Proceedings of The Institute of Acoustics

THE AMPLITUDE DENSITY OF RECORDED MUSIC

A.Barr, M.S.Everett and D.C.Martin

School of Applied Physics, Coventry Lanchester Polytechnic,
Priory Street, Coventry, CV1 5FB, England.

INTRODUCTION

The testing of audio equipment often involves the use of signals, such as sine waves, which bear little resemblance to typical music programmes. For some of the tests a more representative signal in the form of an arbitrarily chosen music programme may be used but results may be misleading due to lack of standardisatic More meaningful results should be obtained using a reproducible standard test signal which has all the salient characteristics of a typical music programme.

To realise such a standard it is necessary to examine both the amplitude and spectral densities of a wide range of music, together with their variations with time. It should then be possible to produce statistical specifications for various music types for which models may then be produced. The present work is aimed at providing the necessary data to enable a range of recorded music to be modelled in terms of its amplitude behaviour and hence provide a basis for a standard signal source.

GENERAL MUSIC SIGNAL CHARACTERISTICS

Although exhibiting a certain degree of periodicity in the short term, music signals could be considered to be both random and stationary over a measurement interval for which the music is continuous. If comparison is made with white noise it is seen that whereas the latter has a Gaussian distribution of instantaneous amplitudes and a flat spectral density, music signals have, in general, a somewhat higher probability of attaining larger amplitudes and a maximum spectral density in the mid-frequency band. Music signals are rarely continuous, however, and may exhibit severe temporal discontinuities which further modify their characteristics.

Although many workers have examined the relative spectral content of music [1,2] and the probabilities of occurrence of signal peaks [3,4], little work has been done in modelling long-time signal behaviour. Indeed, reference is still made to the seminal research of Sivian et al [5] performed in the 1930s. In their work, amplitude distributions in different frequency bands were produced for various solo instruments and orchestras. The results of this study were later modelled by Daugherty and Greiner [6] giving a two element density curve (represented by equation 1) consisting of a central Gaussian component associated with the smaller amplitudes, superimposed on an exponential distribution associated with the larger amplitudes.

$$p(x) = 1.51 \exp(-5.12x^2) + 0.381 \exp(-0.932x) \qquad (1)$$

This model was similar to that reported by Davenport [7] although this was for a study of speech signals. Here, the large amplitude exponential region was ascribed to direct "voiced sounds" whilst the small amplitude Gaussian region was thought to be due to reverberant "unvoiced sounds" and system noise.

THE AMPLITUDE DENSITY OF RECORDED MUSIC

In view of the rather limited choice of signals considered in the above studies and the fact that only "live" signals were analysed, it was felt that an investigation was required into today's wide variety of recorded music.

TEST PROCEDURE

Using the measurement system described elswhere [8], the instantaneous amplitude distributions of a variety of commerically available recordings were obtained. This involved the analysis of over one hundred complete compact discs as well as measurements on a large number of shorter individual tracks or movements from these. A roughly equal split between "popular" and "classical" material was made in the choice of music examined, these loose descriptive terms covering such music types as pop, jazz, heavy rock and popular vocalists in the "popular" category, and orchestral, piano, chamber and choral music in the latter category.

Amplitude density curves were produced for each piece of music analysed. Since many of the observed differences between signals were in the regions of the low-probability peak amplitudes, to facilitate comparison between the curves, they were produced on log probability versus amplitudes axes. In addition, statistical parameters, including long-time rms and crest factor, were calculated from the resulting distributions enabling rather more concise numerical comparisons to be made. It is recognised here that crest factor ratings are not strictly applicable to random signals. In this study, however, it was assumed that the peak amplitude of the signal could be taken as the largest instantaneous amplitude measured. To provide a further measure of the dynamic qualities of the signals, calculations of the kurtosis of the density functions were made.

Table 1 Typical Results of whole album analyses

MUSIC	V(rms) (σ)	Crest Factor	Skewness	Kurtosis
Dire Straits: Brothers in Arms	0.098	19.39	+0.09	9.53
Michael Jackson: Thriller	0.182	8.78	+0.05	6.12
Kiri Te Kanawa: Chants D'Auvergne	0.091	16.61	+0.10	10.03
Bloomstedt Bruckner: Symphony No.4	0.102	15.01	-0.02	14.76
James Newton Howard and Friends	0.114	20.80	+0.36	16.87

Table 2 Typical Results of individual track analyses

MUSIC	V(rms) (σ)	Crest Factor	Skewness	Kurtosis
Dire Straits: So far away	0.114	16.29	+0.15	6.54
Michael Jackson: Billie Jean	0.184	7.87	-0.05	5.27
Kiri Te Kanawa: Lou Coucut	0.132	10.70	-0.04	7.59
Bloomstedt Bruckner: excerpt	0.208	13.60	-0.05	6.74
J.N.H. & Friends: L'daddy	0.128	20.20	+0.86	26.38

THE AMPLITUDE DENSITY OF RECORDED MUSIC

RESULTS

The results given in Tables 1 and 2, and illustrated in Figures 1, 2 and 3 are a selection chosen to be representative of the results as a whole, and not necessarily to show extreme or "worst case" characteristics.

