AN EXPERIMENTAL INVESTIGATION OF VERY LOW FREQUENCY SEMICONDUCTOR NOISE

Thesis by

Robert Alexander Dukelow

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1974

(Submitted March 14, 1974)

ACKNOWLEDGMENTS

I wish to thank Dr. H. C. Martel for his many suggestions as well as his overall influence during my many years under his guidance. Thanks also go to Dr. C. A. Mead, J. M. McCool, and M. S. Ball for their suggestions and aid. Fellowships from Tektronix and RCA, and Teaching Assistantships provided financial support. The Naval Undersea Center provided the use of some essential equipment. Considering the length of this thesis, I must thank Kathy Ellison for her careful and speedy typing. Most of all, however, I wish to thank my wife, Alice, for at least trying to be patient.

ABSTRACT

This thesis represents an attempt to more accurately characterize the statistics of the very low frequency noise of a particular semiconductor device (a grounded input bipolar integrated operational amplifier) than has previously been achieved. Power spectral density estimates are obtained for frequencies ranging from 250 Hz to 10^{-7}Hz. These estimates are based on data recorded over an uninterrupted period of approximately 1 year (355.9 days). Relatively high sample rates are maintained for each source, allowing a high degree of accuracy in these estimates without resorting to the questionable process of averaging the estimates for a number of noise sources[1]. The high sample rate also allows for a reduction in the errors due to aliasing by the use of digital filtering techniques. Spectral density estimates for six separate noise sources are presented.

The preliminary objectives of the experiment were to search for a break in the $f^{-\alpha}$ ($\alpha \sim 1$) spectral density component of semiconductor noise and to attempt to establish the "true" value of α if a unique value exists. Very long time constant "popcorn" noise (ignored by one investigator because it is "as natural as flicker noise itself"[1, page 69]), proved to be an obstacle in measuring the "pure" 1/f noise process and may not only be the reason for the wide range of values for α reported in the literature but, as suggested by some researchers[2], may be the cause of flicker noise. "Popcorn" noise was observed with apparent time constants greater than 10^6 seconds. The statistics of the

observed "popcorn" noise were investigated and showed good agreement with the results of J. N. Puckett, Jr.[3] (who worked primarily with popcorn noise having time constants on the order of a few milliseconds) except for a few cases in which the waveform resembled that which might be expected if one popcorn component were to modulate another. The work of Puckett was also extended in that a test for burst waiting time dependency was performed. No evidence of dependency could be found. Popcorn noise components which were large enough to be clearly identified were removed in the time domain. This technique was found to be quite useful in improving the spectral estimates.

As in previous experiments, power supply regulation and temperature control were found to be essential. In this experiment, however, the temperatures of the noise sources and the power supply fluctuations were measured concurrently with the noise data so that their contribution to the total observed noise could be more accurately ascertained.

Neither a break frequency in the $f^{-\alpha}$ trend nor a unique α were established, although the values of α obtained were all slightly larger than but closer to 1.0 than most reported values. A nearly periodic component (with a period of 1 year) was also observed in two of the noise sources which were contained in a plastic integrated circuit package. Since this component cannot be explained by temperature or power supply voltage it is conjectured that other external parameters such as humidity or barometric pressure (or even cosmic radiation) may account for some of the observed noise, although no such component could be observed in the other four noise sources (which were housed in metal packages).

TABLE OF CONTENTS

Chapter I

INTRODUCTION

1/f noise (also known as "flicker" noise, "excess" noise, or "pink" noise) has been a topic of discussion for many years. The term is used to describe a random noise process which has a power spectral density resembling $f^{-\alpha}(\alpha \sim 1)$ over some range of frequencies.

"Flicker" noise is not unique to semiconductors. It has been reported in such varied "devices" as tubes[4], quartz crystal oscillators[5], the frequency of rotation of the earth[6], and ambient sea noise[16] to name only a few. Due to the wide range of situations under which this type of noise is found, it has been suggested that some unknown fundamental law of physics is responsible. The number of theories which are now available to explain a 1/f power spectral density over finite ranges of frequency in semiconductors, and the seemingly contradictory experimental results reported in the literature, however, suggest not only that different physical processes may be at work in different types of semiconductor devices, but that more than one process may be at work in a single device.

The purpose of this thesis is not to explain the origin of "flicker" noise but to add to the available experimental data concerning the nature of this noise. It is hoped that this data might eventually be useful in understanding the origins of "flicker" noise in a particular device. Another purpose, however, is the hope that improved understanding of the statistics of this type of noise may lead to more

meaningful specifications involving the drift characteristics of DC coupled devices. The many proposed explanations for flicker noise will, therefore, not be discussed except where it seems applicable in the presentation of some experimental result. Discussions of some flicker noise theories can be found in many of the references listed at the end of this thesis.

Chapter II deals with some of the general aspects of power spectral density estimation as well as some particular details concerning 1/f noise. The remainder of the thesis involves the experiment and the analysis of the data derived from that experiment.

The noise sources used are commercial integrated circuit operational amplifiers. This choice was made because of the widespread use of these devices as DC amplifiers and their relative insensitivity to supply voltage and temperature variations (considerable effort was still necessary to achieve sufficient temperature and power supply regulation). The wide use of these devices in the electronics industry may cause the resulting data to have significant value to circuit designers concerned with long term stability.

This is not the first attempt to measure the low frequency noise of this type of noise source. The first observation of excess noise in electronic devices was probably due to Bernamont[7] (1937) over the frequency range of 96 Hz to 162 KHz using thin metallic film resistors. Rollin and Templeton[8,9] (1953) used a scheme in which noise was tape recorded and then played into an audio frequency wave analyzer at a higher speed to estimate the spectral density of noise from carbon and

wire wound resistors, and germanium filaments over the frequency range of 2.5×10^{-4} Hz to 7.5 Hz. Firle and Winston[10](1955) estimated the spectral density of silicon diode noise from 6×10^{-5} Hz to 0.01 Hz using a photographic recording and fast playback technique similar to the tape recorder scheme of Rollin and Templeton. Blakemore[11](1966) made measurements on discrete differential transistor amplifiers down to the microhertz region but concluded that his low frequency estimates were biased by temperature. Baldinger and Nüesch[12](1968) measured silicon transistor noise from 10^{-4} Hz to 0.05 Hz after constructing a temperature control scheme which is quite noteworthy in itself.

Caloyannides[1](1971) made spectral estimates down to $10^{-6.3}$ Hz using integrated operational amplifiers (which greatly reduce temperature and voltage sensitivities). He was limited, however, in the total number of data points which could be recorded in a single run. Ten noise sources were multiplexed allowing ~1000 data points to be recorded for each source at each of several sample rates. The spectral estimates were then averaged to obtain a single spectral density estimate with a, hopefully, reasonable variance. Since an analog low pass filter with a cutoff frequency as low as 5×10^{-5} Hz (for a sample period of $\Delta t = 10^4$ seconds) is impractical, removing the effects of aliasing posed a difficult problem. A method was described which would, in fact, have increased the relative variance of the low frequency estimates near $f = \frac{1}{2\Delta t}$. Apparently, another dealiasing scheme was used which may have introduced a significant bias in these estimates (see Section 2.6).

In order to improve upon these results, a relatively inexpensive (yet reliable) data recording scheme was constructed which allowed more than 400,000 data words (24 bits each) to be stored on each side of a standard 1800 ft, 1/4 inch audio magnetic tape, with a semiconductor buffer memory large enough to allow tapes to be changed without loss of data. This allowed a relatively fast sample rate to be maintained over the entire course of the experiment. The sample rate was reduced (with very little aliasing) with the use of digital filters in a general purpose computer. Spectral density estimates are presented for each of six noise sources (from 10^{-7} Hz to 250 Hz) with calculated relative variances (over most of the frequency range) smaller than in prior work. The additional data also made possible a careful study of the statistics of "popcorn" noise at these low frequencies, its effects on the noise power spectral density estimates, and (in some cases) its removal from the noise data. Other improvements included the placement of the noise source voltage supply regulators inside the temperature controlled oven, recording of the supply voltages and noise source temperatures during the entire experiment (so that the contribution of these terms to the power spectral density could be ascertained), and the incorporation of an uninterruptable power supply to drive not only the noise sources but the entire experiment.

The discovery of popcorn noise with time constants on the order of 10^6 seconds made the hypothesis that 1/f noise is just a sum of these random telegraph waves seem very attractive. Therefore, Appendix A outlines a physically plausible distribution of amplitudes and time

constants which would produce a 1/f spectral density. A computer simulation of such a sum is performed using a random number generator to produce a 1/f spectrum, over 6 decades of frequency, utilizing only 20 random telegraph waves. Plots of the time domain data and spectral estimates of this simulated noise process are presented for comparison with the experimental data.

Chapter II

SPECTRAL ESTIMATION

2.1 Introduction to Spectral Estimation and Notation

The method of power spectral density estimation used in this thesis is that of Blackman and Tukey[13]. A few results will also be used from other sources. Derivation of the important results will be outlined here, although the original sources should be consulted for more detailed or generalized analysis.

The noise voltage, n(t), will be assumed to be a stationary, ergodic, random process. These assumptions are impossible to prove for a finite record length. In fact, since $|f|^{-\alpha}$, $\alpha \geq 1$ is not integrable for any range which includes f = 0, it has been said that there can be no stationary process having such a spectral density. This problem has in the past been circumvented by assuming that the process is stationary but that the true spectral density levels off at some very low frequency ε. C. A. Greenhall[14] objects to what he calls an artificially imposed cutoff (since it has never been observed) and prefers to assume that there is no cutoff. He states, however, that even under this eventuality the derivative, $\frac{dn(t)}{dt}$, may be stationary (n(t) would be said to have stationary increments) and have a spectral density. Although some of the discussions (in this thesis) on various estimates involving a 1/f type process will use the $\varepsilon \to 0$ approach, the issue will be side stepped by pointing out that all estimates actually performed on the experimental data are preceded by a prewhitening filter which performs

as a differentiator for $f \to 0$. Those who wish to deny the existence of a spectral density for n(t) may multiply the spectral estimates by f^2 and consider them to be estimates on $\frac{dn(t)}{dt}$. The importance to the circuit designer is minimal, at any rate, since he can make drift predictions based only on the high passed version of n(t) (or equivalently, the differential behavior of the process over times of the order of $1/f_T$ or less where f_T is the lowest frequency for which the 1/f trend, or f trend for $\frac{dn(t)}{dt}$, is known to continue) until a break frequency ε, is established. The designer might, for example, hope to estimate the standard deviation for $n(T) - n(o)$ where $T < 1/f_T$.

An attempt will be made to use consistent notation throughout the thesis. $S_n(f)$ will denote the true spectral density, and

$R_n(\tau) = \int_{-\infty}^{\infty} S_n(f) e^{i2\pi f\tau}\, df = E\{n(t)n(t+\tau)\} - E\{n(t)\}^2$ is the true autocovariance function of n(t). $C_k = R(k\Delta t)$ and $S_r = S\left(\frac{r}{2M\Delta t}\right)$ will be used in the discussion of sampled data and spectral density. $\hat{X}$ will denote an estimate of the parameter X($\hat{X}$ will in general be a random variable). σ_y^2 or Var{y} will denote the variance of the random variable y ($\sigma_y^2 = E\{y^2\} - E\{y\}^2$). Subscripts may at times be dropped when the meaning is clear. The remaining notation will be defined each time it is used.

2.2 The Spectral Estimator

The intuitive spectral estimator, $\hat{S}(f) = \frac{1}{T}\left|\int_0^T x(t)e^{i2\pi ft}\,dt\right|^2$ has been shown[15] to be a rather poor estimator (at least for a gaussian random variable, and probably for most others). In this case $\lim_{T\to\infty} E\left\{\hat{S}_x(f)\right\} = S_x(f)$ as we would hope, but (under the assumption that $x(t)$ is gaussian) $\mathrm{Var}\left\{\hat{S}_x(f)\right\} \geq \left[E\left\{\hat{S}_x(f)\right\}\right]^2$ for all T.

The Blackman-Tukey type spectral estimator is of the form:

$$\hat{R}_x(\tau) \equiv \begin{cases} \frac{1}{T_n - \tau}\int_0^{T_n-\tau} x(t)\,x(t+\tau)\,dt & 0 \leq \tau \leq T_m \\ 0 & |\tau| > T_m \\ \hat{R}_x(-\tau) & -T_m \leq \tau \leq 0 \end{cases}$$

[Eq. 2.2.1]

$$\hat{S}_x(f) \equiv \int_{-\infty}^{\infty} D(\tau)\,\hat{R}_x(\tau)e^{-i2\pi f\tau}\,d\tau \qquad \text{[Eq. 2.2.2]}$$

which in the case of samples equally spaced in time (sample period = Δt) intuitively reduces to:

$$\hat{C}_k \left(= \hat{R}_x(k\Delta t)\right) \equiv \frac{1}{N-k}\sum_{j=1}^{N-k} x_j\,x_{j+k} \quad (k = 0, \ldots, M)$$

[Eq. 2.2.3]

$$\hat{S}_r \left(= \hat{S}_x\left(\frac{r}{2M\Delta t}\right)\right) \equiv \Delta t\left[\hat{C}_o D(o) + 2 \sum_{j=1}^{M} \hat{C}_j D(j\Delta t) \cos\left(\frac{qr\pi}{M}\right)\right]$$

[Eq. 2.2.4]

where $T_m = M\Delta t$, $x_j = x(j\Delta t)$, and $T_n = N\Delta t$.

The effects of T_n, T_m, and the sampling process on the expectation and variance of the spectral estimator will be discussed in the following sections. The random variable, x(t), may not be the actual noise process but the result of passing the original noise, n(t), through a "prewhitening" filter (see section 2.7).

2.3 Spectral Windows

The estimated autocovariance function must be truncated in time for any real experiment. In the Blackman and Tukey algorithm (Eqs. 2.2.1 and 2.2.2) the maximum lag for which the autocovariance function is estimated is designated by T_m. Taking the expectation of Eq. 2.2.2 yields:

$$E\left\{\hat{S}(f)\right\} = \int_{-\infty}^{\infty} D(\tau)\, R_X(\tau)\, e^{-i2\pi f\tau}\, d\tau$$

where the truncation of the autocovariance function is accounted for in $D(\tau)$ (we assume $D(\tau) = 0$ for $|\tau| > T_m$).

$$E\left\{\hat{S}_X(f)\right\} = \int_{-\infty}^{\infty} Q(f-\alpha)\, S_X(\alpha)\, d\alpha = Q(f) * S_X(f) \qquad \text{[Eq. 2.3.1]}$$

by the convolution theorem of the Fourier transform, where

$$Q(f) = \int_{-\infty}^{\infty} D(\tau)\, e^{-i2\pi f\tau}\, d\tau \quad \text{and } * \text{ denotes convolution.}$$

$$\text{If } D(\tau) = D_0(\tau) \equiv \begin{cases} 1 & |\tau| \leq T_m \\ 0 & |\tau| > T_m \end{cases} \quad \text{then}$$

$$Q(f) = Q_0(f) \equiv \frac{\sin 2\pi f T_m}{\pi f}. \quad \int_{-\infty}^{\infty} Q_0(f)\, df = 1 \text{ and as } T_m \to \infty,\ Q_0(f) \to \delta(f).$$

Therefore, for white noise $E\{\hat{S}(f)\} = S(f)$ for all T_m and if $S(f)$ is "reasonably smooth" over regions of several times $\frac{1}{T_m}$ we would expect $E\left\{\hat{S}(f)\right\} \tilde{=} S(f)$ to be a good approximation. If the details of the spectral density to be estimated are not known then it would seem to be

advantageous to find a $Q(f)$ such that $Q(f) = 0$ for $f > \frac{\beta}{T_m}$ where $\beta \sim 1$. Unfortunately, such a spectral window is inconsistent with a truncated $D(\tau)$. We can, therefore, hope only to find a $Q(f)$ whose envelope goes to zero as quickly as possible for $f \neq 0$. One solution is to increase T_m, which is ultimately limited to T_n (the total time period of experimental observation). No matter how long we have managed to make the experiment, however, we would still like to do as well as we can with those data. It should be noted that T_m will normally be chosen such that $T_m << T_n$ for other reasons.

Blackman and Tukey[13] discuss several possible spectral windows in detail. The only window which will be discussed here is the "hanning" window pair:

$$D_2(\tau) = \begin{cases} \frac{1}{2}\left(1 + \cos\frac{\pi\tau}{T_m}\right) & |\tau| \leq T_m \\ 0 & |\tau| > T_m \end{cases} \qquad \text{[Eq. 2.3.2]}$$

$$Q_2(f) = \frac{1}{2} Q_o(f) + \frac{1}{4}\left[Q_o\left(f + \frac{1}{2T_m}\right) + Q_o\left(f - \frac{1}{2T_m}\right)\right] \qquad \text{[Eq. 2.3.3]}$$

This window is actually twice as wide as the $Q_o(f)$ window if one considers the width to the first zeros. Rewriting the equation for $Q_2(f)$, however, reveals that the envelope for $Q_2(f)$ falls off as $1/f^3$ for large f as compared to $1/f$ for the $Q_o(f)$ window:

$$Q_2(f) = \frac{\sin 2\pi f T_m}{2\pi f(1 + 2T_m f)(1 - 2T_m f)} \quad .$$

It is clear that there is no point in estimating the spectral density at points spaced in frequency much closer than $\Delta f = \frac{1}{2T_m}$. This sample period in the frequency domain also simplifies the windowing procedure since the spectral density estimates, $\hat{S}^{\circ}(f)$, can be obtained at $f_r = \frac{r}{2T_m}$ using the $Q_0(f)$ window and then the $Q_2(f)$ window estimates, $\hat{S}(f)$, may be derived by evaluating:

$$\hat{S}(f_r) = \frac{1}{4}\hat{S}^{\circ}(f_{r-1}) + \frac{1}{2}\hat{S}^{\circ}(f_r) + \frac{1}{4}\hat{S}^{\circ}(f_{r+1}) \quad .$$

As $T_m \to \infty$, $Q_2(f) \to \delta(f)$ and $E\left\{\hat{S}(f)\right\} \cong S(f)$ if $S(f)$ is "smooth" over distances of $\Delta f \sim \frac{1}{T_m}$. This statement, however, is rather imprecise. In section 2.8 the bias will be evaluated for several specific spectral densities which are of interest in this particular thesis.

2.4 Covariance of Spectral Estimates

A reasonably precise statement is needed regarding the accuracy of the experimental spectral estimates. Here, again, the analysis shown is primarily an outline of that done by Blackman and Tukey[13] where the problem is solved in greater generality than is needed or shown here.

The problem is to evaluate $\mathrm{Cov}\left\{\hat{S}(f_1), \hat{S}(f_2)\right\}$ which will yield information concerning both the interdependence of estimates along the frequency axis and the variance of each estimate.

We will, as always, assume that the process to be analyzed, $x(t)$, is a zero mean, stationary, ergodic process. In addition it will be necessary for this calculation to assume that $x(t)$ is a gaussian process. This does not necessarily imply that the results are either correct or incorrect for other types of processes. It will be assumed, without proof, that the results obtained here are reasonably accurate for the random processes observed.

It also proves convenient to assume that the estimator for $R(\tau)$ is:

$$\hat{R}(\tau) = \frac{1}{T_n'} \int_{-\frac{T_n'}{2}}^{\frac{T_n'}{2}} x(t + \tfrac{\tau}{2})\, x\,(t - \tfrac{\tau}{2})\, dt \quad . \qquad \text{[Eq. 2.4.1]}$$

This estimator could actually be used with $T_n' \leq T_n - T_m$ (T_n is the total duration of observation and T_m is the maximum lag used in the autocovariance function) but this would waste some terms which, had they been added in, would tend to decrease the variance of the

autocovariance estimate. Blackman and Tukey suggest that the algorithm of Eq. 2.2.1 be used in practice while using Eq. 2.4.1 for estimating the autocovariance function of the estimates, and using a value of T_n' somewhere between $T_n - T_m$ and T_n. They suggest $T_n' = T_n - \alpha T_m$ with $\alpha = 0.5$ for the $Q_0(f)$ window and $\alpha = 0.3$ for the $Q_2(f)$ window. It is clear, at any rate, that the approximation will be good for any $0 \leq \alpha \leq 1$ if $T_m << T_n$. $T_n' = T_n$ will be assumed from now on although the prime may be reinstated in the final results if desired.

$$\text{Cov}\left\{\hat{S}_x(f_1), \hat{S}_x(f_2)\right\} = E\left\{\hat{S}(f_1)\,\hat{S}(f_2)\right\} - E\left\{\hat{S}(f_1)\right\} E\left\{\hat{S}(f_2)\right\}$$

$$= \int_{-\infty}^{\infty}\int D(\tau_1)D(\tau_2)\cos\omega_1\tau_1\cos\omega_2\tau_2\,\frac{1}{T_n^{\,2}} \int_{-\frac{T_n}{2}}^{\frac{T_n}{2}}\int$$

$$E\left\{x(t_1 + \tfrac{\tau_1}{2})\,x\,(t_1 - \tfrac{\tau_1}{2})\,x\,(t_2 + \tfrac{\tau_2}{2})\,x\,(t_2 - \tfrac{\tau_2}{2})\right\}$$

$$- E\left\{x(t_1 + \tfrac{\tau_1}{2})\,x\,(t_1 - \tfrac{\tau_1}{2})\right\}$$

$$E\left\{x(t_2 + \tfrac{\tau_2}{2})\,x\,(t_2 - \tfrac{\tau_2}{2})\right\}\,dt_1\,dt_2\,d\tau_1\,d\tau_2$$

where $\omega = 2\pi f$ (including subscripted ω and f).

