The Distribution of Peak Energy in Recorded Music, and Its Relation to Magnetic Recording Systems*

JOHN G. MCKNIGHT†

Ampex Corporation, Redwood City, California

This is the author's second paper on this subject. The abstract from the first paper is reprinted from January, 1959, issue of the JOURNAL OF THE AUDIO ENGINEERING SOCIETY.

Measuring procedures and specifications currently used in the sound recording industry are inadequate. A weighting network should be used in objective noise measurement work to evaluate developments toward a lower noise system.

We know that the ear is mcst sensitive in the 1 to 6 kc region, and it is possible to use preemphasis in this region. Better signal-to-noise ratio may be obtained in 15 ips magnetic recorders by utilizing fully their present capabilities. Subjective listening tests show that a system with the 1 to 6 kc region pre-emphasized (the AME curve) is some 7 db quieter than the same system using the NAB curve, but does not show any audible increase in distortion.

Data is needed on the distribution of peak energy in music in order to design preemphasis on the basis of equal probability of overload at all frequencies. Previously published data is discussed, and new data is presented and discussed. The data show that energy distribution is very variable; the energy may be within plus or minus 3 db from 40 c to 16 kc; or the energy may be down as much as 30 db at 40 c and 6 kc.

INTRODUCTION

THERE ARE TWO limiting criteria for designing equalization in a recording system. One criterion is that of making a post-emphasis which will minimize the audible noise from the system, using at the same time a volume indicator to insure that the recording system will not be overloaded.[1]

The second criterion is that of designing the pre-emphasis so that the system will have equal probability of overload at all frequencies; then the volume indicator may be designed primarily for *balancing* levels.

In order to evaluate equalizations on the basis of this second criterion of equal probability of overload at all frequencies, it is necessary to have adequate data on the frequency-distribution of peak energy. In the present case we are interested in music only.

Previous work was reviewed and found to be inadequate for our purposes;[2] therefore, we undertook our own measurements of the distribution of peak energy in recorded music. The apparatus and recordings are described here, and a few of the spectrum analyses presented.

REVIEW OF PREVIOUS STUDIES

A review of the definitive Sivian, Dunn and White analyses is very interesting.

Fig. 1 shows a summary of their spectrum analyses of a 75 piece orchestra in a theatre. The heavy curve shows the only spectrum reprinted in Olson's *Acoustical Engineering*;[3] the implication is that this is a typical spectrum; note, however, that the other spectra have much greater high-frequency content. These upper spectra are similar to some of the data in this new study, and show that the general engineering interpretation of Sivian, Dunn, and White's data (as selected by Olson) may not be correct.

Several problems arise in interpreting the Sivian, Dunn and White data. First, their data is unrelated to VU meter readings, having been done long before the development of

* Received April 20, 1959. Presented February 20, 1959, at the Sixth Annual Western Convention of the Audio Engineering Society, Los Angeles, Calif.

† Senior Engineer, Research Division. Now Manager, Advanced Audio Section, Professional Audio Department.

1. J. G. McKnight, "Signal-to-Noise Problems, and a New Equalization for Magnetic Recording", *J. Audio Eng. Soc., 7*: 5-12 (Jan. 1959)

2. Sivian, Dunn, and White "Absolute Amplitudes & Spectra of Certain Musical Instruments and Orchestras," *J. Acous. Soc. Amer.* 2: 330-371 (January 1931).

3. H. F. Olson, *Acoustical Engineering,* D. Van Nostrand Co., New York, 1957, p 589, Fig. 12. 35.

the VU meter. Second, their spectra are for one orchestra in one pick-up situation, playing one type of musical arrangement. This is a very small sample on which to base conclusions about all recordings of all music. Also, we would like to know the spectrum of the signal from a commercial recording session, where there may be many microphones placed close to the individual instruments, individual microphone equalization, and other techniques for "better-than-being-there" sound.

Analyses were done in 1953 at Ampex Corporation by George Brettell.[4] Fig. 2 shows a summary of his spectra. Again, considerable high-frequency energy is shown, but too few samples are here to draw any general conclusions.

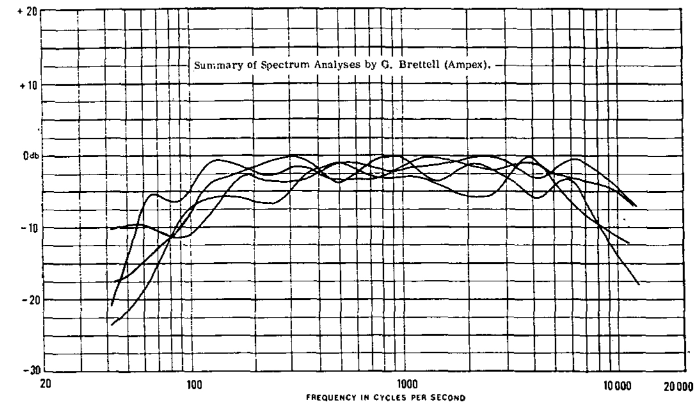


FIG. 2.

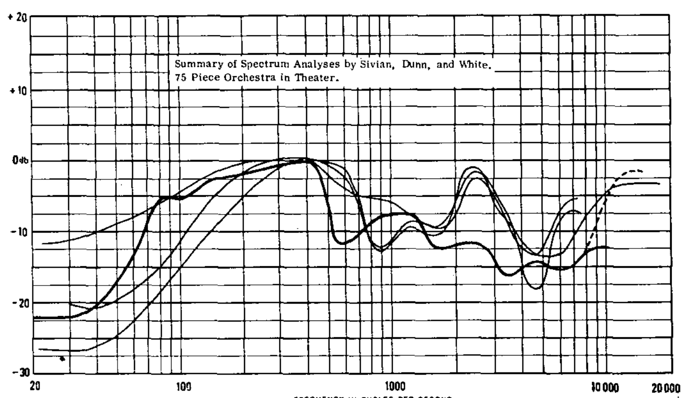


FIG. 1.