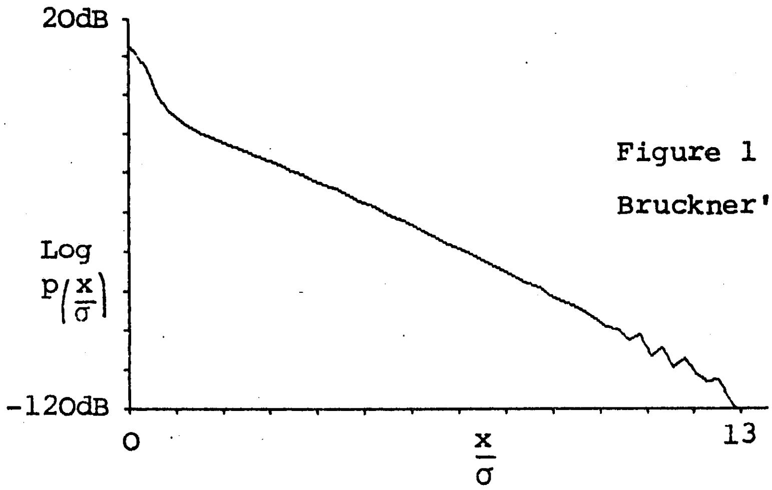

Figure 1

Bruckner's 4th Symphony

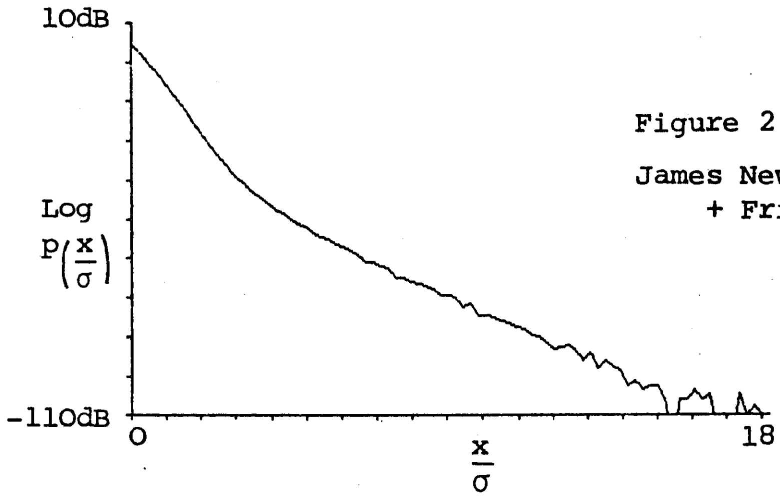

Figure 2

James Newton Howard + Friends

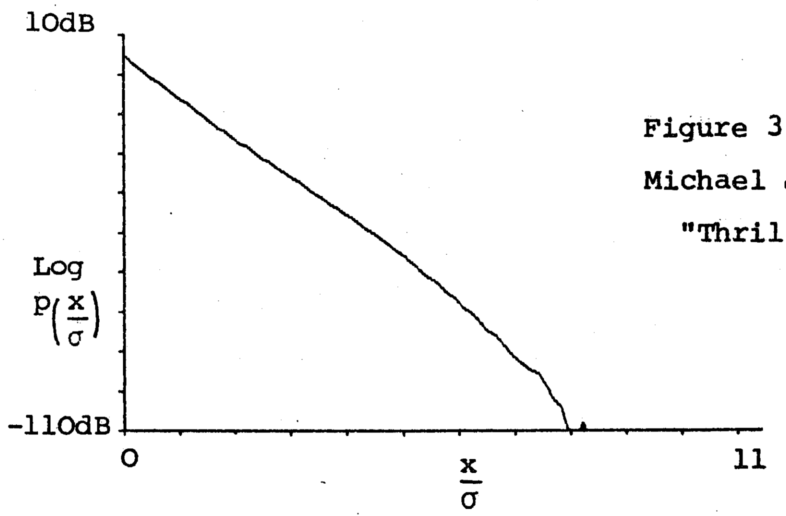

Figure 3

Michael Jackson "Thriller"