It is at this point that the gaussian assumption is used. If x_1, x_2, x_3, x_4 come from a multivariate zero mean gaussian distribution then it can be shown[15, page 168] that:

$$E\{x_1x_2x_3x_4\} = E\{x_1x_3\}\,E\{x_2x_4\} + E\{x_1x_4\}\,E\{x_2x_3\} + E\{x_1x_2\}\,E\{x_3x_4\} .$$

Under this assumption:

$$\mathrm{Cov}\{\hat{S}(f_1), \hat{S}(f_2)\} = \int_{-\infty}^{\infty}\int D(\tau_1)D(\tau_2)\cos\omega_1\tau_1\cos\omega_2\tau_2\,\frac{1}{T_n^2}\int_{\frac{-T_n}{2}}^{\frac{T_n}{2}}\int$$

$$R(t_1-t_2+\tfrac{\tau_2-\tau_1}{2})\,R(t_1-t_2-\tfrac{\tau_1-\tau_2}{2})$$

$$+ R(t_1-t_2+\tfrac{\tau_1+\tau_2}{2})\,R(t_1-t_2-\tfrac{\tau_1+\tau_2}{2})\,dt_1\,dt_2\,d\tau_1\,d\tau_2$$

$$= \int_{-\infty}^{\infty}\int D(\tau_1)D(\tau_2)\cos\omega_1\tau_1\cos\omega_2\tau_2\,\frac{1}{T_n^2}\int_{\frac{-T_n}{2}}^{\frac{T_n}{2}}\int\int_{-\infty}^{\infty}\int$$

$$S(\alpha)S(\beta)\Big\{e^{i2\pi[\alpha(t_1-t_2+\frac{\tau_1-\tau_2}{2})+\beta(t_1-t_2-\frac{\tau_1-\tau_2}{2})]}$$

$$+ e^{i2\pi[\alpha(t_1-t_2+\frac{\tau_1+\tau_2}{2})+\beta(t_1-t_2-\frac{\tau_1+\tau_2}{2})]}\Big\}$$

$$d\alpha d\beta dt_1\,dt_2\,d\tau_1\,d\tau_2 .$$

Letting $\alpha = f' + f$, $\beta = f' - f$, noting that $d\alpha d\beta = 2dfdf'$, and evaluating the t_1, t_2 integrals yields:

$$\mathrm{Cov}\left\{\hat{S}(f_1),\ \hat{S}(f_2)\right\} = \int_{-\infty}^{\infty} \iint_{-\infty}^{\infty} 4D(\tau_1)D(\tau_2) \cos \omega_1\tau_1 \cos \omega_2\tau_2\, e^{-i2\pi f\tau_2}$$

$$\cos(2\pi f\tau_1)\, d\tau_1\, d\tau_2 \int_{-\infty}^{\infty} S(f' + f)\, s(f' - f) \left(\frac{\sin 2\pi f' T_n}{2\pi f' T_n}\right)^2 df'\, df$$

$$\mathrm{Cov}\left\{\hat{S}(f_1),\ \hat{S}(f_2)\right\} = \frac{1}{4} \int_{-\infty}^{\infty} H(f,f_1)\, H(f,f_2)\, \Gamma(f)\, df \qquad \text{[Eq. 2.4.2]}$$

$$H(f,f_1) = 2\int_{-\infty}^{\infty} D(\tau) \cos \omega\tau \cos \omega,\tau\, d\tau = Q(f+f_1) + Q(f-f_1)$$

[Eq. 2.4.3]

$$\Gamma(f) = 4 \int_{-\infty}^{\infty} s(f'+f)\, S(f'-f) \left(\frac{\sin 2\pi f' T_n}{2\pi f' T_n}\right)^2 df' \qquad \text{[Eq. 2.4.4]}$$

It is clear that if $H(f,f_1)$ and $H(f,f_2)$ do not overlap (where they are non-zero) then the convariance of $\hat{S}(f_1)$ and $\hat{S}(f_2)$ will be zero. In practice $Q(f)$ will always extend to infinity but we would expect (assuming $\Gamma(f)$ is reasonably smooth) a very small covariance for $\left|f_1-f_2\right| > \frac{2}{T_m}$ for the $Q_2(f)$ window.

A few approximations now seem in order to arrive at a usable form for $\mathrm{Var}\left\{\hat{S}(f_1)\right\}$. If $S(f)$ is "smooth" over distances of $\Delta f \tilde{=} \frac{1}{T_n}$ then:

$$\Gamma(f) \cong 4\, S^2(f) \int_{-\infty}^{\infty} \left(\frac{\sin 2\pi f' T_n}{2\pi f' T_n}\right)^2 df' \quad .$$

$$r(f) \cong \frac{2}{T_n} S^2(f) \quad .$$

Referring again to Eq. 2.4.2:

$$\mathrm{Var}\left\{\hat{S}(f_1)\right\} = \frac{1}{4} \int_{-\infty}^{\infty} [H(f,f_1)]^2 \, r(f) \, df$$

$$\cong \frac{1}{2T_n} \int_{-\infty}^{\infty} [Q(f+f_1) + Q(f-f_1)]^2 \, S^2(f) \, df \quad .$$

Assuming now that $S^2(f)$ is smooth over distances of $\Delta f \cong \frac{1}{T_m}$ yields:

$$\mathrm{Var}\left\{\hat{S}(f_1)\right\} \cong \frac{1}{2T_m} S^2(f_1) \int_{-\infty}^{\infty} [Q(f+f_1) + Q(f-f_1)]^2 \, df$$

This integral can be evaluated exactly for the $Q_0(f)$ and $Q_2(f)$ windows. For the $Q_0(f)$ window:

$$\mathrm{Var}\left\{\hat{S}(f_1)\right\} \cong \frac{2T_m}{T_n} S^2(f_1) \left[1 + \frac{\sin 4\pi T_m f_1}{4\pi T_m f_1}\right]$$

$$\cong \begin{cases} 4 \dfrac{T_m}{T_n} S^2(f_1) & \text{for } f_1 = 0 \\ 2 \dfrac{T_m}{T_n} S^2(f_1) & \text{for } f_1 = \dfrac{r}{2T_m}, \; r = 1,2,\ldots \text{ (the frequencies at which estimates will actually be made).} \end{cases}$$

For the $Q_2(f)$ spectral window:

$$\mathrm{Var}\left\{\hat{S}(f_1)\right\} \cong \frac{T_m}{4T_n} S^2(f_1) \left[3 + \frac{12\,\pi^4 \sin(4\pi T_m f_1)}{4\pi T_m f_1[(4\pi T_m f_1)^2-\pi^2][(4\pi T_m f_1)^2-4\pi^2]}\right]$$

$$\cong \begin{cases} \frac{3}{2}\frac{T_m}{T_n} S^2(f_1) & \text{for } f_1 = 0 \\ \frac{3}{4}\frac{T_m}{T_n} S^2(f_1) & \text{for } f_1 = \frac{r}{2T_m},\ r = 1,2,\ldots \end{cases}$$

For spectral estimates using the $Q_2(f)$ window we will use:

$$\sqrt{\frac{\mathrm{Var}\{\hat{S}_x(f)\}}{S_x(f)}} \cong \sqrt{\frac{T_m}{T_n}} \quad . \qquad \text{[Eq. 2.4.5]}$$

For frequencies $f = \frac{r}{2T_m}$ $(r = 1,\ldots)$ Eq. 2.4.5 is slightly higher than calculated above. It should be remembered, however, that this is really only a rough approximation even for a gaussian process with a smooth spectral density. More careful evaluations of the effects of non-smooth spectrums may be found in Ref. 13 and 11. It should also be noted that all of the calculations in this section assume well behaved functions allowing complete freedom to interchange the order of integration in the above calculations. This certainly rules out a 1/f type spectral density although, as mentioned previously, the spectral density of a 1/f type process will never be estimated in this thesis without first differentiating it (via the prewhitening filter).

2.5 Sampled Data Spectral Estimator

In practice, it would be convenient to sample a noise source only at discrete time intervals, and to perform the spectral analysis using summations instead of integrals as in Eqs. 2.2.3 and 2.2.4. This is particularly true when it is desired to record data over periods of weeks or months.

Taking the expectation of Eq. 2.2.3 yields $E\left\{\hat{C}_k\right\} = R_x(k\Delta t)$. Therefore, using discrete equispaced samples, $\hat{R}_x(\tau)$ may still be estimated without any bias (although it may be estimated only at discrete lags).

In order to evaluate the effects of using Eq. 2.2.4 in place of Eq. 2.2.2 it is convenient to define the function $\underline{\text{LLI}}(x,\Delta x) \equiv \Delta x \sum_{j=-\infty}^{\infty} \delta(x-j\,\Delta x)$. Although the Dirac delta "function" is not a function in the strict mathematical sense, it has proved to be a useful (and commonly used) tool in communication theory. The following analysis could be performed using true mathematical functions, such as $\frac{\sin ax}{x}$, (with appropriate limits considered in the final step) giving the same results but requiring a great deal more work. It can be shown[13] that the Fourier transform of $\underline{\text{LLI}}(x, \Delta x)$ can be written as:

$$\int_{-\infty}^{\infty} \underline{\text{LLI}}(t, \Delta t)\, e^{-i2\pi ft}\, dt = \Delta t\, \underline{\text{LLI}}(f, \tfrac{1}{\Delta t}) \quad .$$

Consider the following:

$$\int_{-\infty}^{\infty} \text{LLI}\ (\tau,\Delta t)D(\tau)\hat{R}(\tau)e^{-i2\pi f\tau}d\tau = \Delta t \sum_{k=-\infty}^{\infty} D(k\Delta t)\hat{R}(k\Delta t)e^{-i2\pi fk\Delta t}$$

$$= \Delta t[D(o)\hat{R}(o) + 2\sum_{k=1}^{\infty} D(k\Delta t)\hat{R}(k\Delta t)\cos(2\pi fk\Delta t)]$$

$$= \Delta t[D(o)\hat{C}_o + 2\sum_{k=1}^{\infty} D(k\Delta t)C_k \cos(2\pi fk\Delta t)]$$

$$= \hat{S}(f) \text{ of Eq. 2.2.4.}$$

Therefore, for equispaced samples:

$$E\left\{\hat{S}(f)\right\} = \int_{-\infty}^{\infty} \text{LLI}(\tau,\Delta t)D(\tau)R(\tau)e^{-i2\pi f\tau}d\tau$$

$$= \Delta t\ \text{LLI}\ (f, \frac{1}{\Delta t}) * S(f) * Q(f)$$

by the convolution theorem of the Fourier transform (* denotes convolution, $Q(f) \equiv \int_{-\infty}^{\infty} D(\tau)e^{-i2\pi f\tau}d\tau$ and $S(f)$ is the true power spectral density).

If we let $S^A(f) = \Delta t\ \text{LLI}(f, \frac{1}{\Delta t}) * S(f)$ then

$$E\left\{\hat{S}(f)\right\} = S^A(f) * Q(f) \qquad \text{where}$$

$$S^A(f) = \left(\sum_{j=-\infty}^{\infty} \delta(f - \frac{j}{\Delta t})\right) * S(f)$$

$$S^A(f) = \sum_{j=-\infty}^{\infty} S(f - \frac{j}{\Delta t}) \quad \text{which will be referred to}$$

as the true aliased spectral density. If $S(f) = 0$ for $|f| \geq \frac{1}{2\Delta t}$ then

$S^A(f) = S(f)$ in the region $-\frac{1}{2\Delta t} < f < \frac{1}{2\Delta t}$ (although it should be remembered that $S^A(f)$ is periodic in f when windowing is considered).

It should be noted at this point that the equation for the $D_o(\tau)$, $Q_o(f)$ window pair was not exact (although the error was negligible in the continuous case). If one assumes that $Q_o(f) = \frac{\sin 2\pi f T_m}{\pi f}$ then taking the inverse Fourier transform of $Q_o(f)$ reveals that we really require

$$D_o(\tau) = \begin{cases} 1 & |\tau| < T_m \\ \frac{1}{2} & |\tau| = T_m \\ 0 & |\tau| > T_m \end{cases} .$$

Evaluating $\hat{S}_r = \hat{S}(\frac{r}{2m\Delta t})$ for the $Q_o(f)$ spectral window (using Eq. 2.2.4) becomes:

$$\hat{S}_r^{\circ} = \Delta t[\hat{C}_o + 2 \sum_{j=1}^{M-1} \hat{C}_j \cos(\frac{qr\pi}{M}) + \hat{C}_M \cos(r\pi)] \qquad \text{[Eq. 2.5.1]}$$

and the final $Q_2(f)$ windowed estimate will be:

$$\hat{S}_r = \frac{1}{4} \hat{S}_{r-1}^{\circ} + \frac{1}{2} \hat{S}_r^{\circ} + \frac{1}{4} \hat{S}_{r+1}^{\circ} \qquad \text{[Eq. 2.5.2]}$$

as discussed in Section 2.3. The actual spectral estimation program which will be used to analyze the experimental data uses Eqs. 2.2.3, 2.5.1, and 2.5.2.

Blackman and Tukey[13] show that the calculations for the covariance of the $\hat{S}(f)$ estimates are virtually the same for sampled data as for continuous data and so:

$$\frac{\sqrt{\mathrm{Var}\{\hat{S}(f)\}}}{S^{A}(f)} \cong \sqrt{\frac{M}{N'}} \qquad \text{[Eq. 2.5.3]}$$

where we should use $N' = N - \alpha M$ ($\alpha = 0.3$ suggested by Blackman and Tukey for the $Q_2(f)$ spectral window) if the $M << N$ approximation is not valid. $S^{A}(f)$ is the aliased true spectral density.

2.6 Dealiasing Filters

In designing an experiment which records sampled data, one normally attempts to filter the analog data before the sampling process to assure that $S(f) = 0$ for $|f| > \frac{1}{2\Delta t}$. Analog low pass filters with cut off frequencies much lower than 1 Hz are, however, impractical. If spectral density estimates are desired down to 10^{-6} Hz one might consider band limiting the analog noise to $\frac{1}{2}$ Hz, sampling once each second for 10^7 seconds, and using Eqs. 2.2.3, 2.5.1, and 2.5.2 with $M = 5 \times 10^5$ to achieve spectral estimates at $f = r \times 10^{-6}$ Hz ($r = 1,\ldots,5 \times 10^5$) each with a relative standard deviation of $\cong \sqrt{\frac{5 \times 10^5}{10^7}} \times 100\% = 22\%$. Just to calculate the sampled autocovariance function would require approximately 5×10^{12} multiplications (and as many additions) in the digital computer. This amount of computation is clearly out of the question even if an array of 10^7 data points could easily be manipulated in the digital computer. A more practical approach would to use $N \sim 10^4$ and $M \sim 100$ to get $\sim$10% estimates ($\sim \pm$ 0.5 db) over 2 decades of frequency. The data could then be resampled by a factor of 10 (or 100 if no overlap is desired) hoping to remove the effects of aliasing later (one probably only wants a frequency resolution which is approximately constant on the log frequency scale anyway).

Caloyannides[1] used this approach. He pointed out that if the spectral estimations were made using successively longer sample periods, then the spectral density is already known for

$f > \frac{1}{2\Delta t}$ (where Δt is the current sampling period) and the aliasing effect could merely be subtracted out leaving an estimate of the unaliased spectral density. Caloyannides maintains that the resulting estimate then has the same variance as if there were no aliasing. This is, in fact, not true. If $S^A(f) = S(f) + A(f)$, if $A(f)$ (the component due to aliasing) is known exactly, and the dealiased estimator is $\hat{S}(f) = \hat{S}^A(f) - A(f)$ then $\mathrm{Var}\{\hat{S}(f)\} = \mathrm{Var}\{\hat{S}^A(f)\}$ but

$$\frac{\sqrt{\mathrm{Var}\{\hat{S}(f)\}}}{S(f)} = \frac{S^A(f)}{S(f)} \quad \frac{\sqrt{\mathrm{Var}\{\hat{S}^A(f)\}}}{S^A(f)} \cong \frac{S^A(f)}{S(f)} \sqrt{\frac{M}{N'}} \quad \text{(by Eq. 2.5.3).}$$

Thus, if the true aliased spectral density were ten times the true unaliased spectral density (as occurs at some frequencies when 1/f noise is under sampled by several orders of magnitude) a $\pm$ 10% estimate of $S^A(f)$ would become a $\pm$ 100% estimate of $S(f)$. In practice, however, $A(f)$ is never known exactly. $\hat{A}(f)$ is itself only an estimate and is made at a significantly different frequency resolution than needed. At best $\hat{A}(f)$ must be the result of fitting a smooth curve to the high frequency estimates. A very narrow "bump" or "valley" could go undetected in the high frequency estimates, due to the width of the spectral window, yet cause a large bias in the low frequency estimates with their narrower resolution.

It is interesting to note that Caloyannides may not have actually used this scheme at all. His final results do not show a marked increase in relative standard deviation in the regions where the

aliasing is large. From other comments it seems possible that he may have actually used another technique. A technique which appears to have merit, on the surface, (and could easily have been used by Caloyannides judging from his comments and results) is as follows. An idealized curve is fitted to the spectral estimates for $f > \frac{1}{2\Delta t}$. An expected spectral density, $S_E(f_1)$, may then be derived at frequency $f_1 < \frac{1}{2\Delta t}$ by extrapolating the idealized curve (a $f^{-\alpha}$ curve for example). Given this idealized curve $S_E^A(f_1) = A(f_1) + S_E(f_1)$ can be calculated and a constant $C_1 = 10 \log S_E^A(f_1) - 10 \log S_E(f_1)$ arrived at. The final "dealiased" spectral density estimate at f_1 is then $10 \log \hat{S}(f_1) = 10 \log \hat{S}^A(f_1) - C_1$. This is a great morale booster since the size of the wiggles on the log scale remain the same for both the aliased and dealiased estimates (implying an unchanged relative standard deviation). Rewriting these equations, however, yields:

$$10 \log \hat{S}(f_1) = 10 \log \hat{S}^A(f_1) - 10 \log(A(f_1) + S_E(f_1)) + 10 \log S_E(f_1)$$

$$\hat{S}(f_1) = \frac{\hat{S}^A(f_1)\, S_E(f_1)}{A(f_1) + S_E(f_1)} \quad .$$

$$E\left\{\hat{S}(f_1)\right\} = \frac{S^A(f_1)\, S_E(f_1)}{A(f_1) + S_E(f_1)}$$

$$= \frac{[A(f_1) + S(f_1)]\, S_E(f_1)}{A(f_1) + S_E(f_1)}$$

If $S(f) << A(f)$ then $E\{\hat{S}(f_1)\} \cong \frac{A(f_1)S_E(f_1)}{A(f_1) + S_E(f_1)}$ which is independent of $S_1(f_1)$! In fact, if $S_E(f_1) << A(f_1)$ too, then $E\{\hat{S}(f_1)\} \cong S_E(f_1)$. In other words, if the aliasing component is large, this estimator has a large bias which tends to push the final estimate toward the value which was originally extrapolated from the higher frequency estimates. Fortunately, for a 1/f spectral density under sampled by only a few decades, the aliasing is appreciable only in the first decade below $1/2\Delta t$. The lowest decade of frequency estimates (if M = 100) would not be strongly biased.

In this thesis, however, it is hoped to improve upon these techniques. It is already intended to maintain high sample rates over much longer periods of time than Caloyannides, to allow a more careful look at the time domain character of the 1/f noise. If the data are available at a high sample rate, it is only necessary to pass these data through a sampled data low pass filter in the digital computer. What is needed is a low pass filter with a break frequency of $\sim\frac{0.1}{2\Delta t}$ to allow the sample rate to be successively divided by 10. Due to the large number of data points to be processed, the filter should not require a large number of operations per datum point. It should also have a finite (and as short as possible) impulse response so that the transient response of the filter may be discarded without a great deal of lost data. These criteria certainly rule out a transfer function resembling a "brick wall".

