Søren H. Nielsen and Thomas Lund
TC Electronic A/S
DK-8240 Risskov, Denmark

Presented at
the 107th Convention
1999 September 24-27
New York

AUDIO
AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

Level Control in Digital Mastering

SØREN H. NIELSEN AND THOMAS LUND

TC Electronic A/S
Sindalsvej 34, DK-8240 Risskov, DENMARK

Music mastering is becoming a battle for maximum level rather than a quest for audio quality, because counting consecutive samples at 0dBFS is not an adequate restriction of level.

This paper will discuss alternative methods.

New measurement practices may turn into a standard with which recordings have to comply in order to use a proposed Dynamic Range Approval, DRA.

0. INTRODUCTION

Level measurement in CD production has traditionally been a matter of counting consecutive samples at digital full scale, 0 dBFS, and thereby accepting some distortion at a given frequency.

Even though the initial objective of getting adequate digital level and therefore low distortion on the mastertape is understandable, the standard simply does not take into account how modern dynamics processing techniques make it possible to be at full peak level most of the time.

With digital wide- and multiband compression, soft clipping, level mapping and hidden clipping, many recordings are currently distorted and spectrally mistreated during the production and mastering process, just to obtain the loudest possible end result.

Being a supplier of equipment for digital mastering, TC Electronic has a genuine interest in establishing rules and goals for the process with a more pronounced focus on quality, before an even wider dynamic range than the one represented by the current 16 bits is treated the same way.

With large amounts of DSP power now being available, it is not hard to imagine music being put through several stages of "efficient" dynamics processing and clipping before reaching the user. Initial processing with too liberal metering is therefore not in anybody's interest.

1. LEVELS IN BROADCAST

It's surprising how musicians, mastering engineers and record producers have generally accepted the way music is treated on radio, when considering the consciousness of film-creators regarding how a feature film translate to television. When a film is transferred to television, several factors like aspect ratio, color balance, brightness etc. are evaluated, but with a CD on the radio it seems to be only a question of maximum loudness.

Much of the loudness hype in mastering derives from wanting maximum level in radio broadcast, but too much compression in mastering may easily make the material sound worse on the radio or television than a recording with less dynamic processing and more of its original dynamic range preserved.

Is there then a mastering trend of making the CD sound radio processed right from the beginning? Not everywhere. This is a quote from mastering engineer Bob Ohlsson [15]:

"... the biggest problem has been the wholesale replacement of skilled audio operators by an ever heavyer-handed use of compression. Broadcasting is often unlistenable because nobody with their hands on the controls is being paid to listen any more."

1.1 Analog broadcast

The intensive use of compression has also led to problems in the broadcast sector. Similar to the full scale limit in CD production there is a maximum peak level allowed for FM broadcasting, typically a frequency deviation of 75 kHz. The limit is technically spoken not hard in itself - an FM transmitter can be set to generate a larger frequency swing if desired. The larger the frequency swing, the louder is the signal at the listener, but at the same time the bandwidth of the transmitted signal rises. Frequency planning in FM broadcast is based on assumptions about average- (and peak-) levels. Based on these assumptions the interference to neighbor frequencies can be calculated. For the dynamics structure of the signals at the time, where these recommendations were originally written, this was fine. But when the signal level in a larger proportion of the time is higher, the risk of the listener being disturbed from a transmitter at the neighboring frequency has risen.

Therefore, in FM broadcasting a method has been recommended [14] for keeping the long term mean level within certain limits.

1.2 Digital broadcast

With digital broadcasting right around the corner we may be heading for alternatives to transmitting programs with less than 10 dB of dynamic range.

In digital broadcasting the possibility of interference does not depend on signal contents, so here the considerations on level control can be based on more isolated criteria.

Typically, there are provisions for transmitter-guided receiver-side dynamics processing in the digital broadcasting systems. Thereby the audio signal can be sent with minimal

dynamics processing, and the listener can decide whether it is desirable to adapt the signal to the possibly noisy listening conditions. In Eureka 147 DAB [3] a 6 bit control word is transmitted every 24 ms, enabling upto 16 dB gain increment.

High quality audio should always be transmitted, but can be scaled down to the current listening requirements. This strategy will put the production engineer and the end listener in control instead of a one-size-fits-all generic processor.

The analog to digital transition in broadcast may actually help bring down processing in CD mastering: There is **no cure for extensive processing**, but it is easy to apply if the end listener gets the option in his reproduction system.

