(Reprinted from Nature, Vol. 258, No. 5533, pp. 317–318, November 27, 1975)

# '1/f noise' in music and speech

THE power spectrum $S(f)$, of many fluctuating physical variables, $V(t)$, is '1/f-like', varying as $f^{-\gamma}$ ($0.5 \lesssim \gamma \lesssim 1.5$), over many decades of frequency (see ref. 1 for review). We have found that loudness fluctuations in music and speech, and pitch (melody) fluctuations in music exhibit $1/f$ power spectra.

$S(f)$ is related to the autocorrelation function $\langle V(t)V(t+\tau)\rangle$ by the Wiener–Khintchine relationships[2]. If $V(t)$ can be characterised by a single correlation time $\tau_c$, $V(t)$ is correlated with $V(t+\tau)$ for $|\tau| < \tau_c$, and is independent of $V(t+\tau)$ for $|\tau| > \tau_c$. It can be shown that $S(f)$ is 'white' for frequencies $\ll 1/2\pi\tau_c$, and is a rapidly decreasing function of frequency (usually $f^{-2}$) for frequencies $\gg 1/2\pi\tau_c$. A quantity with a $1/f$ power spectrum cannot, therefore, be characterised by a single correlation time. In fact, the $1/f$ power spectrum implies some correlation in $V(t)$ over all time scales corresponding to the frequency range for which $S(f)$ is $1/f$-like[3]. In general, a negative slope for $S(f)$ implies some degree of correlation in $V(t)$ over time scales of roughly $1/2\pi f$. A steep slope implies a higher degree of correlation than a shallow slope.

In our measurements on music and speech, the fluctuating quantity of interest was converted to a voltage whose power spectrum was measured by an interfaced PDP–11 computer

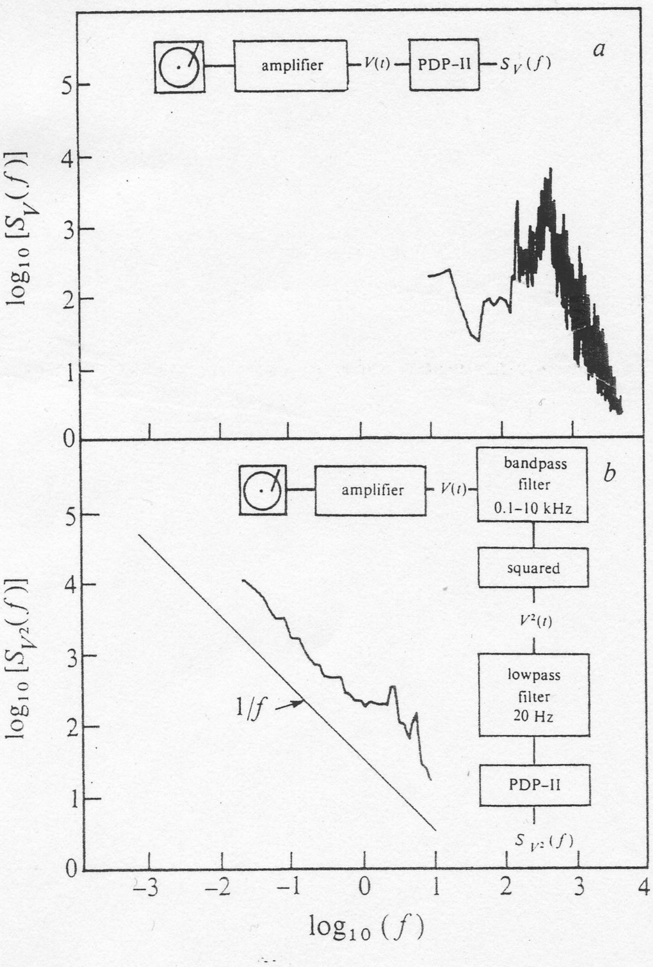


Fig. 1 Bach's 1st Brandenburg Concerto: *a*, $S_V(f)$ against $f$; *b*, $S_{V^2}(f)$ against $f$.

