Title
I/f NOISE"" IN MUSIC: MUSIC FROM I/f NOISE

Author
Voss, Richard F.

Publication Date
1975-08-01

Submitted to Perspectives in
New Music

LBL-4117
Preprint c.1

"1/f Noise" IN MUSIC: MUSIC FROM 1/f NOISE

Richard F. Voss and John Clarke

August 1975

Prepared for the U. S. Energy Research and
Development Administration under Contract W-7405-ENG-48

Submitted to Perspectives in
New Music

LBL-4117

"1/f NOISE" IN MUSIC: MUSIC FROM 1/f NOISE

Richard F. Voss† and John Clarke

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

August 1975

†present address: IBM Thomas J. Watson Research Center
Yorktown Heights, N.Y. 10598

I. Introduction

The time correlations of a fluctuating quantity determine its power spectrum. At low frequencies the power spectrum of a wide range of physical quantities varies approximately as 1/f. Thus, everything from semiconductors[1] to nerve membranes[2] to the annual flood level of the river Nile[3] exhibits what has come to be known as "1/f noise". Although this phenomenon has been extensively studied, there is as yet no single theory that satisfactorily explains its origin. In this paper, we show that the loudness and pitch fluctuations in common types of music also have a power spectrum that varies as 1/f. The 1/f spectrum implies some correlation in these fluctuating quantities over all time scales corresponding to the frequency range for which the power spectrum is 1/f. The observation of the 1/f spectrum in music has implications for music compositional procedures. We have used a 1/f noise source in a simple computer algorithim to produce stochastic music. The results suggest that this technique holds considerable promise for computer composition.

II. Power Spectra and Time Correlations

The power spectrum is an extremely useful characterization of the average behavior of a time-varying quantity. The power spectrum, $S_V(f)$, of a fluctuating quantity, $V(t)$, is a measure of the average "power" $\langle V^2 \rangle$, in a unit bandwidth about the frequency f. $S_V(f)$ may be measured by passing $V(t)$ through a tuned filter of frequency f and bandwidth Δf. $S_V(f)$ is then the average of the squared output of the filter divided by Δf. By using a bank of filters at different frequencies, or by changing the frequency of a single filter, $S_V(f)$ may be measured experimentally as a function of frequency.

Many fluctuating quantities, $V(t)$, may be characterized by a single correlation time, τ_c. In such a case, $V(t)$ is correlated with $V(t + \tau)$ for $|\tau| < \tau_c$, and is independent of $V(t + \tau)$ for $|\tau| > \tau_c$. One example of such a fluctuating quantity is the value of the uppermost face of a die. The correlation time is the average time between throws of the die. Successive throws are uncorrelated, but the value remains the same between throws. For this case, it is possible to show that $S_V(f)$ is "white" (independent of frequency) in the frequency range corresponding to time scales over which $V(t)$ is uncorrelated ($f \ll 1/2\pi\tau_c$); and is a rapidly decreasing function of frequency, usually $1/f^2$, in the frequency range over which $V(t)$ is correlated ($f \gg 1/2\pi\tau_c$). A quantity with a $1/f$ power spectrum cannot, therefore, be characterized 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 $S_V(f)$ is $1/f$. In general, a negative slope for $S_V(f)$ 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. Thus, a quantity with a $1/f^2$ power spectrum is highly correlated. An example of a quantity with a $1/f^2$ spectrum is the distance from the starting point of a person performing a random walk in one dimension. A random decision is made at each step (for example, by tossing a coin) whether the step will be to the right or to the left. The result is a slow erratic motion away from, and possible back to, the starting point. If the person is far to the right of the starting point at some instant of time, he will most likely remain to the right for a long time. Hence, there is a high degree of correlation in his position as a function of time.

Figure 1 shows samples of white, 1/f, and $1/f^2$ noise. Each fluctuating quantity was scaled to cover the same vertical range. The white noise has the most random appearance and shows rapid uncorrelated changes. The $1/f^2$ noise is the most correlated showing only slow changes. The 1/f noise is intermediate, showing structure on all time scales. It is interesting to note that, although simple computer algorithms exist for a white or $1/f^2$ noise source over arbitrarily long time scales, no such algorithm exists to produce 1/f noise. This inability to produce a generating algorithm is related to our incomplete theoretical understanding of 1/f noise. Nature, however, has no such problem: any semiconductor, for example, provides a convenient source of 1/f noise.