2. METERS AND PROCESSING FOR MASTERING

It is common practice to use a mixture of analog and digital level meters for mastering. But there are important differences between them. It is necessary to know what type of meter you have available, and what type of meter your program will be judged with.

Analog peak meters (PPM's normally used in European broadcast and film) [4] do not show very short program peaks, and VU meters (more common in the US) [16] show very little transient information at all. Both kind of meters were designed long time ago with an analog signalpath in mind.

In the digital domain hundreds of samples pass by without the highest value showing up on such a meter, but an overload can produce alias distortion and other unpleasant artifacts in the digital domain.

Consequently, digital meters show a peak read-out 3-5 dB higher than analog PPM's and 16-20 dB higher than VU's, even though a sine tone is shown at a reference level on all instruments.

2.1 Dynamics Processing

Some of the censoring built into a signalpath relying on analog tape and records has been lost in digital production. Most noticeable there is no longer a restriction of high levels at high frequencies.

Multi-band dynamics processors have found their way into mastering mainly because they are good at tackling this problem. Dynamic spectral balancing may be used to compensate for our ear's increasing non-linearity at low levels, and to avoid too much brightness at high levels. Using this technique, it is possible to boost e.g. low and high frequencies at low levels but have the frequency response flat or even cut at high levels.

Spectral balancing is also good at raising the loudness of most program material without raising the peaks. But when the effect is evaluated, engineers often forget to compare the processed signal and the not processed at equal loudness. If not, the loudness dependant non-linearities of our ears will typically fool the engineer to prefer the processed signal.

3

Because of extensive use of spectral shaping, and the fact that a square wave at a specific peak level is louder than a sine wave at the same peak level, counting consecutive samples as a level barrier favors distortion and aggressive spectral shaping in order to achieve loudness.

2.2 Inter-sample precision Digital Peak Meters

In the digital domain the peak level may deviate from the peak level in the analog domain. While this is true in general, it is only significant at high frequencies. There are two possible reasons for this deviation:

1. Basic sampling theory. Sampling occurs at regular intervals, and at frequencies near integer fractions of fs, such as fs/4 and fs/2, the phase of the signal compared to the sampling times may generate a digital peak value somewhat below the analog peak value - at least for a short period of time. If the signal is not exactly at one of the critical frequencies mentioned above, the peak value in the digital domain will get very close to the analog peak value. If the analog signal prior to sampling was properly bandwidth-limited, the output after digital to analog conversion will be substantially equal to the analog input signal. See Figure 1-3. In the IEC standard for digital peak level meters [4, 5] the problem is recognised, but it is accepted that a meter can be inaccurate at certain frequencies related to the sample rate. Newer digital peak meters overcome this problem by estimating the signal between the samples.

2. Gibb's phenomenon [6]. Occurs when limiting the bandwidth of a wide-band signal (or truncating an impulse reponse). This is particularly important when the signal is clipped in the digital domain, but it applies generally. What happens is that a square wave (or hard clipped signal) can be viewed upon as a sum of individual sine waves of frequencies 1, 3, 5,... times the fundamental frequency. The flat top of the square wave depends on the presence of all harmonics at the right levels and phases. If some of the harmonics are removed by lowpass filtering, the peak value of the signal rises. When converting from digital to analog a low pass filter is always applied, so the analog level may be higher than expected.

Not only when converting from digital to analog the signal is reconstructed with a potentially higher peak value as immediately seen, but also when passing the signal through a sample rate converter there is a risk of overload, if the digital peak meter does not take into account the intersample peaks.

2.2 True Loudness Measurements

If the purpose of level metering is to assess the loudness of a signal in an objective way there is a way to do it properly [1, 2]. Use a loudness meter. This is an instrument, which is based on frequency analysis of the signal in a similar way as it happens in the human hearing system. Loudness is a quantity related to perception, not a physical quantity such as sound pressure. For pure tones, sine waves, the well known equal loudness contours express the relative sensitivity of the hearing to tones of different frequencies. This is the background for weighting filter like e.g. "A", "C" and CCIR 468-4 [8] which are intended to be used at a region of levels, where the filter shape is similar to the equal loudness contours.

