

An Integrated Approach to Metering, Monitoring, and Levelling Practices

by Bob Katz, Digital Domain, Inc.

Précis: 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 correlated to a metering level of 0 VU. 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 reached a state of chaos at the end of the 20th century. The author proposes an integrated **system** of metering and monitoring that will encourage more consistent levelling practices among the three disciplines. This system handles the issue of differing dynamic range requirements far more elegantly and ergonomically than in the past. On the threshold of the introduction of a new, high-resolution consumer audio format, we have a unique opportunity to implement a 21st-century approach to levelling. Such a system will also aid production personnel in creating meta-data.

I: The VU Meter

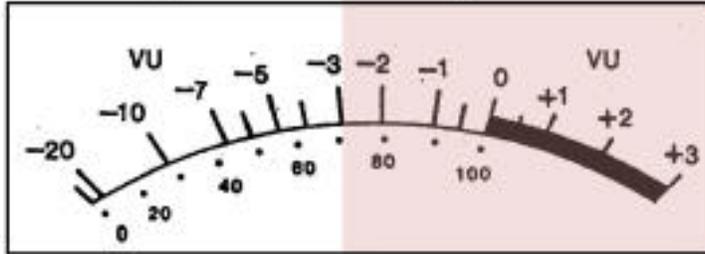
On May 1, 1999, the VU meter celebrates its 60th birthday. 60 years old, but still widely misunderstood and misused. An averaging device, this instrument was intended to help program producers create consistent loudness amongst program elements, but *never* to indicate when the recording medium was begin exceeded, or overloaded. The meter's designers assumed that a recording medium with at least 10 dB headroom over 0 VU would be used. Over 60 years, psychoacousticians have learned how to measure perceived loudness better than a VU, but until we can establish a new averaging-meter standard, and demonstrate that mixing engineers can effectively work with LEQ or Zwicker-based instruments, the VU is the first candidate for program-level measurement.

Summary of VU Inconsistencies and Errors

In General: The meter's ballistics, scale, and frequency response all contribute to a meter which moves far greater than the actual perceived dynamic loudness change of the program material.

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. It's actually too fast to be a loudness meter. 500 milliseconds or greater would correspond better with the ear's integration time. Skilled users soon learned that an occasional “burst” to +3 VU would probably not cause distortion, and usually was meaningless as far as a loudness change.

50% of Scale contains
top 6 dB of range!



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. 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, 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: In current use, there are large numbers of improperly-terminated mechanical VU meters and inexpensively-constructed indicators which are labelled "VU". 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.

Despite all these problems, *the VU meter is a primitive loudness meter*, more effective than any other program level meter in regular use. 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



Outside of the film industry, chaos currently prevails. Above is a waveform taken from a digital audio workstation. The time scale is about 10 minutes total, and the window amplitude is +/- full digital scale. On the left side is a piece of heavily compressed pseudo “elevator music” 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 0 dBFS. The plant manager involved in pressing the 1999 CD remarked “this CD is a lightbulb! 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?

Popular CDs with this problem are becoming increasingly prevalent, coexisting with discs that have beautiful dynamic range and impact, but far lower loudness character. There are many technical, sociological and economic reasons for this chaos that are beyond the scope of this paper. This paper will 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?

In the film world, films are consistent from one to another, because the monitoring gain has been standardized. But in the music world, everyone determines their own average record level, and adjusts their monitor accordingly. The hotter the average level, the more they have to turn down their monitor to obtain the same perceived loudness. Clearly the producers of music discs have the right to mix and master them any way they desire. But in order to save the sound of our music, we must create an integrated metering and monitoring system that will discourage the practice of a “loudness war”, and encourage quality and consistency in music recording. Then, we must educate program producers how to use that system. Manufacturers of consoles, metering and monitoring systems should aid by producing tools that conform with that system, tools which are easy to produce with current digital technology.

