

## 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 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

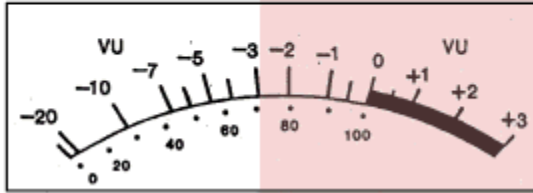
In General, the meter's ballistics, scale, and frequency response all contribute to an inaccurate indicator. The meter approximates momentary loudness changes in program material, but reports that moment-to-moment level differences are greater than the ear actually perceives.

#### 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

50% of Scale contains top 6 dB of range!



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.

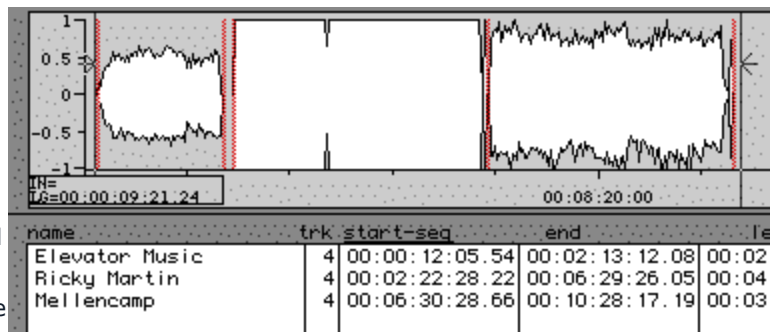
## Lack of conformance to standards

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

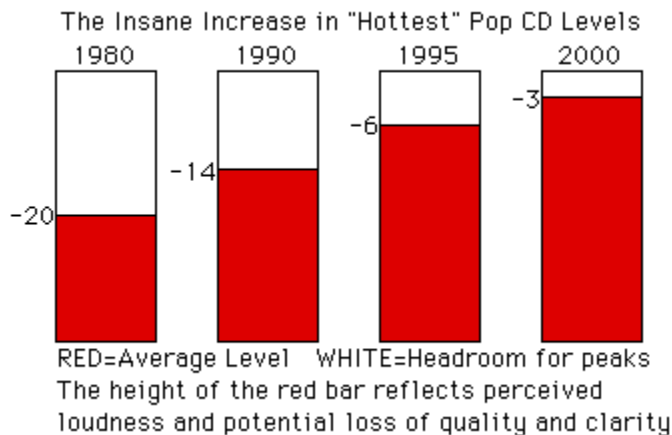
In the music and broadcast industries, chaos currently prevails. Here is a waveform taken from a digital audio workstation, showing three different styles of music recording. The time scale is about 10 minutes total, and the vertical scale is linear, +/- 1 at full digital level, 0.5 amplitude is 6 dB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

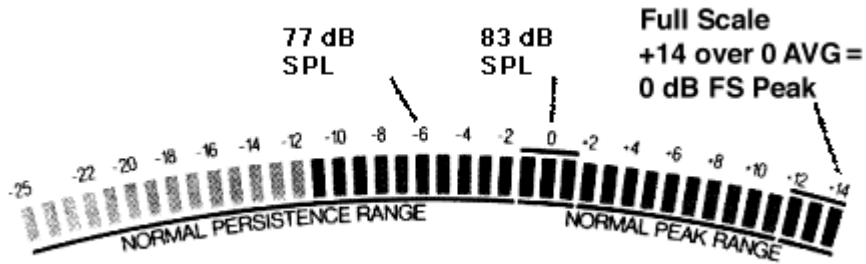
In addition, it should be more widely known **that hyper-compressed recordings do not play well on the radio.** They sound softer and seriously distorted, pointing out that the loudness race has no winners, even in radio airplay. The best way to make a "radio-ready" recording is not to squash it, but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