A simple amplitude or power measurement, however, even with weighting filters, never can give a precise expression of the loudness. The main reason for this is that parts of the signal with different frequencies interact in a complex way, so that signals with the same amplitude or power, but with different bandwidths, are not perceived at the same loudness. The perceived difference may be as big as a factor of 4, corresponding to app. 20 dB in sound pressure level.

As an example of a loudness calculation some stationary sine and square waves at 1 kHz are analysed. Below is shown the relation between sound pressure level and loudness for a sine wave, a square wave with the same peak value (amplitude), and one with the same power (rms value) as the sine wave:

SPL	Loudness	Loudness	Loudness
Sine	Sine wave	Square wave same peak value	Square wave same power
94 dB	49.9 sone	98.4 sone	81.1 sone
84 dB	26.5 sone	52.5 sone	43.6 sone
74 dB	13.5 sone	28.0 sone	23.0 sone

The unit "sone" is expressing loudness on a linear scale, that is, twice as many sone means twice as loud. A rule of thumb from psychoacoustics is that 10 dB of level change corresponds roughly to a factor of two in loudness, and the loudness model confirms this. Notice that the wide-band square wave signal is always significantly louder than a sine wave of either same amplitude or power, upto twice as loud. This is actually one of the reasons why it may make more sense to add overtones to a signal to increase its loudness instead of just increasing its level. Instead of sone the unit "phon" may be used. This is a logarithmic unit similar to dB.

Loudness depends also on the duration of the signal, in a non-linear fashion, so for peak durations below 100 ms loudness decreases, but less than the signal energy would suggest. A 10 ms signal is perceived about half as loud as a 100 ms (or continuous) signal.

5

This relation is similar to the level dependency where a facor of 10 in power (10 dB) is perceived as a factor of 2 in loudness.

There are commercially available instruments for real-time measurement of loudness, e.g. [12, 13]. It is not a technical problem to define calibrated loudness mix environments like in film production.

2.3 Level statistics

Meters like the ones described above indicate instantaneous values, maybe with a short averaging time constant, but which level property should be used for characterising an audio signal?

1. The maximum peak value. This is essentially just a peak-hold function. Or it can be the maximum reading of a quasi-peak meter like in [4, 5].

2. A long time average like the Leq, equivalent level. This measure is often used in environmental and industrial noise measurements where the exposure to a time-varying sound is to be quantified. One hour of exposure to one sound intensity has the same Leq as two hours of a sound half as strong. It has been proposed to use the Leq with a simple filter to estimate the loudness of a cinema audio signal [7].

3. Mean and standard deviation, as known from general statistics. It may express mean level and dynamic range, in dB, sone or phon.

4. The distribution of levels over a time period, a histogram. In short time intervals, the maximum level is determined and classified in intervals of, say, 2dB. The number of occurrences of a certain level is graphically shown on a bar graph, a histogram. This is a very useful tool for characterising an audio signal dynamically [9, 10, 11].
In Figure 4-6 the level histograms of various source material is shown.

3. CONCLUSION

It should be our community's goal to conserve today's music and other audio content in the highest possible quality. Unfortunately several recordings from the past sound better than many recent ones.

Despite low static distortion in some of the current digital multi-band compressors, this paper has shown it is possible to generate very loud and distorted masters by using only the consecutive sample restriction.

Clearly, to obtain the goal of music conservation, new standards are needed.

One way to go would be to establish a set of gentleman processing rules that need to be followed in order for the CD to get a DRA stamp, a Dynamic Range Approval. Such a standard could be made with the help of the level measurement tools described in this paper.

The DRA could also address standardized loudness monitoring in mastering studios. Just like pre-press and film production have their standards for comparing audio and visual content.

We would hope for an audiophile organization like the AES to be initiating the project, but being one of the companies to put multiband processing power in the hands of many users, we feel an obligation to help get the discussion started.

REFERENCES

[1] E. Zwicker & H. Fastl (1990): Psychoacoustics - Facts and Models, Springer-Verlag, Berlin.

[2] International Standardisation Organisation (1975): ISO 532, Method for calculating loudness level, First edition.

[3] World DAB Forum (1999), http://www.worlddab.org.

[4] International Electrotechnical Commision (1991): IEC 268-10, Peak programme level meters, Second edition.

[5] International Electrotechnical Commision (1995): IEC 268-18, Peak programme level meters - Digital audio peak level meter, First edition.

[6] A. Oppenheim & R. Schafer (1975), Digital Signal Processing, Prentice-Hall, Englewood Cliffs, p. 239-241.