We question the data in J. P. Overley's article, "Energy Distribution in Music,"[5] on several accounts. Although he describes the need for peak readings, his instruments would read only average amplitudes. The data was taken from phonograph pressings. And the published data is averaged from many readings. We will see later in the data of this paper that "averages" have little meaning—we are concerned entirely with individuals. The Overley data shows much less energy above 3 kc than any of the other studies discussed in this paper.

Our review has shown us, then, that the general engineering conception of energy falling at high frequencies does not follow from the Sivian, Dunn, and White data; both their and the Brettell spectra show energy beyond 10 kc with amplitude within 3 to 5 db of the mid-band energy.

THE NEW SPECTRUM ANALYSES

The need was now for a large number of spectrum analyses to show the variation and limits of spectra from recordings made of many different types of groups, in many different recording situations, with many different kinds of musical arrangements being played.

First, it was necessary to design and construct equipment which would take the spectrum analysis data accurately and easily. This required a memory device, so that the recording to be analyzed could be played one time, and the complete spectrum read. Fig. 3 shows the block diagram of the audio spectrum analyzing system. The tape reproducer and the line amplifier feed both an attenuator and VU meter; and also through a transformer into a set of filters, which divide the audio spectrum (40 c through 16 kc) into 27 bands, each one-third octave wide. The 27 filters plus one unfiltered channel feed a 28-channel "maximum peak" memory unit, to be described.

In operation, the memory unit was first cancelled; the program to be analyzed was played, during which time the VU attenuator was adjusted. Then the spectrum, the total peak level, and the VU attenuator (giving VU level) were read—and that spectrum analysis was completed.

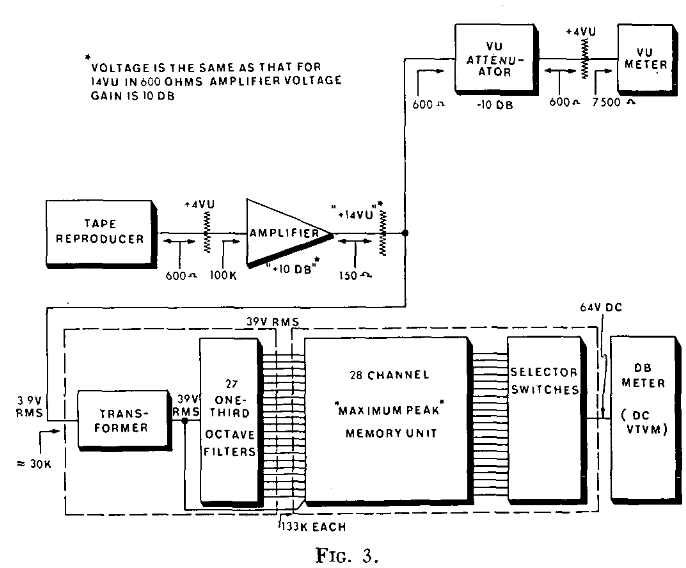


FIG. 3.

4. G. Brettell, Ampex Corp. unpublished engineering data (1953).
5. J. P. Overley, "Energy Distribution in Music", *IRE, Trans. on Audio,* 5: 120-123 (September-October 1956).

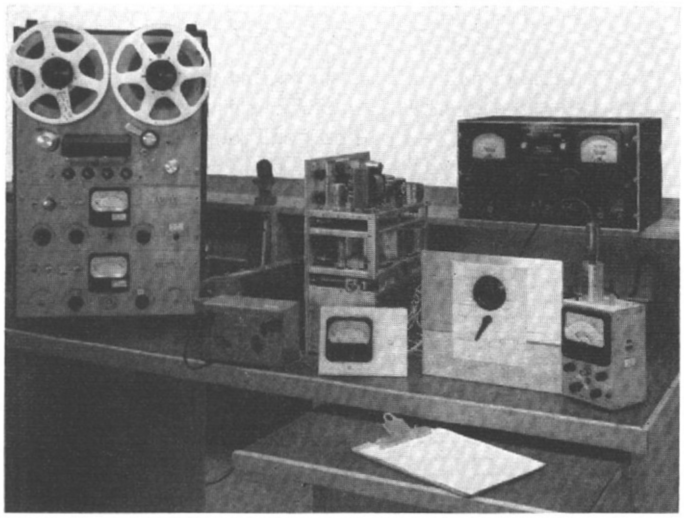


FIG. 4.

Fig. 4 is a photograph of the equipment: tape reproducer, line amplifier, VU attenuator and meter, filter, memory unit and its power supply, and the VTVM used as a db meter from the memory.

The individual "maximum peak" memory units are simple and straight-forward. Fig. 5 shows a typical channel. An input level calibrate control feeds a pentode amplifier which is followed by a cathode follower, which, in conjunction with the 3A2 rectifier, charges the capacitor. The time constant of the charging circuit is such as to charge the capacitor to within 1 db of true peak amplitude in about 1 ms. Following peaks of smaller amplitude have no effect; larger ones charge the capacitor to the new peak value, regardless of previous charge. The back resistance of the 3A2, in parallel with the input resistance of the meter used for readout, is such as to discharge the capacitor about 1 db in 10 minutes.

Fig. 6 shows the memory unit (clip-lead connected to the one-third octave filter), the output meter, and the memory unit power supply. The input level calibrate controls, pentode-triode amplifiers, rectifiers, capacitors, and selector switches are visible on top of the chassis.