EVALUATION OF RESULTS

Initially, it was attempted to fit the results obtained to the Daugherty and Greiner model (equation 1). It was found that for many of the "classical" recordings the model is still largely valid, i.e. the net distribution is the sum of a Gaussian and an exponential term. By analysing the average of the distributions broadly following the form of this model, the values of the coefficients obtained were found to be in agreement with those given in equation 1. However, agreement was poor in the region of transition between the two terms. Replacement of the Gaussian component by a further exponential term improved matters in this respect, the "improved" model being given in equation 2.

$$p(x) = 2.07 \exp(-3.50x) + 0.381 \exp(-0.932x) \qquad (2)$$

It should be stressed that the applicability of this particular model was found to be limited, in general, to distributions of orchestral music. For the total range of music analysed, however, large deviations from this model frequently were observed in both the very small and peak amplitude regions. Such was the spread of results that a single "fixed" equation representing all music types cannot be formulated and hence a three component model with variable coefficients is proposed. This consists of a single exponential component mainly describing small amplitude behaviour, together with two Gaussian components mainly representing medium and large (peak) amplitudes respectively. The form of the proposed model is given by equation 3.

$$p(x) = A \exp(-ax) + B \exp(-bx^2) + C \exp(-cx^2) \qquad (3)$$

A, a etc are the appropriate coefficients for a particular distribution type. It is suggested that each of the three elements in the above equation corresponds to a distinct feature of the music programme, the relative magnitude of each being a characteristic of the particular music type.

a) First element in equation 3.
This element is dominant in the modelling of the spike often found in the probability density curve. This spike is associated with very low signal levels which may be ascribed to the decay and reverberation of music as well as to "direct" low signal levels and noise. It is prominent in "classical" recordings (especially chamber music) and solo instrumental/vocal music. In practice, although sometimes Gaussian in nature, more frequently this spike is observed as a sharper, exponential-type distribution, particularly where fade ins/outs on popular albums enhance the small amplitude region of the overall distribution.

b) Second element
The term $B \exp(-bx^2)$ usually corresponds to the main body of the music. This element dominates in the model for music of limited dynamic range such as most of the rock music analysed.

c) Third element
This element is associated with the largest signal amplitudes. Its relative magnitude gives an indication of the overall dynamic range of the signal.

THE AMPLITUDE DENSITY OF RECORDED MUSIC

When considering how the observed distributions fitted the three-element model it was found useful to identify nine basic types of distribution. These can be arranged into a three by three matrix of shapes as shown in figure 4.