III. 1/f Noise in Music

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 using a Fast Fourier Transform algorithm with acts as a bank of filters. The most familiar fluctuating quantity associated with music is the audio signal, V(t), such as the voltage used to drive a speaker system. Figure 2(a) shows a linear-linear plot of the power spectrum, $S_V(f)$, of the audio signal from J. S. Bach's 1st Brandenburg Concerto (Angel SB-3787) averaged over the entire concerto. The spectrum consists of a series of sharp peaks in the frequency range 100 Hz to 2 kHz corresponding to the individual notes in the concerto and, of course, is far from 1/f. 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 to 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. ($V^2(t)$ is thus similar to the reading given by recording level meters.) The power spectrum of the loudness fluctuations of the 1st Brandenburg Concerto, $S_{V^2}(f)$, averaged over the entire concerto is shown in Fig. 2(b). On this linear-linear plot, the loudness fluctuations appear as a peak close to zero frequency.

Figure 3 is the log-log plot of the same spectra as in Fig. 2. In Fig. 3(a), the power spectrum of the audio signal, $S_V(f)$, is distributed over the audio range. In Fig. 3(b), however, the loudness fluctuation spectrum, $S_{V^2}(f)$, shows the 1/f behavior below 1 Hz. The peaks between 1 Hz and 10 Hz are due to the rhythmic structure of the music.

Figure 4(a) shows the power spectrum of loudness fluctuations for a recording of Scott Joplin piano rags (Nonsuch H-71248) averaged over the entire recording. Although this music has a more pronounced metric structure than the Brandenburg Concerto, and, consequently, has more structure in the spectrum between 1 Hz and 10 Hz, the spectrum below 1 Hz is still 1/f-like.

In order to measure $S_{V^2}(f)$ down to even lower frequencies an audio signal of greater duration than a single record is needed, for example, that from a radio station. The audio signal from an AM radio was filtered and squared. $S_{V^2}(f)$ was averaged over approximately 12 hr, and thus included many musical selections as well as announcements and commercials. Figures 4(b) through (d) show the loudness fluctuation spectra for three radio stations characterized by different motifs. Figure 4(b) shows $S_{V^2}(f)$ for a classical station. The spectrum exhibits a smooth 1/f dependence. Figure 4(c) shows $S_{V^2}(f)$ for a rock station. The spectrum is 1/f-like above 2×10^{-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 sec. Figure 4(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. 4(b) and Fig. 4(d), $S_{V^2}(f)$ remains 1/f-like down to the lowest frequency measured, 5×10^{-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 instanteous pitch. A convenient means of measuring the pitch is by the rate, Z, of zero crossings of the audio signal, V(t). Thus an audio signal of low pitch will have few zero crossings per second and a small Z, while a high pitched signal will have a high Z. For the case of music, Z(t) roughly follows the pitch content. Figure 5 shows the power spectra of the rate of zero crossings, $S_Z(f)$, for four radio stations averaged over approximately 12 hr. Figure 5(a) shows $S_Z(f)$ for a classical station. The power spectrum is closely 1/f above 4×10^{-4} Hz. Figures 5(b) and (c) show $S_Z(f)$ for a jazz and blues station and a rock station. Here the spectrum is 1/f-like down to frequencies corresponding to the average selection length, and is flat at lower frequencies. Figure 5(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 characterized by two correlation times: The average length of an individual speech sound, roughly 0.1 sec., and the average length of time for which a given announcer talks, about 100 sec. For most musical selections the pitch content has correlations that extend over a large range of time scales, and has a 1/f power spectrum. For normal English speech, on the other hand, the pitches of the individual speech sounds are unrelated. As a result, the power spectrum is "white" for frequencies less than about 3 Hz, and falls as $1/f^2$ for $f \gtrsim 3$ Hz. In fact, in Figs. 5(a) through 5(c), one observes shoulders at about 3 Hz corresponding to speech averaged in with the music. The prominence of this shoulder increases as the vocal content of the music increases, or as the commercial interruptions become more frequent.