The material to be analyzed was chosen from what we will call "original recordings" and "studio recordings." We

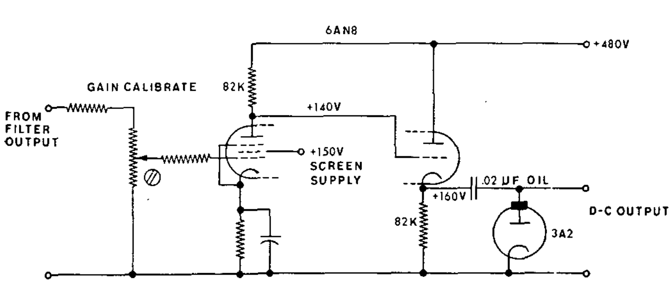


FIG. 5.

recorded most of the "original recordings" with a single condenser microphone, so placed as to make what we considered a good sounding recording. No special "better-than-being-there" techniques were used, and we feel the recordings are representative of the sounds actually made by the instruments.

The "studio recordings" are tapes made at the various major U.S. studios, and represent what may be fed into tape recorders, including the effects of multiple close microphones with individual equalization, echo chambers, etc.

We made analyses of "original recordings" of 34 different sections from 18 entirely different recording situations; and of "studio recordings" of 28 samples from 20 different situations.

We will show 15 analyses here, selected to show the *variation* encountered in "original" and "studio" recordings, for

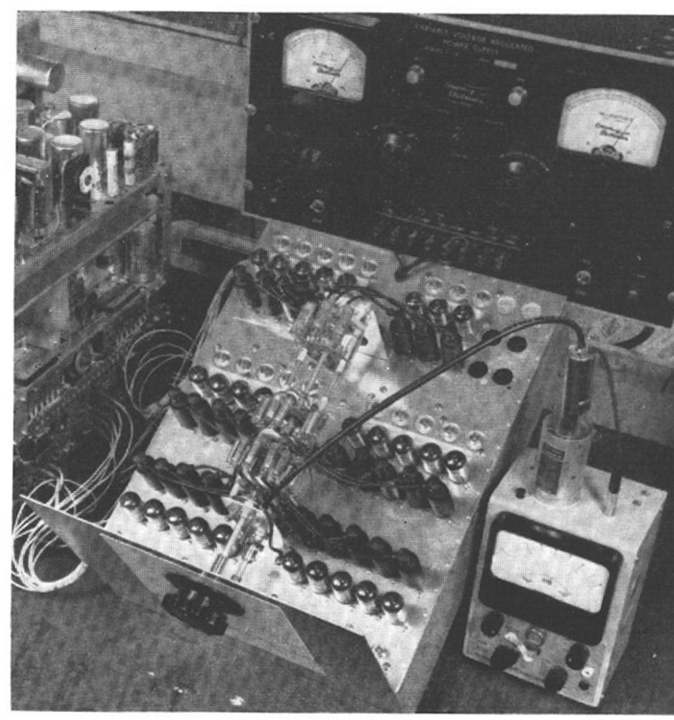


FIG. 6.

symphony orchestra, chamber music, pipe organ, dance, jazz, and popular music. We will also show a summary of the peak factors, and a summary of the limits of the spectra encountered. We feel that the concept of an "average" spectrum is not very useful. If it rains today, you'll get wet even if the "average" weather for today is supposed to be "dry"!

The discussion will be directed toward the high frequencies, but similar variation is found in the low frequency energy.

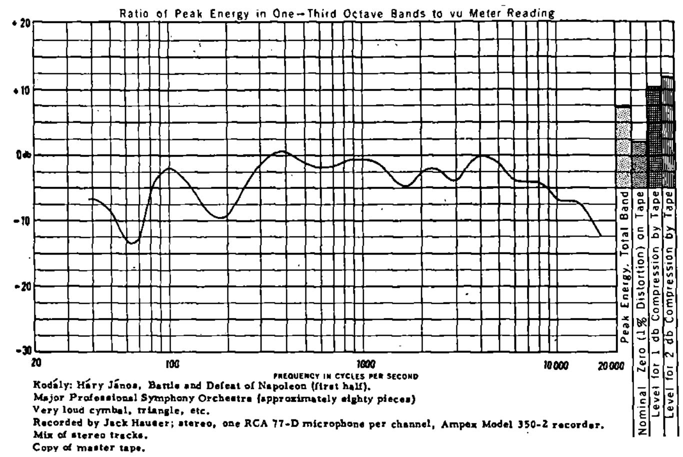


FIG. 7.

Figs. 7, 8, and 9 show three analyses of full symphony orchestras—original recordings of a professional and a university orchestra, and a studio recording of a professional

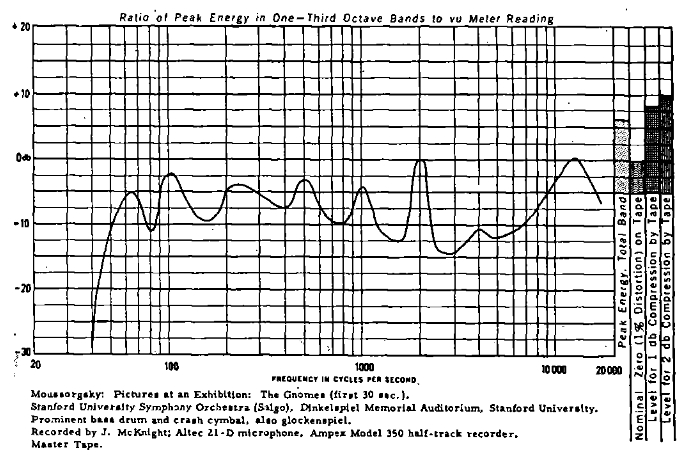


FIG. 8.

