Chapter 15

How To Make Better Recordings in the 21st Century

Part Two: The K-System, an Integrated Approach to Metering, Monitoring, and Leveling Practices

I. History: The VU Meter

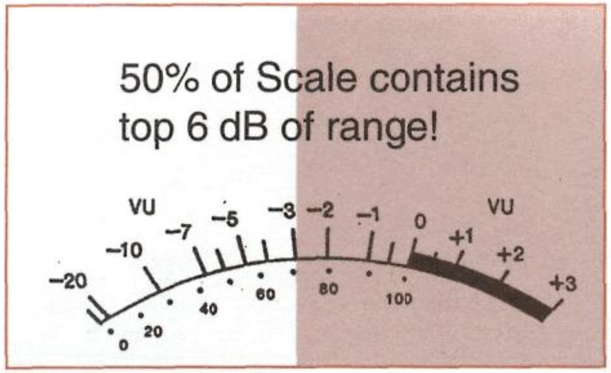
On May 1, 1999, the VU meter celebrated its 60th birthday. 60 years—but still widely misunderstood and misused. The VU meter has a carefully-specified time-dependent response to program material that I call *averaging* to simplify discussion, but really means the particular VU meter response. This instrument was intended to help program producers create consistent loudness amongst program elements, but as it was a poor indicator of recording overloads, the meter's designers depended on the 10 dB or greater *headroom* over 0 VU of 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 does give it a syllabic response, but does not make it accurate. One time constant cannot sum up the complex multiple time constants that make up 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 with regard to loudness change.

Scale: In 1939, logarithmic amplifiers were large and cumbersome to construct, and it was



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

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, as illustrated, 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. The extreme needle movements make it difficult to distinguish compressed from uncompressed material. Soft material may hardly move the meter, but be well within the acceptable limits for the medium and the intended listening environment.[5]

Frequency response: The meter's relatively flat frequency response results in meter deflections that are far greater than the perceived loudness change, since the ear's response is non-linear with respect to frequency. Frequency distribution and average level both affect loudness. 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 adherence to standards: In current use, there are large numbers of improperly-terminated mechanical VU meters and inexpensively-constructed indicators which are labeled "VU." I've seen fights break out amongst program producers reading different "VU" instruments. A true VU meter is a rather expensive device and it can't be called *VU* unless it meets the standard.

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

II. The Magic of *83* with Film Mixes

Unlike music CDs, films are consistent from one to another, because the monitoring gain has been standardized, as we learned in Chapter 14. 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.

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 high-resolution digital. 83 dB also lands on the most effective point on the Fletcher-Munson equal loudness curve, which is where the ear's frequency response is most linear. When digital technology reached the large theatre, the SMPTE attached the SPL calibration to a point 20 dB below full scale digital instead of 0 VU.* When we converted to digital technology, the VU meter was rapidly replaced by the peak program meter, which didn't faze the film world, but definitely caused the music industry to suffer, as we shall see.

III. United We Stand At Home

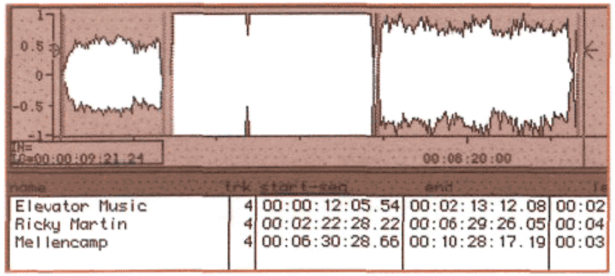
As we saw in Chapter 14, with the integration of media into a single system, it is in the direct interest of music producers to think holistically and unite with video and film producers for a more consistent consumer audio presentation. New program producers with little experience in audio production 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, recording engineer's experience is low, and the monitors they use to make program judgments are less than ideal. It is our responsibility to educate new engineers on how to make loudness and quality 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.

* See Appendix 9 for discussion on how "85" became "83".

