphone should be in the exact listening position. Wear earplugs to keep your ears fresh when you're not required to listen. After moving the woofer, don't forget to readjust the crossover gain and phase with our listening technique.

If all goes well, *Spanish Harlem* will be even better adjusted and we can rest assured that our system is **really really tweaked**. Now sit back and enjoy. Oops, your work is never done. Now that you've adjusted your system, I'll let you in on one more secret: Servo amplifiers have internal adjustments that affect woofer damping, make the bass "tighter" or "looser." but that's another story.

#### **Acknowledgments:**

Jon Marovskis of Janis Subwoofers introduced me to the concept of a pitch detection technique many years ago. This article refines and expands on his original idea.

Many thanks to Dave Moulton for insightful technical and editorial comments.

Also, thanks for manuscript review and suggestions by Johnson Knowles of the Russ Berger Design Group, Eric Bamberg, Greg Simmons and Steven W.Desper.

-----

Acoustician Johnson Knowles suggests a viscoelastic polymer pad material like EAR Isodamp C1002 or C1000. The internal damping characteristics of the viscoelastics are exceptionally effective as a speaker to stand interface material.

U.S. Consultant Steve Desper recommends STIK-TAK by Devcon Corporation, available at your local hardware store. It's a cheap solution and works well. Australian Greg Simmons has found a similar product--marketed as *Blue Tak*: "Use enough of it relative to the weight of your speakers. For a small monitor weighing just over 20kg, I used four balls about 15mm in diameter (one under each corner). With 20kg on top of them, these balls squashed down to about 4mm or 5mm thickness, and held the monitor very firmly."

---

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

<  
Prev

Next  
>

[ Back ]

Popular

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)

- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us

© 2007 Digital Domain - All Rights Reserved


[Home](#)
[About Us](#)
[Products](#)
[Services](#)
[Main Menu](#)
[Articles and](#)
[Home Demo's](#)
[Contact Us](#)
[About Us](#)
[Downloads](#)
[Delivery & FTP](#)
[Services](#)
[Products](#)
[Mastering Audio - The Book](#)
[K-Stereo](#)
[Articles and Demo's](#)
[Honor Roll](#)
[Artist's Page](#)
[Audio FAQ](#)
[News and Press](#)
[Links](#)
[Contact Us](#)
[Home](#)

## We Have Lift-off!!



### THE SOUND OF LIFTOFF!

Thanks to the help of a friend at NASA, I was able to record the liftoff of the Space Shuttle Discovery on March 8, 2001 from the Kennedy Space Center VIP viewing site at 3.1 miles from the launchpad (as close as they let anyone get during the launch except for crew members). Also with the courtesy and help of Gary Baldassari, Mike Morgan, and Andy DeGanahl, who supplied some of the equipment used. Andy and I braved the all-nighter and captured the launch at 6:42:09.059 am EST.

**10/20/06**

**We are very pleased to present to you a quick time video of the shuttle launch with 16 bit 48 K Apple Lossless Sound. PC Users: If the movie will not run when you click on it, then download Quicktime 7.** The music is provided by Orlando's superb [Sovereign Brass](#), whose album I mastered. Audio engineer Andy DeGanahl took the images with his trusty 8mm handcam, Andy, don't quit your day job! Maybe shortly I'll upload a 96 kHz/24 bit stereo audio only of the last 10 seconds of the launch. Write if you'd like me to do that.

#### Technical specs of the recording

Four microphones and two independent hard disc recorders at 24 bits, 96 kHz were used, which were sync'd up later to produce a fantastic surround recording. Two spaced omnis at 6 foot left and right distance were DPA 4041s, and on the same stands, "synchronous" Sennheiser MKH-30 figure 8's. When decoded via dual MS decoders to surround, the outdoor enthusiastic audience should subtend an angle from about 45 degrees left or right all the way around and behind the listener, with the NASA announcements to the right and behind you. The shuttle liftoff commands stage front center, but with doppler waves and echos throughout the front soundstage and distant echoes behind you. Playing this back in surround is a true "environmental experience."

OOPS. Your Flash player is missing or outdated. [Click here](#) to update your player so you can see this content.