orchestra. These were chosen to show a large amount of high frequency energy. (The curves are all plotted relative

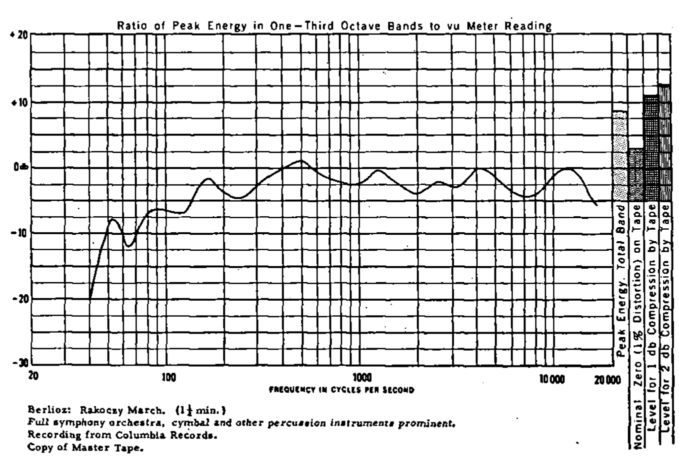


FIG. 9.

to the VU meter reading for each selection as "zero." The first bar at the right of each graph shows the peak energy in the total band. The others show where the 1% total harmonic distortion level was, and where 1 and 2 db of compression in mid-band by the tape would occur. These will be discussed in a future paper.)

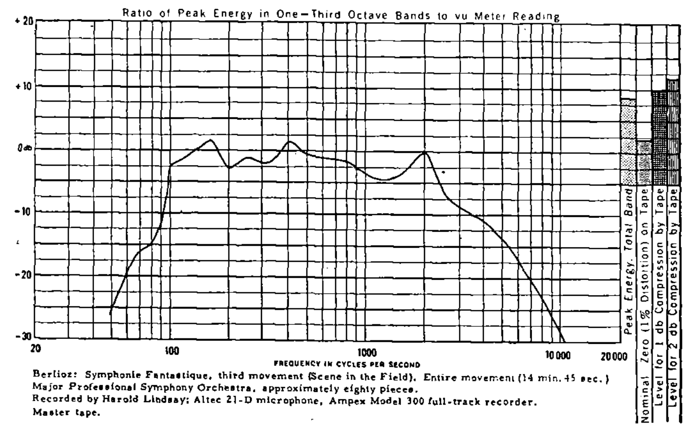


FIG. 10.

On the other hand, Figs. 10 and 11 show analyses of professional full symphony orchestras. Here we see very little high-frequency energy.

Figs. 12 and 13 show two original recordings of chamber music; Fig. 12, with a tambourine, shows appreciable high-frequency energy; Fig. 13, a baroque orchestra of strings with some winds, shows relatively little high-frequency energy.

Fig. 14 shows a church pipe organ. Note the large amount

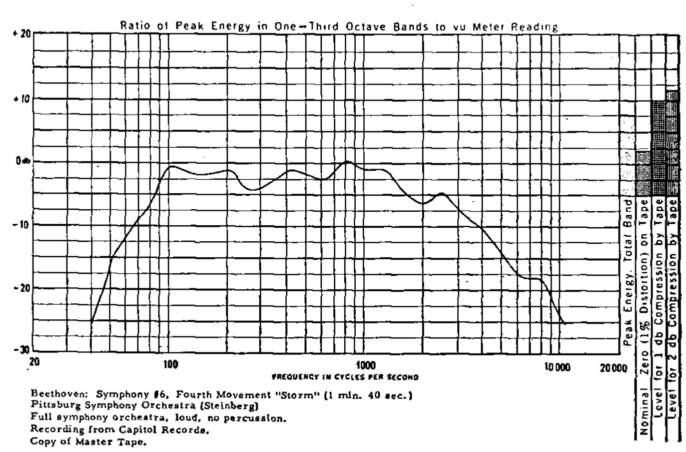
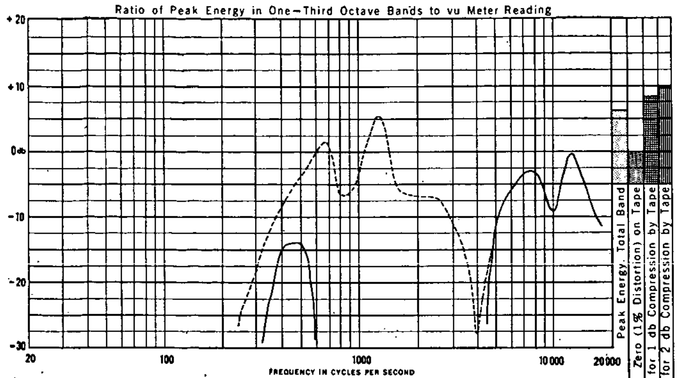


FIG. 11.

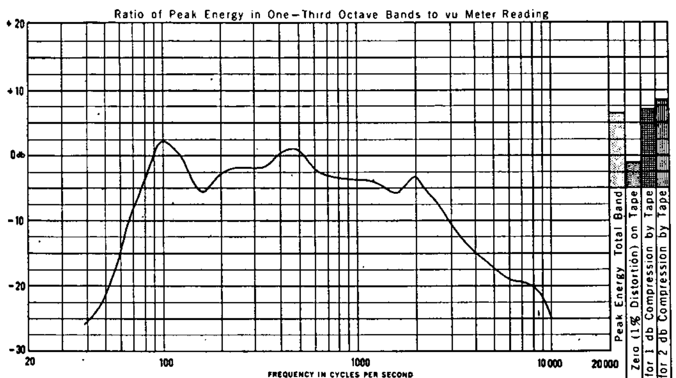
of low-frequency energy, with a steady slope downward toward higher frequencies. Recordings of a theater pipe organ, despite the percussion and other effects, had similar spectra.