The 1/f spectrum for quantities associated with music and speech is, perhaps, not so surprising. We speculate that measures of "intelligent" behavior should show a 1/f-like power spectrum. Whereas a quantity with a white power spectrum is uncorrelated with its past, and a quantity with a $1/f^2$ power spectrum depends very strongly on its past, a quantity with a 1/f power spectrum has an intermediate behavior, with some correlation on all time scales, yet not depending too strongly on its past. Human communication is one example where correlations extend over various time scales. In music much of the communication is directly by the pitch content which exhibits a 1/f spectrum. In English speech, on the other hand, the communication is not directly related to the pitch of the individual sounds. The ideas communicated may have long time correlations even though the pitches of successive sounds are unrelated.

IV. 1/f Noise and Stochastic Composition

The observation of 1/f power spectra for the loudness and pitch fluctuations in music has implications for stochastic music composition. In the past, stochastic compositions have been based on a random number generator (white noise source) which is uncorrelated in time. In the simplest case the white noise source can be used to determine the pitch and duration (quantized in some standard manner) of successive notes. The resulting music is and sounds structureless. (Figure 6 shows an example of this "white music" which we have produced using a white noise source.) Most work on stochastic composition has been concerned with ways of adding the time structure that the random number generator could not provide. Low level Markov processes (in which the probability of a given note depends on its immediate predecessors) were able to impose some local structure but lacked long time correlations. Attempts at increasing the number of preceeding notes on which the given note depended gave increasingly repetitious results rather than interesting long term structure.[4] By adding rejection rules for the random choices (a trial note is rejected if it violates one of the rules), Hiller and Isaacson were also able to obtain local structure but no long term correlations.[5] J. C. Tenney has developed an algorithm that introduces long term structure by slowly varying the distribution of random numbers from which the notes were selected.[6] Thus, although it has been possible to impose some structure on a specific time scale, the stochastic music has been unable to match the correlations and structure found in music over a wide range of time scales.

We propose that the natural means of adding this structure is with the use of a 1/f noise source rather than by imposing constraints upon a white noise source. The 1/f noise source itself has the same time correlations as we have measured in various types of music. To illustrate this process at an elementary level, we present short typical selections composed by white, 1/f, and $1/f^2$ noise.

In each case a physical noise source was used to produce a fluctuating voltage of the desired spectrum. The voltage was sampled and digitized by the PDP-11 computer to produce a series of random numbers stored in the computer whose power spectrum was the same as that of the noise source. The series was then scaled so that successive numbers determined the pitch of successive notes over a two octave range:

 .

A high number specified a high pitch and vice versa. This process was then repeated to produce an independent series of stored random numbers whose value corresponded to the duration of successive notes.

The PDP-11 was then used to "perform" the stochastic composition by controlling a single amplitude modulated voltage controlled oscillator. The computer was also used to put the stochastic compositions in more conventional form. Samples of these computer "scores" are shown in Figs. 6 through 8. Accidentals apply only to the notes they precede. In Fig. 6 a white noise source was used to determine pitch and duration. In Fig. 7 a 1/f noise source was used, while in Fig.8 a $1/f^2$ noise source was used. Although Figs. 6 through 8 are not intended as complete formal compositions, they are representative of the types of correlation that can be achieved when the three types of noise sources of Fig. 1 are used to control various musical parameters. In each case the noise sources were "Gaussian" implying that values near the mean were more likely than extreme values.

Our 1/f music was judged by most listeners to be far more interesting than either the white music (which was "too random") or the scale-like $1/f^2$ music (which was "too correlated"). Indeed the surprising sophistication of the 1/f music (which was close to being "just right") suggests that the 1/f noise source is an excellent method of adding time correlations.

