"1/f noise" in music: Music from 1/f noise

Richard F. Voss[a)] and John Clarke

Department of Physics, University of California and Materials and Molecular Research Division, Lawrence Berkeley Laboratory, Berkeley, California 94720
(Received 30 January 1976; revised 8 September 1977)

The spectral density of fluctuations in the audio power of many musical selections and of English speech varies approximately as $1/f$ (f is the frequency) down to a frequency of 5×10^{-4} Hz. This result implies that the audio-power fluctuations are correlated over all times in the same manner as "$1/f$ noise" in electronic components. The frequency fluctuations of music also have a $1/f$ spectral density at frequencies down to the inverse of the length of the piece of music. The frequency fluctuations of English speech have a quite different behavior, with a single characteristic time of about 0.1 s, the average length of a syllable. The observations on music suggest that $1/f$ noise is a good choice for stochastic composition. Compositions in which the frequency and duration of each note were determined by $1/f$ noise sources sounded pleasing. Those generated by white-noise sources sounded too random, while those generated by $1/f^2$ noise sounded too correlated.

PACS numbers: 43.75.Wx, 43.60.Cg, 43.75.−z, 43.70.−h

INTRODUCTION

The spectral density of many physical quantities varies as $1/f^{\gamma}$, where f is the frequency and $0.5 \lesssim \gamma \lesssim 1.5$, over many decades. Thus vacuum tubes,[1] carbon resistors,[2] semiconducting devices,[3] continuous[4,5] or discontinuous[6] metal films, ionic solutions,[7] films at the superconducting transition,[8] Josephson junctions,[9] nerve membranes,[10] sunspot activity,[11] and the flood levels of the river Nile[11] all exhibit what is known as "$1/f$ noise."[12] Although this phenomenon has been extensively studied, there is as yet no single theory that satisfactorily explains its origin. In this paper,[13] we show that the audio power and frequency fluctuations in common types of music also have spectral densities that vary as $1/f$. The $1/f$ behavior implies some correlation in these fluctuating quantities over all times corresponding to the frequency range for which the spectral density is $1/f$. The observation of the $1/f$ spectral density in music has implications for music compositional procedures. We have used a $1/f$ noise source in a simple computer algorithm to produce stochastic music. The results suggest that $1/f$ noise sources have considerable promise for computer composition.

I. SPECTRAL DENSITY AND TIME CORRELATIONS

Although frequently used in the analysis of random signals or "noise," the spectral density (power spectrum) is an extremely useful characterization of the average behavior of any quantity varying in time. The spectral density $S_V(f)$ of a quantity $V(t)$ fluctuating with time t is a measure of the mean squared variation $\langle V^2\rangle$ in a unit bandwidth centered on the frequency f. The average is taken over a time that is long compared with the period; in practice, we usually average over at least 30 periods. $S_V(f)$ 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. Thus, if $V(t)$ is a voltage, $S_V(f)$ is in units of volt squared per hertz.

A second characterization of the average behavior of $V(t)$ is the autocorrelation function, $\langle V(t)V(t+\tau)\rangle$. $\langle V(t)V(t+\tau)\rangle$ is a measure of how the fluctuating quantities at times t and $t+\tau$ are related. For a stationary process $\langle V(t)V(t+\tau)\rangle$ is independent of t and depends only on the time difference τ. $S_V(f)$ and $\langle V(t)V(t+\tau)\rangle$ are not independent, but are related by the Wiener–Khintchine relations[14]

$$\langle V(t)V(t+\tau)\rangle = \int_0^\infty S_V(f)\cos(2\pi f\tau)\,df \tag{1}$$

and

$$S_V(f) = 4\int_0^\infty \langle V(t)V(t+\tau)\rangle\cos(2\pi f\tau)\,d\tau\,. \tag{2}$$

Many fluctuating quantities may be characterized by a single correlation time τ_c. In such a case, $V(t)$ is correlated with $V(t+\tau)$ for $|\tau| \ll \tau_c$, and is independent of $V(t+\tau)$ for $|\tau| \gg \tau_c$. Usually, $\langle V(t)V(t+\tau)\rangle = \langle V^2\rangle \times\exp(-|\tau|/\tau_c)$. From Eq. (2), it is then possible to show that $S_V(f)$ is "white" (independent of frequency) in the frequency range corresponding to times over which $V(t)$ is independent ($f \ll 1/2\pi\tau_c$); 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 $1/f$ spectral density cannot, therefore, be characterized by a single correlation time. In fact, the $1/f$ spectral density implies some correlation in $V(t)$ over all times corresponding to the frequency range for which $S_V(f)$ is $1/f$-like.[15] In general, a negative slope for $S_V(f)$ implies some degree of correlation in $V(t)$ over times of roughly $1/2\pi f$. A steep slope implies a higher degree of correlation than a shallow slope. Thus a quantity with a $1/f^2$ spectral density is highly correlated.