Figs. 15 and 16 show original recordings of a large pro-



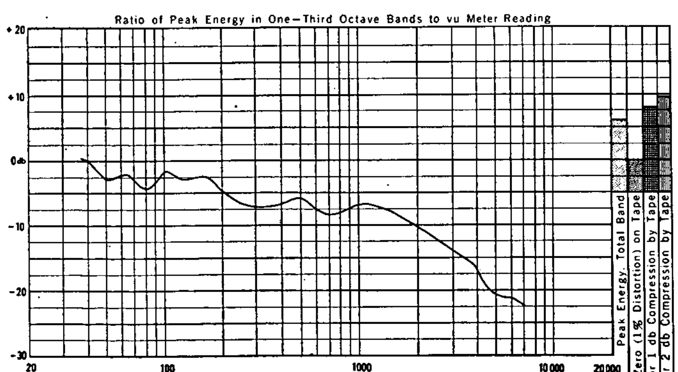
Eleanor Ogden: Disappointment, or When is Whitsunday Anyway? Entire selection (1 min. 30 sec.)
Chamber group at Stanford University. Alto, alto recorder, harpsichord, tambourine.
Solid curve, opening 10 seconds, tambourine only. Broken curve, entire selection.
Recorded by J. McKnight; Altec 21-D microphone, Ampex Model 350 full-track recorder.
Master Tape.

FIG. 12.



J. S. Bach: Brandenburg Concerto #2. First movement (last 3½ min.).
Soloists and Chamber Orchestra of the Telemann Society, New York.
Concertino: alto recorder, oboe, modified clarinet for clarino, solo violin.
Ripieno: one each, first and second violins, viola, cello, bass. Carl Fischer Hall, New York City.
Recorded by J. McKnight; Telefunken U-47 microphone, Ampex Model 600 full-track recorder.
Copy of the 7½ ips Master Tape.

FIG. 13.

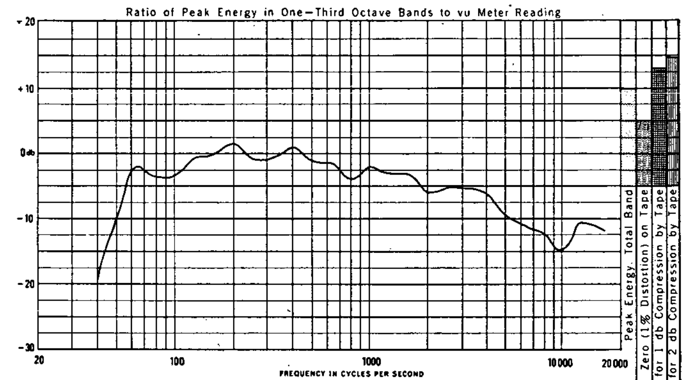


Karg Elert: Now Thank We All Our God. Entire selection (3-1/4 min.).
Herbert Nanney, pipe organ, Memorial Church, Stanford University.
Recorded by Harold Lindsay; Altec 21 microphone, Ampex Model 300 full-track recorder.
Copy of Master Tape.

FIG. 14.

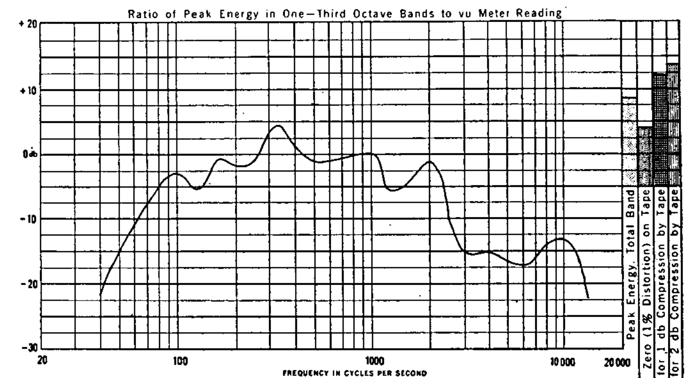
fessional jazz band in concert (including cow bells and all!) and a small non-professional dance band. Fig. 15 shows a medium amount of high-frequency energy, Fig. 16, less high-frequency energy.

Fig. 17 and 18 show two studio recordings of professional dance bands. Fig. 17 shows a strong peak in the 5 kc region, from a muted trumpet solo. Fig. 18 is of a dance band with solos by piano, muted trumpets, saxophones and strings. The spectrum is practically flat to 10 kc.



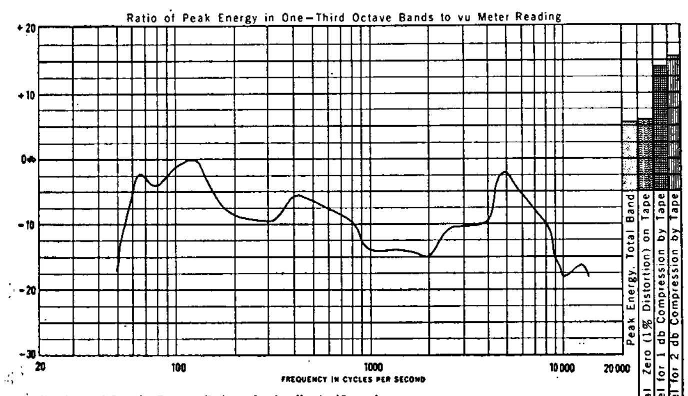
Concert Jazz, Entire selection (4 min.)
Large professional dance band, recorded in concert.
Stereo, one Altec 21-C microphone per channel, Ampex 350-2 recorder. Right hand channel only.
Master tape.