V. Discussion

There is, however, more to music than 1/f noise. Although our simple algorithms were sufficient to demonstrate the superiority of a 1/f noise source over a white noise source in stochastic composition, the variation of only two parameters (pitch and duration of the notes of a single voice) can, at best, produce only a very simple form of music. More structure is needed, not all of which can be provided by 1/f noise sources. We improved on this elementary composition by using two voices that were either independent or partially correlated (notes having the same duration but independent pitches or vice versa), and by varying the overall loudness with an additional 1/f noise source. We added more structure to the music by introducing either a simple, constant rhythm, or a variable rhythm determined by another 1/f noise source. We suggest the use of 1/f noise sources on various structural levels (from the characterization of individual notes to that of entire movements) coupled with external constraints (for example, rhythm or the rejection rules of Hiller) as offering promising possibilities for stochastic composition. A further possibility is the use of 1/f noise sources to control various synthesizer inputs providing correlated but random variations.

Acknowledgements

We are pleased to acknowledge helpful conversations with and encouragement from Drs. M. Mathews and A. Moorer. We are particularly indebted to Professor Edwin E. Dugger for his advice, comments, and a critical reading of the manuscript.

This work was supported by the U. S. Energy Research and Development Administration.

References and Notes

1. For a review, see Aldert van der Ziel, Noise: Sources, Characterization, Measurement (Prentice-Hall, Inc., Englewood Cliffs, NJ, 1970), p. 106.
2. A. A. Verveen and H. E. Derkson, Proc. IEEE 56, 906 (1968).
3. B. B. Mandelbrot and J. R. Wallis, Water Resources Research 5, 321 (1969).
4. See, for example, R. C. Pinkerton, Sci. American 194, 77 (1956) or H. F. Olson and H. Belar, JASA 33, 1163 (1961).
5. L. A. Hiller, Jr. and L. M. Isaacson, Experimental Music (McGraw-Hill Book Co., NY, 1959).
6. Discussed by J. R. Pierce, M. V. Mathews and J. C. Risset, Gravesaner Blätter, 27/28, 92 (1965).

Figure Captions

Fig. 1. Samples of white, 1/f, and $1/f^2$ noise.

Fig. 2. Bach's 1st Brandenburg Concerto (linear scales): (a) power spectrum of audio signal, $S_V(f)$ vs f; (b) power spectrum of loudness fluctuations, $S_{V^2}(f)$ vs f.

Fig. 3. Bach's 1st Brandenburg Concerto (log scales): (a) $S_V(f)$ vs f; (b) $S_{V^2}(f)$ vs f.

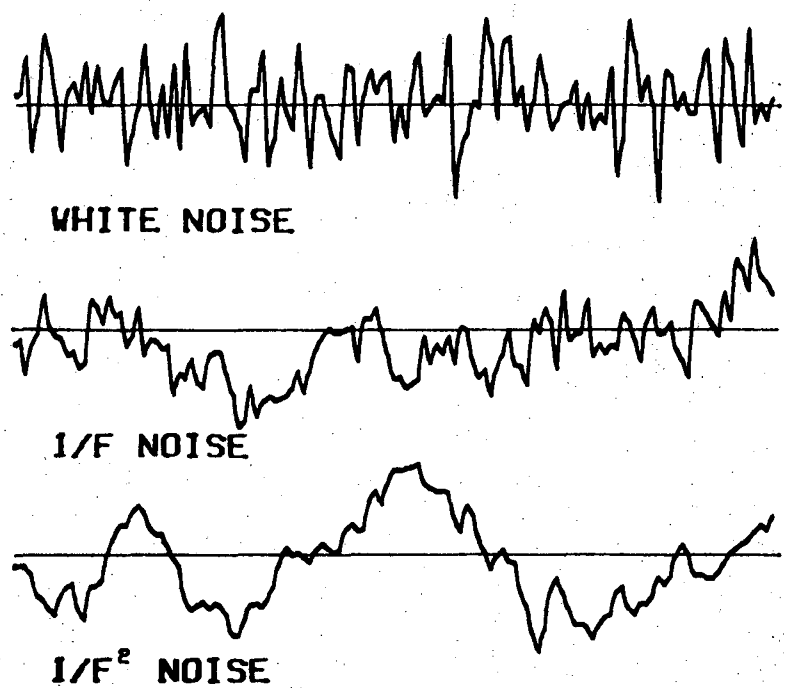
Fig. 4. Loudness fluctuation spectra, $S_{V^2}(f)$ vs f for: (a) Scott Joplin piano rags; (b) classical radio station; (c) rock station; (d) news and talk station.