Current-day leveling problems: The Loudness Race

The loudness race is not new; in the days of vinyl, mastering engineers competed to produce the loudest LP. But what is new is the fantastic magnitude of the problem: due to the nature of the digital medium, there is no longer the physical limit which was previously imposed by analog mechano-electrical systems and magnetic analog recording. Without that limit it is possible to produce CDs whose average level is almost the same as the peak level, an incredible 20 dB above the old average levels! Powerful digital compressors and limiters enable mastering engineers to produce a distorted signal for which there is no precedent in over 100 years of recording.[1] So, as we converted to digital technology, the result became chaos, yielding unprecedented differences in loudness between recordings.

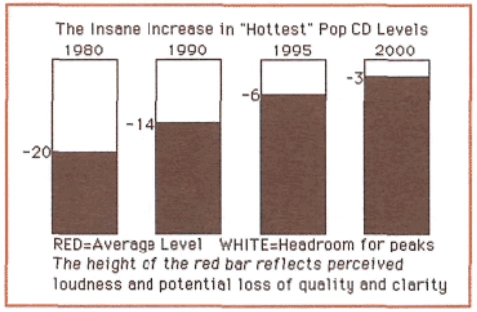
On the next page is a waveform taken from a digital audio workstation, showing three different styles of music recording. The time scale is about 10 minutes total, and the vertical scale is linear, +/- 1 at full digital level, 0.5 amplitude is 6 dB below full scale. The "density" of the waveform gives a rough approximation of the music's dynamic range and crest factor. 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 song from a popular compact disc produced in 1999. 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



On the left, moderately compressed "Elevator Music." In the Middle, a "top of the pops" selection from the year 1999. At right, a rock and roll record from 1990. Vertical and horizontal scales are identical.

greater than 6 dB, though both peak to full scale! Auditioning the 1999 CD, one mastering engineer remarked, "this CD is a light switch?! The music starts, all the meter lights 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? Why has the average sound quality of popular music CDs gone downhill since the introduction of the digital medium, and what can we do to fix the problem?[2]

The psychoacoustic problem is that when two **identical** programs are presented at slightly differing loudness, the louder of the two often appears "better," **but only 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 (illustrated below). And why there is a remarkable (and unnacceptable) 15 dB difference in average level among pop CDs! Remember that the loudness "race" has always been an artificial one, since the consumer adjusts their



Is this what will happen to the next generation carrier? (e.g. DVD-A, SACD). It will, if we don't take steps now to stop it.

volume control according to each record anyway. This uncontrolled situation is 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?

There are, of course, specific places where heavy compression is needed: background listening, parties, bar and jukebox playback, car stereos, headphone-wearing joggers, the loudspeakers at the record stores, headphone auditioning at the record store *kiosk*, and so on. In each of these cases, it should be possible to either produce a custom-compressed CD just for the purpose, or to install a compressor in the jukebox, CD changer, or reproduction system. Certainly this is a lot less damaging than compromising recorded music for all listeners. What we wish for is a low-fidelity replacement for the analog cassette. Ironically, the compact disc has become its own worst enemy, for it cannot be different things to different needs.[3] I dream of a perfect world where all the MP3 singles are heavily compressed and all the CD albums undamaged.

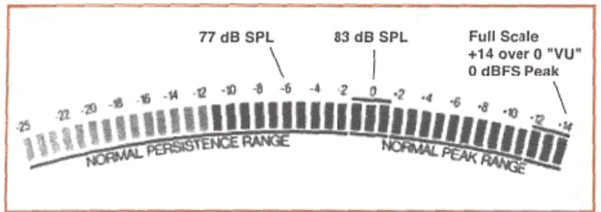
IV. The relationship between SPL and 0 VU

Around 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 many pop productions, our calibrated monitor control sits at -6 dB (which yields 77 dB SPL with -20 dBFS RMS pink noise).