FIG. 15.



Dance Music (4 min.).
Nine piece dance band, non-professional.
Recorded by C. Posey; Altec 21-C microphone, Ampex 350 half-track recorder.
Master Tape.

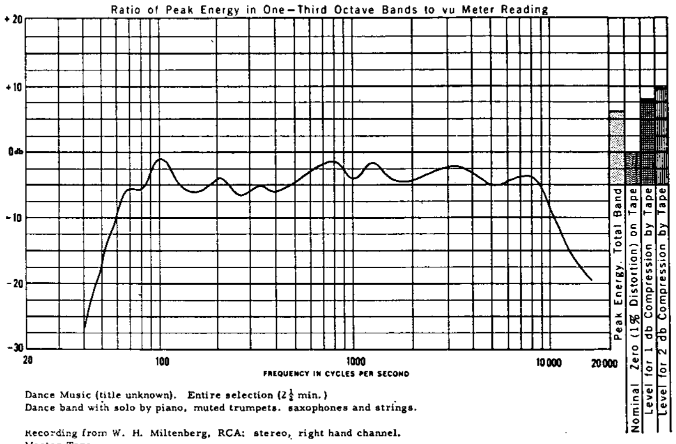
FIG. 16.



You Stepped Out of a Dream. Entire selection (1 min. 15 sec.)
Ray Anthony Dance Band.
Low level, with muted trumpet solo.
Recording from Capitol Records.
Copy of master tape.

FIG. 17.

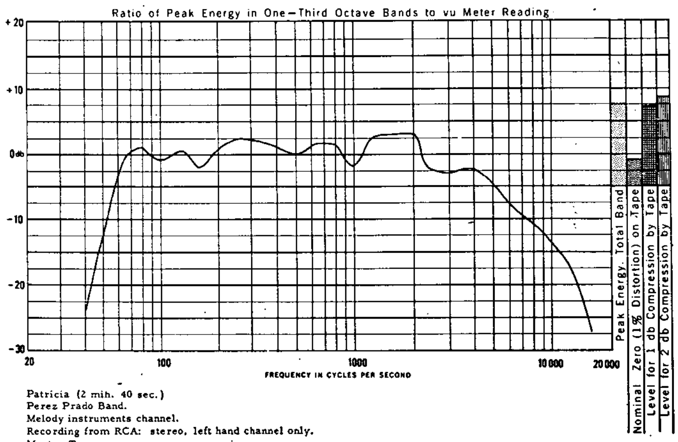
Figs. 19 and 20 show the two different tracks of a stereo recording (neither track contained any sound from the other) of a studio master tape of the Perez Prado band, playing "Patricia." The melody instruments channel (Fig. 19)



Dance Music (title unknown). Entire selection (2½ min.)
Dance band with solo by piano, muted trumpets, saxophones and strings.

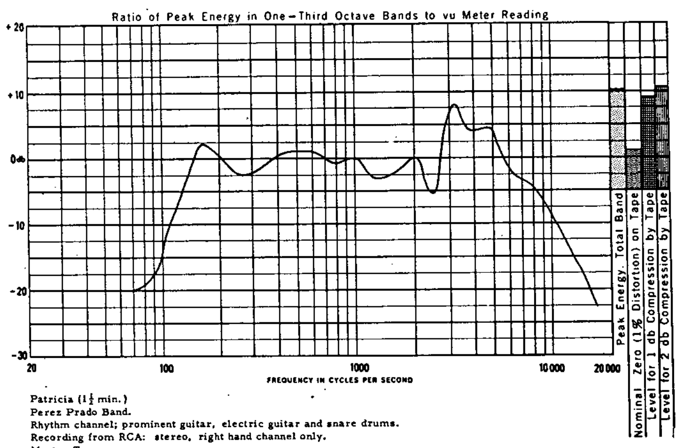
Recording from W. H. Miltenberg, RCA: stereo, right hand channel.
Master Tape.

FIG. 18.



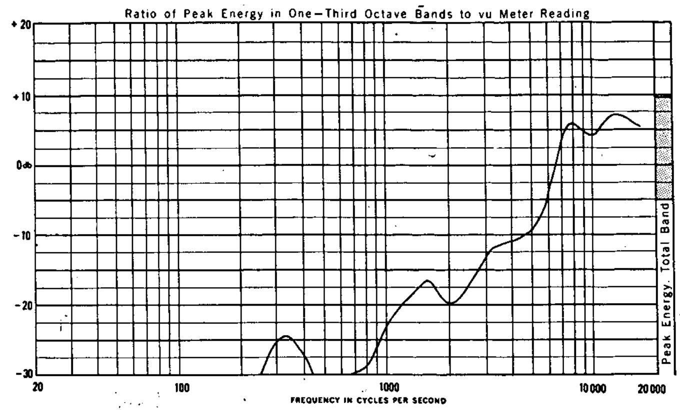
Patricia (2 min. 40 sec.)
Perez Prado Band.
Melody instruments channel.
Recording from RCA: stereo, left hand channel only.
Master Tape.

FIG. 19.



Patricia (1½ min.)
Perez Prado Band.
Rhythm channel; prominent guitar, electric guitar and snare drums.
Recording from RCA: stereo, right hand channel only.
Master Tape.

FIG. 20.



Keys jingling 1 ft. from Altec 21-D microphone. Analyzed directly, not recorded.

FIG. 21.