Figure 1 shows samples of white, $1/f$, and $1/f^2$ noise voltages versus time. Each fluctuating voltage was amplified to cover the same range, and each had the same high-frequency cutoff. 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.

[a)]Present address: IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598.

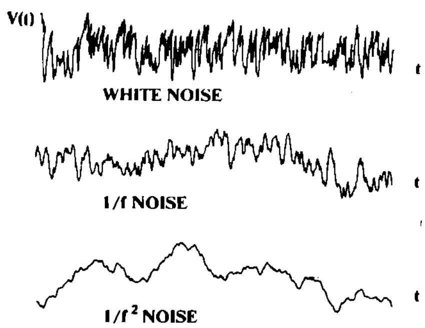

FIG. 1. Samples of white, $1/f$, and $1/f^2$ noise voltages, $V(t)$ versus time t.

II. $1/f$ NOISE IN MUSIC

In our measurements on music and speech, the fluctuating quantity of interest was converted to a voltage whose spectral density was measured by an interfaced PDP-11 computer using a Fast Fourier Transform algorithm that simulates 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 plot of the spectral density $S_V(f)$ of the audio signal from J. S. Bach's First Brandenburg Concerto averaged over the entire concerto. The spectral density 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.

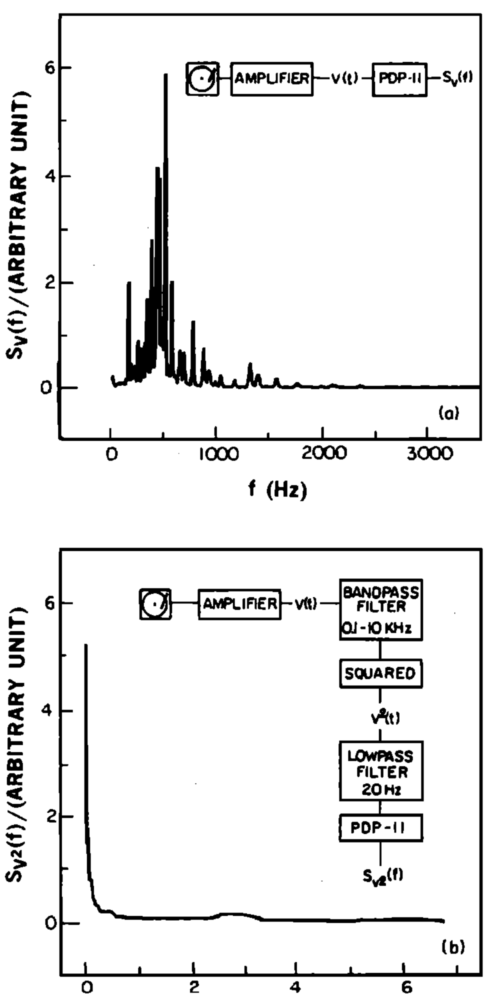

FIG. 2. Bach's First Brandenburg Concerto (linear scales). (a) Spectral density of audio signal, $S_V(f)$ vs f; (b) spectral density of audio power fluctuations, $S_{V^2}(f)$ vs f.