Since on the meter, -20 dBFS reads -6 AVG, then 6 dB higher, or 0 AVG must be 83 dB SPL. This means we're really running average SPLs similar to the theatre standard (our sound quality is not as



The Dorrough Meter. With the monitor control's position set to 6 dB below the film reference, 77 dB SPL lands at -20 dBFS, or -6 AVG on the meter. Not by coincidence, this corresponds with 83 dB SPL at the meter's 0 AVG point, revealing the obvious correlation between a mastering engineer's meter ZERO and 83 dB SPL.

clear as that of the theatre, and our loudness is probably slightly lower because some high-frequency transients have been clipped by 6 dB of compression). Our "pop studio" headroom is only 14 dB above 83 instead of 20. The absolute loudness* of our pop presentation is nominally 6 dB louder than a film in the theatre, necessitating turning down the monitor gain by 6 dB.

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 he 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 by designing a loudness meter with a sliding scale, where the moveable 0 point is tied to the same monitor SPL. Regardless of the scale, production personnel would tend to place music near the 0 point on forte passages.**

V. The K-System Proposal

This leads us to my K-System proposal, 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. It also develops a common *language of levels*, so that engineers can properly communicate.

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

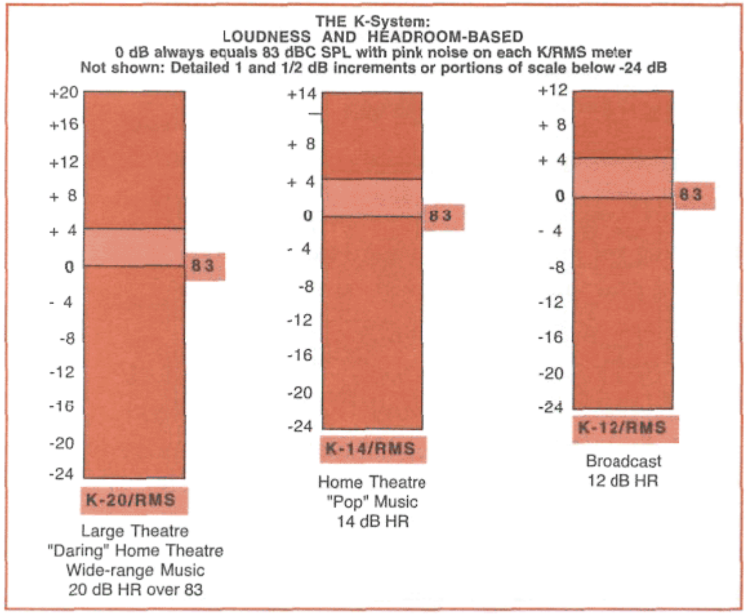
* **ABSOLUTE LOUDNESS:** A term I use when comparing the apparent loudness of different sources **without moving the monitor control.**

† Linear until things get so squashed that the increasingly compressed sound is not equally louder for the same measured increase in the flat meter's average level.

> *"The K-system is not just a meter scale, it is an integrated system tied to monitoring gain."*

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

*[**K-System Meter**. For a color image, please see the Color Plates section, **Figure C15-01**.]*



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.

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 that resembles the classic VU meter.

Note that full scale digital peak level is **always** at the top of each K-System meter, it does not change. Only the average level calibration slides, 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, making this **the first integrated metering and monitoring system.***

Simplified Explanation

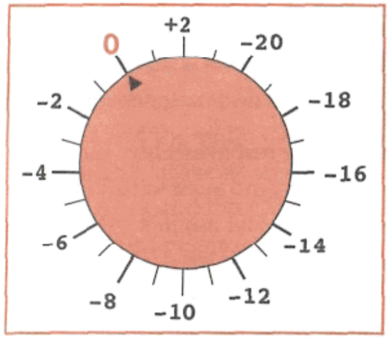
Many mastering engineers have recognized that the peak meter is inadequate for judging loudness, so they use a traditional analog VU meter. But because of the wide range of average levels on current pop CDs, they use a variable VU meter attenuator to prevent the VU from pinning or reading out of range. Think of the K-System as a **coordinated attenuator for both the averaging meter and the monitor gain**. The principle is that as we attenuate the average meter while going from K-20 to K-14 we must also turn **down** the monitor gain, to arrive at the same loudness to the ear. If the monitor gain were **not attenuated**, then K-14 material reaching 0 dB average on its scale would

* I invented these K-(N) terms because it was getting very awkward to describe the crest factor or loudness of music in a simple but useful way.

sound 6 dB **louder** than K-20 material going to 0 dB average on its scale.