[7] I. Allen (1997): Are Movies Too Loud?, Presented at SMPTE Film Conference, March 22, 1997. http://www.dolby.com

[8] International Telecommunication Union (1986): Rec. 468-4, Measurement of audio-frequency noise voltage level in sound broadcasting.

[9] M. Ankerman et al. (1992): Aussteuerungsmesser mit Anzeige der Kurzzeit-Abtastwerte-Verteilung, ITG-Fachbericht 118, VDE-Verlag GmbH, p. 163-169.

[10] H. Jakubowski (1971). Analyse des Programmaterials, Nachrichtentechnische Fachberichte Band 41, VDE-Verlag GmbH, Berlin, p.194-204.

[11] TC|Works (1999), TC|Native L User Manual, TC|Works GmbH, Hamburg, Item no. 606077011.

[12] DK Audio MSD600C (1999), http://www.dk-audio.dk.

[13] Neutrik NC10 (1999), http://www.neutrik.com.

[14] International Telecommunication Union (1995): Rec. ITU-R BS.412-7, Planning standards for FM sound broadcasting at VHF, section 2.3.

[15] Bob Ohlsson in a reply to the pro-audio@pgm.com list
Subject: Loudness Metering, June 1st 1999.

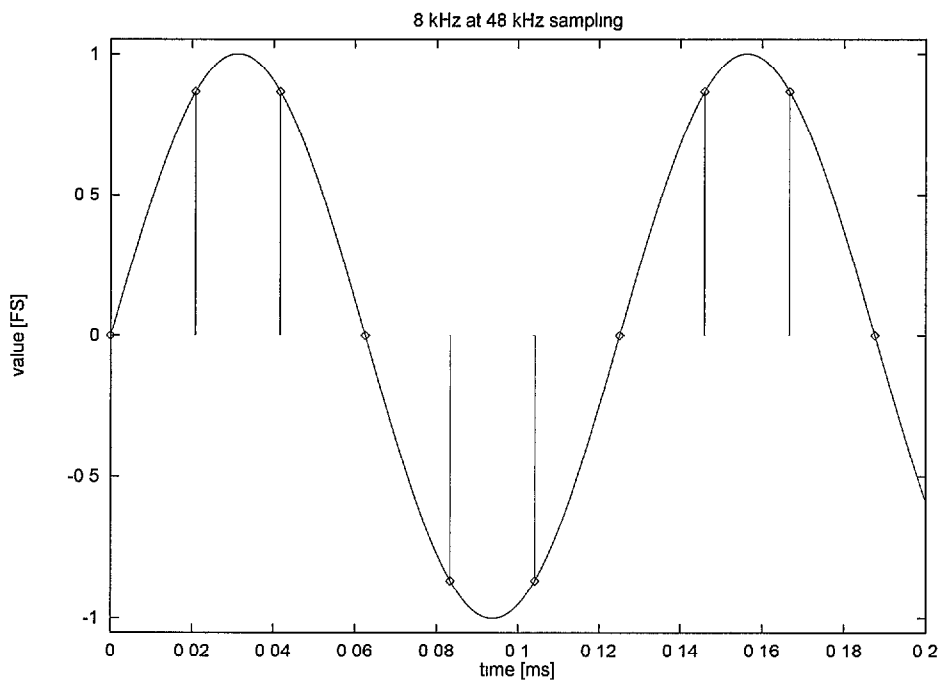[16] International Electrotechnical Commision (1990): IEC 268-17, Standard volume indicators, first edition.

Figure 1

An 8 kHz tone sampled at 48 kHz. Notice the repeated pattern of sample values. The digital peak level is 1.3 dB below the true peak level.