Although it is expected that the noise to be processed will usually have the 1/f characteristic (the aliasing essentially occurs

prior to the prewhitening filter), it was decided to specify a filter which would perform satisfactorily even for white noise. It would then perform even better for any spectral density which increases monotonically with decreasing frequency.

Let $G(f) = |H(f)|^2$ where $H(f)$ is the transfer function of the filter. The specifications for the filter are as follows:

1) $|10 \log G(f)| \leq 0.1$ db for $|f| \leq \frac{0.01}{2\Delta t}$ (so that accumulative bias will not occur due to repeated sample rate reductions)
2) $10 \log G(\frac{0.1}{2\Delta t}) \sim -3$ db and $|10 \log G(f)| \leq |10 \log(\frac{0.1}{2\Delta t})|$ for $|f| < \frac{0.1}{2\Delta t}$ (and should be as "smooth" as possible in this region).
3) $10 \log G(f) \leq -30$ db for $f \geq \frac{0.2}{\Delta t}$ (at each sample rate only the frequencies in the range $\frac{1}{2\Delta t} \leq f \leq \frac{2}{\Delta t}$ need be considered as significant contributors to aliasing)
4) $10 \log G(f) \leq 10 \log G(\frac{0.1}{2\Delta t})$ for $\frac{0.1}{2\Delta t} \leq f \leq \frac{0.2}{2\Delta t}$ (the aliasing can never, even for white noise, more than double the expected value of the spectral density).
5) $10 \log G(f) \leq -20$ db for $f \geq \frac{0.19}{2\Delta t}$ (even after 10 decades of sample rate reduction, the cumulative effects of aliasing cannot exceed ~ 0.5 db).

A rather complete discussion of digital filtering is given in Appendix C along with the description of the filter actually used. This filter meets or exceeds these specifications. A dealiasing filter for $\frac{1}{5}$ sample rate reduction is also given.

2.7 Prewhitening

The "smoothness" conditions of Sections 2.3 and 2.4 lead one to the conclusion that if the general shape of the spectral density is known, then a less biased and more accurate estimate may be obtained by first passing the noise through a "prewhitening" filter (which tends to flatten the power spectral density), estimating the spectral density of this new function, and correcting (in the frequency domain) for the known transfer function of the filter used. This could be taken as an iterative scheme where several filters must be "designed" for each set of spectral estimates made. In practice, however, one need only make the spectral density "reasonably" smooth and only one prewhitening filter will be used for all of the estimates of this thesis.

This filter was actually designed by Blakemore[11] although one significant improvement has been made to his design. A filter with transfer function $H(f)$, $G(f) = |H(f)|^2$ is needed with $G(f) \sim kf$ over as large a range as possible but with $G(f) \to f^{2n} (n \geq 1)$ as $f \to 0$ to insure that the resulting processes have a chance of being well behaved.

For any non-recursive sampled data filter (a non-recursive filter is preferable because of its finite impulse response), $G(f)$ may be written as:

$$G(f) = \sum_{k=0}^{N} A_k \cos(2\pi f k \Delta t)$$

where $N + 1$ is the number of terms in the filter and A_k are real constants. This filter is realizable if and only if $G(f) \geq 0$ for all f (see Appendix C).

Blakemore[11] suggests expanding the desired function, $G(f) = |f|$ for $|f| \leq \frac{1}{2\Delta t}$ in an infinite cosine series to find the A_k's. This results in:

$$|f| = \frac{1}{2} - \frac{4}{\pi^2} \sum_{\substack{k=1 \\ (k \text{ odd})}}^{\infty} \frac{1}{k^2} \cos(2\pi f k \Delta t) \quad \text{for} \quad |f| \leq \frac{1}{2\Delta t}$$

Truncating this series to any finite number of terms, however, does not give $G(0) = 0$. Blakemore suggests adding a constant to force $G(f) = 0$ for $f = 0$:

$$G(f) = \frac{4}{\pi^2} \sum_{\substack{k=1 \\ (k \text{ odd})}}^{N} \frac{1}{k^2} - \frac{4}{\pi^2} \sum_{\substack{k=1 \\ (k \text{ odd})}}^{N} \frac{1}{k^2} \cos(2\pi f k \Delta t)$$

implies:

$$A_k = \begin{cases} \dfrac{4}{\pi^2} \displaystyle\sum_{\substack{n=1 \\ (n \text{ odd})}}^{N} \dfrac{1}{n^2} & \text{for } k = 0 \\ -\dfrac{4}{\pi^2} \dfrac{1}{k^2} & \text{for } k \text{ odd} \\ 0 & \text{for } k \text{ even} \end{cases}$$

The actual implementation of the filter is $y_j = a_0 x_j + a_1 x_{j-1} + a_2 x_{j-2} + \ldots + a_N x_{j-N}$ where x_j and y_j are the input and output respectively and the a_k's satisfy:

$$A_0 = a^2_0 + a^2_1 + \ldots\ldots\ldots\ldots\ldots\ldots + a_n^2$$

$$A_1 = a_0a_1 + a_1a_2 + \ldots\ldots\ldots\ldots + a_{N-1}a_N$$

$$A_2 = a_0a_2 + a_1a_3 + \ldots\ldots + a_{N-2}a_N$$

" "

" "

" "

" "

" "

$$A_N = a_0a_N$$

Blakemore obtained an approximate solution to the above set of equations by the use of an iterative digital computer program. The approximate solution for N = 9 (as well as some other filters designed by Blakemore) was tested and found to behave as $|f|$ for the higher frequencies and as f^2 for a wide range of frequencies below $\frac{0.1}{2\Delta t}$. The values published by Blakemore, however, do not exactly satisfy

$\sum_{k=o}^{N} a_k = 0$ as was originally intended and cause $G(f)$ to level out

before reaching $f = 0$ (i.e., $G(0) \neq 0$). This is an important criterion when one is dealing with a $f^{-\alpha}$ process. The values, for use in this work, have been adjusted very slightly to assure that

$\sum_{k=o}^{N} a_k = 0$. The values actually used are $a_o = 1.2463822$,

$a_1 = -0.8866939$, $a_2 = -0.0962213$, $a_3 = -0.1223435$, $a_4 = 0.0376163$, $a_5 = -0.0471037$, $a_6 = -0.0172859$, $a_7 = -0.0221598$, $a_8 = -0.0070502$, $a_9 = -0.0099076$. The frequency response for this filter is shown in Fig. 2.1. It actually behaves like $|f|$ only for the highest decade of frequency and continues as f^2 to $f = 0$. This response is found to give a sufficiently smooth spectral density when the filter is applied to a 1/f type process (see Section 2.8).

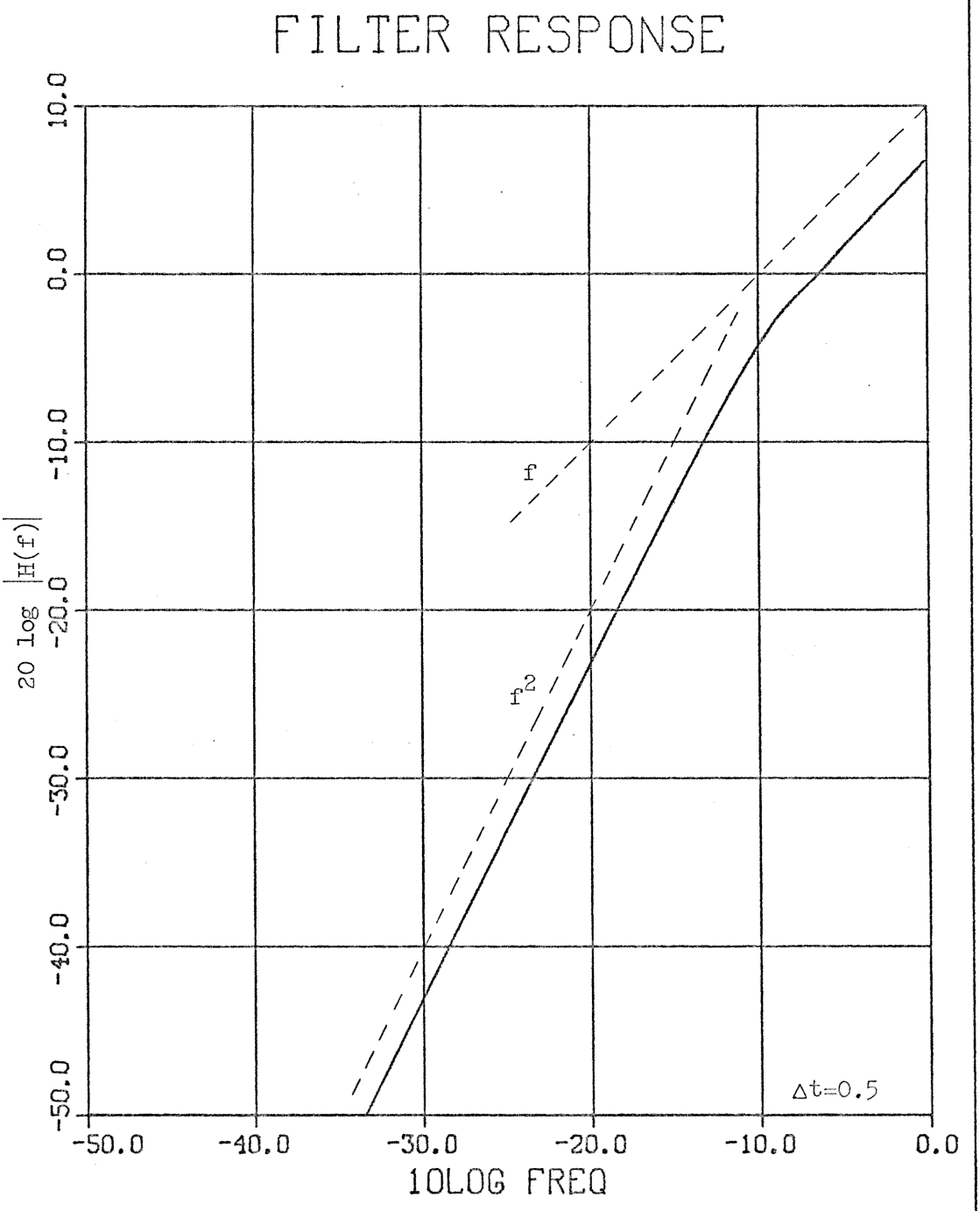


Figure 2.1 Transfer function of prewhitening filter (used in all spectral estimates).

2.8 Bias Due to "Hanning" Window Given Specific True Spectral Densities.

The specific spectral window used in all of the experimental estimates shown later is the "hanning" window: $Q_2(f) = \frac{\sin(2\pi T_m f)}{2\pi f(1+2T_m f)(1-2T_m f)}$, where $T_m = M\Delta t$ (the maximum time lag used in estimating the autocovariance function). Clearly, if the data had been exactly "prewhitened" then there would be no bias in the spectral estimate at any frequency. In practice, however, exact prewhitening would not only require very long filter lengths and much computation (a different filter would have to be designed for each estimate), but foreknowledge of the true power spectral density (requiring an iterative procedure of spectral estimation and filter design). It is, therefore, important to understand the bias which might be expected due to the spectral window.

If noise having a true spectral density of $S(f)$ is prewhitened by a filter having a magnitude squared transfer function $G(f)$, then the expected value of the estimate is:

$$E\left\{\hat{S}(f)\right\} = \frac{1}{G(f)} \int_{-\infty}^{\infty} G(\alpha)\, S(\alpha)\, Q_2(f-\alpha)\, d\alpha \quad .$$

For convenience, a value of $M = 100$ will be used in all of the experimental power spectral density estimates, resulting in spectral density estimates at $f = \frac{n}{200\Delta t}$ $(n = 1, \ldots, 100)$. The "prewhitening" filter shown in Fig. 2.1 is also used in all of the estimates.

Since $G(f) \sim f^2$ (see Fig. 2.1) for $f < \frac{10}{200\Delta t}$ it would seem useful to calculate the bias introduced at the low frequencies with $G(f) = f^2$ and several simple forms of spectral density which might be encountered.

Suppose $S(f) = A$ (white noise), then

$$\frac{E\{\hat{S}(f)\}}{S(f)} \simeq \frac{1}{f^2} \int_{-\infty}^{\infty} \alpha^2 \, Q_2(f-\alpha) \, d\alpha \qquad \text{for} \quad f << \frac{1}{2\Delta t}$$

$$\simeq \frac{1}{f^2} \int_{-\infty}^{\infty} (f-\alpha)^2 \frac{\sin(2\pi T_m \alpha)}{2\pi\alpha(1 + 2T_m\alpha)(1 - 2T_m\alpha)} \, d\alpha$$

$$\simeq \frac{1}{f^2} \, \mathrm{Im} \left\{ \int_{-\infty}^{\infty} (f-z)^2 \frac{e^{i2\pi T_m z}}{2\pi z(1 + 2T_m z)(1 - 2T_m z)} \, dz \right\}$$

$$\simeq \frac{1}{f^2} \left[\frac{f^2}{2} + \frac{\left(f + \frac{1}{2T_m}\right)^2}{4} + \frac{\left(f - \frac{1}{2T_m}\right)^2}{4} \right] \qquad \text{by contour integration.}$$

$$\frac{E\{\hat{S}(f)\}}{S(f)} \simeq 1 + \frac{1}{2(2T_m f)^2} = 1 + \frac{1}{2n^2} \qquad \text{for} \quad f = \frac{n}{2T_m}, \; n << 100.$$

Therefore, the estimates for white noise "prewhitened" by an f^2 filter would have biases of +1.76 db at the lowest frequency estimated ($f > o$) decreasing to 0.51, 0.23, 0.13, 0.09, 0.06, ... for successively higher frequency estimates.

For $S(f) = 1/f$ ($f > o$) this integral becomes more difficult. Another technique was, therefore, applied to a somewhat more general power spectral density. An important goal of this research is to search for a possible break in the 1/f characteristic. Although the

number of possible forms that such a breakdown could assume is infinite, one rather obvious possibility will be investigated.

$$\text{Suppose } S(f) = \begin{cases} \dfrac{1}{f_c + |f|} & \text{for } |f| \leq \dfrac{1}{2\Delta t} \\ 0 & \text{for } |f| > \dfrac{1}{2\Delta t} \end{cases}$$

(a 1/f characteristic with a "reasonable" form of low frequency break and bandlimited to allow sampling without aliasing). The true auto-covariance function for this process after it has been "prewhitened" by a $G(f) = f^2$ filter is $C(\tau) = 2\int_0^{1/2\Delta t} \dfrac{f^2}{f_c + |f|} \cos(2\pi f\tau)\,df$.

Letting $f_c = \dfrac{\beta}{2M\Delta t}$ and $C_q = C(\tau = q\Delta t)$ yields:

$$C_q = \frac{1}{(\Delta t)^3} \begin{cases} \dfrac{1}{4} - \dfrac{\beta}{2M} + \dfrac{1}{2}\left(\dfrac{\beta}{M}\right)^2 \ell n\left(1 + \dfrac{M}{\beta}\right) & \text{for } q = 0 \\ \dfrac{(-1)^q - 1}{2(\pi q)^2} + \dfrac{1}{2}\left(\dfrac{\beta}{M}\right)^2 \left[\cos\left(\beta\dfrac{q\pi}{M}\right)\left\{ci\left(\beta\dfrac{q\pi}{M} + \pi q\right) - ci\left(\beta\dfrac{q\pi}{M}\right)\right\}\right. & \\ \left. + \sin\left(\beta\dfrac{q\pi}{M}\right)\left\{si\left(\beta\dfrac{q\pi}{M} + \pi q\right) - si\left(\beta\dfrac{q\pi}{M}\right)\right\}\right] & \text{for } q = 1,\ldots,M \end{cases}$$

Note that for $\beta \to 0$:

$$C_q = \frac{1}{(\Delta t)^3} \begin{cases} \dfrac{1}{4} & \text{for } q = 0 \\ \dfrac{(-1)^q - 1}{2(\pi q)^2} & \text{for } q = 1,\ldots,M \end{cases}$$

and for $\beta \to \infty$:

$$C_q = \frac{1}{(\Delta t)^3} \frac{M}{\beta} \begin{cases} \frac{1}{6} & \text{for } q = 0 \\ (-1)^q \frac{1}{(\pi q)^2} & \text{for } q = 1,\ldots,M \end{cases}$$

A computer program was used to evaluate the autocovariance function for various values of β, discrete Fourier transform the truncated autocovariance function, perform the Q_2 window convolution, and divide out the effects of the prewhitening filter (in other words, this is just the power spectral density estimation algorithm applied to an ideal autocovariance function). The results are shown in Table 2.1a and 2.1b for β = 0(1/f noise), ∞(white noise), 1.0, 2.0, 3.0, 4.0, and 5.0 for the lowest decade of frequencies which would normally be estimated by this algorithm. The third column represents the bias produced by the estimation procedure and the fourth column is (where applicable) the difference (on the log scale) between the extrapolated 1/f trend of the unbiased spectral density and the biased estimate. These data are used in Section 7.5. In all of the cases shown, there is appreciable bias due to the spectral window only at the lowest frequencies estimated by a given run.

β	n	$10 \log \frac{E\{\hat{S}(f)\}}{S(f)}$	$10 \log \frac{E\{\hat{S}(f)\}}{(1/f)}$
0.0 ($\frac{1}{f}$ noise)	1	+ 0.174 db	+ 0.174 db
	2	- 0.006	- 0.006
	3	+ 0.001	+ 0.001
	4	- 0.001	- 0.001
	5	0.000	0.000
∞ (white noise)	1	1.76 db	
	2	0.51	
	3	0.23	
	4	0.13	
	5	0.09	
	6	0.06	
	7	0.04	
	8	0.03	
	9	0.03	
	10	0.02	
1.0	1	0.69 db	- 2.32 db
	2	0.07	- 1.70
	3	0.02	- 1.23
	4	0.01	- 0.96
	5	0.00	- 0.79
	6	0.00	- 0.67
	7	0.00	- 0.58
	8	0.00	- 0.51
	9	0.00	- 0.46
	10	0.00	- 0.41
2.0	1	0.98 db	- 3.79 db
	2	0.14	- 2.87
	3	0.04	- 2.18
	4	0.02	- 1.75
	5	0.01	- 1.45
	6	0.00	- 1.25
	7	0.00	- 1.09
	8	0.00	- 0.97
	9	0.00	- 0.87
	10	0.00	- 0.79

Table 2.1a : given $S(f) = \frac{1}{|f| + \frac{\beta}{2M\Delta t}}$, $E\left\{\hat{S}(f)\right\}$ is the expected value of the estimate using a f^2 prewhitening filter where M = 100, $f = \frac{n}{2M\Delta t}$.