**By logging in, you'll be able to download our mastering demo's. In order to retrieve a username and password, contact us [here!](#)**

Username

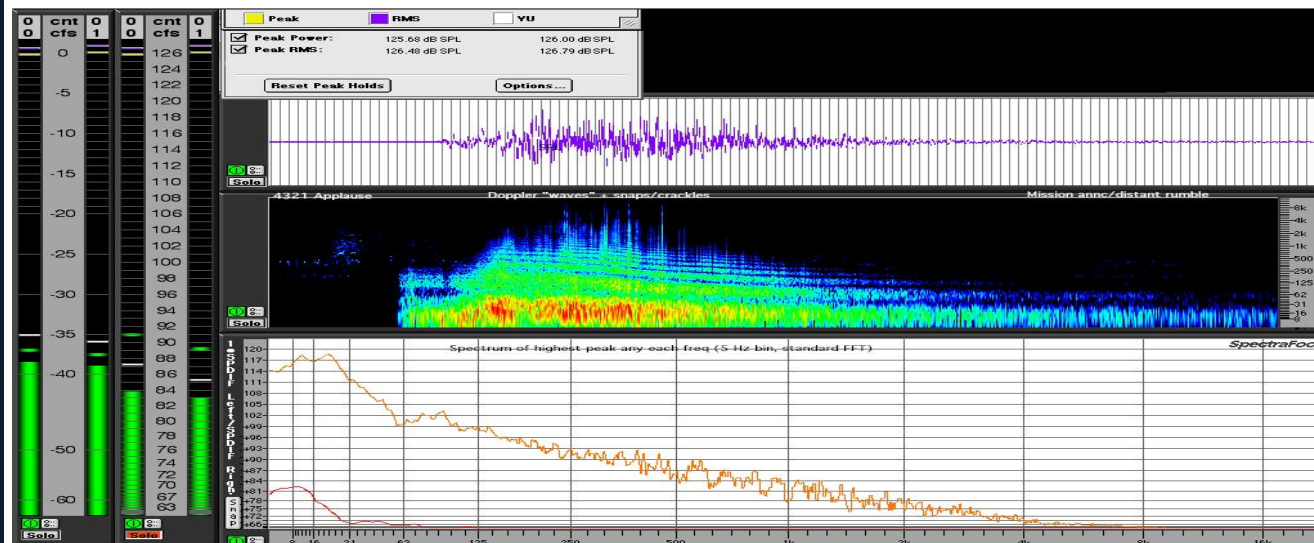
Password

Remember me

## On the Spectrafoo Audio Analysis

Through the magic of **Spectrafoo's** audio analysis tools, the audio "portrait" below demonstrates that there's nothing like being there. The spectrogram runs from T minus 4 seconds to about T plus 2 minutes. I don't think there's anything on earth that compares with the sound and sight of that fire-breathing monster on liftoff. If you study these incredible specs, including a spectrogramic timeline of the liftoff, you will see that to do justice to the experience, you will need a low-distortion subwoofer system capable of producing up to ~119 dB SPL on peaks at 25 Hz and ~116 dB SPL at 16 Hz and below! If not, then you will not be able to feel the chest-thumping, clean solid bottom that is produced. Ironically, the shuttle liftoff from the VIP site is "just loud enough" in person, a pleasant and not ear-damaging experience. Think of it as an 8.3 GWatt amplifier/loudspeaker with zero percent distortion and response down to DC! Running at say, 40% efficiency, that would take 20 thousand megawatts from the breaker box! Those figures are calculated by Dick Pierce from the comparable Saturn 5 moon rocket. These are the figures at 0 foot distance. Of course, some power has been dissipated at 3.1 miles, but examine the astonishing figures below.

By the way, the accompanying FFT illustrates that a point 1 (0.1) channel will serve well. Because in my standard stereo system, the woofers are properly calibrated, but the FFT shows there is far more peak energy below 100 Hz than I can achieve with a stereo system calibrated to Dolby standard level at 1 kHz. If I engineer a surround version of this, I will cross over the bass so that an ordinary system can allow bringing up the dialogue and mid frequency material to Dolby standard gain (as it stands, I can only reproduce this recording at levels about 20 dB below the actual measured acoustic levels without damaging my satellite speakers with too much low frequency information). But if I cross over the excess bass to a .1 channel with 10 dB more headroom, and then raise the gain of the recording, we should get a reasonable result with Dolby Standard monitor gain.





Thanks for reading!

Bob Katz

---

Copyright Digital Domain, Inc. We invite you to link to our site, which will be periodically revised.

<  
Prev

[ Back ]

Popular

- CD Mastering & Replication Services
- Mastering Audio - The Book
- Honor Roll
- K-Stereo
- Level Practices (Part 2) (Includes the K-System)

- Home
- About Us
- Products
- Services
- Delivery & FTP
- Articles and Demo's
- Contact Us

© 2007 Digital Domain - All Rights Reserved