In this particular case a long decay time constant in the digital peak detector will not help getting the true peak level.
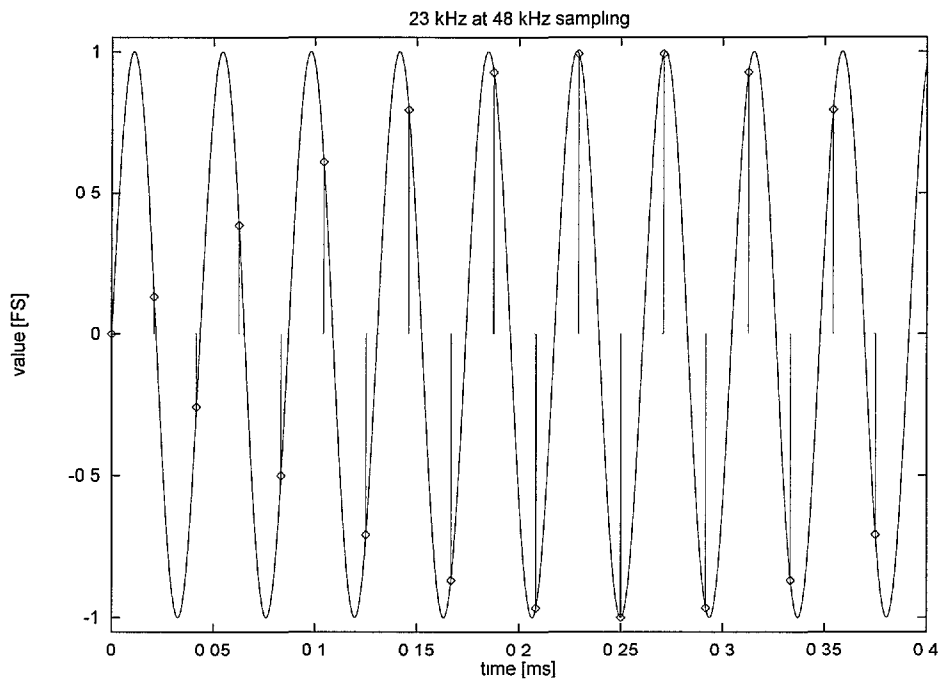
23 kHz at 48 kHz sampling

Figure 2

A 23 kHz tone sampled at 48 kHz. With this high input frequency the digitally read peak level varies periodically between zero and the true peak level.
A long decay time constant of the digital peak detector can compensate in this case.
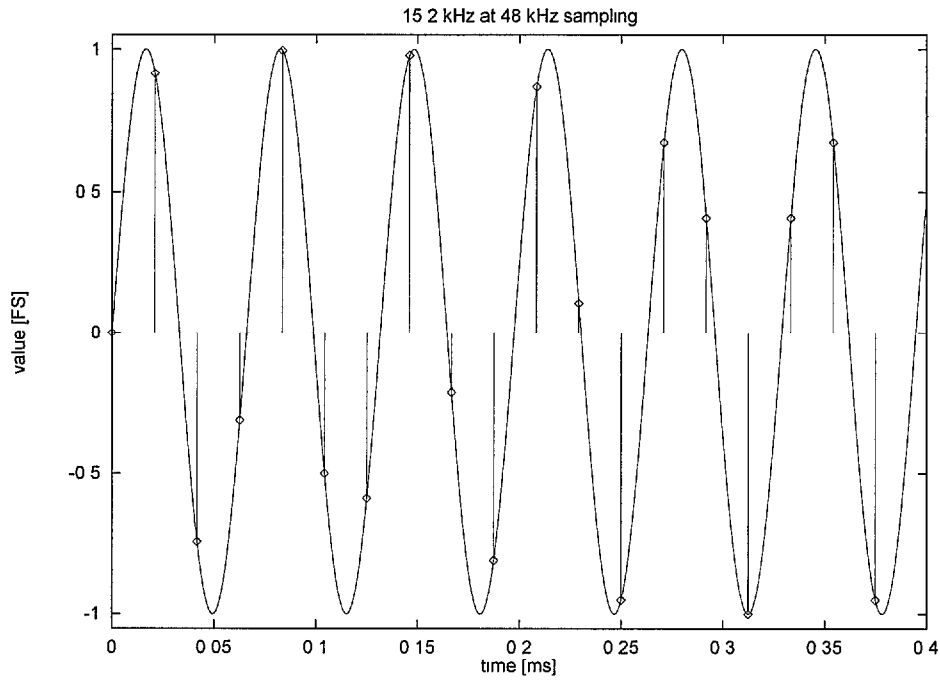
Figure 3