III. The Magic of 85 with Film Mixes

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 4 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 85 dB at 0 VU standard originally proposed by Dolby's Ioan Allen in the mid-70's. It's stood the test of time. 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. The meter becomes simply a guide. In fact, working with a fixed monitor gain is *liberating, not limiting*. When digital technology reached the large theatre, Dolby attached the 85 dB calibration to a point 20 dB below full digital scale (abbreviated **-20 dBFS, referred to as the standard Dolby Cal point in this document**). This calibration must be measured with pink noise, with an averaging meter (not a peak meter), playing one channel (loudspeaker) at a time, and the SPL meter set to slow, C weighting. Tom Holman has recently proposed more refined ways of making this measurement, but the basic principle remains. The 85 dB SPL/0 VU/-20 dBFS standard has also stood the test of time, as digital productions can be created with excellent headroom.

When AC-3 and DTS became available for home theatre, Dolby recommended that the monitor calibration standard be lowered by 6 dB to 79 dB SPL (at -20 dBFS average). This is because mixes originally geared for large theatres do not totally translate to the small venue. There is often so much dynamic range and impact from loudspeakers in a small space, that even high-powered home theatre systems (and tolerant listeners) have trouble bearing the loudness if reproduced at the Dolby 85 monitor calibration. It's admirable that certain program producers are preserving the original 85 dB large theatre mix for posterity, but many home listeners may complain that "this DVD is too loud", or "I lose some of the dialogue when things are soft". This is because they turn down their monitor gains without the mix being changed, and soft passages may become too soft. To make the 85 dB-calibrated presentation palatable for such listeners, the dynamic range may have to be reduced by 6 dB (6 dB upward compression) in order to be reproduced at a Dolby cal of 79.

In the future, metadata (dialnorm with optional compression) may reduce this conflict, but currently there are only two solutions: a) to compromise the audio soundtrack by using permanent dynamic range compression during DVD production, or better, b) for the home system to be able to insert a compressor. The latter gives us the best of both worlds. Not all film studios making DVDs are using the purist approach, and this explains some of the variability among DVD-video soundtracks.

Note that when Dolby recommended 79 dB for the home, they connected it to the same Dolby Cal point of -20 dBFS, but the VU meter seems to have been forgotten. More on this in a moment.

IV. The Magic of 79 for Home Music Productions

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 (MP3?). Given the current state of the music industry, music-only discs may end up being used as background music---and if it doesn't have a moving picture, lose the interest of young fans. But this paper specifically refers to foreground music that the discerning consumer will play at normal “enjoyment” loudness.

With the integration of media into a single system, finally it is in the direct interest of music producers to unite with video and film producers for a more consistent consumer audio presentation. This will eventually happen, but not if music producers experimenting with 5.1 surround pay only casual attention to monitor level calibration. New program producers with little experience in audio production are also coming into our 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 quality of monitors used for program production may be less than ideal. It is definitely time to educate, and establish a standard, before new chaos reigns. The current lack of VU meters, and the plethora of peak-only meters on every computer, DAT machine and digital console will definitely not help. Engineers must be trained to realize that the peak meter is for one purpose only: to protect the medium.

As a music mastering engineer, I have been studying the perceived loudness of music compact discs for 11 years. Around 1993 I installed a 1 dB per step monitor control in my system 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 master the disc to work well at that monitor gain. In 1996, out of curiosity, I measured that monitor gain (at -20 dBFS) to determine how far I was working from Dolby level, and I was pleased to find that I've been working at 79 dB, the Dolby *home theatre* standard. Even though I master stereo material, this discovery was very gratifying.

By now, I've mastered over 100 CDs at the 79 dB calibration, with very satisfied clients. In 1994 I installed a pair of Dorrough meters, in order to view the simultaneous average and peak level on the same scale. These meters use a scale with 0 average (Dorrough's version of VU) at 14 dB below full digital scale. Full scale is marked as +14 dB. This scale is useful for music mastering engineers, since a conservatively-recorded stereo 1/2" 30 IPS analog tape rarely has a crest factor greater than 14 dB. For most individuals with normal hearing working in a small room, 79 dB calibration (at -20 dBFS) results in a 14 dB crest factor, or perhaps you may choose to look at it the other way (a 14 dB crest factor will naturally lead to a 79 dB calibration). They're both sides of the same coin.

As monitor gain is reduced, average recorded level must come up. This is because the mastering engineer always seeks the same loudness to the ears. And because the medium has a fixed maximum peak level, the crest factor must be reduced, and more compression/limiting must be used to keep the system from overloading.