β	n	$10 \log \frac{E\{\hat{S}(f)\}}{S(f)}$	$10 \log \frac{E\{\hat{S}(f)\}}{(1/f)}$
3.0	1	1.14 db	- 4.88 db
	2	0.20	- 3.78
	3	0.06	- 2.95
	4	0.03	- 2.41
	5	0.01	- 2.03
	6	0.01	- 1.75
	7	0.00	- 1.55
	8	0.00	- 1.38
	9	0.00	- 1.25
	10	0.00	- 1.14
4.0	1	1.25 db	- 5.74 db
	2	0.24	- 4.53
	3	0.08	- 3.60
	4	0.04	- 2.98
	5	0.02	- 2.54
	6	0.01	- 2.21
	7	0.01	- 1.96
	8	0.00	- 1.76
	9	0.00	- 1.59
	10	0.00	- 1.46
5.0	1	1.33 db	- 6.45 db
	2	0.27	- 5.17
	3	0.10	- 4.16
	4	0.04	- 3.48
	5	0.02	- 2.99
	6	0.01	- 2.62
	7	0.01	- 2.33
	8	0.01	- 2.10
	9	0.00	- 1.92
	10	0.00	- 1.76

Table 2.1b

2.9 Mean Removal

Until now, the processes for which the spectral density estimates are to be made have been assumed to have zero means. In fact, since the noise sources are dc coupled to the voltage sampling device with the noise sources initially set to some small but arbitrary voltage, the means will never be zero. It would, therefore, seem necessary to remove the mean at some point in the estimation procedure. If the observed process is $x(t) = y(t) + c$, where c is a constant and $y(t)$ is a zero mean process, then one might estimate c (by taking the time average of $x(t)$ for example) and subtract $\hat{c}$ from all of the $x(t)$ sample values. Blakemore[11] suggests that a more efficient scheme is to subtract a constant from each of the autocovariance estimates (there are usually much fewer terms involved). For either the sampled data or continuous estimation schemes:

$$E\left\{\hat{R}_x(\tau)\right\} = E\left\{(y(t+\tau) + c)(y(t) + c)\right\}$$

$$= E\left\{y(t+\tau)y(t)\right\} + c^2$$

$$= R_y(\tau) + c^2 \quad .$$

Therefore, if $\hat{R}_y(\tau) \equiv \hat{R}_x(\tau) - c^2$ then $E\{\hat{S}_y(f)\}$ will have no additional bias. Unfortunately, in practice, c is not known exactly. Blakemore shows that if $\hat{c} = \frac{1}{N}\sum_{j=1}^{N} X_n$ then, if the $Q_o(f)$ spectral window is used, the spectral window is modified, only for the $f = 0$ spectral density estimate, to $Q^{\circ}_{o}(f) = \left(\frac{\sin 2\pi fM\Delta t}{2\pi fM\Delta t}\right) - \left(\frac{\sin \pi fN\Delta t}{\pi fN\Delta t}\right)^2$. For estimates

using the $Q_2(f)$ spectral window the spectral windows are modified for the $f = 0$ and $f = \frac{1}{2M\Delta t}$ estimates to $Q_2^{\circ}(f) = Q_2(f) - \frac{1}{2}\left(\frac{\sin \pi f N\Delta t}{\pi f N\Delta t}\right)^2$ and $Q^1{}_2(f) = Q_2(f) - \frac{1}{4}\left(\frac{\sin \pi(f - \frac{1}{2M\Delta t} N\Delta t}{\pi(f - \frac{1}{2M\Delta t}) N\Delta t}\right)^2$. If $N >> M$ and if $\int_0^{\frac{1}{N\Delta t}} S_y(f)\, df << \int_{\frac{1}{N\Delta t}}^{\frac{1}{M\Delta t}} S_y(f)\, df$ then the bias will be negligible and, at any rate, restricted to the $f = 0$ and $f = \frac{1}{2M\Delta t}$ estimates. One might, however, wish to consider the $f = 0$ estimate to actually be an estimate for some frequency in the range $0 < f < \frac{1}{2M\Delta t}$.

In this thesis, there are no spectral density estimates shown for which the prewhitening filter was not used. This filter acts as a differentiator and automatically removes any mean. If z_i represents the output of the prewhitening filter with an input of $n_i = y_i + c$ then:

$$z_i = \sum_{k=0}^{N'} a_k x_{i-k} \qquad (N' \text{ is the length of the filter})$$

$$= \sum_{k=0}^{N'} a_k y_{i-k} + c \sum_{k=0}^{N'} a_k$$

$$= \sum_{k=0}^{N'} a_k y_{i-k}$$

just as if the mean had never been there.

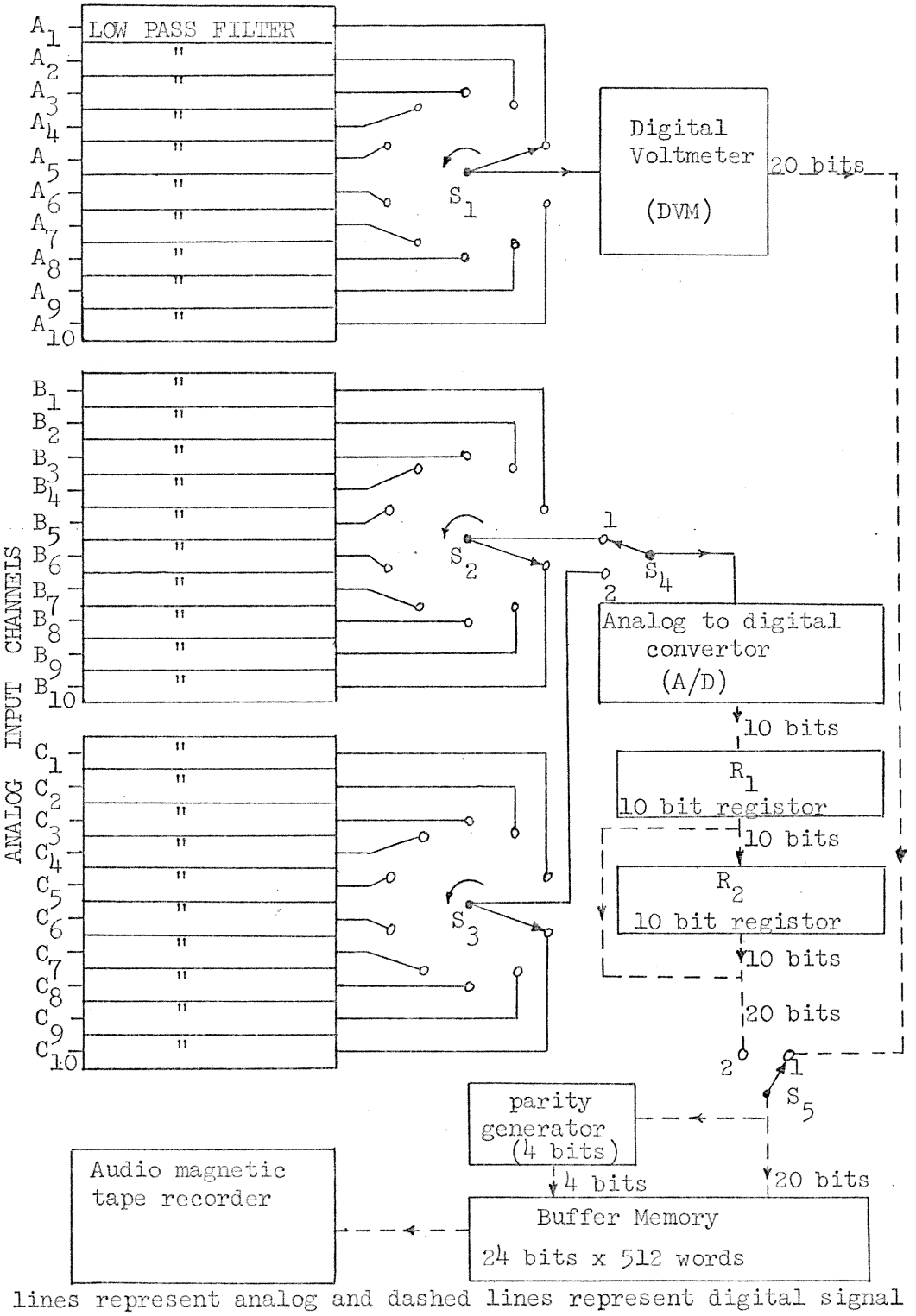
Caloyannides[1] is concerned about the effects of an additive linear term ($n(t) = y(t) + c_2 t$). In this case the output of the prewhitening filter would have a non-zero mean due to differentiating the linear term. For this reason, the mean removal was included in the spectral estimation program used in this thesis. The bias is minimal even for a 1/f process since the output of the prewhitening filter would have a spectral density $S_z(f) \propto |f|$. No provision, however, was made in the program for linear trend removal since this would amount to removing a t^2 term in the original data.

Chapter III

DATA RECORDING SYSTEM

3.1 The Data Recording System

A system was required to measure and record the various voltages available in a form which could be conveniently transferred to the digital computer for analysis. Since the system was to be operating for nearly a year, it was desirable that the equipment require attention no more than once every few days. The block diagram of the system used is shown in Fig. 3.1 with the appropriate timing shown in Fig. 3.2. There are 30 analog channels labeled $A_1,\ldots,A_{10}$, $B_1,\ldots,B_{10}$, $C_1,\ldots,C_{10}$. Each of these channels passes through a low pass presampling filter and unity gain buffer amplifier (Fig. 3.3). Multiplexing switches S_1, S_2, and S_3 are comprised of reed relays, while switch S_4 is an analog FET switch (siliconix, DG181BA). Individual filters prior to the multiplexing switch are necessary since the time for which a given channel is connected to the voltage measuring device is very nearly equal to the time constant of the desired filter. The noise which is introduced by the buffer amplifiers should be negligible based on calculations involving typical noise characteristics of the amplifiers available from the manufacturers (current noise at the input to the device is the dominant component here due to the large input resistances involved). Grounding data channels A_1, B_1, and C_1 (see Table 5.1) during the experiment yields data to support this assumption.



Solid lines represent analog and dashed lines represent digital signal paths.

Figure 3.1 Block diagram of data recording system.

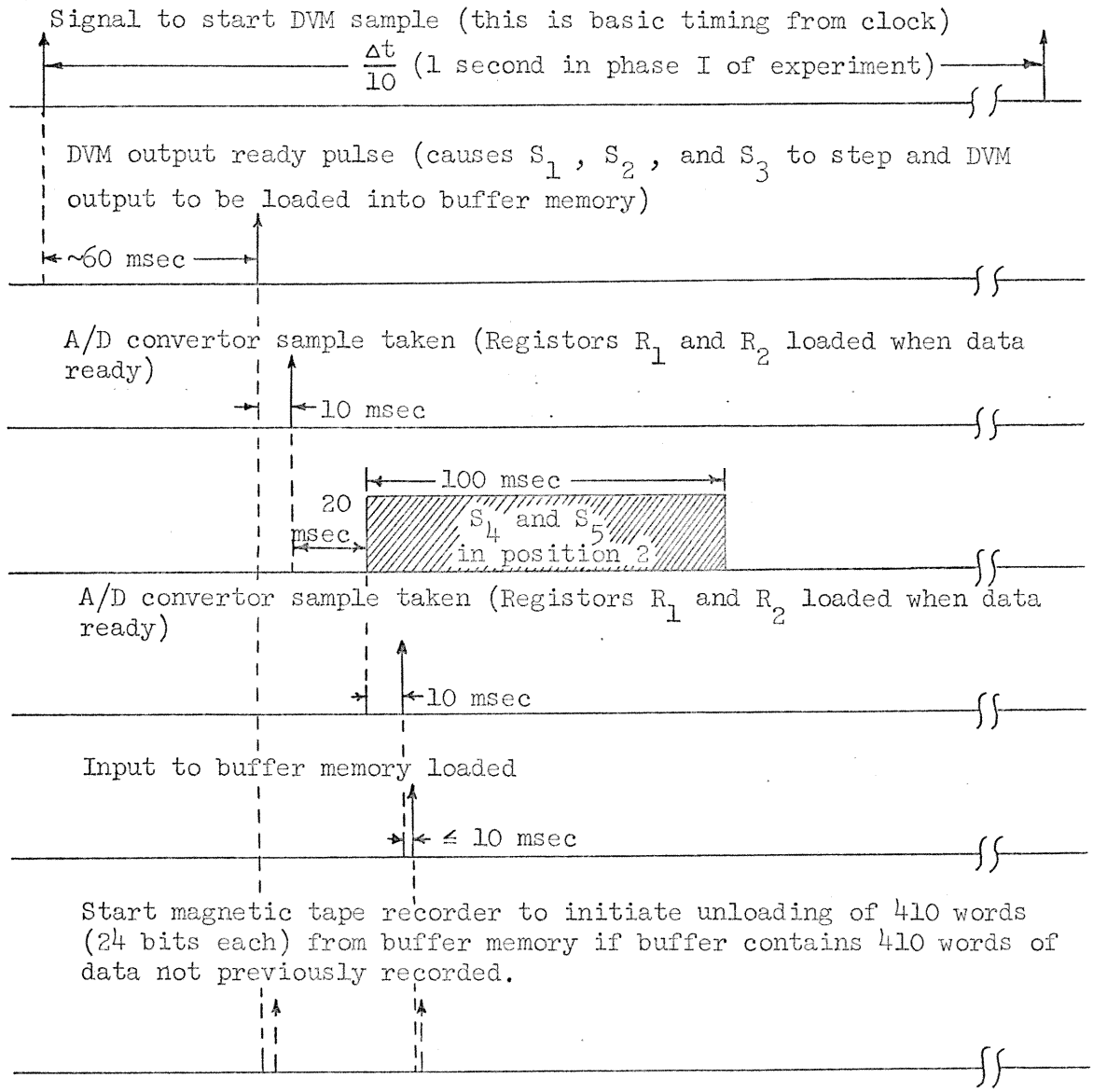
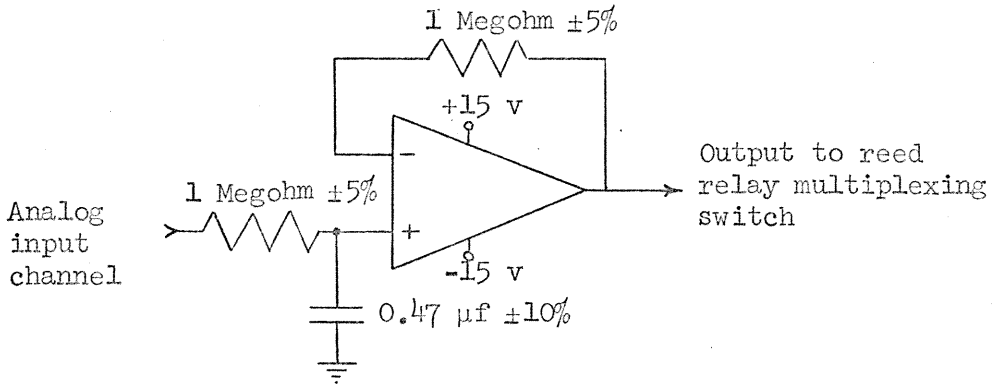


Figure 3.2 Timing diagram of data recording system.



Amplifier is SN72308 for channels $A_1, \ldots, A_{10}$ and one half of a SN72747 for channels $B_1, \ldots, B_{10}, C_1, \ldots C_{10}$.

Figure 3.3 Low pass filter/buffer amplifier used prior to multiplexing switches.

Input channels $A_1,\ldots,A_{10}$ are measured by a 4 digit digital voltmeter (DVM) which has either ±.9999 volt full scale range or ±9.999 volt full scale range when used with the external autoranging circuitry to be discussed later (the external autoranging circuitry is designed to achieve the maximum sampling rate possible under these multiplexed conditions). The DVM has a 5% "overrange" capability which extends its range to approximately ±1.2 volts or ±12 volts. The maximum sample rate of the DVM is ∿1 sample per second under multiplexed conditions (maximum sample rate of 1 sample per 10 seconds for each of the 10 channels $A_1,\ldots,A_{10}$).

Channels $B_1,\ldots,B_{10}$, $C_1,\ldots,C_{10}$ are sampled by a ten bit analog to digital converter (A/D) with a ±5 volt input range (Analog Devices model ADC10QM). This configuration of a DVM and A/D converter operating in parallel was chosen so that various voltages for which 0.01 volt resolution is satisfactory could be measured without sacrificing sampling rate for each channel. The high speed of the A/D converter allows channels B_i and C_i to be sampled within 0.04 seconds after channel A_i. The displacement of switches S_2 and S_3 relative to S_1 in Fig. 3.1 is intentional. DVM samples occur immediately before switching and A/D samples occur immediately after switching.

The DVM digital output requires 20 binary bits (external logic transforms the output of the DVM to the BCD format shown in Table 3.1). The digital output of the A/D for channels B_i and C_i are packed together to form one 20 bit word. Four parity bits are added to form 24 bit words. The interlacing nature of the 4 bit parity check yields a

Bits	Interpretation
$(b_{23}, b_{22}, b_{21}, b_{20})$	most significant digit of DVM in BCD
$(b_{19}, b_{18}, b_{17}, b_{16})$	2nd digit of DVM
$(b_{15}, b_{14}, b_{13}, b_{12})$	3rd digit of DVM
$(b_{11}, b_{10}, b_9, b_8)$	4th digit of DVM
(b_7, b_6)	(0,0)⟹+ sign, (1,0)⟹- sign
	(1,1)⟹overrange (sign unknown)
(b_5, b_4)	(1,0)⟹1 volt range
	(0,1)⟹10 volt range
	(0,0)⟹100 volt range

Table 3.1: DVM data word format

system in which the probability of missing an even multipled bit error (errors often occur in continuous groups due to tape dropout) is significantly reduced. The offending bit position(s) which is in error may often be ascertained since the magnitude of each channel often does not change by large increments between successive samples. The disallowed states in BCD coding give a small degree of additional error detection capability. Over all, recording errors did not prove to be a problem. Less than 1 word in 10^5 showed parity errors which in most cases appeared to occur as a single bit error which could be corrected with reasonably high certainty.

Consider each 24 bit word to be composed of bits $b_{23},\ldots,b_0$. For a word representing channels B_i and C_i, bits $b_{23},\ldots,b_{14}$ represent the B_i channel voltage and $b_{13},\ldots,b_4$ represent the C_i channel voltage (each channel is represented by a 10 bit unsigned integer, I, where the channel voltage is $\frac{I-512}{1024} \times 10$ volts). The DVM word format is described by Table 3.1. The parity bits b_3, b_2, b_1, and b_0 are generated by the equation $b_j = \left(\sum_{k=1}^{5} b_{j+4k}\right) \bmod 2$ for $j = 0,1,2,3$.

The 24 bit data words are stored in a buffer memory composed of 24 circulating shift registers each containing 512 bits shifted at a 64 kHz rate. A counter running at the same rate as the shift registers identifies the word address available at any instant of time. Reading and writing are independent and may occur simultaneously. Maximum access time for any word position is 0.008 second. A number $N \leq 512$ (wired into the buffer memory control) determines the word length of

each "record" which is transferred onto an audio magnetic tape recorder in a manner to be discussed later. Transfer is automatically initiated as soon as the Nth word past the end of the previous N word record has been stored in the buffer memory. The memory controller actually transfers two additional words past the end of the record. These additional words contain the same data as the first one or two words of the following "record" if the rate at which words are being stored in the buffer memory is greater than the reciprocal of the total time required to transfer a "record" to magnetic tape. This overlap provides a method for testing the continuity of "records" after they have been transferred into the digital computer.

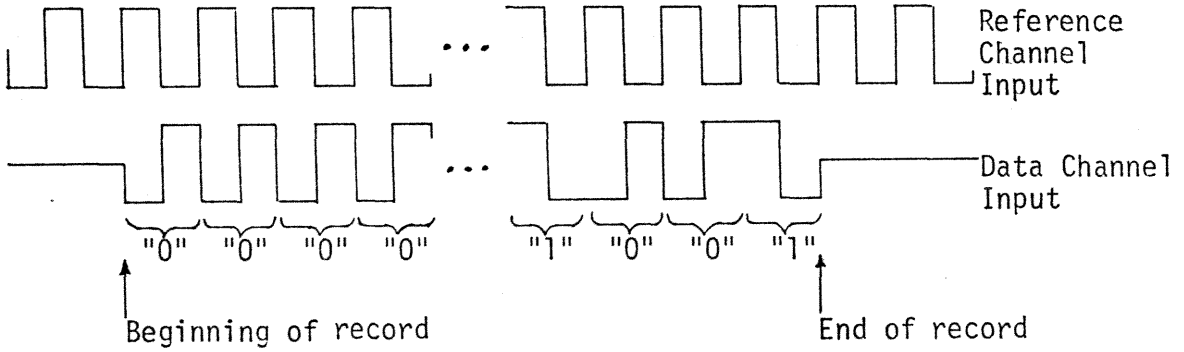
After the Nth word is stored in the buffer memory, the tape recorder motor is turned on for 1 second before two 24 bit words containing all zeros and then the data record are written onto the tape at a rate of 10^3 bits/second. The constraint that a word in the buffer memory which has not been transferred out onto tape should never be written over yields a theoretical maximum rate at which words may be stored in the buffer memory of

$$R_{max} \cong \begin{cases} \dfrac{N}{1.048 + N \times .024} & \text{for } N \leq 472 \\ \dfrac{513 - N}{1.072} & \text{for } N \geq 473 \end{cases} .$$

The peak R_{max} is 38.1 words per second at N = 472. The actual value of N (number of words per record) used in the experiment is 410 giving a maximum allowable storage rate in the buffer memory of $R_{max} \cong 37.6$ words/

second (the maximum storage rate actually used in the data recording experiment is 2 words/second).

The tape recording scheme was contrived to make use of a relatively inexpensive audio tape recorder (which is readily available and is thought to be highly reliable relative to the more expensive digital incremental tape systems available). The data are recorded with minimal modification to the tape recorder so as not to undermine its future usefulness. Two such tape decks (Sony Model TC-352D and TC-353D) are used (one for recording and the other for playback). The heads were aligned to insure that the relative phase between the two channels would be the same for both recorders. The only modification made was a triac switch allowing the motor (but not the electronics) of the record unit to be remotely switched on and off. Recording is performed at 1 7/8 i.p.s. tape speed. A 1 kHz square wave is constantly applied to the reference channel. A 1 kHz square wave whose phase is set to either 0° or 180° over 1 cycle of the reference channel (to represent a logical "1" or "0") is applied to the data channel of the recorder.