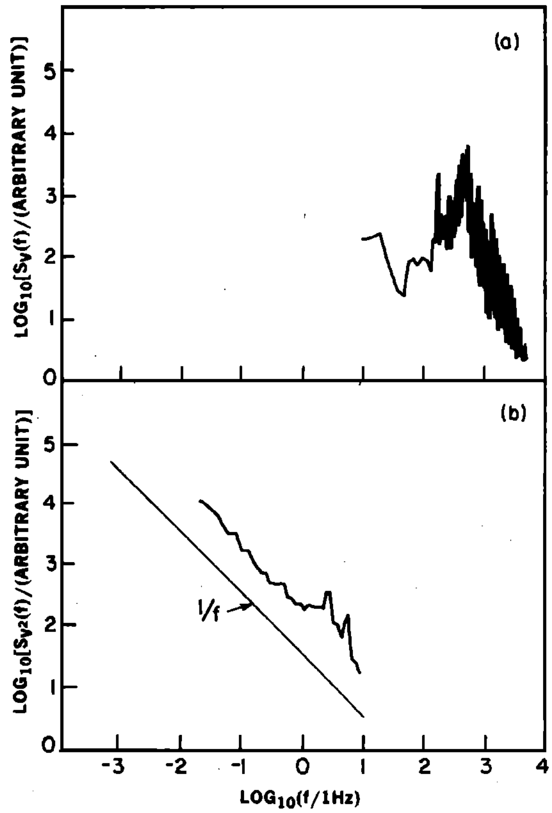

FIG. 3. Bach's First Brandenburg Concerto (log scales). (a) $S_V(f)$ vs f; (b) $S_{V^2}(f)$ vs f.

One such quantity we call the audio power of the music, which is proportional to the electrical power delivered to a loudspeaker by an amplifier and hence to $V^2(t)$. $V^2(t)$ varies monotonically with the loudness of the music. In order to measure $V^2(t)$ the audio signal $V(t)$ was amplified and passed through a bandpass filter in the range 100 Hz to 10 kHz. The output voltage was squared, and filtered with a 20-Hz low-pass filter. This process produced a slowly varying signal, $V^2(t)$, that was proportional to the "instantaneous" audio power of the music. Correlations of $V^2(t)$ represent correlations of the audio power of successive notes. The spectral density of the audio power fluctuations of the First Brandenburg Concerto, $S_{V^2}(f)$, averaged over the entire concerto is shown in Fig. 2(b). On this linear–linear plot, the audio power fluctuations appear as a peak close to zero frequency.

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

Figure 4(a) shows the spectral density of audio power fluctuations for a recording of Scott Joplin piano rags 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 spectral density between 1 and 10 Hz, the spectral density below 1 Hz is still $1/f$-like.

In order to measure $S_{V^2}(f)$ at frequencies below 10^{-2} Hz 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 h, and thus included many musical selections as well as announcements and commercials. Figures 4(b) through (d) show the spectral densities of the audio power fluctuations for three radio stations characterized by different motifs. Figure 4(b) shows $S_{V^2}(f)$ for a classical station. The spectral density exhibits a smooth $1/f$ dependence. Figure 4(c) shows $S_{V^2}(f)$ for a rock station. The spectral density is $1/f$-like above 2×10^{-3} Hz, and flattens for lower frequencies, indicating that the correlation of the audio power fluctuations does not extend over times longer than a single selection, roughly 100 s. 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 spectral density 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 times of a least 5 min.

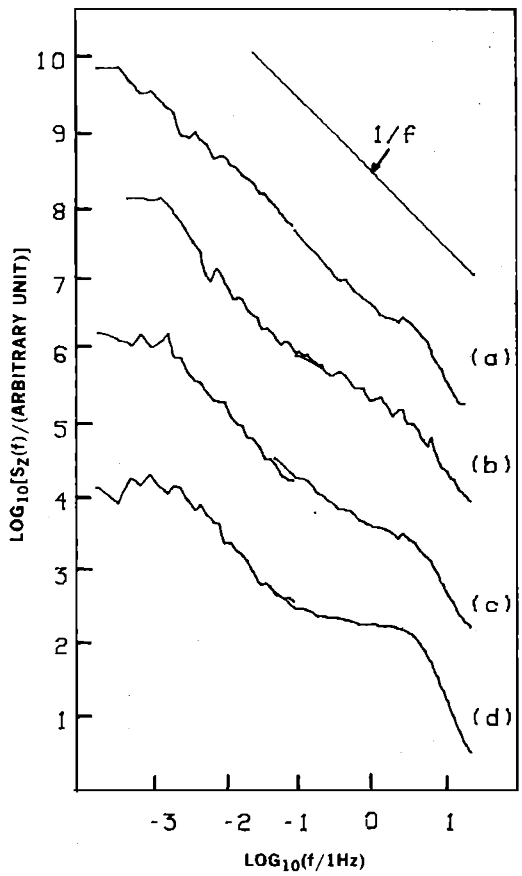

FIG. 5. Spectral density of frequency fluctuations, $S_Z(f)$ vs f for four radio stations (a) classical; (b) jazz and blues; (c) rock; and (d) news and talk.