CD Changers and Mastering. The current chaos has resulted in CDs with a wide variance of average level and perceived loudness, a disservice to the consumer. Do we want history to repeat itself when we move to DVD-A or SACD? Currently, we have to train consumers that adjusting their volume control is a perfectly normal part of the process of changing from CD to CD. But they expect the CD changer to be like the radio, which is why the CD changer is the enemy of the mastering engineer. And our clients, even professionals, are often just as ignorant of the situation and its consequences. We end up mastering to the lowest common denominator, and fight desperately to avoid that situation. Manufacturers take note: all CD changers should include a compressor that can be user engaged.

Mastering and The Loudness Race. By 1997, some of my music clients were complaining that their reference CDs were “not hot enough”, a tragic testimony on the loudness race which has hit the industry. Powerful digital compressors and limiters now 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. Each client wants his CD to be as loud as or louder than the previous transgressor, and we must spend a lot of time showing them that the sound quality suffers as the average density goes up. The psychoacoustic problem is that even when two identical programs are presented at slightly differing loudness, the louder of the two appears “better”. This explains the creeping increment of average program loudness, until sound quality is so bad that everyone can perceive it. When the mechanical VU meter ruled, it was difficult for engineers to ignore the warning sign of the needle banging against the peg, but today there is no such warning. This is why monitor and meter calibration is absolutely essential for the new digital consumer formats.

Trying to “hold the fort”, I’ve raised the average level of the CDs I was producing only when requested, which of course forced me to reduce my monitor gain, often using 77 dB (Dolby cal). For every dB of increased perceived level, considerably more compression must be applied. The sound gets quite squashed, and processor distortion severe when the monitor falls below 77. Other popular CDs are considerably compressed (and often distorted), and must be reproduced with monitor gain at 75-74 dB (Dolby cal) for comfortable listening levels. Many consumers are finding their volume controls at the bottom of their travel, where they have the least movement resolution.

V. 79 is really 85!

The next step is to realize a simple but important arithmetic relationship. We've been discussing dual-function meters where the 0 VU point is at 14 dB below full scale. We've also been talking about the Dolby calibration method, which uses a

reference 20 dB below full scale. 20 dB below full scale falls at -6 VU on this meter. Which leads to the simple equation:

Since $-20 \text{ dBFS} = -6 \text{ VU} = 79 \text{ dB SPL}$

then $-14 \text{ dBFS} = 0 \text{ VU} = 85 \text{ dB SPL}$

In other words, we are mastering at $0 \text{ VU} = 85$, the same as the Dolby large theatre standard. If the monitor gain is reduced by 2 dB to 77, we tend to increase the average level, and the VU meter creeps up by 2 dB. It's a linear proportion. It's clear that the ear likes 0 VU (the average level) to end up at 85 dB, even when we work in a reduced headroom, or compressed structure. This leads us to the logical conclusion that we should unify production practices at the 0 VU point by using a sliding meter scale, where the moveable VU is always tied to the 85 dB monitor calibration point.

VI. The K-System Proposal

With apologies to the late Marshall McLuhan, in digital audio, the recording *medium* is *not* the *message*.

In the 20th Century we concentrated on the *medium*. In the 21st Century, we should concentrate on the *message*. The peak meter tells us nothing about the *message*. The peak meter takes care of the recording *medium*. If we remember the duality of these two terms, we will never get confused.