The circuit driving the data channel is a tri-level DC coupled circuit to remove transient effects at the beginning of the record due to the AC coupling of the tape recorder system. The record levels are in the normal "linear" range of the recorder.

The tape must be brought to nominal speed (requiring about 1 second) before recording can begin. The efficiency of tape use is, therefore, approximately $\frac{M \times 10^{-3}}{M \times 10^{-3} + 1} \times 100\%$ where M is the number of binary bits in the data record. 410 data words per record gives a tape efficiency greater than 90%. The bit density on the magnetic tape is ~530 bits/inch within the record interval. Because of the 1 second required to achieve nominal tape speed (and tape stretch), approximately 20.5 inches of tape are used in writing each 410 word record. This means that approximately 1053 records may be written on each side of an 1800 ft. roll of tape. If each of the 30 channels is sampled with a period of Δt = 10 seconds (this is the sample rate for phase I of the experiment which lasts for ~3 months), each record represents 3 minutes 25 seconds. This is sufficient time to change tapes without loss of data (and even clean the tape heads occasionally). At this sample rate each side of a tape could store the data for ~ 2.5 days. During phase I of the experiment, 2 days of data were stored on each side of a tape for operator convenience. This phase of the experiment required 7.5 rolls of magnetic tape per month. Performing the same experiment with paper tape would require (assuming 4 punched rows per word, 10 rows/inch) 173 rolls (1000 ft each) of paper tape per month (assuming 100% utilization of the tape). The paper tape would require changing once each

4.16 hours resulting in considerable operator fatigue over a period of a few months (even if one assumes that the mechanical paper tape punch could survive the ordeal). In Phase II of the experiment (which lasted for ~9 months) the sample period was increased to Δt = 100 sec. During this phase the magnetic tape was changed once each 20 days whereas the paper tape would still be limited to 1.74 days.

The tape is decoded at 7 $\frac{1}{4}$ i.p.s. allowing 2 days of data (20 days in Phase II) to be transferred onto an IBM compatible tape in about 40 minutes. The block diagram of the decoder is shown in Fig. 3.4. The multiplier-integrator scheme is used to decrease errors due to tape stretch (causing relative phase shifts between channels) and dropout. The data are transferred to a 7 track IBM compatible tape for use in the computer. The relatively short time periods for which the incremental tape transport was needed allowed this piece of equipment to be borrowed (and properly maintained) during the course of the experiment. Reliability and cost were not, therefore, factors for this piece of equipment.

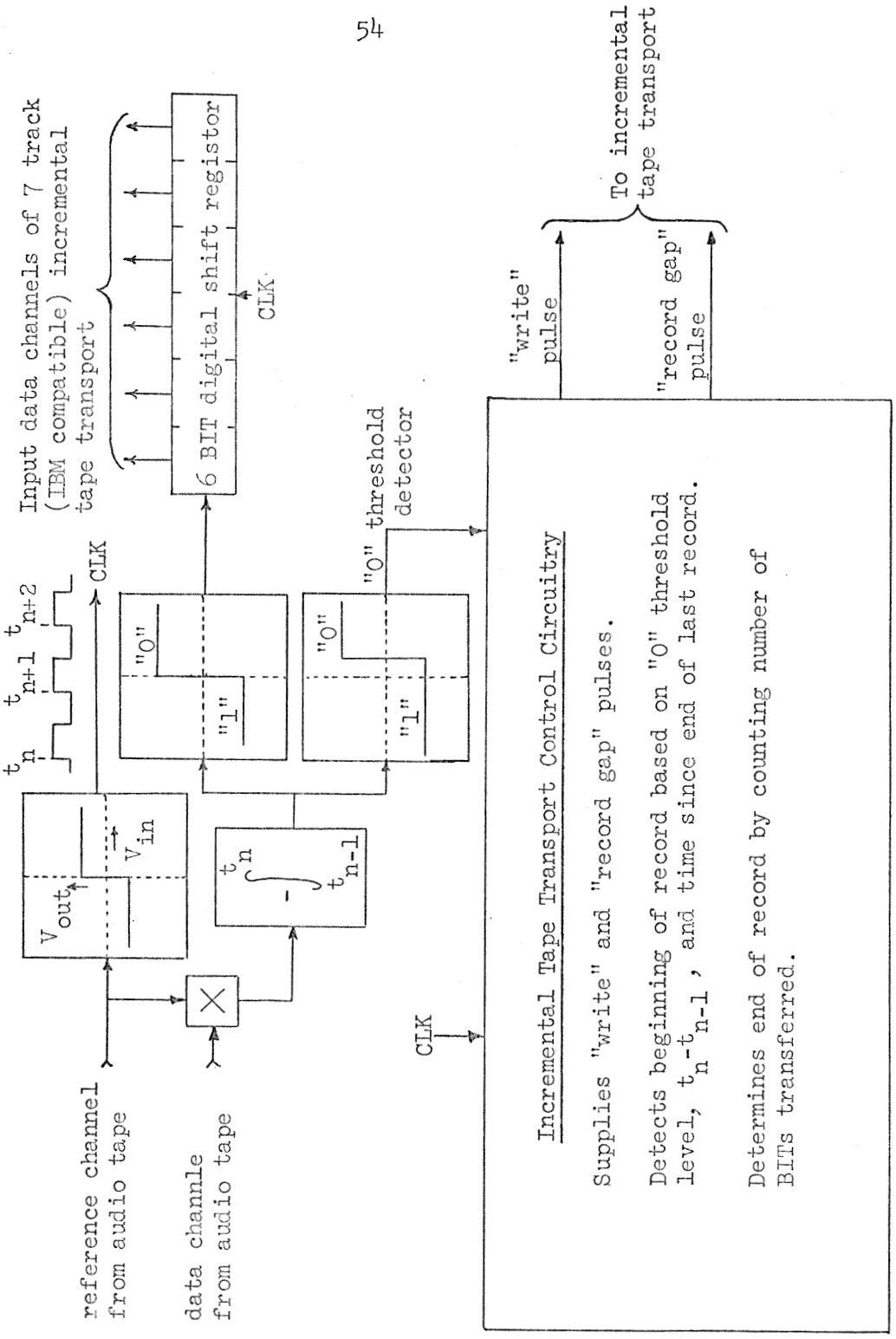


Figure 3.4 Block diagram of audio tape to IBM compatible tape data Transfer system.

3.2 Digital Voltmeter Autoranging

The digital voltmeter (DVM) used in this experiment (Hewlett-Packard 3440A with 3443A plug-in unit) provides five ranges of sensitivity (from ±.09999 volts to ±999.9 volts full scale). It was desired to use the ±.9999 volt range to give a resolution of 0.1 millivolt although the nature of the waveforms to be measured (random processes with unknown statistics) demanded that the ±9.999 volt range (~±12 volts with overrange) be available. The DVM provides for an autoranging capability but does not guarantee an accurate reading within the 1 second desired for this experiment.

The specifications of the DVM state that an accurate reading may be made within 0.45 seconds on the ±9.999 volt scale and 1 second on the ±.9999 volt scale. The autoranging algorithm (performed by a specially constructed controller) is as follows. As soon as the DVM completes a measurement of one analog channel, the DVM is set to the ±.9999 volt range at the same time as the analog channels are switched. This guarantees ~1 second settling time if this scale is maintained. At least .65 seconds (1 second in Phase II of the experiment) before the final reading is made, a preliminary reading is taken with the DVM. If the most significant digit is "9" (or the DVM indicates overrange) the range is set to ±9.999 volts allowing > 0.45 seconds settling time at this range setting before the final measurement is taken. Since the settling time allowed before the preliminary reading may be under .33 seconds, the unit occasionally fails to uprange for a voltage greater in magnitude than 1 volt. It was found experimentally, however, that the 5%

overrange capability of the DVM is sufficient to allow for a correct reading under these marginal circumstances although the sign is lost under overrange conditions. The sign is assumed by the computer software to be the same as the previous sample of the same channel when overrange is detected.

3.3 Uninterruptible Power Supply

Preliminary runs indicated that momentary losses in line voltage (due to switching in the power system) are not uncommon. These momentary line failures not only produced glitches which could be observed in the noise data, but occasionally caused such confusion in the data recording controllers as to effectively terminate the run.

One solution which was considered was to provide individual battery supplies for the +5, ±15, ±20 volts needed to operate the noise sources, ovens, and various controlling circuitry which had been constructed. An AC supply would still be required for the tape recorder motor, however, and to supply the electronics of the commercial DVM and tape recorder unless rather severe surgery was performed on these units. It was decided, therefore, to build an inverter capable of operating all of the equipment (utilizing the 115 volt (60 Hz) power supplies already in use to provide the various DC voltages).

Since no switching transients could be tolerated, the inverter was required to supply power during the entire course of the experiment. The inverter (Figs. 3.5 and 3.6) is supplied by a DC power supply in parallel with two 12 volt automobile storage batteries so that failure of the line voltage results in only the slight drop in voltage due to the internal resistance of the storage batteries. The inverter must supply a nearly constant load current of ~2.5 amps (r.m.s.) with a peak voltage of ~115 x $\sqrt{2}$ volts (163 volts) to satisfy the DC supplies which primarily use capacitor input filters. A tri-level waveform (shown in

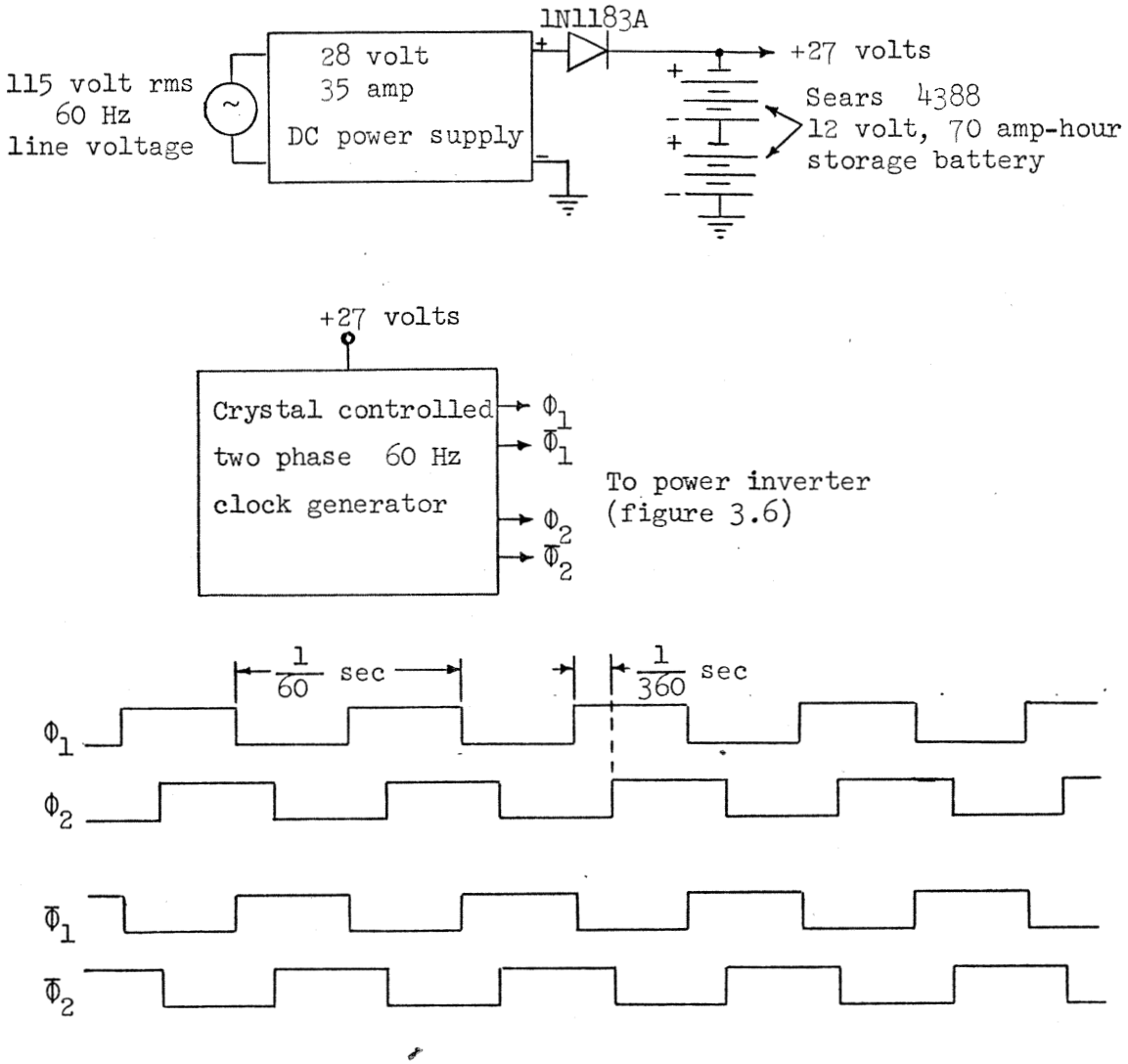


Figure 3.5 Uninterruptible power source and clock for power inverter.

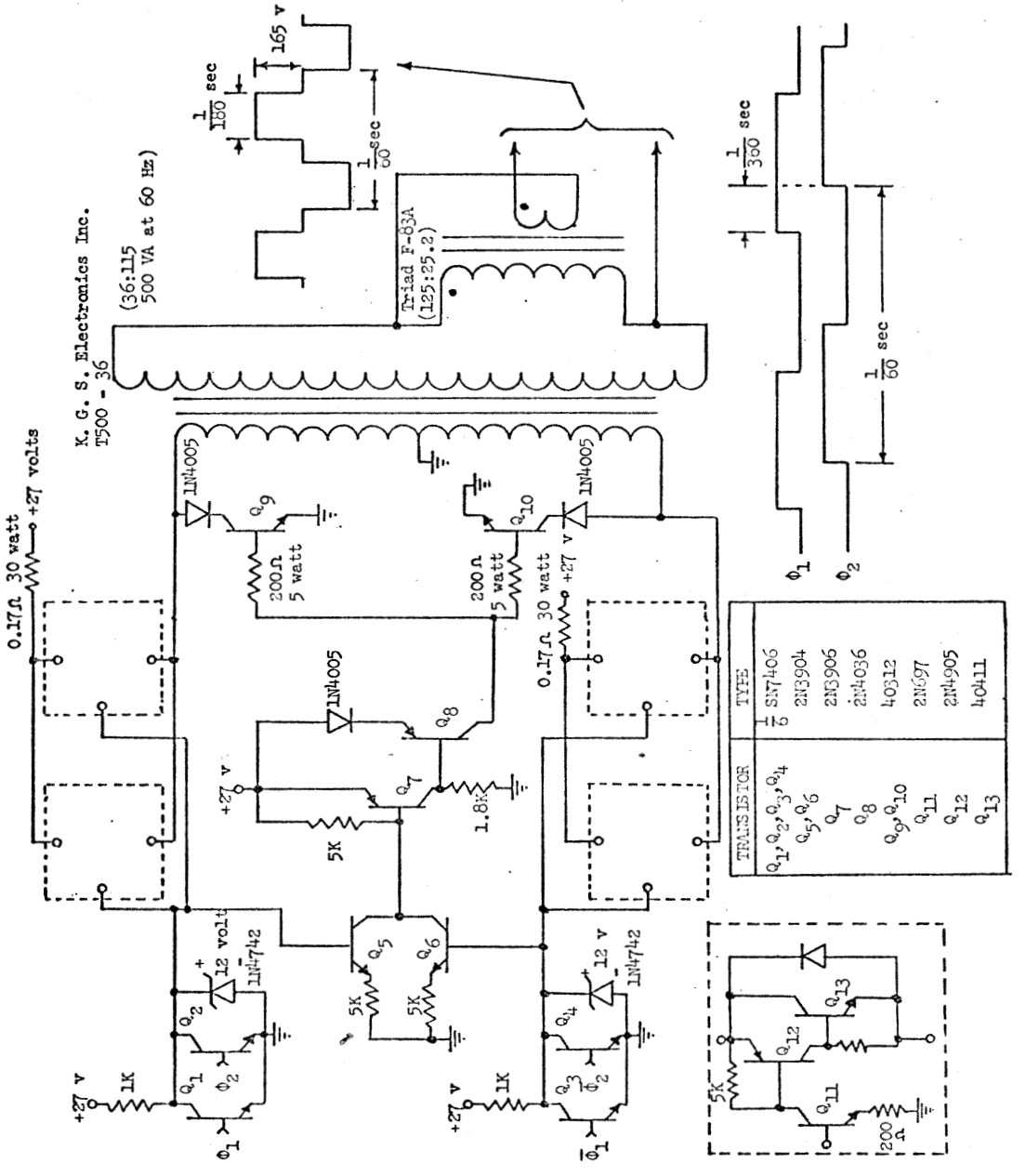


Figure 3.6 Schematic of power inverter.

Fig. 3.6) which is "on" 2/3 of the time and "off" 1/3 of the time provides a voltage integral over one-half cycle of $\frac{163}{120} \times \frac{2}{3}$ (as compared to $\frac{163}{120} \times \frac{2}{\pi}$ for a sine wave) to insure that the power transformers in the system not saturate and that the tape recorder motor operate properly.

The battery capacity is sufficient to provide reasonably constant voltage for at least ten minutes. This time period is more than sufficient to allow for the transition to an alternate power source during planned building power interruptions (as occurred once during the course of the year long experiment).

The frequency of the power inverter is crystal controlled so that the timing for the data sampling system may be derived from this source.

Chapter IV

NOISE SOURCES

4.1 Introduction

The purpose of obtaining experimental data on low frequency semiconductor noise is to extend the present knowledge of the statistics of this type of noise process. It is hoped that these data will eventually be useful in verifying (or at least eliminating) various hypotheses as to the physical process which generates a $f^{-\alpha}$ spectral density for frequencies ranging down to the microhertz region. Past experiments[1,11] have indicated that this trend may continue down to $\sim 10^{-6.3}$Hz. It is hoped that this experiment might not only extend the frequency range of available data but, in addition, increase the accuracy of the spectral estimates in the frequency ranges already investigated and allow for a closer look at the data to see if any components of the noise process can be identified (either coming from the noise source itself or induced from external influences) which might give a clue as to the physical process involved in the 1/f noise or indicate experimental errors which might have been made now or in the past.

4.2 The Noise Sources

The noise sources chosen for this experiment are bipolar integrated operational amplifiers. This choice was made due to the inherently low temperature and power supply sensitivities of these amplifiers (relative to their noise) as well as their widespread use. These devices are, unfortunately, rather complex structures having many interior parameters which are either difficult to obtain or may, in fact, be unknown even by the manufacturer. The circuit diagram of each device used suggests that the differential input pair of transistors should be the largest contributor to the output noise of the device if each of the transistors in the circuit has reasonably high current gain and similar noise properties. The extent to which this is true is not important here since we are primarily concerned only with characterizing the statistical behavior of a common semiconductor 1/f noise process even though it may be the result of adding the noise of several of the transistors in the operational amplifier. The actual circuit components will, however, be identified as completely as possible. Negative feedback is placed around each amplifier so that the magnitude of the total gain (and, therefore, the magnitude of the observed noise referenced to the input) is both known and reduced to a level such that the probability of saturation of the amplifier due to its own noise is small. The linearity of the amplifier is also improved by the feedback.