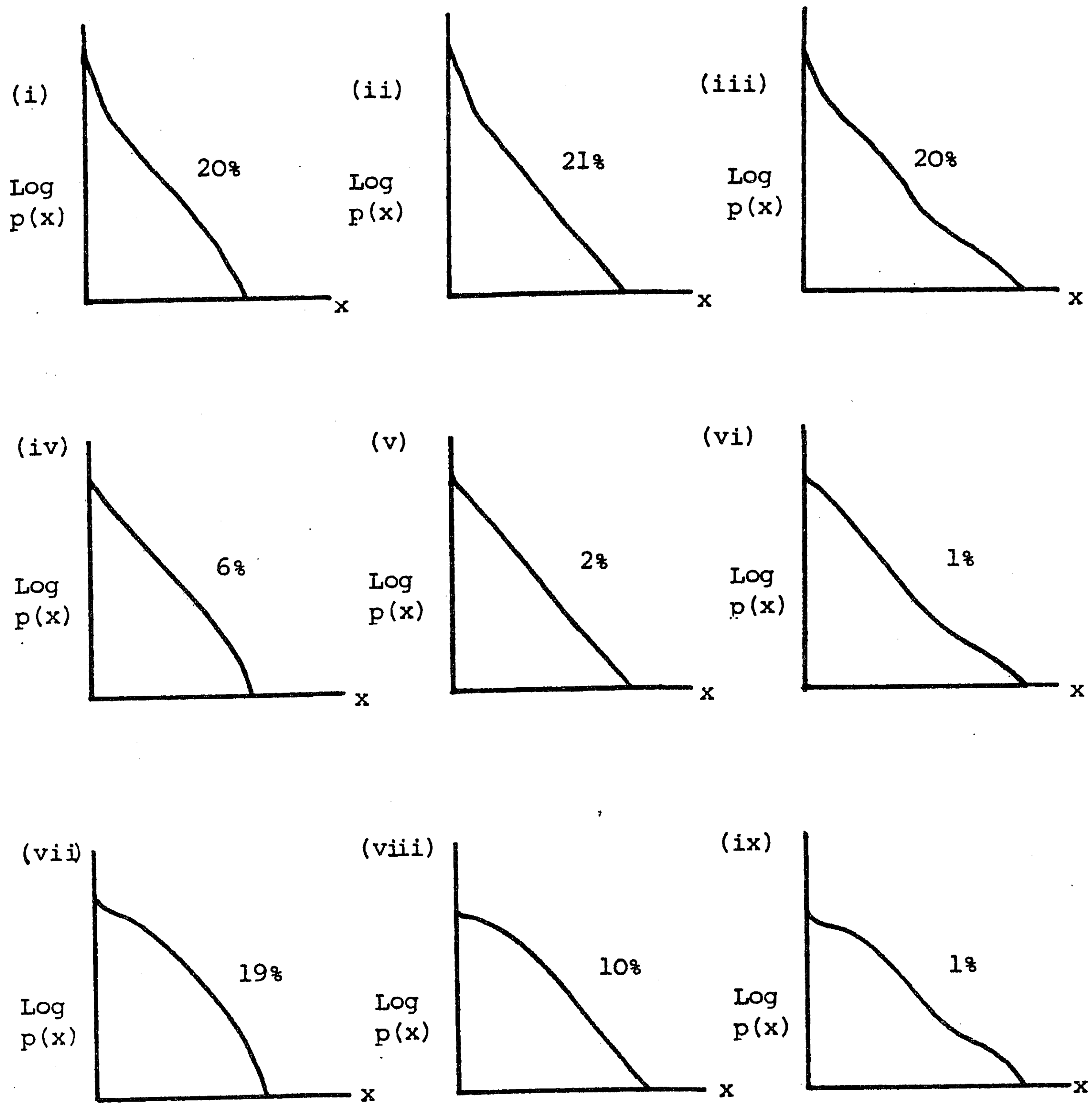

Figure 4. Log probability versus amplitude curves

Distributions in a given vertical column show variations in the relative magnitude of the small amplitude region, whereas those in a given horizontal row show variations in the relative magnitude of the peak amplitude region. The percentages to the right of each curve show the abundance of each distribution in the total set of recordings analysed. It can be seen that over 60% of the recordings demonstrated a dominant small amplitude region ie. a large relative magnitude of the first element (curves (i), (ii) and (iii)). A further 35% of the recordings had distributions which were associated with a dominant second element (curves (iv), (vii) and (viii)).

With respect to music type, it was found that music in the "classical" category was almost exclusively described by curves (ii) and (iii), the remainder of this category being described by curve (i). The music in the "popular" category, however, was very diverse in its characteristics. This was reflected by the fact that all nine shapes of the matrix were needed to describe the various recordings in this category (although half of these could be described by distributions (vii) and (viii).

By taking an average of the measured distributions for each of the nine categories it has been possible to produce quantitative data for each curve in figure 4. Table 3 gives a summary of these results. Ultimately, coefficients for the general model will be derived from these data.

Table 3 Summary of curve data

Curve no.	p(o)	Central Region Gradient*	x/σ for $p(x/\sigma) = 10^{-5}$	x/σ for $p(x/\sigma) > n$ n=99.9%	n=99.99%
i	7.1	0.475	10.1	5.65	7.59
ii	6.5	0.400	12.8	6.04	8.47
iii	6.3	0.360	15.2	6.60	9.73
iv	4.6	0.600	8.2	4.94	6.55
v	3.7	0.620	8.9	4.94	6.63
vi	4.0	0.630	10.7	4.95	6.81
vii	4.8	0.695	7.7	4.69	6.07
viii	4.6	0.675	8.4	4.45	5.84
ix	2.2	0.705	9.1	4.48	6.07

*Semi-logarithmic gradient of the central "linear" region of the distribution curve. ie. gradient = $\text{Log}\left[p\left(\frac{x}{\sigma}\right)\right] / \frac{x}{\sigma}$

VARIATION OF THE MEASUREMENT INTERVAL

In the work carried out so far, emphasis has been placed on the analysis of "complete music programmes" rather than musical excerpts ie. measurements on CD's in their entirety. A number of measurements on shorter signal lengths have shown, however, that significant differences in the distributions from those of the complete discs may be observed. This can be seen in tables 1 and 2 where the results of measurements on individual tracks can be compared with those of their parent albums. In general it is found that there is a decrease in the dynamic qualities of the signal as the measurement interval is decreased. By reducing the measurement interval still further, to a period over which the music is subjectively "continuous", a near-Gaussian distribution may be observed for almost any excerpt of music. This effect is produced irrespective of the excerpt type, except for those having strong percussive or staccato-type natures where large amplitude peaks are frequent. It could be argued, therefore, that the major differences between the different music types are those associated with their long-term dynamic characteristics.