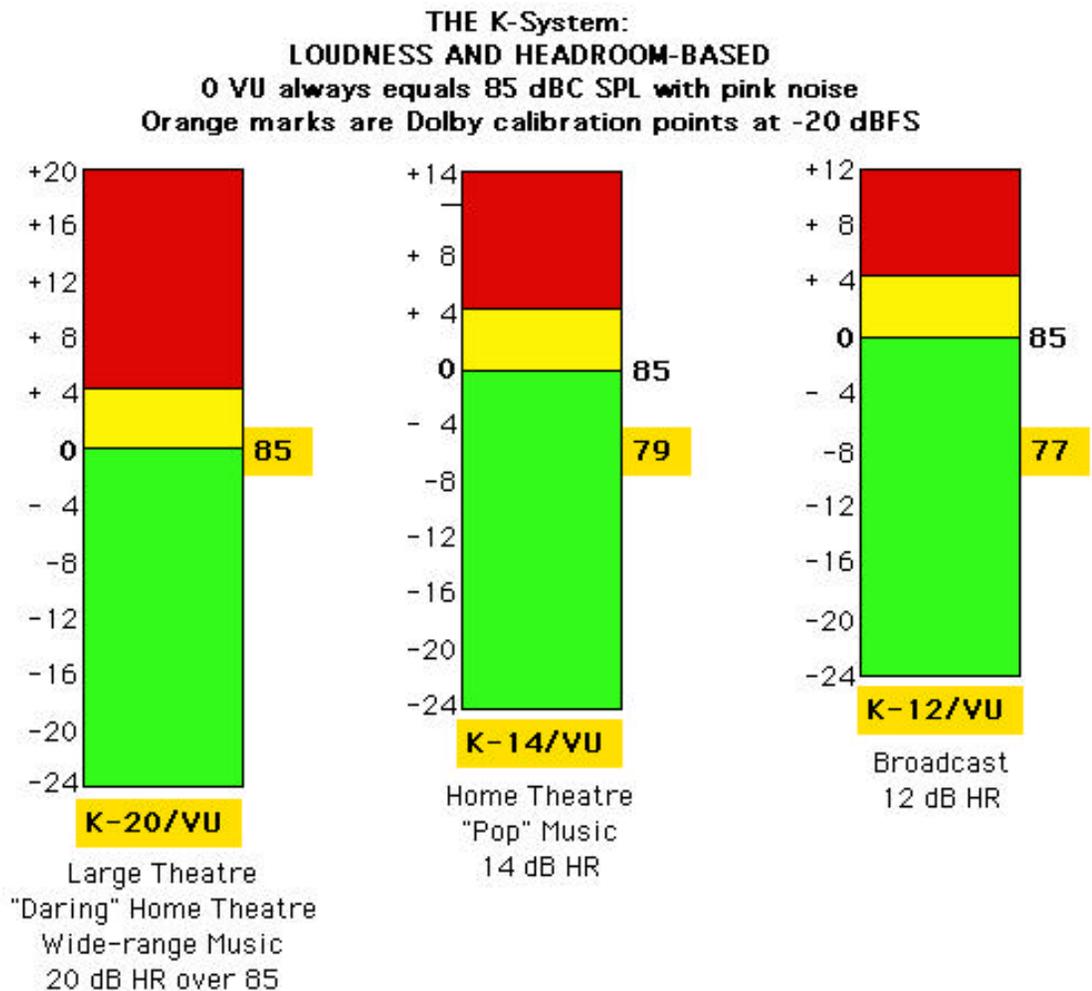
If only the designers of the compact disc system had foreseen the chaos that would result from the loss of an average metering standard or a monitoring standard. To avoid the same problem with the DVD and other new high resolution media, we must unite behind a new standard that integrates metering, monitoring and levelling practices, called the K-system.

First, because the medium is *not* the message, we should stop using meters which have 0 dB at the top---this discourages operators from understanding where the message is. This 21st century meter should be tied to a calibrated monitor gain, with the averaging meter's 0 set to 85 dB SPL. The scale must be linear-decibel through at least a 24 dB range, dual characteristic, peak and average, with the average (VU) level being the most important part of the display. The averaging portion of the meter should be the bar, with a moving line above the bar representing the instantaneous (1 sample) peak level.

Peak section of the meter: Highest Peak in the last 10 seconds should be available, as well as infinite high peak hold (user choice). The peak line should have a slow fall time such as 2 seconds/24 dB. An adjustable and resettable OVER counter is highly recommended. A tri-color system is suggested, with green below 0 VU, amber to +4 VU, and red above that to the top of scale (see fig below). **Averaging section:** The highest average level in the last ten seconds should be available, or the highest long-term average level, (user choice) resettable. The 0 point (always 85 dB SPL) slides depending on the venue of interest, resulting in a change in headroom and the amount of compression required. *The K-system is not just a*

meter scale, it is an integrated system tied to monitoring gain.

The three K-System meter scales are officially known as K-20, K-14, and K-12, but colloquially will be known as the papa, mama, and baby meters. The K-20 meter is for use with wide dynamic range material, e.g., large theatre mixes, “daring home theatre” mixes, audiophile music, classical (symphonic) music, hopefully future “audiophile” pop music mixed in 5.1, and so on. The K-14 meter is for the vast majority of high-fidelity productions for the home, e.g. home theatre, and pop music (which includes the wide variety of moderately compressed music, from folk music to hard rock). And the K-12 meter is for productions to be dedicated for broadcast.