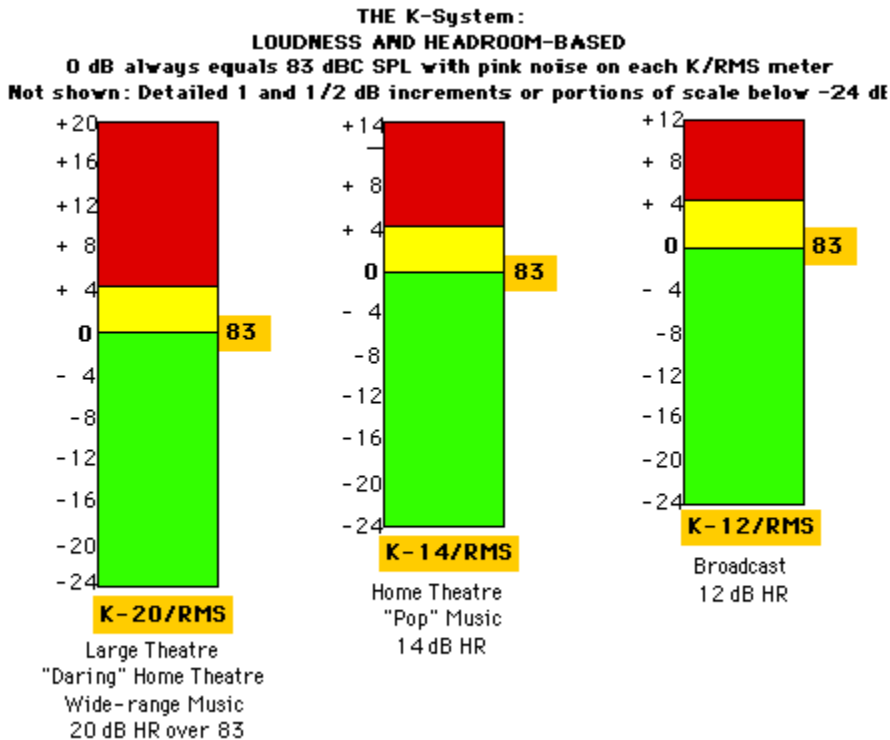
## VI. The K-System Proposal

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.

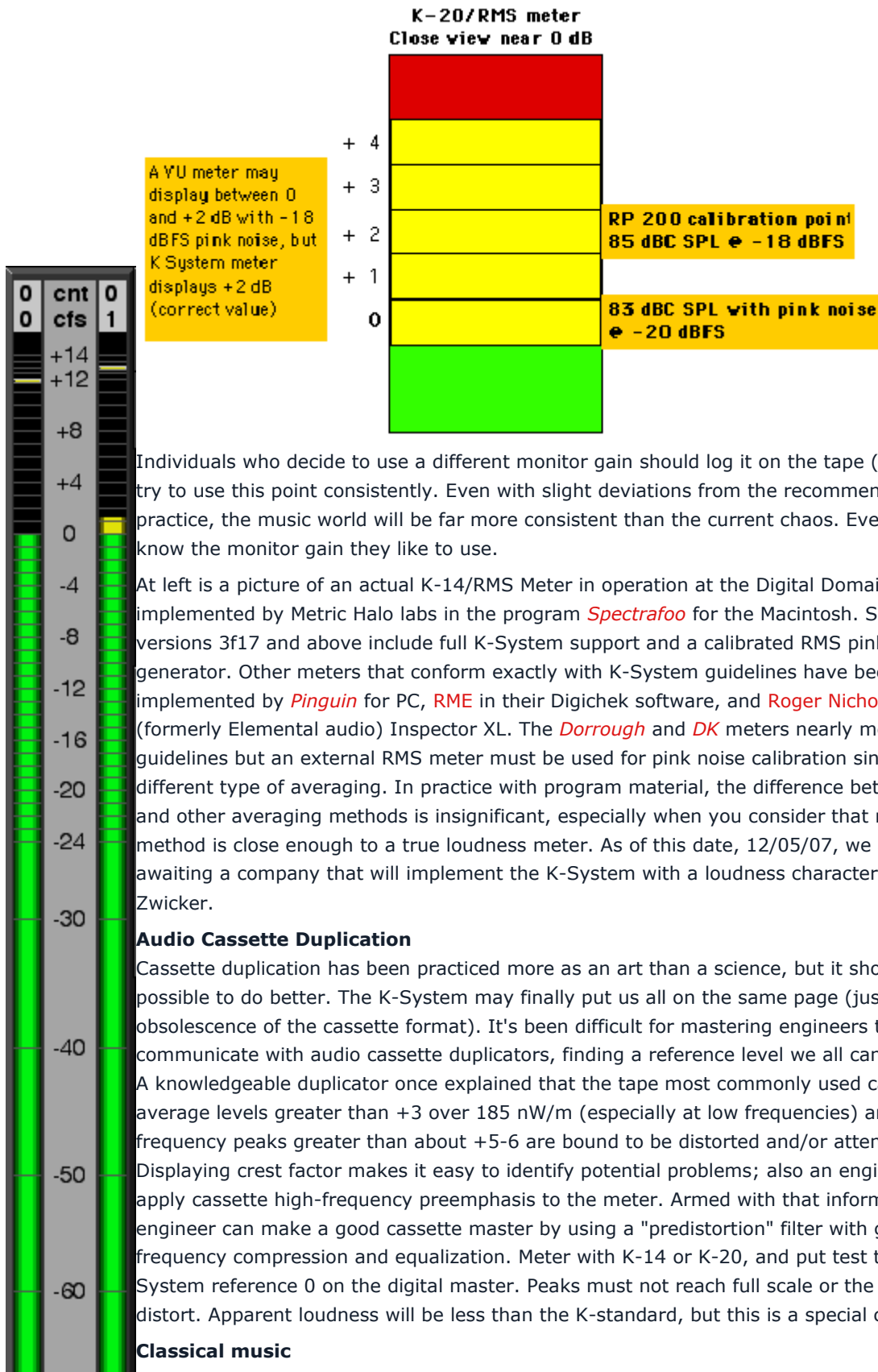
### **Multipurpose Control Rooms**

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.



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 Penguin 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.