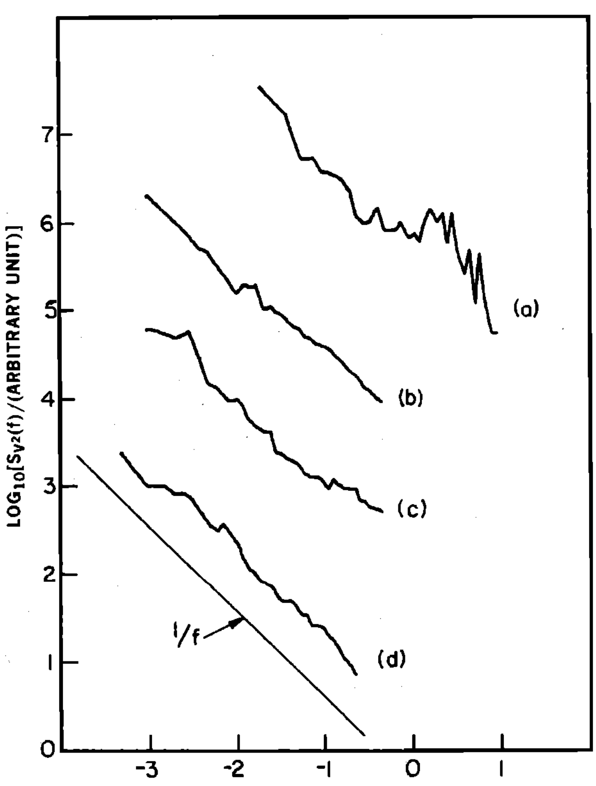

FIG. 4. Spectral density of audio power fluctuations, $S_{V^2}(f)$ vs f for (a) Scott Joplin piano rags; (b) classical radio station; (c) rock station; and (d) news and talk station.

Another slowly varying quantity in speech and music is the "instantaneous" frequency. A convenient means of measuring the frequency is by the rate Z of zero crossings of the audio signal, $V(t)$. Thus an audio signal of low frequency will have few zero crossings per second and a small Z, while a high-frequency signal will have a high Z. For the case of music, $Z(t)$ roughly follows the melody. Correlations in $Z(t)$ represent correlations in the frequencies of successive notes. Figure 5 shows the spectral density of the rate of zero crossings, $S_Z(f)$, for four radio stations averaged over approximately 12 h. $Z(t)$ was also smoothed by a 20-Hz low-pass filter before the spectral density was measured. Figure 5(a) shows $S_Z(f)$ for a classical station. The spectral density varies closely as $1/f$ above 4×10^{-4} Hz. Figures 5(b) and 5(c) show $S_Z(f)$ for a jazz and blues station and a rock station. Here the spectral density 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 spectral density. The spectral density is that of a quantity characterized 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. For most musical selections the

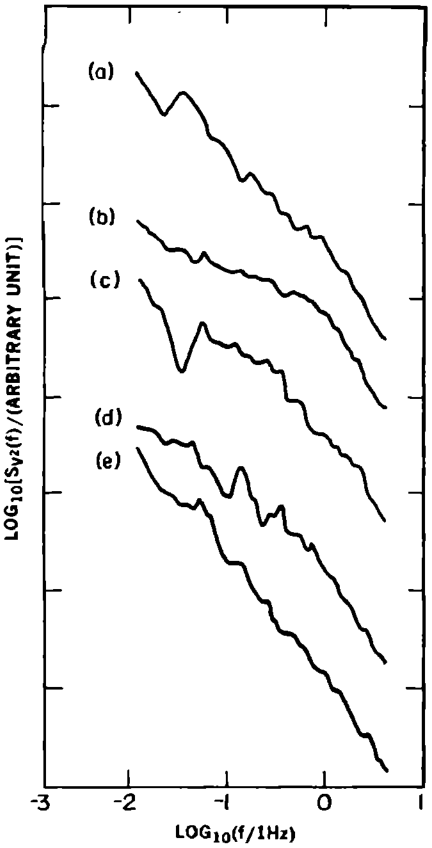

FIG. 6. Audio power fluctuation spectra densities, $S_{V^2}(f)$ vs f for (a) Davidovsky's Synchronism I, II, and III; (b) Babbit's String Quartet number 3; (c) Jolas' Quartet number 3; (d) Carter's Piano concerto in two movements; and (e) Stockhausen's Momente.

frequency content has correlations that extend over a large range of times and, consequently, has a $1/f$-like spectral density. For normal *English* speech, on the other hand, the frequencies of the individual speech sounds are statistically unrelated. As a result, the spectral density is "white" for frequencies less than about 2 Hz, and falls as $1/f^2$ for $f \gtrsim 2$ Hz. In fact, in Figs. 5(a)–5(c), one observes shoulders at about 2 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.