Note that full scale digital is always at the top of the meter. Note how the 85 dB point slides, and its relationship to Dolby's 85, 79 and 77 recommendations at -20 dBFS. The Dolby points remain anchored to 20 dB below full scale, whereas the K-system point is a sliding 0 VU. They are functionally equivalent, but new users should be told “always calibrate 0 VU to 85 dB---no matter which scale you are using.” Using the term K-(N) defines simultaneously the 0 VU point (relative to full scale), the monitoring gain, and the approximate monitoring sound pressure level

(assumed to be 85).

The meter calibration method is similar to AES-17, where the 0 dB reference point is identical for the average and peak sections of the meter with a sine wave signal. In other words, a sine wave whose peak level is -14 dBFS will simultaneously read 0 dB on the K-14 scale's average meter and peak meter. Of course, any wave with a different crest factor will cause the peak and average scales to read differently, but this is the whole purpose of the meter.

Europeans have been using a -18 dBFS reference for a while, perhaps to improve the signal to noise ratio of low-bit systems. But this is not a conflict with K-20. The European -18 dB reference is to a specific analog voltage level, and K-20 is referenced to a specific sound pressure level. The two systems can coexist. If there is some technical need to generate pink noise at -18 dBFS instead of -20, then the corresponding SPL to calibrate the monitor gain should be 87 dB for K-20.

VII. Production Techniques with the K-System

A manual for a digital limiter reads: "For best results, start out with a threshold of -6 dBFS". This is like saying "always put a teaspoon of salt and pepper on your food before tasting it." This kind of widespread misinformation does not encourage proper production practice. A gain reduction meter is not an indication of loudness.

If console and workstation designers standardize on the K-System it will make it easier for engineers to move from studio to studio. By anchoring operations to a consistent average reference, operators will produce more consistent output, and everyone will recognize what the meter means. The process is simple, upon beginning a production, the operator chooses his goal: If making an audiophile recording, then he uses K-20, if making "typical" pop or rock music, or audio for video, then he probably uses K-14. K-12 should be reserved strictly for audio to be dedicated to broadcast, and broadcast recording engineers may certainly choose K-14 if they feel it fits their program material. Audio for video studios who've been working without an averaging meter should convert to the K-System right away.

Recognize that each "lower" meter scale requires more dynamic range compression than the preceding higher version. For example, translating material mixed in a large theatre using K-20 to use in the home with K-14 will probably require about 6 dB compression. To look at it another way, the average program level may have to be raised 6 dB to accommodate the smaller, noisier venue and home equipment with less headroom. Producers must choose a method for adapting material that was mixed for the larger scale to the smaller one. The most desirable way to accomplish this will be to remix the elements (stems or submixes) using careful manual compression of the elements and probably some peak limiting. In this way, 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. This "secondary" or "home" mix may be able to fit on the same DVD with or without metadata.

The second method will be when the stems are not available---a re-mastering

engineer working in a superior acoustic environment can apply carefully-selected combinations of levelling, equalization (if necessary), compression and limiting, searching for the least damage to the material and the truest translation of the original soundtrack. Perhaps the home version may end up sounding better than the theatre version, when the remastering engineer does his job with excellent monitors and if the original theatre mix was done in a hurry. The third method (“quick and dirty”) will be “semi-automatic”, where an operator working under inferior acoustics applies some kind of generic compression/limiting to the material. The third method may be used in broadcasting (no offense intended to quality-oriented broadcasters), and obviously will do some damage to the program producer's original intent.

The existence of two prime meters, the K-20 and K-14, will create a cluster around two different monitor 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 chores. Of course, as we progress towards metadata it may be possible for the consumer to hold one monitor position for all foreground program material. The K-System will improve sound quality by uniting the steps of pre-production (recording and mixing), post-production (mastering) and metadata (authoring) with a common audio language.