A 15.2 kHz tone sampled at 48 kHz. The digital peak level approaches the true peak level within a relatively short time window. This is typical for most real-life signals.
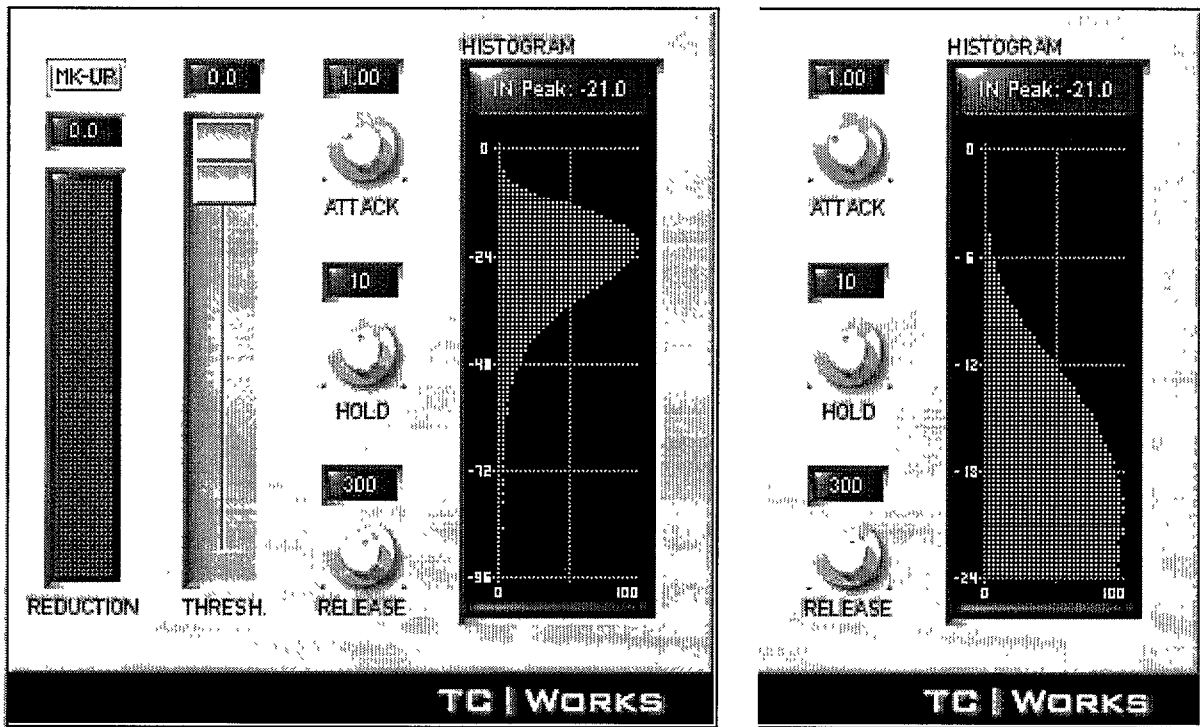
Figure 4

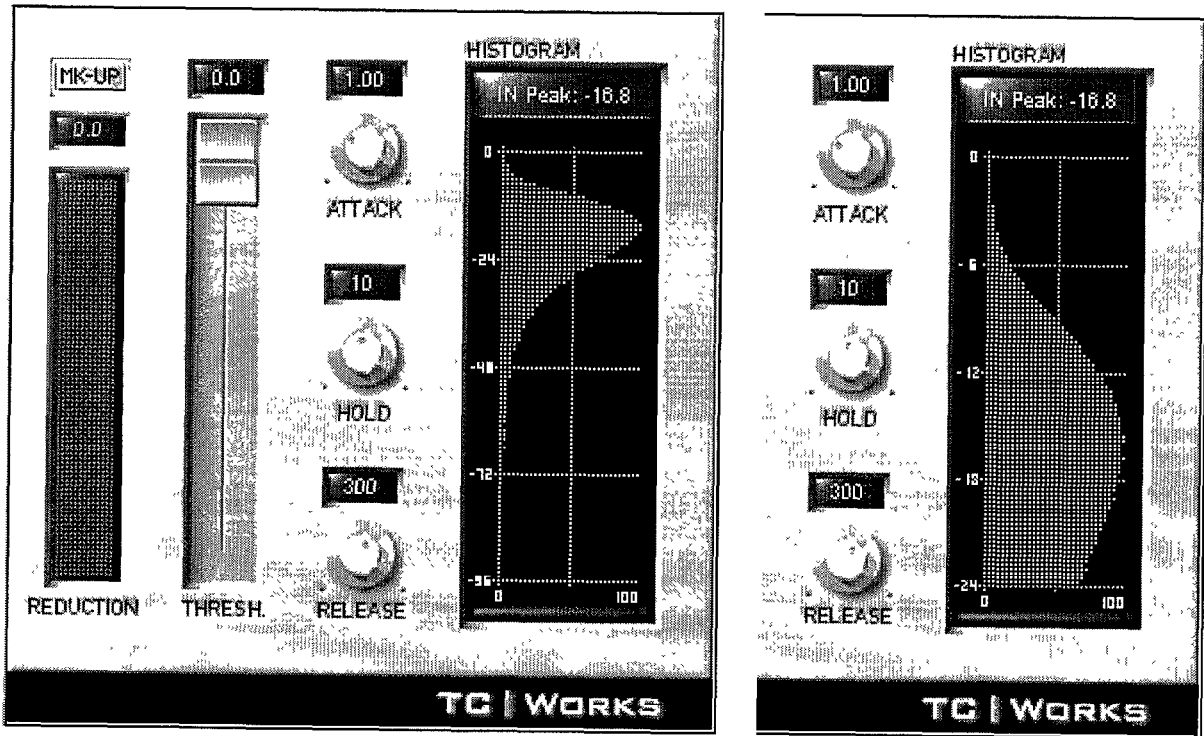Level histograms from normalized acoustic piano recording, 1992.