SPECTRAL DEVIATIONS

The above analyses have permitted modelling of the behaviour of different, unfiltered music signals. It is often necessary, however, to know the characteristics of signals within various frequency bands, a prime example being in the design of loudspeaker systems. Hence the signals from a selection of individual CD tracks and passages were analysed after being passed through a range of low- and high-pass filters. These were chosen in preference to octave or $\frac{1}{3}$-octave filters in order to preserve the amplitude characteristics of the signals [1]. Second order responses for filters having fixed cut-off frequencies of 0.2, 0.3, 0.5, 1, 2, 3 and 4kHz were chosen to be representative of those used in commercial hi-fi loudspeakers.

In contrast to the distributions of unfiltered music signals, many of those of the filtered signals displayed a significant asymmetry or skewness. This effect was most pronounced at the frequency extremes, especially at the low frequencies in pop music of the type in which bass drums, guitars or synthesizers feature prominently. There seemed to be no significant trend in the sign of the skewness values.

Although a decrease in the signal bandwidth brought about the expected decrease in the long-time rms value (and hence average power) of the signal, in the majority of cases the peak instantaneous amplitudes did not show a similar reduction. Hence, for these cases, increases in the crest factor and kurtosis values of the measured distributions were observed. The data in Table 4 illustrates this effect, for an excerpt from "Brass Splendour" by the Philip Jones Brass Ensemble when the bandwidth was restricted using a series of high-pass filters.

Table 4

High Pass Filter Cut-off Frequency (kHz)	LIN	0.2	0.3	0.5	1.0	2.0	3.0	4.0
Relative rms/dB	0.0	-0.1	-0.3	-0.9	-3.3	-7.6	-11.8	-15.4
Kurtosis	7.8	7.9	8.0	8.3	8.9	9.7	10.6	11.7
Crest Factor	8.5	8.5	9.3	8.9	9.7	9.6	10.3	13.9

In terms of the 3-element model, the simplest way of interpreting this effect is as a decrease in the relative magnitude of the second element.

The degree of increase in kurtosis was similar in the majority of cases, except for solo vocal and solo instrumental types of music, for which a much larger increase was observed as the high pass filter cut-off frequency was increased.

CONCLUSIONS

Instantaneous amplitude density distributions of various recorded music signals have been measured. Initially attempts were made to fit these distributions to a previous model, good agreement being found for most orchestral music. However, it was found necessary to introduce a variable coefficient three-element model to describe the behaviour of the wide range of music analysed. In this model, low level signals are mainly associated with an exponential element whilst Gaussian terms describe average and peak level signals. By adjusting the coefficients of this model it is possible to describe quantitatively the nine basic distributions observed for the recordings analysed.

It was found that the distributions could not be classified according to the commonly accepted categories of music, necessarily, but more according to its temporal character or long time dynamic behaviour. For example, "classical" music types all had similar amplitude distributions (for complete works), whereas "popular" music took on a wide variety of different distributions ranging from near-Gaussian to those observed in the "classical" case. Also, by reducing the signal bandwidth it was found that the peak:mean ratios of the distributions generally increased at the higher frequencies.

REFERENCES

[1] B.B.Bauer, Octave-band spectral distributions of recorded music, Jnl.Audio Engng.Soc.,Vol.18, No.2, 165-72, (1970)

[2] D.Huszty, Long-time spectra of radio broadcast programme signals, Archives of Acoustics, Vol.5, No.2, 125-41, (1980)

[3] J.G.McKnight, The distribution of peak energy in recorded music, and its relation to magnetic recording systems, Jnl.Audio Engng.Soc., Vol.7, No.2, 65-80, (1959)

[4] G.Penkov, Power and real signals in an audio system, JNL.Audio Engng.Soc., Vol.31, No.6, 423-9, (1983)

[5] L.J.Sivian, H.K.Dunn and S.D.White, Absolute amplitudes and spectra of certain musical instruments and orchestras, J.A.S.A., Vol.2, No.1, 330-71 (1931)

[6] D.G.Daugherty and R.A.Greiner, Some design objectives for audio power amplifiers, IEEE Trans.Audio & Electroacoustics, Vol.AU-14, No.1, 43-8 (1966)

[7] W.B.Davenport, An experimental study of speech wave probability distributions, J.A.S.A., Vol.24, No.4, 390-9, (1952)

[8] A.Barr, M.S.Everett and D.C.Martin, Measurement of instantaneous amplitude distributions of audio frequency signals, PROC. I.O.A., (1986)