Fig. 5. Power spectra of pitch fluctuations, $S_Z(f)$ vs f, for four radio stations: (a) classical; (b) jazz and blues; (c) rock; (d) news and talk.

Fig. 6. Pitch and duration determined by a white noise source.

Fig. 7. Pitch and duration determined by a 1/f noise source.

Fig. 8. Pitch and duration determined by a $1/f^2$ noise source.



XBL757-6718

Fig. 1

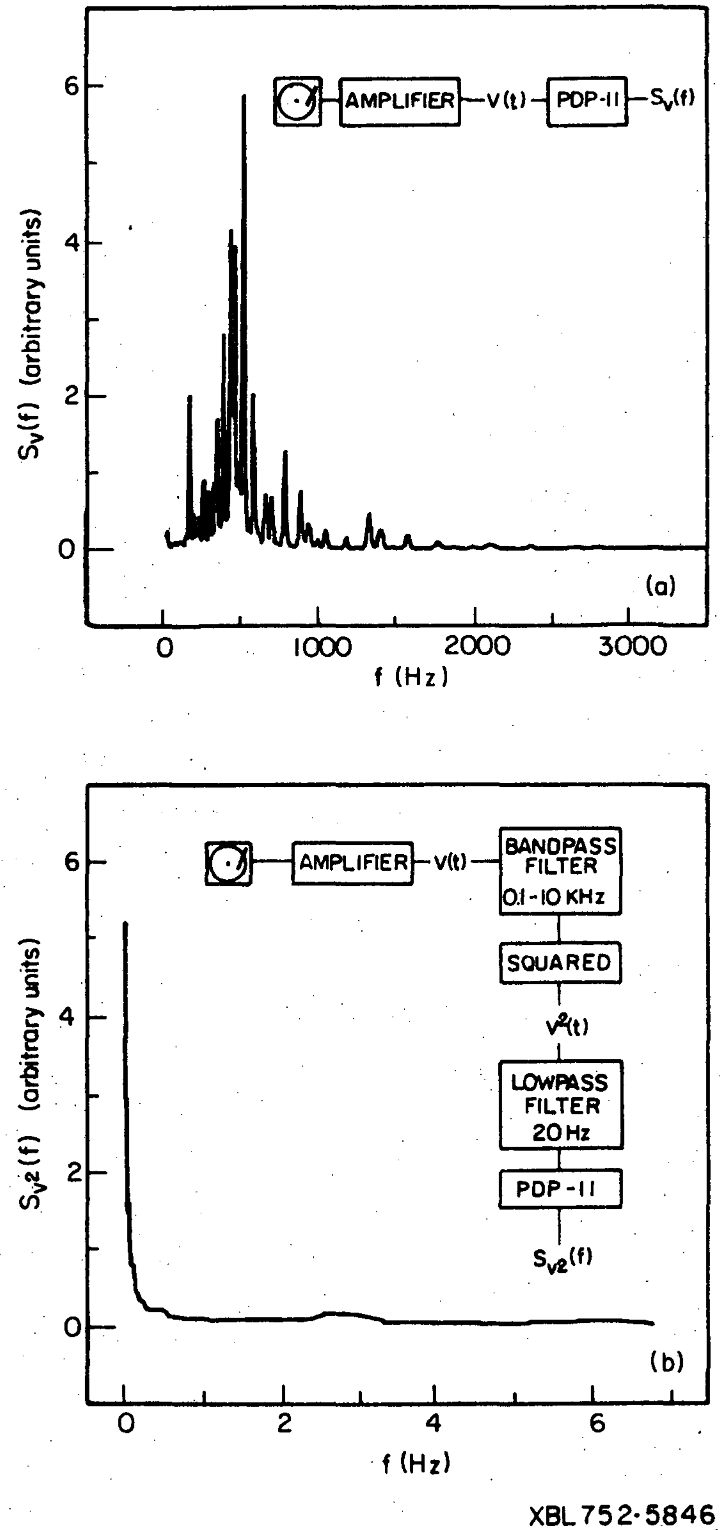
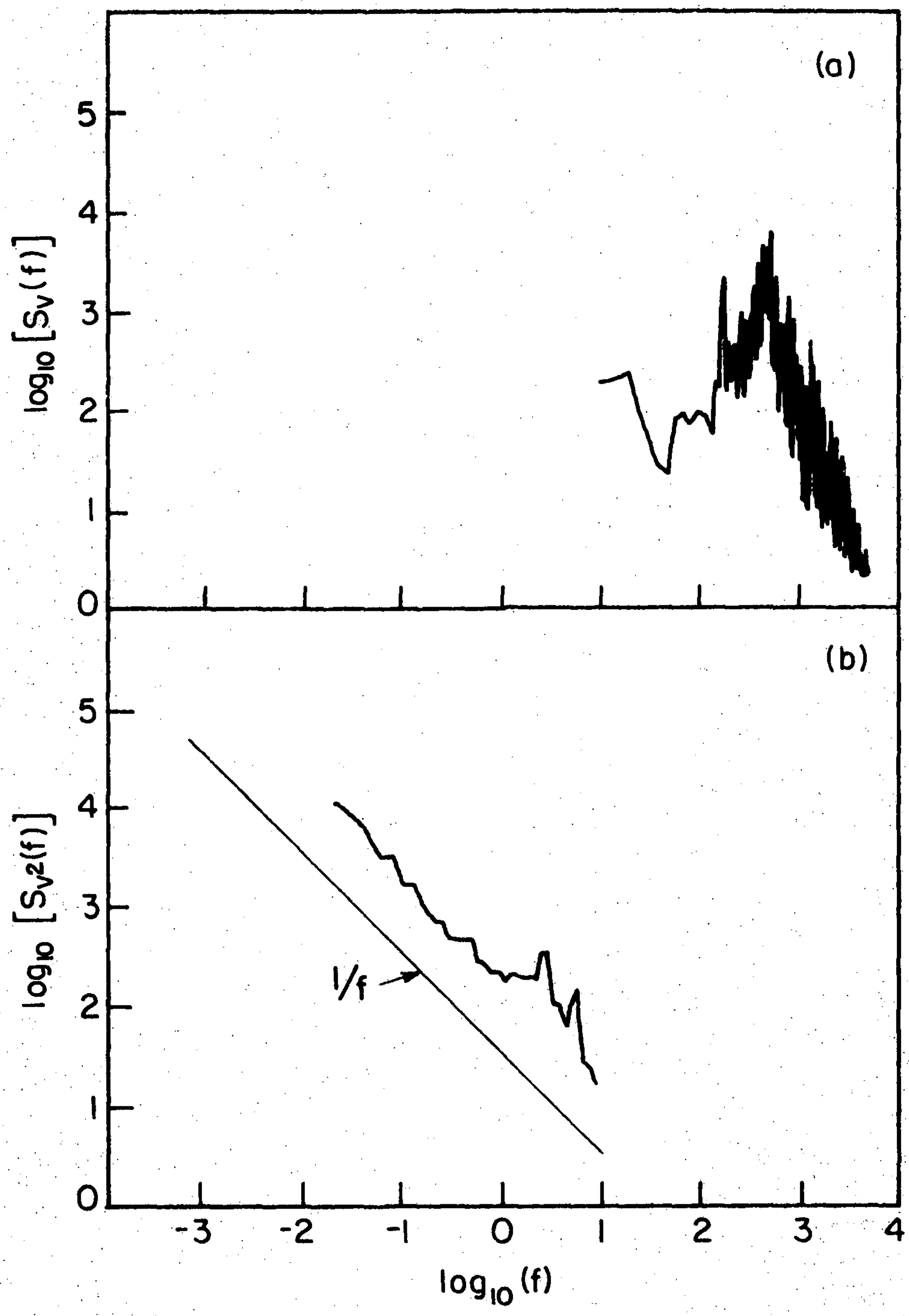
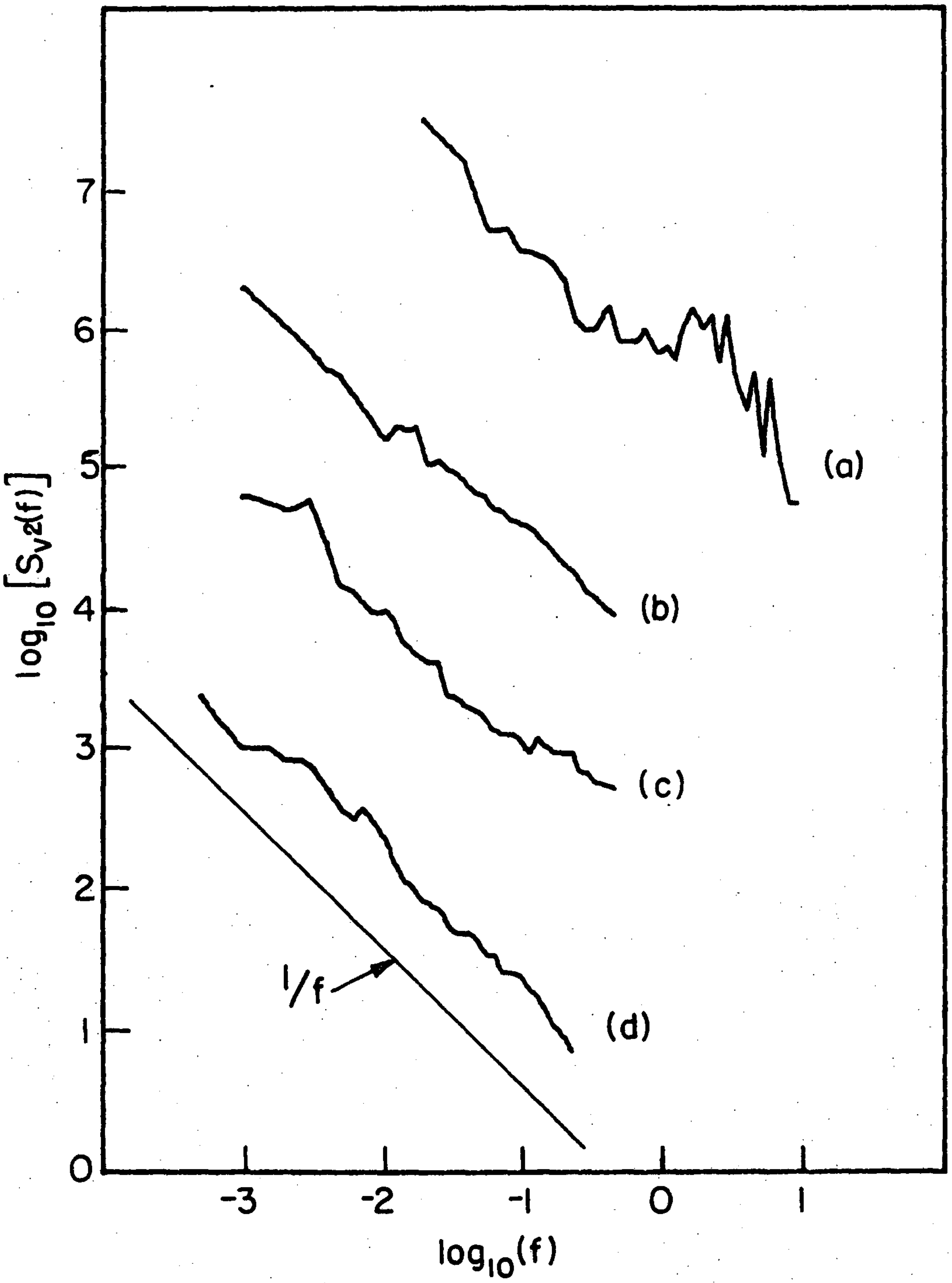

