Energy Distribution in Music*

JOHN P. OVERLEY†

Summary—A knowledge of the manner in which the acoustic power encountered in music varies with respect to frequency can be a useful tool in the design of components to be used in audio reinforcement or reproduction systems. This paper deals with the amplitude of fractional-second energy peaks, without reference to the rate of their occurrence. It is these peaks which must be considered when distortion is of primary consideration; average power is useful only in predicting temperature rise (where applicable) of signal-handling components. Throughout the discussion emphasis is placed upon the difference between average and peak energy consideration.

The source material from which the distribution analysis is drawn consisted of recent commercial vinyl recordings played on a carefully equalized reproducing system. Ten various types of music are classified and a distribution curve for each is drawn. The methods used in arriving at a typical curve are shown by breaking the spectrum into octaves with a band-pass filter.

The distribution information mentioned above is applied to the design of a three-channel loudspeaker system as an example of use. Other possible applications are mentioned.

Present-day audio systems designed for voice and music reproduction vary greatly in specifications and application, but all have in common the requirement to respond to more than one frequency. The frequency pass band may range from the narrow limits of 200-3,000 cps, typical of a voice communication system, to the extended range of 20-20,000 cps or better, achieved only in certain high-fidelity systems. Each component in a system should be capable of delivering the required power without exceeding the maximum permissible distortion or risking damage due to overload.

In the design and testing of various audio components it is helpful to know the expected signal energy distribution with respect to frequency. In other words, because the energy in typical speech and music is not uniformly distributed throughout the audio frequency spectrum, design compromises may be effected to reduce the possibility of overload at any frequency. The distribution curves developed in this article were intended primarily for use in the design of loudspeaker systems, but are applicable to other components.

Before proceeding with an explanation of the recorded data, let us examine a few of the energy characteristics of typical human speech and, in particular, music. Consider, for a moment, the sound of an orchestra. The dynamic level may vary over an extreme range of values, depending upon how many instruments are playing, the loudness of each, and acoustics of the room or auditorium. Usually, the maximum sound energy at any frequency will occur during the loud musical passages when most players are active. An instrument played loudly not only produces the greatest level of fundamental, but

its tone may be considerably richer in harmonic content than when it is played softly. Thus, if we are interested in finding the maximum energy present at any frequency, the investigation, for the most part, may be narrowed to a study of the apparent loudest passages. Certain exceptions to this generalization are recognized: a solo instrument or voice may be recorded at a higher than normal level with a separate microphone for emphasis; certain combinations, such as a choir of women's voices, require unusual sound handling ability, as we shall see later. Now consider a sustained chord played by the orchestra at a constant, high volume level. Although no audible variations exist, a volume indication will exhibit continual fluctuation over a substantial range. Because the phase of each instrument bears a random relationship to every other one, their vector sum (the resultant sound intensity) is not a constant. At a certain time when several instruments are "in phase," very high instantaneous values of sound amplitude may result. In this manner a series of peaks is generated whose amplitudes are many times that of the average. If an audio system is to give distortionless reproduction, it must be capable of passing, without clipping, the highest peaks which have a time duration sufficiently long, and occur frequently enough, to be perceived by the human ear. Such peaks may be due not only to several instruments playing in unison or at harmonically related frequencies as described above, but are also influenced by the reverberation of the chamber and by the harmonic structure of each single instrument.

It is, therefore, necessary to distinguish between average energy and peak energy. To illustrate this concept, compare two electrical signals of the same peak amplitude, but one which is sinusoidal in nature, and the other which is a pulse of short duration (see Fig. 1).

Average or total energy is proportional to the shaded area under the curves. It is approximately the quantity which would be measured by a volume indicating meter of the conventional type. Obviously, the sine wave represents a much greater average energy than the pulse. Peak energy is a function of maximum amplitude, however, and is seen to be identical for the two signals. Hence, for distortion-free reproduction, the power handling requirements of any component (covering the full frequency range) would be the same in each case even though a conventional VU meter would register

---