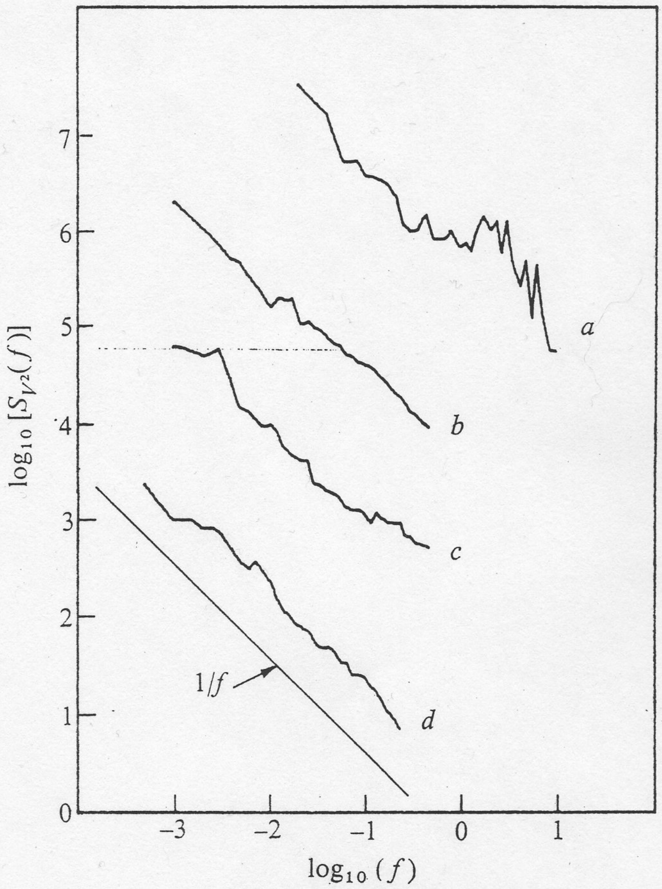


Fig. 2 Loudness fluctuation spectra, $S_{V^2}(f)$ against $f$ for: *a*, Scott Joplin Piano Rags; *b*, classical radio station; *c*, rock station; *d*, news and talk station.

using a fast Fourier transform algorithm. Our first measurements were on the output voltage, $V(t)$, of an audio amplifier. Figure 1*a* shows a log–log plot of the power spectrum, $S_V(f)$, of the audio signal from J. S. Bach's 1st Brandenburg Concerto averaged over the entire concerto. The spectrum consists of a series of sharp peaks in the frequency range 100 Hz–2 kHz corresponding to the individual notes in the concerto and, of course, is not $1/f$-like. Although this spectrum contains much useful information, our primary interest is in more slowly varying quantities.

One such quantity is the loudness of the music. The audio signal, $V(t)$, was amplified and passed through a bandpass filter in the range 100 Hz–10 kHz. The filter output was squared and the audio frequencies filtered off to give a slowly varying signal, $V^2(t)$, proportional to the instantaneous loudness of the music. The power spectrum of the loudness fluctuations of the 1st Brandenburg Concerto, $S_{V^2}(f)$, averaged over the concerto is shown in Fig. 1*b* on a log–log plot. Below 1 Hz the spectrum is $1/f$. The peaks between 1 and 10 Hz are due to the rhythmic structure of the music. Figure 2*a* shows the power spectrum of loudness fluctuations for a recording of Scott Joplin Piano Rags averaged over the whole recording. Although this music has a more pronounced rhythm than the Brandenburg Concerto, and consequently has more structure in the spectrum between 1 and 10 Hz, the spectrum below 1 Hz is still $1/f$-like.

To measure $S_{V^2}(f)$ down to even lower frequencies, we analysed

20/

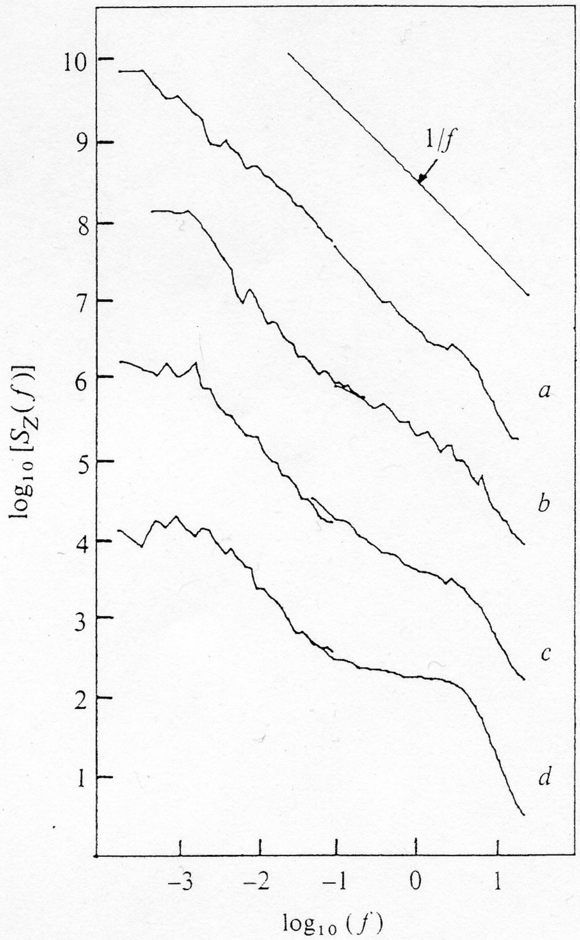


**Fig. 3** Power spectra of pitch fluctuations, $S_Z(f)$ against $f$, for four radio stations: *a*, classical; *b*, jazz and blues; *c*, rock; *d*, news and talk.