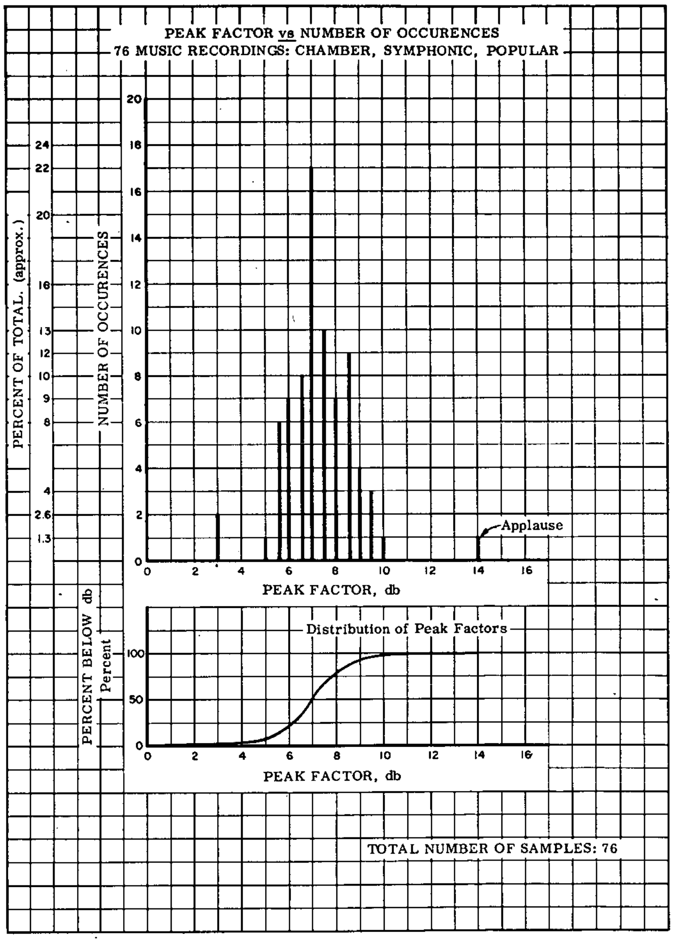


FIG. 22.

shows almost constant energy to 5 kc, falling sharply above. But the rhythm instruments channel (Fig. 20) has an 8 db peak at 3 kc, from a prominent electric guitar, and is down only 5 db at 8 kc.

Fig. 21, for interest, is keys jingling, one foot away from an Altec 21-D microphone, analyzed directly (not recorded). It is easy to see why key jingling is hard to record and reproduce.

Although this paper does not discuss peak factors individually, Fig. 22 shows a summary of the peak factors of 76 music recordings. (The peak factor is the amount by which the true peak energy exceeds the VU meter reading.) The values of peak factor are seen to be from 3 to 10 db for music, with 7 db being the most common value. If in our equipment design we allow for peaks 10 db over the VU meter "zero" level, we will include all the music studied here. (This assumes one adjusts the level so that the VU meter usually swings to zero, with occasional peaks to plus one or two. Nothing is left for those who "pin the meter" frequently, as too many do.)

Fig. 23 shows several of the analyses drawn on one graph, to re-emphasize the variability in spectra from various recordings. *Realize especially that all of these give the same reading on a VU meter.*

Suppose we were to draw *all* of the analyses, in similar fashion, on one graph. Then we would get an area, as shown in Fig. 24. Again, all of the curves which would fall in this area give the same reading of zero on a VU meter. (The Sivian, Dunn, and White data comes near to the top of this area, also).

What do we conclude from all of this data? Mainly, that a usable "average" spectrum for music does not exist. If we have a record/reproduce system using a VU meter for its level indicator and if we are to allow for the maximum high- and low-frequency energy found in these studies, then the system must be able to handle all signals from at least 40 c to 16 kc at the VU "zero" level. In general, this means little use of pre-emphasis, if we are to be able to pass all frequencies at zero level.

Why does our present equipment work at all, having been designed on the assumption of much less high-frequency energy than this data shows? First, not all material has the maximum energy—many recordings will be in the middle to bottom of the area. Second, a fair amount of overload may exist without becoming objectionable. Third, it is common practice to listen to tapes, to see if overload does exist; if it does, the level is lowered. Fourth, we *don't* always get away with it! Fifteen-ips NAB equalized tape is almost always satisfactory, since it will pass all frequencies at zero level. But disc and 7½ ips tape often have noticeable high-frequency breakup. And 3¾ ips tape for music can be very

(*Continued on Page 80*)

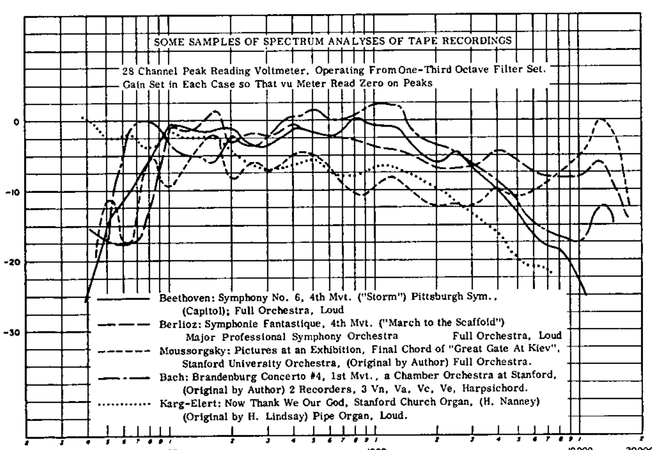


FIG. 23.