Past work in this field indicates the necessity of regulating the power supply voltages and the temperature of the noise source. A great deal of effort was expended to reduce the temperature

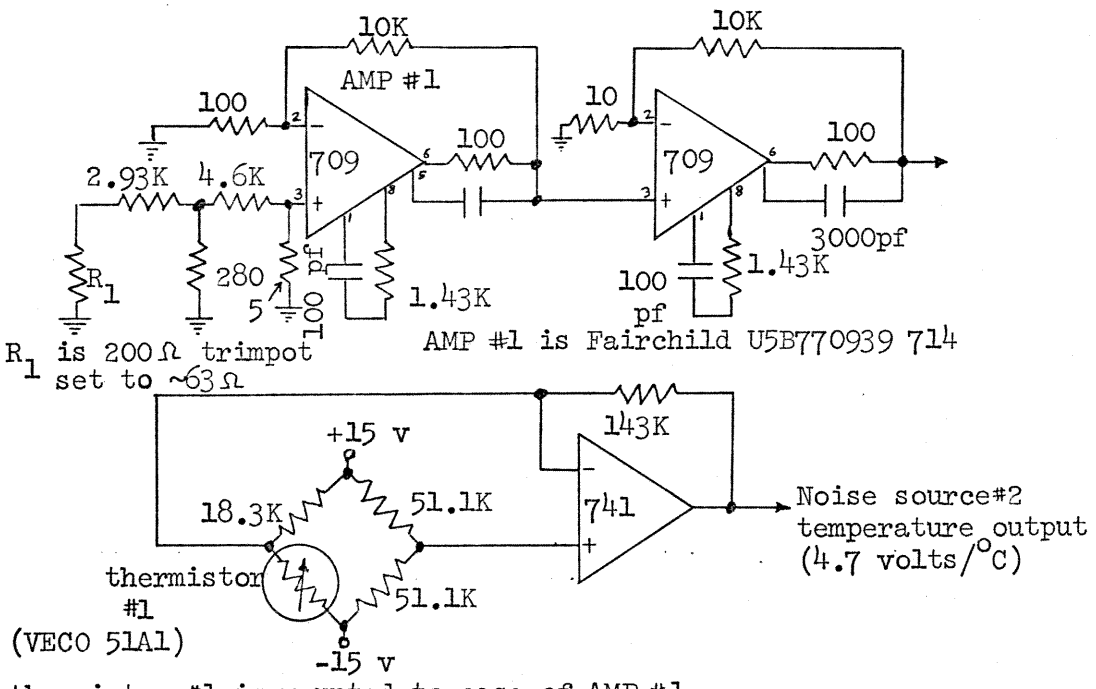
variations of both the noise sources and the final power supply regulators. This involved a reasonably large volume of circuitry which somewhat complicated the task. It was desired to limit the time spent on constructing an oven so as to allow for more time (giving greater frequency resolution and statistical accuracy) for recording data. A compromise was reached in that only moderately sophisticated ovens were constructed which hold the temperature fluctuations to about 100 millideg centigrade peak to peak but where the temperatures of the noise sources are recorded during the course of the experiment with very fine resolution. The ovens filter out any high frequency components of room temperature to reduce the possibility of non-negligible temperature gradients across the integrated circuit chip as the temperature changes. It is then assumed that the effect of the temperature fluctuations on the output of each noise source can be considered as a linear temperature coefficient. The temperature of each source is recorded along with the noise data by the use of a thermistor mounted to the case (as near as possible to the silicon chip). The temperature component may then be subtracted from the data as they are analyzed. Functions of the power supply voltages are also recorded for 5 of the 6 noise sources from which the final data were taken so the effects of power supply noise can be investigated (See Fig. 4.4, $V_{\pm check}$).

Many calculations were performed using "typical" and "worst case" specifications supplied with the components to be used to try and insure that the external influences on the noise sources would be negligible. Various relatively short preliminary data recording runs

were also made. In all cases, however, there were insufficient data available concerning the very low frequency noise characteristics of the various components. The final analysis, therefore, depends almost entirely on the parameters measured during the course of the experiment rather than guesses made from attempting to interpret manufacturers specifications or trends observed over short experimental runs. The calculations leading to the design goals, therefore, will be omitted for the most part and their sufficiency will be verified by experimental data.

Two separate ovens (one containing a single noise source which will be referred to as noise source #2, and the other containing 5 noise sources referred to as T1, T2, T3, T4, and T5) were used in obtaining the final experimental data. The circuit diagrams of the 6 noise sources are shown in Figs. 4.1 and 4.3. The resistance values indicated are the final values chosen to set all output voltages to ~0 volts at the beginning of the experiment. A single resistor may actually represent a network of resistors in order to achieve a reasonable operating point. The values were often measured to the implied accuracy since the components could not be substituted while the ovens were operating and new values had to be extrapolated from precisely known previous values when operating points were being established. All resistors used inside the ovens are either 1% metal film (RN60C or RN60D) or wire wound with precision of 1% or better (except for the frequency compensation resistors, used only for the 709 operational amplifiers, which are 5% carbon). Each amplifier (Figs. 4.1,...,4.6)

labeled 741 is actually one half of a Fairchild U7A7747393 dual operational amplifier (except the 741 amplifier of Fig. 4.1 which is a Fairchild U9T7741393). The amplifiers serving as final noise source amplifier and thermistor amplifier for each noise source (T1,...,T5) are contained in a single 747 package. Amp #1 (the amplifier considered to be the actual source of noise) of Fig. 4.3 is a 709 for T3, T4, and T5 noise sources, and is a shared 747 dual operational amplifier for T1 and T2.



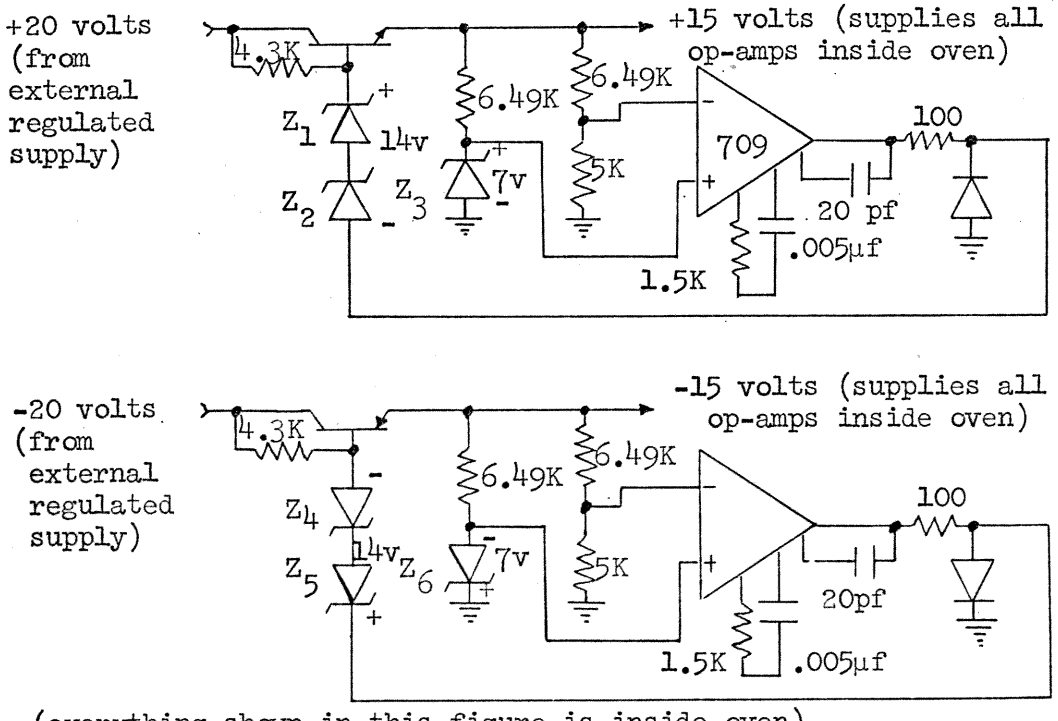
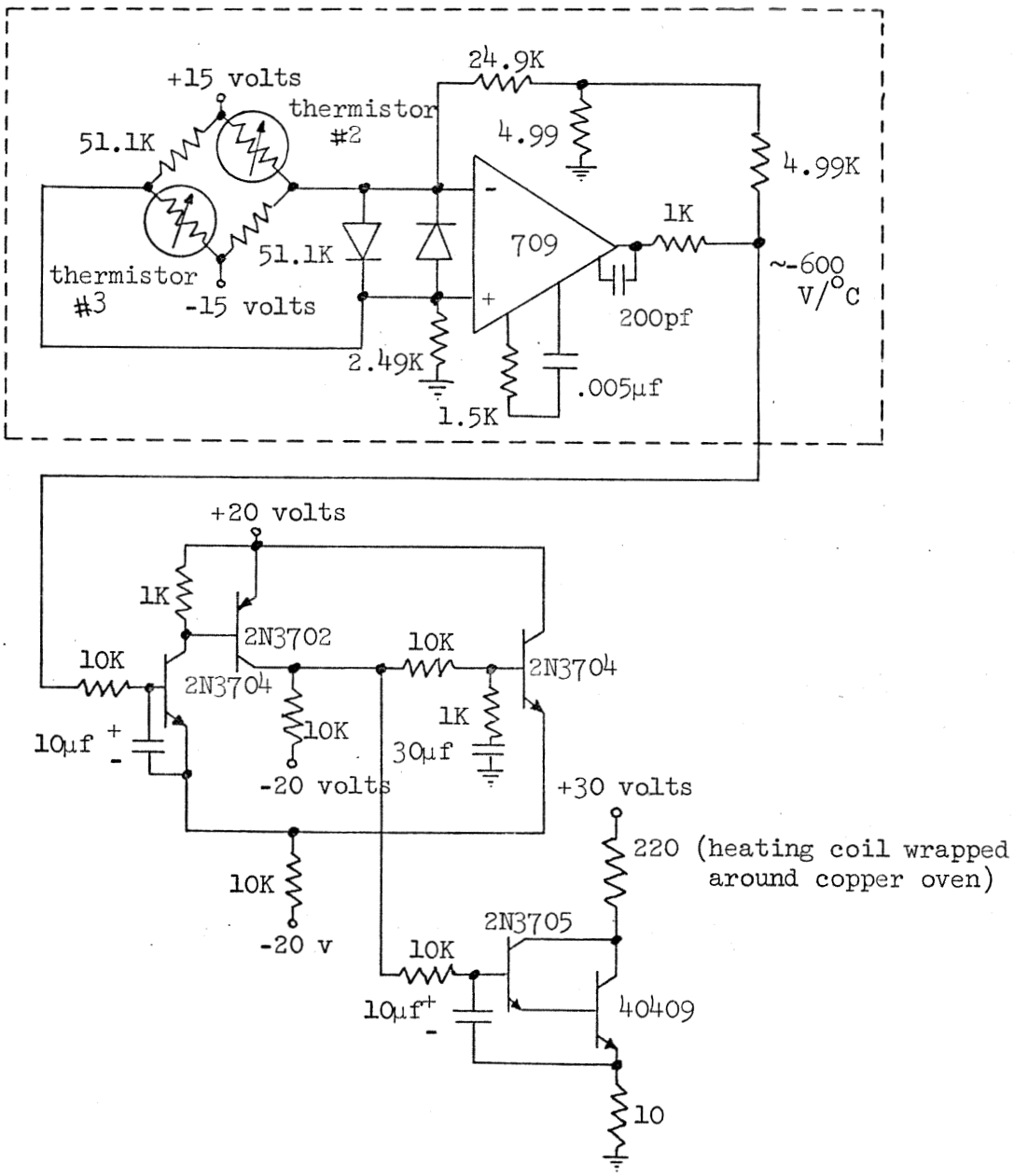


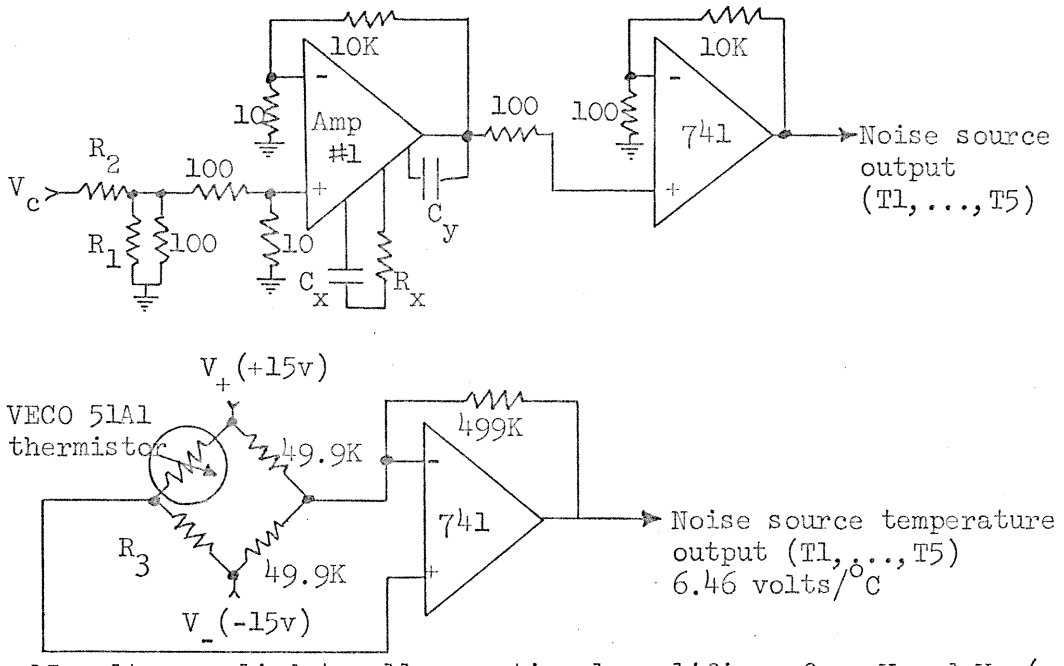
Figure 4.1 Schematic diagram of #2 noise source and voltage regulators.



Thermistors #2 and #3 mounted along diagonally opposing interior edges of copper oven (VECO 51A1).

Components inside dotted line are inside oven along with those shown in Figure 4.1.

Figure 4.2 Schematic diagram of #2 noise source oven heater control.



+15 volts supplied to all operational amplifiers from V_+ and V_- (see figure 4.4).

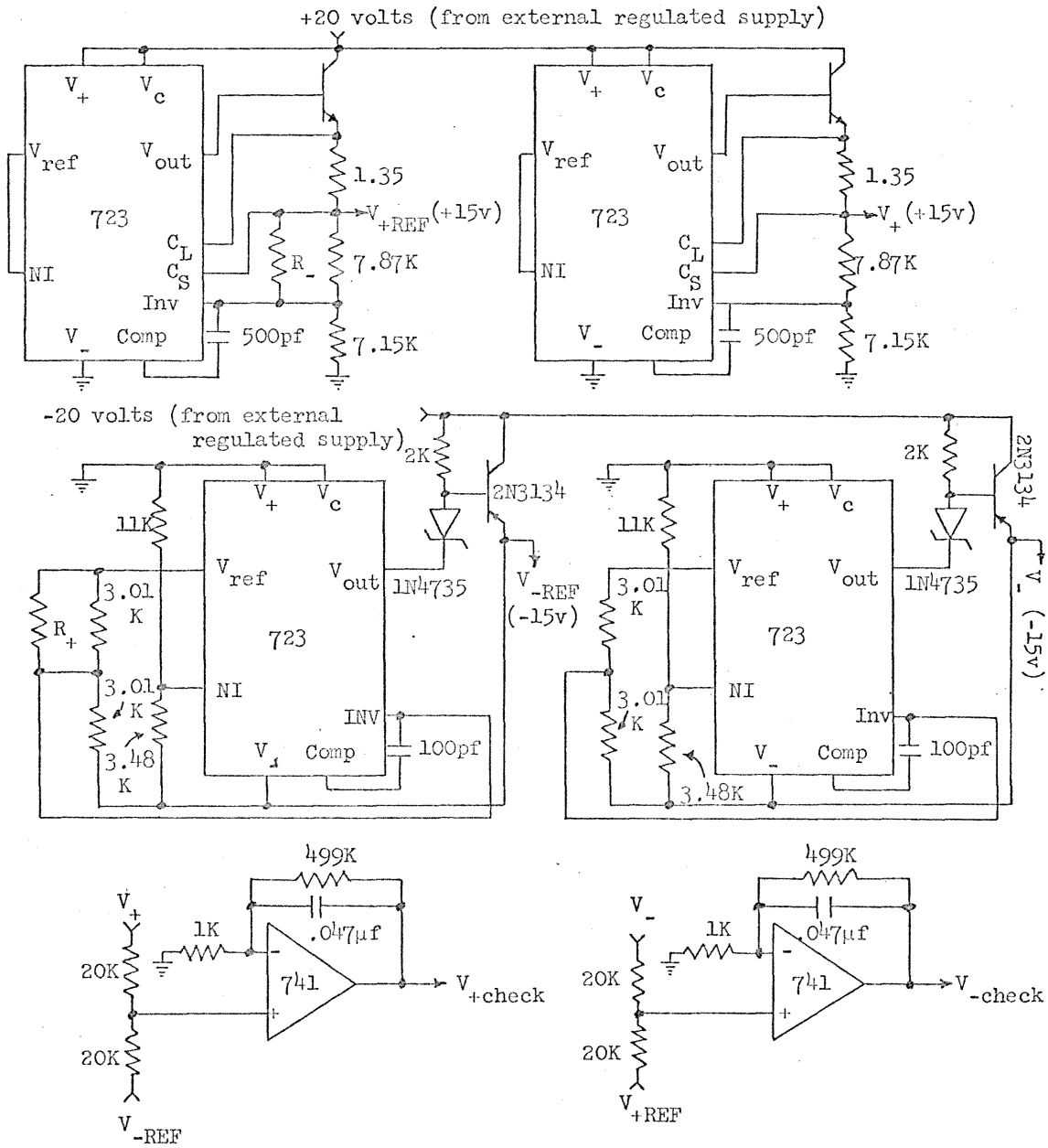
Components C_x, R_x, and C_y apply only to T3, T4, and T5 sources.

Thermistor mounted to case of Amp #1.

Noise source	Amp #1 (Fairchild part #)	R_1	R_2	V_c	implied V_{offset} of Amp #1	R_3
T1	U7A7747393	∞	74.41KΩ	-15 volts	-.96 mv	53.76KΩ
T2	U7A7747393	24.9Ω	135.65KΩ	-15 volts	-.17 mv	53.97KΩ
T3	U5B770939 634	24.9Ω	139.59KΩ	+15 volts	+.16 mv	55.04KΩ
T4	U5B770939 634	∞	173.78KΩ	+15 volts	+.41 mv	55.13KΩ
T5	U5B770939 634	24.9Ω	172.64KΩ	-15 volts	-.13 mv	54.57KΩ

Amp #1 for sources T1 and T2 are actually contained in a single dual operational amplifier integrated circuit chip.

Figure 4.3 Schematic diagram for noise sources T1, ..., T5.



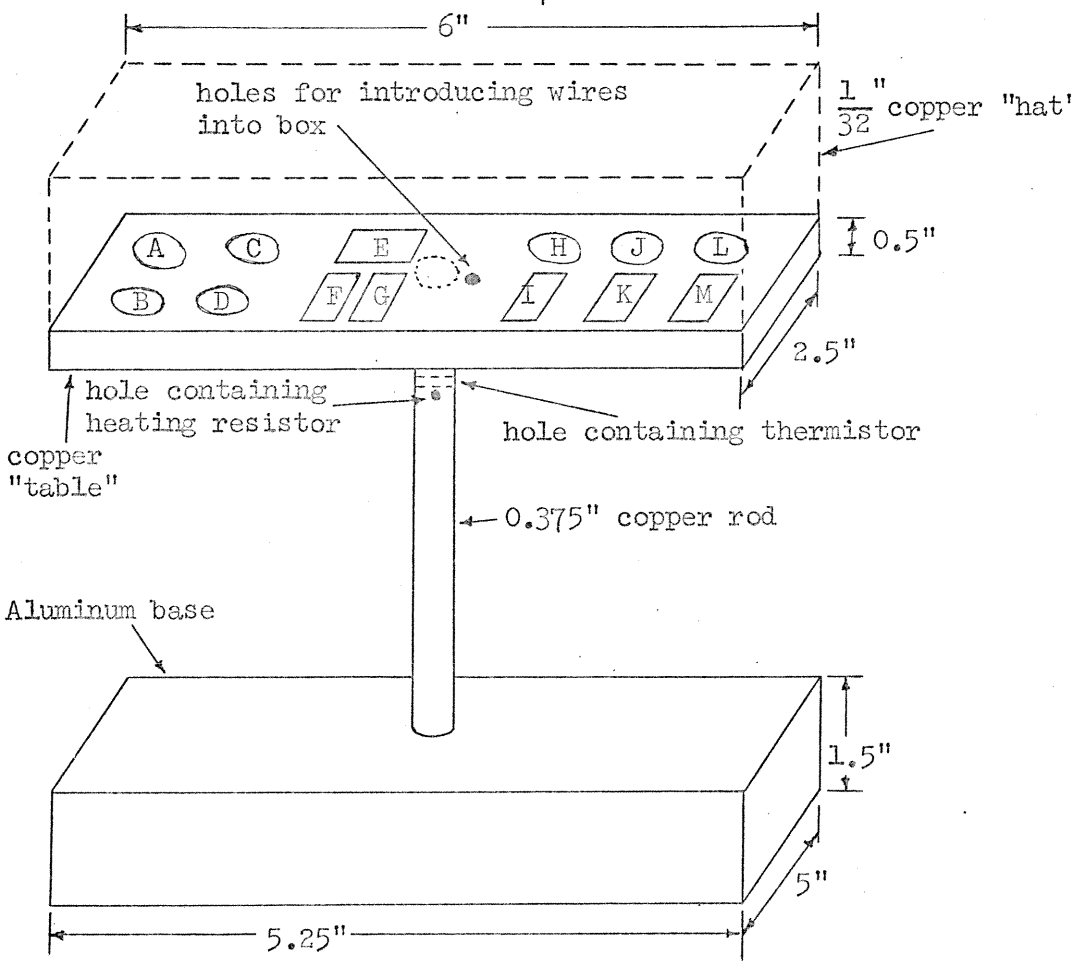
All components shown are inside temperature regulated oven. All operational amplifiers inside oven (including noise sources) are supplied by V_+ and V_-. Amplifiers labled 741 are actually two halves of a Fairchild U7A7747393. Voltage regulators labled 723 are Fairchild U5R7723393. R_+ and R_- are adjusted for $V_{\pm check} \sim 0$ volts.