AMPLIFIER
V(t)
PDP-11
$S_v(f)$
$S_v(f)$ (arbitrary units)
(a)
f (Hz)
AMPLIFIER
V(t)
BANDPASS FILTER 0.1-10 KHz
SQUARED
$V^2(t)$
LOWPASS FILTER 20 Hz
PDP-11
$S_{v^2}(f)$
$S_{v^2}(f)$ (arbitrary units)
(b)
f (Hz)
XBL 752-5846

Fig. 2



XBL752-5845

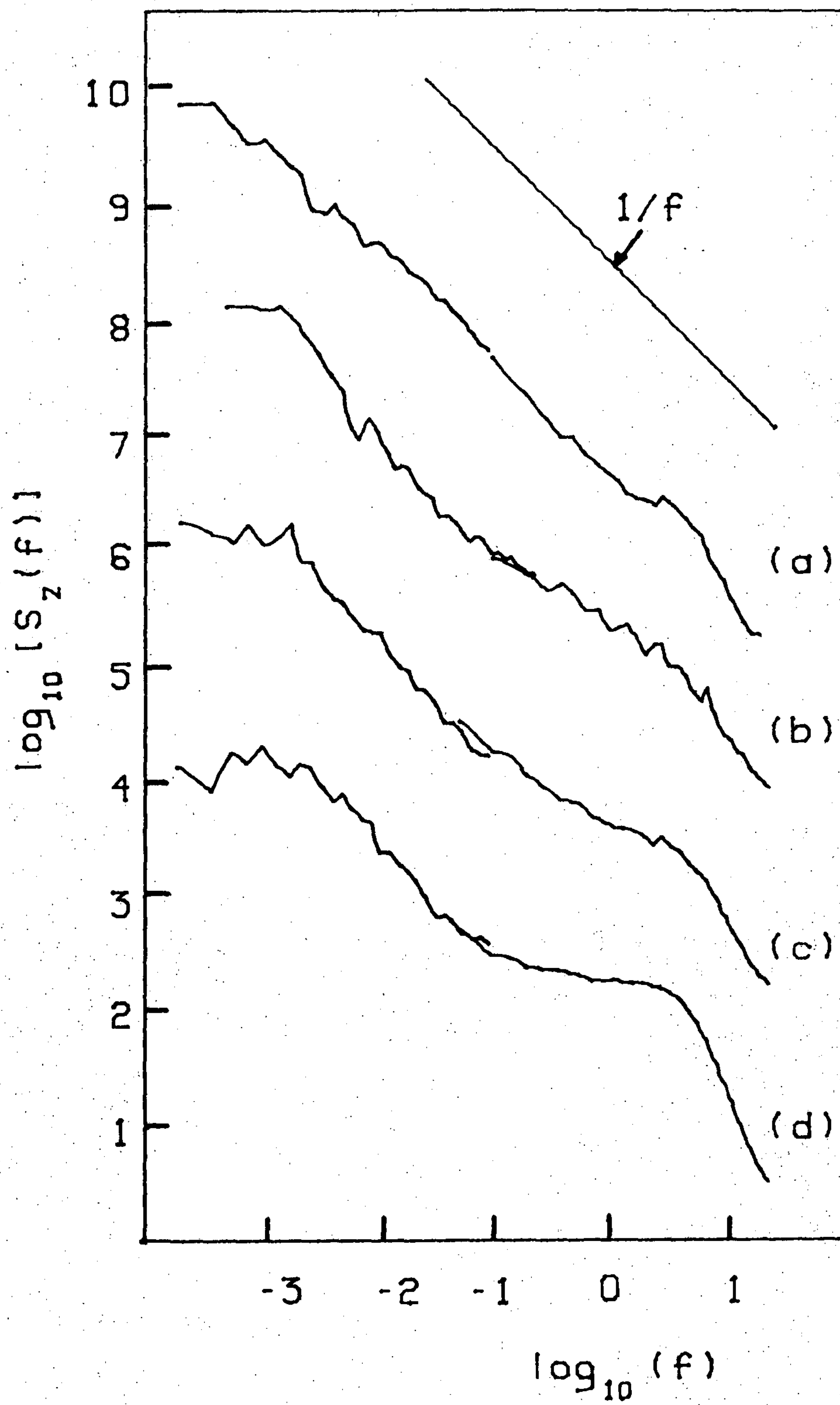
Fig. 3



XBL 752-5844

Fig. 4

0 0 0 0 4 3 0 8 3 8 7



XBL757-6728

Fig. 5

WHITE MUSIC

XBL757-6722

Fig. 6

1/F MUSIC

XBL757-6724

Fig. 7

1/F² MUSIC

XBL757-6727

Fig. 8

U U 0 4 3 0 3 5 3 9

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720