The 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 VU on loudest passages. Rock (and some electric pop music) must take calculated advantage of this “loud zone”, since with this type of music, climaxes, loud choruses and peak moments in the music sound incorrect if they only reach 85. This is musically equivalent to *fortissimo*, which composers have equated to 88-90+ dB since the time of Beethoven. The key word is *occasionally*, for the use of this part of the loudness range is probably musically incorrect (and ear-damaging) if sustained for long periods; the sustained (LEQ-longterm average) average for *forte* passages usually sounds uncomfortable if not maintained around 85 for the majority of musical genres. If engineers find themselves using the red zone all the time, then either the monitor gain is not calibrated to 0 VU=85, the music is extremely unusual (e.g. “heavy metal”), or the engineer needs to use more monitor gain to correlate with his personal sensitivities. If not, then the recording will end up overcompressed, with squashed transients, and its loudness quotient out of line with the K-System guidelines.

Compression is a powerful esthetic tool. However, it should be recognized the higher the monitor gain, the less compression is needed to make material sound good or “punchy”. “Use less processing, just turn up the monitor,” is good advice for clients. However, since K-14 is where most of today's (yesterday's) pop CDs are at, we tend to master them into that ballpark. Plus, a K-14 presentation can sound better than a K-20 (when listened at equal average loudness), with skillfully-applied dynamics processing by a mastering engineer working in a calibrated room. Even at

K-12, after a lot of work, it is possible to produce a somewhat clean master with some punch and transient clarity. But clearly, the higher the K-number, the more freedom we have and the easier it is to make it sound open and clean. Monitor systems with good headroom must be used to make value judgments---if our monitors compress, how can we tell if the program material has problems?

For highest sound quality, digital multitrack recording engineers should always use K-20, and if the mix is going to be later mastered, stick with K-20 for the mixdown. Save K-14 for the calibrated mastering suite. If recording to analog tape, continue to meter and monitor at K-20, and realize that the tape will act like a compressor with a threshold 12-14 dB above 0 VU. K-20 doesn't prevent the mix engineer from using compressors during mixing, but engineers will gravitate towards using compression as an esthetic device rather than a "loudness-maker". With 20-24 bit converters, the mix does not have to reach full scale (peak). Use the VU and your ears as you normally would, and with K-20, even if the peaks don't hit the top, the tape (or file) is still considered normal and ready for mastering, with no meaningful loss of SNR.

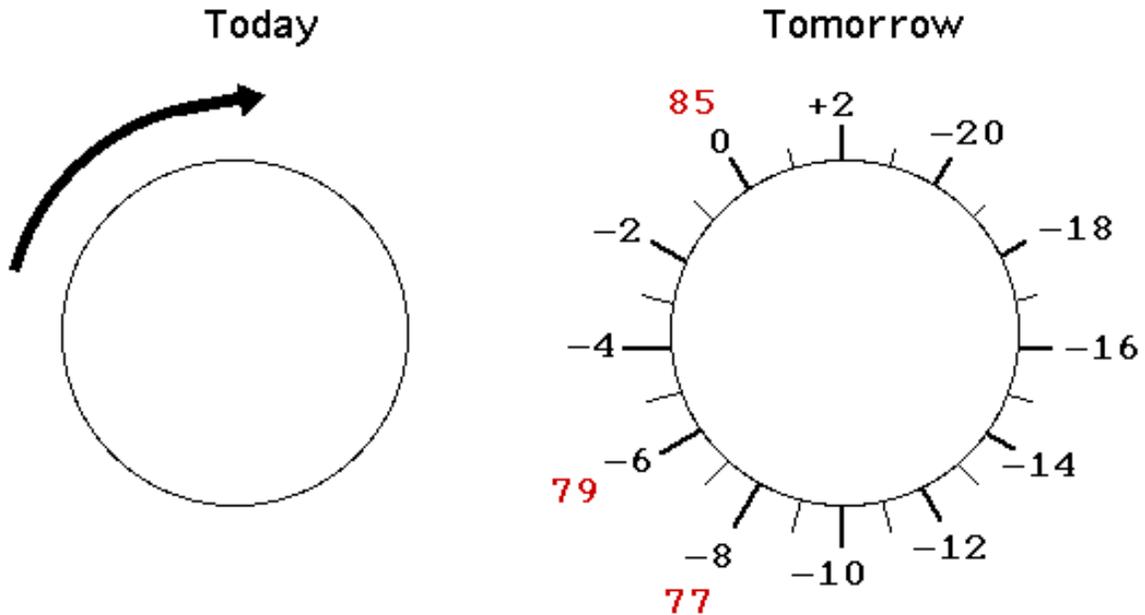
Using K-20 assures a clean-sounding mix with some "meat" that the mastering engineer can grab onto when he gets the tape or file. At that point, the producer should discuss with the mastering engineer if he will convert this program to K-14, or keep it at K-20. The K-System becomes the lingua franca of interchange within the industry. It helps avoid the problem where three mix engineers work on parts of an album, and the mastering engineer has to deal with mating three mixes each produced to an unknown standard.