Figure 4.4 Schematic diagram of voltage regulators supplying noise sources T1, T2, T3, T4, and T5.

4.3 The Temperature Ovens

The oven for noise source #2 consists of a $\frac{1}{32}$" thick copper box 3.5" x 2.6" x 2.1" (with one removable end having a small hole for passing wires through). Electrically insulated Nichrome wire is wrapped around the length of the box as a heating element. The temperature is sensed by a pair of thermistors mounted in opposing interior edges of the box using silicon thermal compound (and tape) to insure good contact with the metal box (this is done to minimize the time delay in the feedback control loop). The ±15 volt regulators, noise source #2, #2 noise source temperature sensing thermistor, and thermistor amplifiers (Figs. 4.1 and 4.2) are enclosed in the copper box (insulated from the box by packing material). The copper box was then placed in a 5" x 4" x 3" aluminum box (insulated by several small styrofoam spacers). The outer aluminum box provides passive attenuation of the high frequency components of room temperature. The feedback network which drives the heater coil is external to the oven except for the 709 operational amplifier immediately following the thermistors (Fig. 4.2) and provides frequency compensation to allow for an increased loop gain at DC. The walls of the copper box are held at a temperature of ~40°C. The internal oven temperature (measured at the case of noise source #2) is ~68°C due to the power dissipated inside the box.

The oven for noise sources T1,...,T5 consists of a 6" x 2.5" x 0.5" copper "table" (see Fig. 4.5) supported in the center by a 0.375" copper rod. The copper rod has been forced into a tapered hole in the "table" and then soldered to assure a good thermal contact. The other end of



Areas marked by letters represent position of major components.

A - V_{+ref} regulator
B - V_{-ref} regulator
C - V_{+} regulator
D - V_{-} regulator
E - T1 and T2 noise sources
F - T1 noise source amplifier
G - T2 noise source amplifier
H - T3 noise source
I - T3 noise source amplifier
J - T4 noise source
K - T4 noise source amplifier
L - T5 noise source
M - T5 noise source amplifier

Figure 4.5 Thermal oven for noise sources T1, T2, T3, T4, and T5.

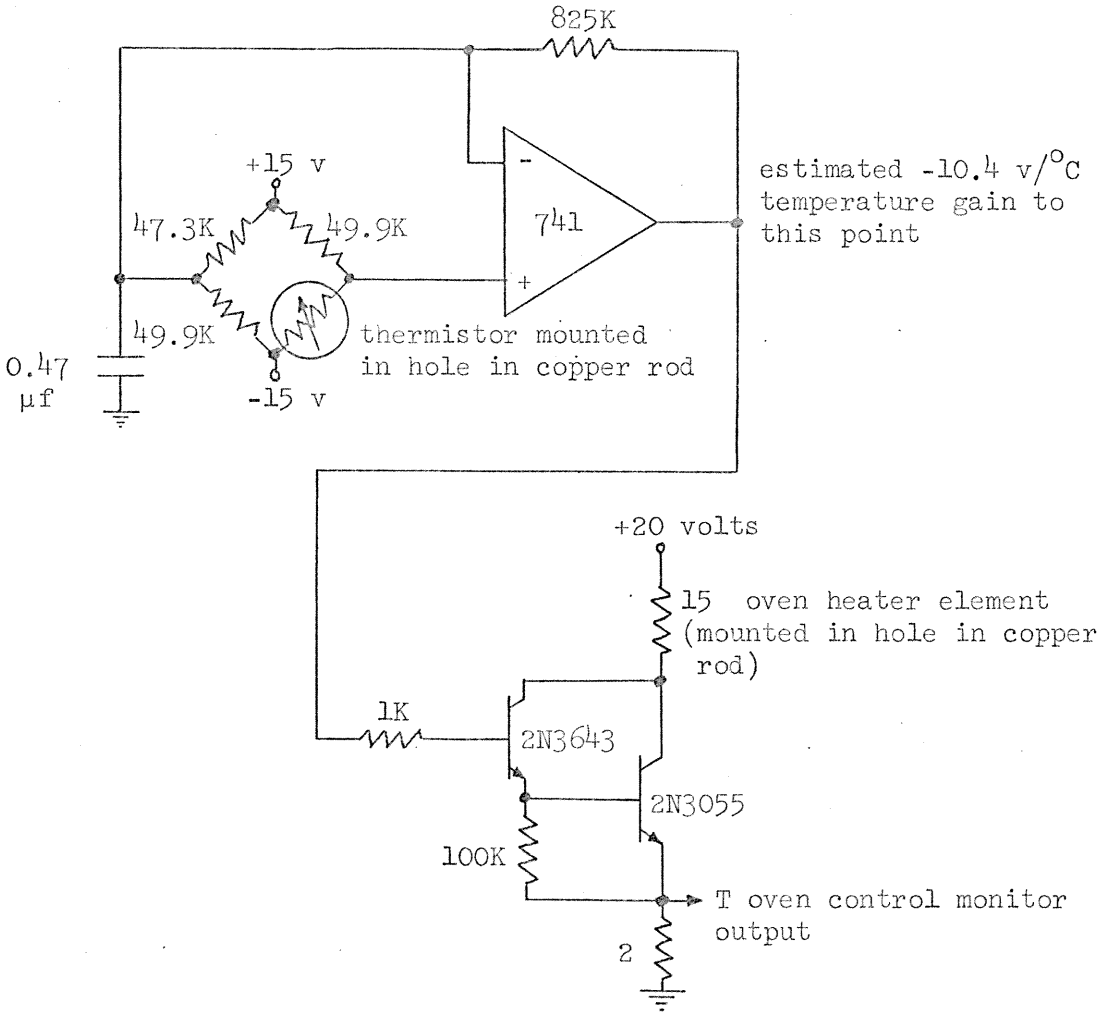


Figure 4.6 Schematic diagram of heater control circuitry for oven containing noise sources T1, T2, T3, T4, and T5.

the rod is inserted into an aluminum base for support (and heat dissipation). The temperature of a small region of the copper rod is sensed (as near to the "table" as possible) by a thermistor which has been inserted into a hole in the rod (the remaining space is filled with a compound, Astrodyne Thermal Bond 312, having a thermal conductivity of 1.07×10^{-2} watts/cm-°C). A 15 ohm resistor in a hole immediately below the thermistor provides a source of heat. The close spacing between the heating element and the thermistor simplify the control problem. The components (noise sources T1,...,T5, voltage regulators, and temperature sensors) are mounted on the "table" in the relative positions shown in Fig. 4.5. The integrated circuits and transistors are mounted with the thermally conductive adhesive to the "table" with the leads pointing up. The thermistors used to sense the noise source temperatures are mounted on the integrated circuit packages with the thermally conductive adhesive (the high electrical resistivity, $\sim 10^{13}$ ohm-cm, of this compound allows the thermistor to be mounted at a point where the leads are emerging from the case of the integrated circuit). A $\frac{1}{32}$" thick copper "hat" is fitted over the top of the "table" so that all of the circuitry is enclosed in a copper box. Wires are brought into the oven through a small hole in the "table". This entire assembly is then housed in a 8" x 7" x 10" aluminum box to provide attenuation of high frequency room temperature variations. The aluminum base rests on the bottom of the aluminum box (providing a path for the heat flowing down the copper rod). The heater control thermistor was not calibrated

but statistical data gathered on later calibration of thermistors of the same type indicate that the control thermistor temperature is ~40°C ± 3°C. The thermistors used in sensing the noise source temperatures were calibrated and indicate that the temperatures of noise sources T1,...,T5 are all 42°C ± 0.3°C (the 0.3°C uncertainty is due to the fact that two of the resistors in each of the thermistor bridge circuits is known only to ±1%).

Recorded data indicate that the temperature fluctuations at the control points of both ovens are at least an order of magnitude smaller than the temperature fluctuations at the noise source amplifiers. This indicates that an improved oven structure, rather than increased loop gain, would be necessary to improve the temperature control of the noise source amplifiers. Later data will show that the effects of temperature variations are negligible due to the low temperature coefficient of the amplifiers, so that improved ovens are not required.

Chapter V

EXPERIMENTAL DATA - PHASE I

5.1 Multiplexing

As previously mentioned, the multiplexer provided for multiplexing 30 input channels. Ten channels were output to the Digital Voltmeter (DVM) with a resolution of 0.1 millivolts on the most sensitive scale. Twenty channels were output to the Analog to Digital Converter (A/D converter) with a resolution of approximately 10 millivolts. Each input channel was low pass filtered by its own single pole RC filter with a 3 db point at 0.34 Hz. The output of each low pass filter was buffered by a unity gain amplifier prior to the multiplexing switch to eliminate the transient loading effects of the multiplexer on the filters. Individual filters were used on each channel rather than one on the measuring instrument (DVM or A/D converter) since the time constant of the filter was to be of the same order as the multiplexer switching period.

The multiplexer is logically divided into thirds since each time the multiplexer is switched one DVM channel and two A/D channels are sampled. The DVM channels were labeled as $A_1,\ldots,A_{10}$ and the two sets of A/D channels were labeled $B_1,..,B_{10}$ and $C_1,\ldots,C_{10}$. The DVM channels are sampled sequentially, separated by a period of time referred to as the multiplexer sample period (this is $\frac{1}{10}$ the sample period of each channel which will be the sample period most often referred to). Channel B_i ($i = 1,\ldots,10$) is sampled 10 millisec after

the DVM signals completion of the A_i channel sample. Channel C_i is sampled 30 millisec after the B_i channel sample (see Figs. 3.1 and 3.2).

It was decided to devote one input channel in each third of the multiplexer to a grounded input as a check on the offset voltage drift of the appropriate measuring device (the measuring device here is the low pass filter, buffer amplifier, and DVM or A/D converter). This does not yield information on the drift of all 30 channels but does provide one example for each set of channels which is hopefully typical of the others.

5.2 Data Channel Allocation

The six noise sources (which will be referred to as #2, T1, T2, T3, T4, and T5) were considered to be the most important values to be measured and each was allocated a DVM channel. Even though the input offset voltages of the noise sources had been compensated for, noise sources #2 and T1 had drifted giving output voltages of greater than 1.0 volt in magnitude. This caused the DVM to change scale on every multiplexer cycle resulting in a factor of 10 decrease in resolution for these two sources. Rather than open the ovens at this point it was decided to insert a voltage divider having a gain of 0.5 into each of these two channels prior to the low pass filter. This decreases the resolution by a factor of two. In all results which follow, this factor of two in gain has been accounted for so that when a sample voltage for #2 noise source is quoted it is 2.0 times the value actually measured on the appropriate channel.

The remaining three DVM channels were allocated to the temperature sensors for T2, T3, and T5. Five A/D channels were allocated to noise source temperature sensors in such a way as to insure that these temperatures were sampled no more than 0.1 sec after the corresponding noise source (independent of the multiplexer sample period). A/D channels were then allocated for a monitor of the peak value of the line voltage (from the power inverter), the room temperature (at an approximate gain of 2.2°C/volt using a glass bead thermistor suspended in air by its connecting wires), V_{+check} and V_{-check} (as discussed

earlier), and the two oven control voltages.

The remaining seven A/D channels were allocated to the sources already connected to the corresponding DVM channels. This was done to allow a check to be made, if later desired, on the linearity and drift of the A/D converter (the author having much more faith in the accuracy of the DVM). Table 5.1 is a complete list of the channel allocations.

All the results, unless otherwise stated, in the remainder of the thesis result from the data recorded through channels A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_{10} B_2, B_3, B_5, B_6, B_7, B_8, B_9, and B_{10}. To reduce processing time and expense these were the only channels processed.

A_8 and A_9 were used for T2 and T5 temperature data because of the higher resolution and accuracy of the DVM channels. During the first phase of the experiment these temperature samples lag their corresponding noise source samples by as much as 4 seconds and in the second phase by as much as 40 seconds. This lag is not considered to be significant due to the relatively low power spectral density of temperature at these higher frequencies ($\frac{1}{4}$ Hz or $\frac{1}{40}$ Hz). See Figs. 5.1, and 5.14,...,5.19).

Channel	Source	Channel	Source	Channel	Source
A_1	Ground	B_1	Ground	C_1	Ground
A_2	$\frac{1}{2}$ (noise source #2)	B_2	#2 temperature	C_2	#2 oven control
A_3	$\frac{1}{2}$ (noise source T1)	B_3	T1 temperature	C_3	T oven control
A_4	noise source T2	B_4	T2 temperature	C_4	noise source T2
A_5	noise source T3	B_5	T3 temperature	C_5	noise source T3
A_6	noise source T4	B_6	T4 temperature	C_6	noise source T4
A_7	noise source T5	B_7	line monitor	C_7	noise source T5
A_8	T2 temperature	B_8	room temperature	C_8	T2 temperature
*A_9	T3 temperature	B_9	V_{-check}	C_9	T3 temperature
A_{10}	T5 temperature	B_{10}	V_{+check}	C_{10}	T5 temperature

Table 5.1 Channel Allocations

*Shortly after the beginning of the experiment, the data recorded from channel A_9 revealed a highly erratic behavior not characteristic of the other temperatures and not agreeing with channels B_5 and C_9. Data from channel A_9 have, therefore, been discarded.

5.3 Experimental Data

The first phase of the experiment involved taking data with a source sample period of Δt = 10 sec (multiplexer sample period = 1 sec) for a total period of approximately 8×10^6 seconds (93 days). During this phase, all of the spectral estimates will be biased by the aliasing of spectral components at frequencies of $\frac{1}{2\Delta t}$ = 0.05 Hz to approximately 0.34 Hz (the cutoff frequency of the analog filter was not reduced due to the awkward value of capacitance required in the 30 channels). From previous discussions, however, we see that for flicker noise this aliasing will appreciably effect only the first decade of frequency which we will later correct by sampling each of the sources at Δt = 1 sec for the relatively short period of 10^4 sec (2.8 hours).

As has been previously mentioned, the data, as soon as they are converted to a digital form, have four parity bits added to each word. In addition to this there are disallowed states from the DVM since it has a BCD format. Without going into great detail, both of these facts proved useful. There were often as many as four or five parity errors detected per side of the magnetic tape (~2 days data). Each of these errors was considered individually. Most of the time the actual data word could be reconstructed with almost 100% certainty since a single bit error could be isolated to a particular quarter word (as discussed previously) and the data were often (especially in the case of temperature data) changing very slowly from sample to sample. Occasionally the errors were bad enough to necessitate simply extrapolating between good data points, but the number of times this was done

was insignificant when compared to the 1.6×10^7 data words. The parity test also detected the few times that the magnetic tape decoder did not get proper synchronization at the beginning of a record (or lost synchronization during a record) causing right or left shifts in the data words. This was easily corrected in the computer once the proper number of bit positions to shift was determined by the somewhat painstaking process of examining the shifted data by eye. The third error that sometimes occurred was an error in the buffer memory causing a temporary or permanent shift in the multiplexing sequence in the records. In this case the parities would all be right but the demultiplexer would attempt to convert the A/D codes (binary) assuming them to be DVM codes (BCD). This would indicate errors which could also be corrected with slight loss of data. In such cases it was always possible to find the probable malfunction in the appropriate circuitry (most often a digital counter which missed an increment pulse). By comparison of the time each tape should have taken to be recorded (knowing the sample rate, number of samples per record, and number of samples actually recorded) and the actual recording time (usually measured to an accuracy of ±30 seconds for a 2 day record as determined by the Pacific telephone time recording), any possibility of gross data was discounted.

After these errors were corrected, the data were assumed to accurately represent the voltage as seen by the DVM or A/D converter and were stored on IBM compatible tape (7 track, 800 BPI). The data were demultiplexed and divided into files. Each file is the result of

50 records as originally recorded on the magnetic tape recorder. Each data file consists of 20 blocks each 1025 words long (6 tracks per word to provide the 36 bit words required by the Univac 1108 computer). The first 10 blocks contain the data from channels $A_1,\ldots,A_{10}$. This is in integer 1's compliment format at a gain 10^4(i.e., a voltage of 0.1 millivolt is recorded as 1 on the tape). Block 10 + i contains the data from channels B_i and C_i still coded as for the magnetic tape recorder. Bits 3,...,0 contain the parity for the word. Bits 13,...,4 contain the binary equivalent of C_i at that sample time. Bits 23,...,14 contain the binary equivalent of B_i at that sample time. Therefore, to recover a datum point corresponding to the B_i channel in volts we would select the appropriate word from block 10 + i, shift right by 14 places (with zero fill), subtract 512, and multiply by $\frac{10}{1024}$. The data for the first phase of the experiment are now stored on 6 such tapes (782 files). The tapes are designated as Spectral Data-1,...,Spectral Data-6. The following list indicates the number of files on each of these tapes.

Tape	Number of Files
Spectral Data - 1	142
Spectral Data - 2	135
Spectral Data - 3	134
Spectral Data - 4	135
Spectral Data - 5	135
Spectral Data - 6	101

Table 5.2

Using this information we may refer to sample number 500,000 of channel A_3 as A_3 (500,000), or as A_3 (825) of data file 488, or as A_3 (825) of data file 77 of tape Spectral Data-4. As a further point of information, the first sample was taken at approximately 16:14 (PST) November 29, 1972. The final point of the first phase was taken at approximately 10:47 (PST) March 2, 1973. Since this represents 782 data files at 10250 seconds per data file the predicted end point would be 10:46 March 2, 1973. This is an error of only 1 minute in 3 months (a 0.00075% error) which is considerably better than had been hoped to set the timing oscillator frequency and is just about within the accuracy of the start and stop time measurements.

At this point, one might assume that the data are ready to be analyzed. Much care was taken in constructing the equipment and it was hoped that there would be no glitches. A glitch is taken to mean any datum point (or set of data points closely spaced in time) which result from some process other than that which we are attempting to measure. From the past experiences of others[1,11] and my own trial runs, there was reason to expect some glitches and a method was needed, therefore, to detect any that might occur. In practice, we can only hope to detect glitches which deviate in some significant way from the normal noise characteristics (and since we have no a priori way of knowing what the "normal" noise characteristics are, we must hope that the glitches do not occur too frequently).

The thought which occurs immediately is to plot all of the data. If we were to do this at a density of 1025 points per line and 10 lines per page as in Fig. 5.1 we would require 780 pages to plot just

those sources plotted in Fig. 5.1 (there are still a few more sources used in later analysis which were not plotted in Fig. 5.1 for lack of space). Of course the data could be plotted at 10 times this density to yield only 78 pages as in Fig. 5.5. An estimate of the cost of making this plot, from the costs incurred in making similar (but somewhat shorter) plots, was about $500 (assuming the lowest priority computer rates available at the time with about a 1 week turn around time for any given computer run). This cost was considered out of line, particularly since this involved a density of more than 1000 data points per inch (in which individual datum points certainly could not be resolved).

The method chosen to examine the data in the time domain was to first plot the maximum and minimum data values over one 1025 word block for each channel to be retained for analysis. This resulted in a plot for which 1025 data point segments having significantly larger than average peak to peak deviations could be immediately identified. The suspicious segments (for all channels) were then plotted to a density of 100 points per inch which allowed "glitches" to be located to about ±5 sample points. The method for determining whether or not to classify a given datum point as a glitch was somewhat subjective. I will only note the philosophy which I attempted to follow. Spikes which were clearly correlated in two or more sources (not different channels carrying the same source) were considered to be glitches. Single spikes of more than about 3 times the normal high frequency peak to peak variations of the source were assumed to be glitches. Spikes in

V_{+check} and V_{-check} were tolerated at somewhat larger levels than for the other sources if they were not clearly correlated with spikes in any of the other sources. After a glitch was determined to exist from the plot, 10 data points for each source were printed in the vicinity of the glitch to establish the exact datum point (or points) involved. A new value was then assigned which was thought to be a reasonable approximation to the lost data points and a card punched for use by the computer in inserting this correction. The cards designate the tape, file number, channel to be corrected, word number in the block considered to be the first point of the glitch, word number in the block considered to be the last point in the glitch (if more than one point involved), and the new value to be assigned to these points. In addition, for most points the value of the first word of the glitch is given on the card so that the computer can check that an error has not been made in specifying the location of the glitch.