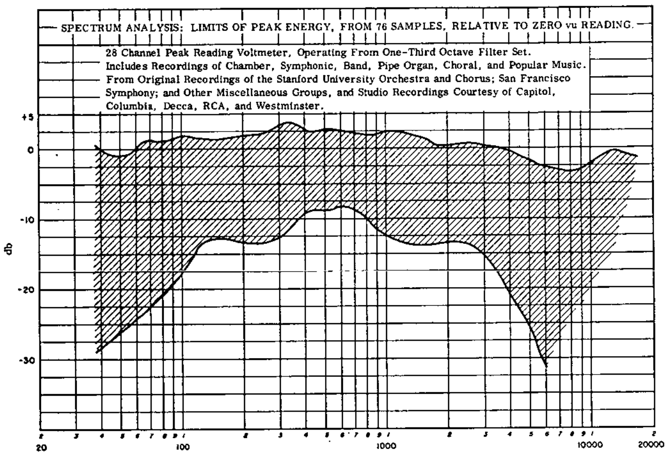


FIG. 24.

THE AUTHOR

John G. McKnight, who was born in Seattle in 1931, studied at Stanford University and received his B.S. in Electrical Engineering there in 1952.

In 1953 he worked with Ampex Corporation on the development of cinemascope-stereophonic sound equipment. After several years of other work, he returned to Ampex, where he is now Senior Engineer in the Research Division. He spent the years 1953-1956 on the engineering staff of the Armed Forces Radio Service in New York. At the same time he worked as development engineer for the Gotham and the Narma Audio Development Companies. He has always been interested in the problems of magnetic recording, specifically as they concern music.

Mr. McKnight is a member of the Audio Engineering Society and an affiliate member of the IRE.

From (9): $A' = \frac{A_{1.2}}{1 + A_{1.2}\beta} \sim \frac{1}{\beta}$ if $A_{1.2}\beta >> 1$.

The gain of the tube has to be 50 in this case.

Therefore, $\beta = \frac{R_0}{R_0 + R_2} = \frac{1}{50}$ and $R_2 = 112.10^3$ ohm.

We used: $R_2 = 120.10^3$ ohm.

With the given values and equation (2), we may solve $A_{1.2} \sim 1170$ and from (8) the Nyquist factor $N = 26.4$. From the equations (7) and (9), we may solve T_1 and T_2 if the cutoff frequency and N are known, making $Q = 1$.

$$T_1 = \frac{1}{4\pi f_o}\left[1 \pm \sqrt{1 - \frac{4}{N}}\right] \quad (7a)$$

$$T_2 = \frac{1}{4\pi f_o}\left[1 \mp \sqrt{1 - \frac{4}{N}}\right] \quad (9a)^2$$

Substituting $f_o = 25$ c/s and $N = 26.4$ we find:
$T_1 = 0{,}25.10^{-3}$ sec. (or $6{,}1.10^{-3}$ sec.)
$T_2 = 6{,}1.10^{-3}$ sec. (or $0{,}25.10^{-3}$ sec.)
From equations (3) and (4), C_1 and C_2 are solved:
$C_1 = 222\ pF$. We used: $C_1 = 220\ pF$
$C_2 = 38{,}5\ nF$. We used: $C_2 = 39\ nF$
R_3 consists of the bass potmeter of 2 MOhm and the resistors of 150 and 39 KOhm in series. Substituting this value for R and $\omega_o = 2\ \pi\ 25$ in (12) we may solve C_3:
$C_3 = 29\ nF$. We used: $C_3 = 22\ nF$, being the nearest practical value.

[2] Because a time constant should be a real number the factor N should comply with the condition $N \geq 4$.

NOTE: It is desirable to use the smallest T in the first triode, for when using it in the second triode, the voltage on the anode of this triode will increase too much, as the output voltage of the second stage as to be constant up to 25 c/s causing distortion. The voltage on the anode of the first triode is smaller.

THE AUTHORS


J. RODRIGUES DE MIRANDA


H. VAN DEN KERCKHOFF

J. Rodrigues de Miranda, who was born in 1905, received his schooling in Holland. In 1927 he graduated from the technical University of Delft as an electrotechnical engineer. He spent the next eleven years managing the technical department of a wholesale importing company. He joined N. V. Philips' in 1938, and he has remained there doing development work in acoustics. He is now the Senior Engineer and Manager of the Philips' Reproduction Advisory Group. His memberships in numerous professional societies include the AES and IRE in the United States.

H. van den Kerckhoff, born in 1926 in Indonesia, joined Philips' and Mr. Miranda in 1956, the year he graduated from the Technical University at Delft. In his laboratory investigations on acoustics, he has concentrated on the development of amplifiers and preamplifiers.

(Continued from Page 71)

bad, if a great deal of high-frequency energy is present, or very good if high-frequency content is controlled.

CONCLUSION

Does this data indicate the need for any action on the part of magnetic recording equipment design engineers? If our present recording equipment had signal-to-noise ratio to spare, we could simply allow for the worst case at all times. Since we do not have this spare signal-to-noise ratio, we can pursue the alternate course mentioned at the beginning of this paper: namely, design a post-emphasis for minimum audible noise; then use a level indicator which will insure that the system will not be overloaded. Such a device *might* be the equalized, peak-reading volume indicator. This, however, has the disadvantage that the meter does not indicate level balance. But the VU meter does not indicate overload—and we must either decide which is more important or use two meters, one for balance, one for overload. (What we propose is not new, but has not been widely applied as yet.) This would in any case require additional complication and work on the part of the manufacturer and the operators, but we feel that this might lead to a method for increasing the signal-to-noise ratio without necessarily increasing the distortion of the recording. Considerable evaluation work is needed to see if this method of metering will fulfill its promise.