Peak and Average calibrated to same decibel value with sine wave

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.



For medium-size control rooms, typical monitor gain (control position) will be 0 dB with the K-20 meter, -6 dB with the K-14 meter, and –8 dB with the K-12 Meter. 0 dB monitor gain is the calibration point that corresponds with the RP200 standard (see Chapter 14).

VI. 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 to RP200 as in Chapter 14. 0 dB always represents the same calibrated(83 dBC) SPL on all three scales, unifying production practices worldwide. 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. It will be hard for current pop mastering engineers to convert to K-14 or even K-12 in some cases, because much of today's damaged pop music is significantly hotter than even K-12—but we must find a way to back off from the loudness race. Ideally, K-12 should be reserved strictly for audio to be dedicated to broadcast; broadcast recording engineers may 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 (Fortissimo) 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 0 dB *(forte)* on any K-system meter. Use the *fortissimo* 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 music mixing engineers often work at a higher SPL, and vary their monitor gain to check the mix at different SPLs. I recommend that mix engineers calibrate your monitor attenuators so you can easily 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 sufficient. For highest sound quality, use K-20 while mixing and save K-14 for the calibrated mastering suite. If mixing to analog tape, K-14 may prove more appropriate. K–20 doesn't prevent the mix engineer from using compressors during mixing, but I hope that engineers will return to using compression as an esthetic device instead of trying to win the loudness race.

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 (sine wave), and mix normally with the analog console and VUs. However, mixing consoles should be retrofitted with calibrated monitor attenuators so the mix engineer can repeatably return to the same monitor setting.

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 master, or better, remix the entire program from the multitrack 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. We should try to fit this reduced-range mix on the DVD with 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 (absolute loudness) as many current pop CDs, it will probably be more dynamic and enjoyable, and importantly there will not be a serious loudness jump compared to K-20 DVDs in the same player. If the producer insists on a hotter CD, try to make it no louder than K-14, so there will be no more than a 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. The hypercompressed CD is the one that's out of line, not the K-14.

Full scale peaks and SNR. As we've discussed (Chapters 5 and 14) it is not necessary to peak a 24-bit recording to full scale. Another good reason is that a program's signal-to-noise ratio is determined by its **actual loudness**, 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. 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 considered normal and ideal for mastering.

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. Operators should be trained to change the monitor gain according to the K-standard.

In Color Plate ***Figure C15-02*** is a picture of the K-20/RMS meter in close detail, with the calibration points. Individuals who wish 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 practice, the music world will be far more consistent than the current chaos. Everyone should know the monitor gain they like to use.

In Color Plate ***Figure C15-03*** 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 computer. SpectraFoo versions 3f17 and above include full K-System support and a calibrated RMS pink noise generator. On the PC, Pinguin has implemented meters that conform exactly with the K-System. The Dorrough and DK meters nearly meet K-System guidelines but be sure to use an external RMS meter for calibration since they use a different type of averaging. In practice with program material, the difference between RMS and other meter averaging methods is imperceptible. I hope soon a company will implement the K-System with a *truer* loudness characteristic.

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, ironically just in time for the cassette's obsolescence. It's been difficult for mastering engineers to communicate with cassette duplicators, finding a reference level we all can understand. The cassette 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. An engineer can make a good cassette master by using a "predistortion" filter with gentle high-frequency compression and equalization. Use 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. 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. For example, dedicated classical producers have avoided mastering their harpsichord recordings to full scale, or they sound unnaturally loud at standard monitor gains. It's hard to get out of the habit of peaking our recordings to the highest permissible level. I strongly feel it is much better for the consumer to have a consistent monitor gain than to peak every recording to full scale digital. Attentive listeners prefer auditioning at or near the natural sound pressure of the original classical ensemble.[4]

Classical engineers should 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 at the 0 dB position 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 all reconsider this thought when 24-bit media reach the consumer. Until then chaos will remain in the classical field, and perhaps only metadata will sort out the classical music situation at the listener's end.