Current-day analog mixing consoles equipped with VUs are far less of a problem than the digital models with only peak meters. Calibrate A/D gain to -20 dBFS at 0 VU, and mix normally with the analog console. However, isn't it surprising to find a \$750,000 console with a \$10.00 monitor pot? How does the mix engineer repeatably return to the same monitor setting? A simple retrofit will help engineers worldwide to produce more consistent work.

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 pop music mixing the next. Because all producers in all disciplines are working with the same averaging meter at the same calibrated SPL, work will be more consistent, and also translate easier from discipline to discipline. Possibly the 5.1 mix of a project will be produced at K-20, and the stereo version at K-14, to be more compatible with current CDs. A simultaneous meter scale change and monitor gain change accomplishes the job. It would seem intuitive to have a single button that changes the meter scale and 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 labelled similar to the following, and the operator must shift the meter scale manually.

Monitor Controls



As per the figure, the SPL labels on the manually-adjustable monitor gain control should be permanently anchored to the consistent internal reference -20 dBFS (avg). Operators should be trained that 85 conforms with meter K-20, 79 with K-14, 77 with K-12.

Operators may personally desire to run their monitors slightly hotter or lower than the standard, but realize that the equal-loudness contours imply their mixes will be bass-shy or bass heavy when reproduced at the calibrated level. Individuals who prefer to listen off the standard should find their ideal monitor gain point, log it on the box, and try to use it 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 what monitor gain they like to use.

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). Over the years, it's been difficult for mastering engineers to communicate with audio cassette duplicators, finding a reference level we all can understand, since many of them do not really understand the difference between VU and peak level, nor can they speak intelligently in terms of the maximum operating level and harmonic distortion of their equipment and tapes at their operating levels. A knowledgeable duplicator once explained that the tape most commonly used cannot tolerate average levels greater than +3 VU (especially at low frequencies) and high frequency peaks greater than about +5-6 over 0 VU are bound to be distorted and/or attenuated. Armed with that information, an engineer can make a good cassette master by passing through a "predistortion" filter with gentle high-

frequency compression and equalization. Meter with the K-System (K-14 is suitable) and put test tone at 0 VU on the master (usually on DAT). Peaks will never reach full scale or the cassette will distort. The compressed material will require using a higher monitor gain, but this is a special case.

Margin of Error. Given variations in speaker bandwidths, interpretation of meter readings (especially the shaky VU ballistic), and user loudness preferences, there may be a variability in performance of 2-3 dB. But this will still be far more consistent than the current chaos. An example of a typical error mechanism: operator #1 works with K-14 and allows his VU to go to +3 or +4 quite often. Operator #2 works with K-12, but is religious about not exceeding 0 VU. Both programs may end up having the same apparent loudness and be reasonably open and not too compressed. An experienced mastering engineer can put on any CD, adjust the monitor gain to his ears, look at the stepped attenuator setting, and predict the VU meter excursion within a couple of dB. Everyone should look forward to working in an environment with that degree of precision.

Classical music. It's very desirable that the classical music listener be able to work with a consistent monitor gain. Modern-day symphonic music conforms well to the K-20 standard. But renaissance music, softer ensembles and string quartets produce lower native SPLs, exhibit lower crest factors, and if peaked to full scale on the peak meter, sound considerably louder than the symphony. The listener has no choice but to turn his monitor gain down. In live concerts, when the symphony performs before intermission and the string quartet after intermission, do we have to turn down the gain of our ears? In an ideal world, the classical recording engineer could set the gain of his microphone with a calibrated noise source, and record all the music in the world at a fixed record and monitor gain.