Figure 5

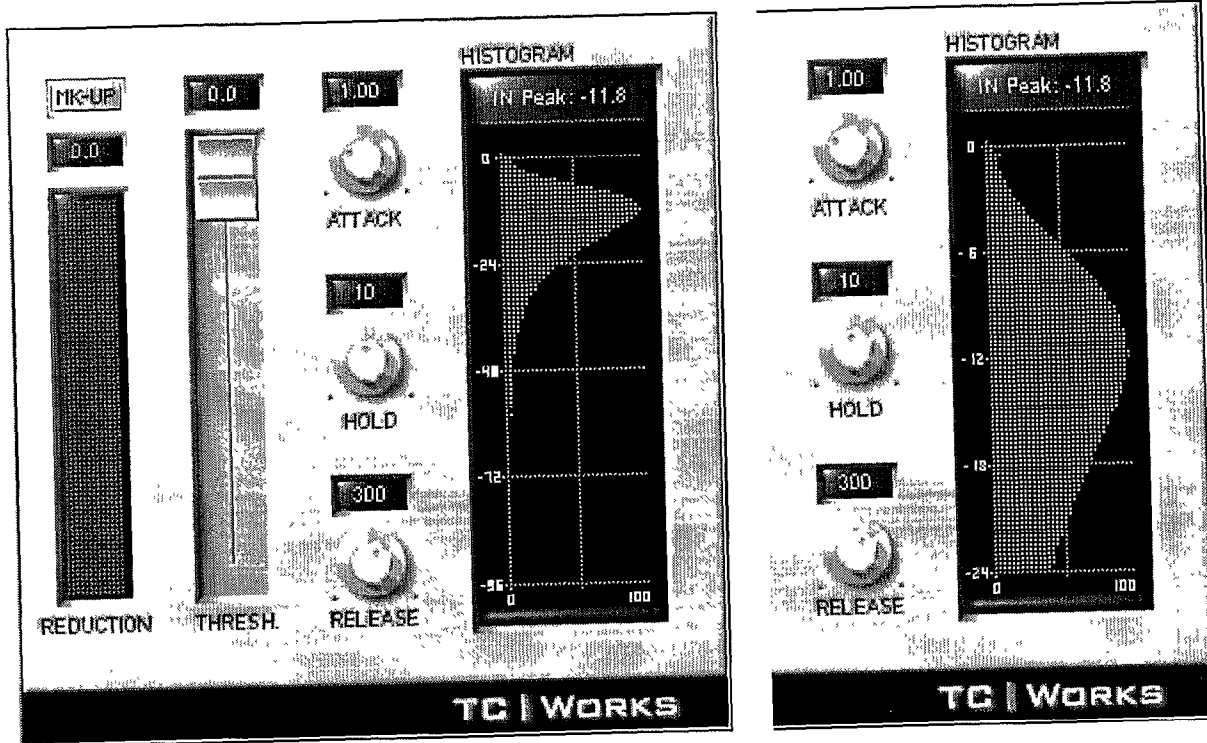Level histograms from normalized pop/rock recording, 1985.

Figure 6

Level histograms from normalized pop/rock recording, 1998.