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 or even K-20.
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 by the loudness race. Some mastering engineers already have made archives without severe processing.

Multichannel

There's good news for audio quality: 5.1 surround sound. Current 5.1 mixes of popular music sound open, clear, beautiful, yet also impacting. Six speakers provide much more headroom and sound output than two, so if you work by the monitor gain, the channel meter levels will tend to run a bit lower. 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, the way it should be. System hiss is not evident at 0 dB monitor position with long-wordlength recording, good D/A converters, modern preamps and power amplifiers.

Labeling The Boxes

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.

VII. Metadata and the K-System

Metadata is *data within data*, that is, control data embedded in the digital audio stream. Dolby Digital, MPEG2, AAC, and hopefully MLP will take advantage of metadata control words (defined below); note that standard PCM, as used in the Compact Disc, has no provision for metadata, and to the best of my knowledge, neither does SACD. Pre-production with the K-System will speed up the authoring of metadata for broadcast and digital media. Music producers must become familiar with how metadata affects the listening experience.

Metadata Control Words

Dialnorm, *dialogue normalization*, also known as *volume 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 and carried as a command on the metadata word. At each program change, 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 a somewhat unnatural manner, like the radio, average levels of sports broadcasts, rock and roll, newscasts, commercials, quiet dramas, soap operas, and classical music all end up at the loudness of dialogue, a rather strange effect, but no different loudness-wise than standard radio today. The listener can turn his receiver up and experience the intended loudness—without the noise modulation and squashing of current analog broadcast techniques. Or, he can choose to turn off the dialnorm on some receivers, and hear a loudness variance from program to program.

Dialnorm is a simple gain change, without compression, and maintains the crest factor and dynamic range of the studio mix. For example, in variety shows, the music group will sound pleasingly louder than the presenter. Sports crowds will be excitingly loud, and the announcer 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. 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. Attentive listeners using *mixlev* will no longer have to adjust monitor gains for different music types.

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

In Summary

The designers of the compact disc never anticipated that an all-digital recording system would yield an alarming loudness race and seriously distorted music, worse than ever took place in the days of the LP. I propose a new system with a common language, integrating monitoring and loudness metering to produce more consistent masters, and move audio practice into the 21st century. Teach everyone how—the Rosetta stone is in this chapter.

---

1 Ironically, current-day compression practices (especially in pop music) are far more aggressive than necessary, even stronger than our approach to the noisier analog medium of the past! CDs can and should be produced to the same audio quality standard as the DVD, but I'd be satisfied with the leveling practices that made good LPs.

2 I see an interesting analogy of the loudness race and the migration of pitch since the 16th century. Music seems to be racing to be just a little more sharp than the previous generation, so that an A played on an instrument tuned to previous standards is now the G or G# of today, so it ultimately turns into a problem of transposition. Unfortunately, audio systems cannot accomodate an infinite loudness rise. We must voluntarily "transpose" back, or go deaf.

3 This is what the DVD and DVD-A proclaim to be, a single audio medium for all needs, because the table radio or the car can contain built-in compression, following the metadata coefficients laid down by the program producer. Let's meet again in 20 years and see if that promise has been met.

4 The late Gabe Wiener produced classical recordings noting in the liner notes the SPL of a short passage. He encouraged listeners to adjust their monitor gains to reproduce the "natural" SPL which arrived at the microphone. I used to second-guess Wiener by first adjusting monitor gain by ear, and then checking against Wiener's number. Each time, I found my monitor gain was within 1 dB of Wiener's recommendation. Thus demonstrating that the natural SPL is desirable for attentive, foreground listeners.

5 One of my first lessons in the inaccuracy of the VU meter was in 1972, when I heard William Pierce, voice of the Boston Symphony, clearly and distinctly in the noisy control room at Channel 24, yet he hardly moved the needle. The trained operator must use his ears and learn how to interpret this instrument.