Figures 6 and 7 show the measured audio power and frequency fluctuation spectral densities for several pieces by different composers. In each case the spectral density was averaged over the length of the piece in the manner described above. Although all of the pieces show an increasing spectral density at lower frequencies, individual differences can be observed. For the audio power fluctuations (Fig. 6), the selections by Davidovsky [Fig. 6(a)] and Stockhausen [Fig. 6(e)] show the correlations characteristic of the $1/f$ spectral density while those by Babbit [Fig. 6(b)], Jolas [Fig. 6(c)], and Carter [Fig. 6(d)] show decreasing correlations at times longer than several seconds. For the frequency fluctuations (Fig. 7), Davidovsky's Synchronism [Fig. 7(a)] remains closest to the $1/f$ spectral density, while the other pieces show decreased correlations for times longer than about 10 s.

The $1/f$ behavior of quantities associated with music and speech is, perhaps, not so surprising. We speculate that measures of "intelligent" behavior should show a $1/f$-like spectral density. Whereas a quantitiy with a white spectral density is uncorrelated with its past, and a quantity with a $1/f^2$ spectral density depends very strongly on its past, a quantity with a $1/f$ spectral density has an intermediate behavior, with some correlation over all times, 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 conveyed by frequency changes that exhibit a $1/f$ spectral density. In English speech, on the other hand, the communication is not directly related to the frequencies of the individual sounds: Successive sounds may convey related ideas even though their frequencies are statistically uncorrelated. In other words, the ideas communicated may have long time correlations even though the frequencies of successive sounds are unrelated. It would also be of interest to investigate the music and speech of other cultures, such as Chinese, in which pitch plays an important role in communication.

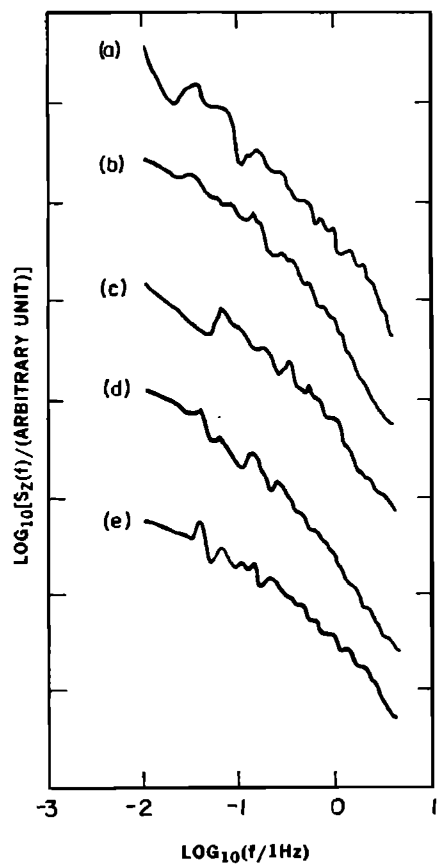

FIG. 7. Frequency fluctuation spectral densities, $S_Z(f)$ vs f for (a) Davidovsky's Synchronism I, II, and III; (b) Babbit's String Quartet number 3; (c) Jolas' Quartet number 3; (d) Carter's Piano concerto in two movements; and (e) Stockhausen' Momente.

III. 1/f NOISE AND STOCHASTIC COMPOSITION

The observation of $1/f$ spectral densities for audio power and frequency fluctuations in music has implications for stochastic music composition. Traditionally, 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 frequency and duration (quantized in some standard manner) of successive notes. The resulting music is and sounds structureless. (Figure 8 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 correlations 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 preceding notes on which the given note depended gave increasingly repetitious results rather than interesting long term structure.[16] 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.[17] 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.[18] 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 times.

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 on 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 with the desired spectrum. The white noise was obtained from the Johnson noise

WHITE MUSIC

FIG. 8. Frequency and duration determined by a white noise source.

1/f MUSIC

FIG. 9. Frequency and duration determined by a $1/f$ noise source.

voltage across a resistor. The $1/f$ noise was obtained from the voltage fluctuations across a current-biased transistor. The $1/f^2$ noise was obtained by filtering the white noise source by a 6 dB per octave low-pass filter where the cutoff frequency was less than the inverse of the length of the piece of music. The noise voltage was sampled, digitized, and stored in a PDP-11 computer as a series of numbers whose spectral density was the same as that of the noise source. These numbers were rounded and scaled to represent the notes of a standard musical scale (pentatonic, major, or 12 tone chromatic) over a two-octave range: A high number specified a high frequency and vice versa. This process was then repeated with another noise source to produce an independent series of stored numbers whose values corresponded to the durations of successive notes.