So, should we encourage recording engineers to peak their string quartets at, say 8 dB below full scale and symphonies to full scale? The answer is yes, and no. We've all been trained to peak our recordings to the highest permissible level, and it's very hard to get out of that habit, even with 20-bit systems, that have 24 dB better signal-to-dither-ratio than 16-bit. Of course, metadata can solve that discrepancy. It is always desirable to maximize our encoders, record to (reasonably) full level to stay away from quantization noise, but once the program has been encoded, it is not essential to peak to full level on the decode (D/A) side---especially when the monitor gain has been set with the K-System. In other words, as long as we use all the bits of the encoder, there will be no compromise in the decoder if we lower the digital gain in post production or mixing (and properly redither to the final wordlength).

String quartets, harpsichords and solo instruments that produce lower SPL are potentially a problem fitting into the K-System unless we can educate engineers to follow their ears (aided by the VU) and not the peak. At a producer's request, I once mastered a solo pennywhistle record peaked to full scale that monitors correctly at 20 dB below 85---it was the loudest record I've ever made! Instead, after A/D conversion to the max, I recommend these instruments be mixed/mastered using the K-20 scale, work to the ear and not the peak meter, and allow that the peaks will never reach 0 dBFS. That way symphony and solo instrument discs can coexist

without causing the listener to radically move his volume control. Currently, such uncompressed classical recordings contribute to a similar chaos as overcompressed rock. They both sound too loud if peaked to 0 dBFS. These last discrepancies may never be settled without metadata, unless program producers are willing to accept peak levels that do not reach 0 dBFS.

If classical program producers insist on peaking to full scale, they should note the meter scale and monitor gain they used on the box, to aid in creating metadata during the authoring step.

VIII. An Extendable System

The K-System is extendable. Since the VU is a rough approximation of loudness, it is desirable to advance to other methods, such as LEQ and Zwicker. More advanced methods will also result in more consistency, both because they compensate for Fletcher-Munson, and because the advanced meters have much slower “ballistics” than the VU. We have yet to see if operators can get used to a slow, non-syllabic meter. The VU's original designers settled on a rather fast meter because operators could not get used to a slower one, but the peak line on the top portion of the dual-function meter should provide the syllabic response while they rely on the slower average bar for loudness.

Program producers should mark their tape boxes or digital files with an indication that the K-system was used, and which meter was used (which automatically specifies the calibrated monitor gain). For example, a box could be labelled “K-14/VU”. In this way, succeeding operators and producers will know that the program creator used the K-14 scale with VU characteristics, and calibrated his monitors to 85 at 0 VU. When (if) we move into other loudness-measuring systems, the boxes could be marked “K-20/Zwicker”, for example. These labels will become as common as listings of nanowebers per meter and test tones for analog tapes. If the mix engineer worked at a different monitor gain than the K-standard (e.g., 88 dB), then that fact should be noted on the box, as an aid to authoring engineers trying to insert metadata.

A new zero reference for the decibel be created for each psychoacoustic loudness implementation. For example, **0 dB (Zwicker)** shall be defined to correspond psychoacoustically with 0 VU = 85 dB SPL (pink noise). Acousticians should reach a consensus on the best method of calibrating the new zero reference point. In other words, any future meter 0 reference should correspond psychoacoustically to the same loudness with pink noise as 0 VU does with 85. That way, the transition will only be to a slower meter, not a new monitor calibration point! Of course, it's a lot easier to say “zero VU” than “zero Zwicker”...someone will have to come up with a new term!

IX. Single Channel or Multiple Channel Loudness?

The Dolby calibration method is by individual channels, which is probably a good idea. A K-system meter is needed on each channel, but depending on the coherence between left and right loudspeakers, total SPL can rise from 3 to (rarely)

6 dB when each channel is fed equal signal, typically 4-5 dB. With this variance from room to room, we can only hope that each engineer works with a “typical” loudspeaker system. Perhaps we should recommend coherence specifications (which are also affected by angle and distance between loudspeakers). Perhaps we need a K-System meter on the aggregate (mono sum) of all channels. Time will tell.

X. In Conclusion

Let's bring audio into the 21st century. The K-system, which integrates monitoring and metering practice, will take us a long way.