the output from an a.m. radio. $S_{V^2}(f)$ was averaged over approximately 12 h, and thus included many musical selections as well as announcements and commercials. Figure 2*b*–*d* shows the loudness fluctuation spectra for three different radio stations. Figure 2*b* shows $S_{V^2}(f)$ for a classical station. The spectrum exhibits a smooth $1/f$ dependence. Figure 2*c* shows $S_{V^2}(f)$ for a rock station. The spectrum is $1/f$-like above $2\times10^{-3}$ Hz, and flattens for lower frequencies, indicating that the correlation of the loudness fluctuations does not extend over time scales longer than a single selection, roughly 100 s. Figure 2*d* shows $S_{V^2}(f)$ for a news and talk station, and is representative of $S_{V^2}(f)$ for speech. Once again the spectrum is $1/f$-like. In Fig. 2*b* and *d*, $S_{V^2}(f)$ remains $1/f$-like down to the lowest frequency measured, $5\times10^{-4}$ Hz, implying correlations over time scales of at least 5 min. In the case of classical music this time is less than the average length of each composition.

Another slowly varying quantity in speech and music is the instantaneous pitch. A convenient means of measuring the pitch is by the rate, $Z$, of zero crossings of the audio signal, $V(t)$. For the case of music, $Z(t)$ roughly follows the melody. Figure 3 shows the power spectra of the rate of zero crossings, $S_Z(f)$, for four radio stations averaged over approximately 12 h. Figure 3*a* shows $S_Z(f)$ for a classical station. The power spectrum is closely $1/f$. Figure 3*b* and *c* shows $S_Z(f)$ for a jazz and blues station and a rock station. Here the spectra are $1/f$-like down to frequencies corresponding to the average selection length, and are flat at lower frequencies. Figure 3*d*, however, which shows $S_Z(f)$ for a news and talk station, exhibits a quite different spectrum. The spectrum is that of a quantity characterised by two correlation times: the average length of an individual speech sound, roughly 0.1 s, and the average length of time for which a given announcer talks, about 100 s. Communication, like most human activities, has correlations that extend over all time scales. For most musical selections the communication is through the melody and $S_Z(f)$ is $1/f$-like. On the other hand, for normal English speech the pitches of the individual speech sounds are uncorrelated: the ideas are not directly related to the pitches of syllables but rather to the meanings attached to them. As a result, the pitch power spectrum for speech is 'white' for frequencies less than about 3 Hz, and falls as $1/f^2$ for $f\gtrsim 3$ Hz. In fact, in Fig. 3*a*–*c* one observes these speech shoulders which become more pronounced as the vocal content of the music increases or the commercial interruptions become more frequent.

The observation of $1/f$-like power spectra for various musical quantities also has implications for stochastic music composition. Most stochastic compositions are based on a random number generator (white noise source), which produces unrelated notes, or, on a low level Markov process, in which there is correlation over only a few successive notes. Neither of these techniques, however, approximates the $1/f$ spectrum and the long time correlations reported here in music. We have used independent $1/f$ noise sources in a simple algorithm to determine the duration (expressed as half, quarter, or eighth notes) and pitch (expressed in various standard scales) of successive notes of a melody. The music obtained by this method was judged by most listeners to be much more pleasing than that obtained using either a white noise source (which produced music that was 'too random') or a $1/f^2$ noise source (which produced music that was 'too correlated'). Indeed, the sophistication of this '$1/f$ music' (which was 'just right') extends far beyond what one might expect from such a simple algorithm, suggesting that a '$1/f$ noise' (perhaps that in nerve membranes[4]?) may have an essential role in the creative process.

This work was supported by the USERDA.

RICHARD F. VOSS
JOHN CLARKE

*Department of Physics,*
*University of California and*
*Inorganic Materials Research Division,*
*Lawrence Berkeley Laboratory,*
*Berkeley, California 94720*

Received July 7; accepted October 23, 1975.

1 van der Ziel, A., *Noise: Sources, Characterization, Measurement*, 106 (Prentice-Hall, Englewood Cliffs, New Jersey, 1970).
2 See, for example, Reif, F., *Fundamentals of Statistical and Thermal Physics*, 585 (McGraw-Hill, 1965).
3 van der Ziel, A., *Physica*, 16, 359 (1950).
4 Verveen, A. A., and Derkson, H. E., *Proc. IEEE*, 56, 906 (1968).