The PDP-11 was then used to "perform" the stochastic composition by controlling a single amplitude modulated voltage controlled oscillator. We used sinusoidal, square, and triangle waveforms and a variety of attack and decay rates. The computer was also used to put the stochastic compositions in more conventional form. Samples of these computer "scores" are shown in Figs. 8–10. Accidentals apply only to the notes they precede. The scores are presented without bars since no constraints were imposed on durations of successive notes. In Fig. 8 a white noise source was used to determine frequency and duration. In Fig. 9 a $1/f$ noise source was used, while in Fig. 10 a $1/f^2$ noise source was used. Although Figs. 8–10 are not intended as complete formal compositions, they are representative of the correlations between successive notes 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.

Over a period of about two years we have played samples of our music to several hundred people at nine universities and research laboratories. The listeners ranged from those with little technical knowledge of music to professional musicians and composers. We played selections of white, $1/f$, and $1/f^2$ music varying in length from one to ten minutes. 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

$1/f^2$ MUSIC

FIG. 10. Frequency and duration determined by a $1/f^2$ noise source.

scalelike $1/f^2$ 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 for adding time correlations.

IV. DISCUSSION

There is, however, more to music than $1/f$ noise: Rhythm, for example, represents a periodic constraint on the duration of successive notes. Although our simple algorithms were sufficient to demonstrate that $1/f$ noise is a better choice than white noise for stochastic composition, the variation of only two parameters (frequency 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 these elementary compositions by using two voices that were either independent or partially correlated (notes having the same duration but independent frequencies or vice versa), and by varying the overall intensity 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.

ACKNOWLEDGMENTS

We are pleased to acknowledge helpful conversations with and encouragement from Dr. M. Mathews and A. Moorer. We are particularly indebted to Edwin E. Dugger for his advice, comments, and a critical reading of the manuscript. Finally, we are grateful to Dr. Robert W. Young for his helpful comments during the revision of the manuscript.

[1]J. B. Johnson, Phys. Rev. **26**, 71–85 (1925).

[2]C. J. Christenson and G. L. Pearson, Bell Syst. Tech. J. **15**, 197–223 (1936).

[3]For a review, see A. van der Ziel, *Noise: Sources, Characterization, Measurement* (Prentice-Hall, Englewood Cliffs, NJ., 1970).

[4]F. N. Hooge and A. M. H. Hoppenbrouwers, Physica **45**, 386–392 (1969).

[5]R. F. Voss and J. Clarke, Phys. Rev. **B13**, 556–573 (1976).

[6]J. L. Williams and R. K. Burdett, J. Phys. Chem. **2**, 298–307 (1969).

[7]F. N. Hooge, Phys. Lett. **33A**, 169–170 (1970).

[8]J. Clarke and T. Y. Hsiang, Phys. Rev.. **B13**, 4790–4800 (1976).

[9]J. Clarke and G. Hawkins, Phys. Rev. **B14**, 2826–2831 (1976).

[10]A. A. Verveen and H. E. Derkson, Proc. IEEE **56**, 906–916 (1968).

[11]B. B. Mandelbrot and J. R. Wallis, Water Resour. Res. **5**, 321–340 (1969).

[12]The spectral density of voltage fluctuations has been measured down to 5×10^{-7} Hz in semiconductors [M. A. Caloyannides, J. Appl. Phys. **45**, 307–316 (1974)]; in the case of the river Nile, the spectral density of the annual flood levels extends down to 3×10^{-11} Hz.

[13]A preliminary account of these measurements appeared in Nature, **258**, 317–318 (1975).

[14]F. Reif, *Fundamentals of Statistical and Thermal Physics*, (McGraw-Hill, New York, 1965), pp. 585–587.

[15]The fact that $1/f$ noise contains a distribution of correlation times is discussed by Aldert van der Ziel, Physica **16**, 359–372 (1950).

[16]See, for example, R. C. Pinkerton, Sci. Am. **194**, 77–86 (1956) or H. F. Olson and H. Belar, J. Acoust. Soc. Am. **33**, 1163–1170 (1961).

[17]L. A. Hiller, Jr. and L. M. Isaacson, *Experimental Music* (McGraw-Hill, New York, 1959).

[18]Discussed by J. R. Pierce, M. V. Mathews, and J. C. Risset, Gravesaner Blätter 27/28, 92–97 (1965).