The final set of correction data used in the first phase is shown in Table 5.4. It will be noted that channel A_2 (noise source #2) has some very long strings of corrections (nearly 3000 data points in one case). This is thought to be a loose connection in the measuring channel (the data from this channel often jump to ~0 volts for long periods of time). One will immediately notice (see Table 5.1) that this is one of the few sources for which a redundant channel was not allocated. These data were, therefore, used in the spectral analysis. The following precautions were observed. The set of points chosen to make spectral estimates at Δt = 10 sec avoided all known glitches for

all the channels. The set of points chosen for estimates at Δt = 100 sec avoided regions where a large number of glitches had been detected. Estimates made at Δt = 1000 sec and Δt = 10,000 sec used all available data. To give some additional perspective, Table 5.3 indicates the total number of points in each channel which were replaced because of glitches and the percent of the total number of points available per channel (782 x 1025 = 801,550) which this number represented.

Figures 5.1,...5.7 show some of the data actually used in the spectral analysis. Figure. 5.5 shows all 10250 data samples used in the spectral analysis using a 10 sec sampling period (spectral estimates for $\frac{1}{2000}$ Hz < f < $\frac{1}{20}$ Hz). These data come from data files 63-72 of magnetic tape Spectral Data-3. Only one of the temperature data channels from noise sources T1,..,T5 was used due to space limitations and their similarities. They do, in fact, look the same as the one given (T2 temperature) for all of the different scales represented in Figs. 5.1, ...,5.7. Figure 5.1 is the first tenth (in the time sense) of the data in Fig. 5.5 expanded in time.

Figure 5.6 shows the 10250 data samples used in the spectral analysis using a 100 sec sampling period (spectral estimates for $\frac{10^{-4}}{2}$ Hz < f < $\frac{10^{-2}}{2}$ Hz). These data result from passing the original data through the low pass dealiasing filter described earlier and resampling by retaining only every tenth sample. The set of data in Fig. 5.6 corresponds (in time) to data files 51 through 150 of the original data. Figure 5.2 expands the first tenth of Fig. 5.6 in time.

Table 5.3

Channel	Source	total number of points replaced = n	$\frac{n}{801550} \times 100\%$
A_2	$\frac{1}{2}$ x (noise source #2)	7506	0.94 %
A_3	$\frac{1}{2}$ x (noise source T1)	1078	0.13 %
A_4	noise source T2	1100	0.14 %
A_5	noise source T3	11	0.0014%
A_6	noise source T4	25	0.003 %
A_7	noise source T5	23	0.003 %
A_8	T2 temperature	0	0 %
A_{10}	T5 temperature	0	0 %
B_2	#2 temperature	0	0 %
B_3	T1 temperature	0	0 %
B_5	T3 temperature	0	0 %
B_6	T4 temperature	0	0 %
B_7	line monitor	0	0 %
B_8	room temperature	0	0 %
B_9	V_{-check}	64	0.008 %
B_{10}	V_{+check}	61	0.008 %

```
DATAFILE   8   OF TAPE  2578 (SPECTRAL DATA-1)    4 CARDS
A 2( 934)          CHANGE TO  -4726    FROM   1175
A 3( 934)          CHANGE TO  -5738    FROM  -7003
A 4( 934)          CHANGE TO  -3916    FROM  -6506
B 9( 279)          CHANGE TO    570    FROM    213
DATAFILE  25   OF TAPE  2578 (SPECTRAL DATA-1)    6 CARDS
A 4( 428)          CHANGE TO  -2700    FROM   -477
A 4( 429)          CHANGE TO  -2900    FROM   -179
B 9( 427)          CHANGE TO    564    FROM    535
B 9( 428)          CHANGE TO    564    FROM    527
B10( 427)          CHANGE TO    433    FROM    388
B10( 428)          CHANGE TO    433    FROM    373
DATAFILE  32   OF TAPE  2578 (SPECTRAL DATA-1)    6 CARDS
A 4( 223)THRU  225 CHANGE TO  -2900    FROM   -165
A 4( 399)THRU  402 CHANGE TO  -2500    FROM   2945
B 9( 222)THRU  224 CHANGE TO    568    FROM    520
B 9( 397)THRU  399 CHANGE TO    560    FROM    547
B10( 222)THRU  224 CHANGE TO    448    FROM    391
B10( 398)THRU  399 CHANGE TO    442    FROM    297
DATAFILE  41   OF TAPE  2578 (SPECTRAL DATA-1)    3 CARDS
A 4( 694)THRU  699 CHANGE TO  -1900    FROM  -1664
B 9( 696)THRU  699 CHANGE TO    570    FROM    409
B10( 696)THRU  699 CHANGE TO    406    FROM    217
DATAFILE  75   OF TAPE  2578 (SPECTRAL DATA-1)    2 CARDS
B 9( 965)THRU  968 CHANGE TO    583    FROM    521
B10( 965)THRU  968 CHANGE TO    425    FROM    343
DATAFILE  78   OF TAPE  2578 (SPECTRAL DATA-1)    5 CARDS
A 3( 753)          CHANGE TO  -7070    FROM  -9996
A 4( 753)          CHANGE TO  -1900    FROM -76900
A 5( 753)          CHANGE TO  -2700    FROM -21970
A 6( 753)          CHANGE TO  -4200    FROM   8290
A 7( 753)          CHANGE TO  -4800    FROM -15000
DATAFILE  82   OF TAPE  2578 (SPECTRAL DATA-1)    2 CARDS
A 2(  68)          CHANGE TO  -4700    FROM  -2311
A 7( 520)          CHANGE TO  -5000    FROM -10808
DATAFILE 118   OF TAPE  2578 (SPECTRAL DATA-1)    2 CARDS
A 4( 386)          CHANGE TO  -1300    FROM -11537
A 5( 386)          CHANGE TO  -4800    FROM  -6740
DATAFILE   3   OF TAPE  2866 (SPECTRAL DATA-2)    1 CARDS
A 2( 422)          CHANGE TO  -5107    FROM  -3143
DATAFILE  22   OF TAPE  2866 (SPECTRAL DATA-2)    2 CARDS
A 3( 620)          CHANGE TO  -8600    FROM -10850
A 4( 620)          CHANGE TO  -1600    FROM  -8093
DATAFILE  50   OF TAPE  2866 (SPECTRAL DATA-2)    1 CARDS
A 2(  50)THRU   51 CHANGE TO  -5100    FROM  -4183
DATAFILE 125   OF TAPE  2866 (SPECTRAL DATA-2)    1 CARDS
A 2( 220)THRU 1025 CHANGE TO  -4897    FROM
DATAFILE 126   OF TAPE  2866 (SPECTRAL DATA-2)    1 CARDS
A 2(   1)THRU 1025 CHANGE TO  -4897    FROM
DATAFILE 127   OF TAPE  2866 (SPECTRAL DATA-2)    1 CARDS
A 2(   1)THRU   90 CHANGE TO  -4995    FROM
DATAFILE  32   OF TAPE  4329 (SPECTRAL DATA-3)    1 CARDS
A 2( 470)THRU 1025 CHANGE TO  -5142    FROM  -4875
DATAFILE  33   OF TAPE  4329 (SPECTRAL DATA-3)    1 CARDS
```

Table 5.4 a Correction data.

```
A 2(   1)THRU 1025 CHANGE TO  -5142    FROM
DATAFILE  34   OF TAPE  4329 (SPECTRAL DATA-3)    1 CARDS
A 2(   1)THRU 1025 CHANGE TO  -5142    FROM
DATAFILE  35   OF TAPE  4329 (SPECTRAL DATA-3)    1 CARDS
A 2(   1)THRU  385 CHANGE TO  -5142    FROM
DATAFILE  41   OF TAPE  4329 (SPECTRAL DATA-3)    1 CARDS
A 2( 700)THRU  775 CHANGE TO  -4897    FROM  -4322
DATAFILE  68   OF TAPE  4329 (SPECTRAL DATA-3)    4 CARDS
A 4( 236)THRU  237 CHANGE TO    500    FROM   1837
B 9( 237)          CHANGE TO    497    FROM    456
B10( 237)          CHANGE TO    450    FROM    402
B10( 849)          CHANGE TO    457    FROM    490
DATAFILE  85   OF TAPE  4329 (SPECTRAL DATA-3)    1 CARDS
A 2( 643)          CHANGE TO  -5415    FROM  -3268
DATAFILE 109   OF TAPE  4329 (SPECTRAL DATA-3)    4 CARDS
A 3( 231)THRU  234 CHANGE TO -11900    FROM -12570
A 4( 231)THRU  233 CHANGE TO    700    FROM   2118
B 9( 229)THRU  232 CHANGE TO    473    FROM    465
B10( 229)THRU  232 CHANGE TO    473    FROM    463
DATAFILE 111   OF TAPE  4329 (SPECTRAL DATA-3)    4 CARDS
A 3(  10)THRU   11 CHANGE TO -11900    FROM -12450
A 4(  10)THRU   11 CHANGE TO    400    FROM   2091
B 9(   9)THRU   11 CHANGE TO    479    FROM    427
B10(   9)THRU   11 CHANGE TO    460    FROM    386
DATAFILE 119   OF TAPE  4329 (SPECTRAL DATA-3)    2 CARDS
A 3( 291)          CHANGE TO -12300    FROM -13250
A 4( 291)          CHANGE TO    -25    FROM  -4180
DATAFILE 123   OF TAPE  4329 (SPECTRAL DATA-3)    3 CARDS
A 3( 345)          CHANGE TO -12100    FROM -16510
A 4( 345)          CHANGE TO     75    FROM -11040
A 5( 345)          CHANGE TO  -4340    FROM  -8367
DATAFILE  19   OF TAPE  2969 (SPECTRAL DATA-4)    1 CARDS
A 3( 378)          CHANGE TO -12150    FROM -14050
DATAFILE  26   OF TAPE  2969 (SPECTRAL DATA-4)    2 CARDS
A 3( 699)THRU  701 CHANGE TO -12000    FROM -13440
A 4( 699)THRU  701 CHANGE TO   -880    FROM  -3069
DATAFILE  34   OF TAPE  2969 (SPECTRAL DATA-4)    5 CARDS
A 3( 185)THRU  267 CHANGE TO -12220    FROM -14390
A 4( 185)THRU  267 CHANGE TO  -2000    FROM  -6700
A 4( 141)THRU  142 CHANGE TO  -1200    FROM   1203
B 9( 141)THRU  142 CHANGE TO    491    FROM    379
B10( 141)THRU  142 CHANGE TO    485    FROM    361
DATAFILE  36   OF TAPE  2969 (SPECTRAL DATA-4)    2 CARDS
A 3( 611)THRU  780 CHANGE TO -11982    FROM -12510
A 4( 611)THRU  780 CHANGE TO  -2400    FROM  -3061
DATAFILE  78   OF TAPE  2969 (SPECTRAL DATA-4)    1 CARDS
A 2( 270)          CHANGE TO  -5235    FROM   2460
DATAFILE  97   OF TAPE  2969 (SPECTRAL DATA-4)    1 CARDS
A 2( 529)THRU 1025 CHANGE TO  -5386    FROM  -5358
DATAFILE  98   OF TAPE  2969 (SPECTRAL DATA-4)    2 CARDS
A 2(   1)THRU  191 CHANGE TO  -5386    FROM
A 2( 421)THRU  426 CHANGE TO  -5386    FROM  -4432
DATAFILE 103   OF TAPE  2969 (SPECTRAL DATA-4)    1 CARDS
A 2( 695)THRU  800 CHANGE TO  -5142    FROM  -4742
DATAFILE 105   OF TAPE  2969 (SPECTRAL DATA-4)    1 CARDS
A 2( 770)THRU  969 CHANGE TO  -5386    FROM  -5307
DATAFILE 106   OF TAPE  2969 (SPECTRAL DATA-4)    2 CARDS
```

Table 5.4 b Correction data.

```
A 2( 310)THRU  385 CHANGE TO  -5142    FROM  -5296
A 2( 765)THRU 1025 CHANGE TO  -5264    FROM  -4905
DATAFILE 107   OF TAPE  2969 (SPECTRAL DATA-4)    1 CARDS
A 2(   1)THRU   15 CHANGE TO  -5264    FROM
DATAFILE 109   OF TAPE  2969 (SPECTRAL DATA-4)    1 CARDS
A 2( 745)THRU 1025 CHANGE TO  -5264    FROM  -5401
DATAFILE 110   OF TAPE  2969 (SPECTRAL DATA-4)    1 CARDS
A 2(   1)THRU  530 CHANGE TO  -5264    FROM
DATAFILE  18   OF TAPE   905 (SPECTRAL DATA-5)    2 CARDS
A 3( 510)THRU  551 CHANGE TO -12715    FROM -12690
A 4( 510)THRU  551 CHANGE TO   -700    FROM   -884
DATAFILE  19   OF TAPE   905 (SPECTRAL DATA-5)    1 CARDS
A 2( 780)THRU 1025 CHANGE TO  -5142    FROM  -4902
DATAFILE  20   OF TAPE   905 (SPECTRAL DATA-5)    1 CARDS
A 2(   1)THRU   90 CHANGE TO  -5142    FROM
DATAFILE  26   OF TAPE   905 (SPECTRAL DATA-5)    9 CARDS
A 4( 842)THRU  849 CHANGE TO   -500    FROM    135
A 4( 872)          CHANGE TO   -300    FROM   1664
A 4( 881)THRU  882 CHANGE TO   -400    FROM   3911
B 9( 843)THRU  848 CHANGE TO    545    FROM    537
B 9( 872)          CHANGE TO    545    FROM    437
B 9( 880)THRU  882 CHANGE TO    545    FROM    436
B10( 842)THRU  848 CHANGE TO    456    FROM    442
B10( 872)          CHANGE TO    457    FROM    422
B10( 880)THRU  882 CHANGE TO    450    FROM    320
DATAFILE  27   OF TAPE   905 (SPECTRAL DATA-5)    3 CARDS
A 4( 484)          CHANGE TO   -500    FROM   2332
B 9( 484)          CHANGE TO    543    FROM    440
B10( 484)          CHANGE TO    447    FROM    338
DATAFILE  35   OF TAPE   905 (SPECTRAL DATA-5)    2 CARDS
A 4(1013)          CHANGE TO   -800    FROM   2511
B10(1012)          CHANGE TO    459    FROM    358
DATAFILE  36   OF TAPE   905 (SPECTRAL DATA-5)    2 CARDS
A 3( 134)THRU  290 CHANGE TO -12471    FROM -14090
A 4( 134)THRU  290 CHANGE TO   -500    FROM  -7698
DATAFILE  44   OF TAPE   905 (SPECTRAL DATA-5)    4 CARDS
A 3( 853)          CHANGE TO -12590    FROM -18980
A 3( 888)THRU  971 CHANGE TO -12620    FROM -14860
A 4( 853)          CHANGE TO   -400    FROM -12580
A 4( 888)THRU  971 CHANGE TO   -300    FROM  -4699
DATAFILE  52   OF TAPE   905 (SPECTRAL DATA-5)    4 CARDS
A 3( 512)THRU  661 CHANGE TO -12715    FROM -14100
A 3( 713)THRU  796 CHANGE TO -12715    FROM -13800
A 4( 512)THRU  659 CHANGE TO   -600    FROM  -3223
A 4( 713)THRU  796 CHANGE TO   -600    FROM  -1386
DATAFILE  65   OF TAPE   905 (SPECTRAL DATA-5)    1 CARDS
A 2( 932)          CHANGE TO  -2962    FROM  -4577
DATAFILE  85   OF TAPE   905 (SPECTRAL DATA-5)    5 CARDS
A 3( 798)THRU  805 CHANGE TO -13300    FROM -14240
A 4( 798)THRU  805 CHANGE TO  -2000    FROM   1915
A 5( 798)THRU  805 CHANGE TO  -4600    FROM  -6138
B 9( 797)THRU  804 CHANGE TO    547    FROM    502
B10( 797)THRU  804 CHANGE TO    518    FROM    454
DATAFILE  90   OF TAPE   905 (SPECTRAL DATA-5)    1 CARDS
A 4( 728)THRU  732 CHANGE TO  -3000    FROM    387
DATAFILE 102   OF TAPE   905 (SPECTRAL DATA-5)    3 CARDS
A 2( 451)THRU  455 CHANGE TO  -2738    FROM  -4553
```

Table 5.4 c Correction data.

```
A 3( 892)          CHANGE TO -12650    FROM -11779
A 4( 892)          CHANGE TO  -3400    FROM -22420
DATAFILE   4   OF TAPE  1031 (SPECTRAL DATA-6)    4 CARDS
A 3( 344)THRU  345 CHANGE TO -12550    FROM -15420
A 4( 344)          CHANGE TO  -3300    FROM  -5010
A 7( 342)          CHANGE TO  -6600    FROM -27790
A 7( 344)THRU  345 CHANGE TO  -6600    FROM  -9950
DATAFILE  18   OF TAPE  1031 (SPECTRAL DATA-6)    3 CARDS
A 3( 125)THRU  142 CHANGE TO -12500    FROM -13370
A 6( 125)THRU  142 CHANGE TO   -800    FROM  -1883
A 7( 125)THRU  142 CHANGE TO  -7200    FROM  -8372
DATAFILE  44   OF TAPE  1031 (SPECTRAL DATA-6)    9 CARDS
A 3( 599)THRU  600 CHANGE TO -12900    FROM -13210
A 3( 624)THRU  653 CHANGE TO -12715    FROM -14620
A 4( 599)THRU  600 CHANGE TO  -3200    FROM  -2503
A 4( 624)THRU  649 CHANGE TO  -3100    FROM  -3832
A 6( 625)          CHANGE TO  -1400    FROM  -3377
B 9( 599)          CHANGE TO    517    FROM    450
B 9( 625)          CHANGE TO    518    FROM    461
B10( 599)          CHANGE TO    491    FROM    414
B10( 625)          CHANGE TO    494    FROM    459
DATAFILE  46   OF TAPE  1031 (SPECTRAL DATA-6)    8 CARDS
A 3( 405)THRU  425 CHANGE TO -12900    FROM -13720
A 3( 441)THRU  442 CHANGE TO -12800    FROM -15260
A 3( 517)THRU  519 CHANGE TO -12700    FROM -14680
A 4( 405)THRU  425 CHANGE TO  -3100    FROM  -4817
A 4( 437)THRU  442 CHANGE TO  -3100    FROM  -8157
A 4( 517)THRU  522 CHANGE TO  -3100    FROM  -6878
B 9( 419)THRU  425 CHANGE TO    521    FROM    489
B10( 420)THRU  421 CHANGE TO    540    FROM    465
DATAFILE  59   OF TAPE  1031 (SPECTRAL DATA-6)    8 CARDS
A 2( 564)THRU  569 CHANGE TO  -5200    FROM  -5882
A 3( 327)THRU  457 CHANGE TO -12500    FROM -16480
A 3( 581)THRU  626 CHANGE TO -12300    FROM -14140
A 4( 327)THRU  462 CHANGE TO  -2900    FROM -13400
A 4( 581)THRU  626 CHANGE TO  -2900    FROM  -5561
A 6( 177)          CHANGE TO  -2500    FROM    426
A 6( 327)THRU  330 CHANGE TO  -3000    FROM  -4766
B10( 327)THRU  328 CHANGE TO    505    FROM    457
DATAFILE  68   OF TAPE  1031 (SPECTRAL DATA-6)    4 CARDS
A 3( 760)THRU  771 CHANGE TO -12600    FROM -12860
A 4( 760)THRU  766 CHANGE TO  -3600    FROM  -3089
B 9( 758)THRU  766 CHANGE TO    570    FROM    534
B10( 758)THRU  766 CHANGE TO    484    FROM    450
DATAFILE  95   OF TAPE  1031 (SPECTRAL DATA-6)    2 CARDS
A 3( 506)THRU  521 CHANGE TO -12471    FROM -13410
A 4( 505)THRU  524 CHANGE TO  -3700    FROM  -5613
```

Table 5.4 d Correction data.

Figure 5.7 shows all of the approximately 8000 pts remaining after the data are again low pass filtered and resampled at a tenth rate to Δt = 1000 sec. Figure 5.3 shows the first tenth of these data expanded in time.

Figure 5.4 shows the 800 data points remaining after one more filtering and resampling process to Δt = 10,000 sec. For obvious reasons this is shown at one scale only.

Notice that Figs. 5.2 and 5.5, 5.3 and 5.6, and 5.4 and 5.7 are drawn to the same scale (in seconds/inch) but with different bandwidths (and sample rates). Figures 5.4 and 5.7 show the same segments in time whereas the other pairs mentioned do not (although the end of Fig. 5.3 overlaps the beginning of 5.6 in time). The general time domain characteristics of the various noise sources can be observed(as well as a feeling for the effect of the low pass filter) in these plots.

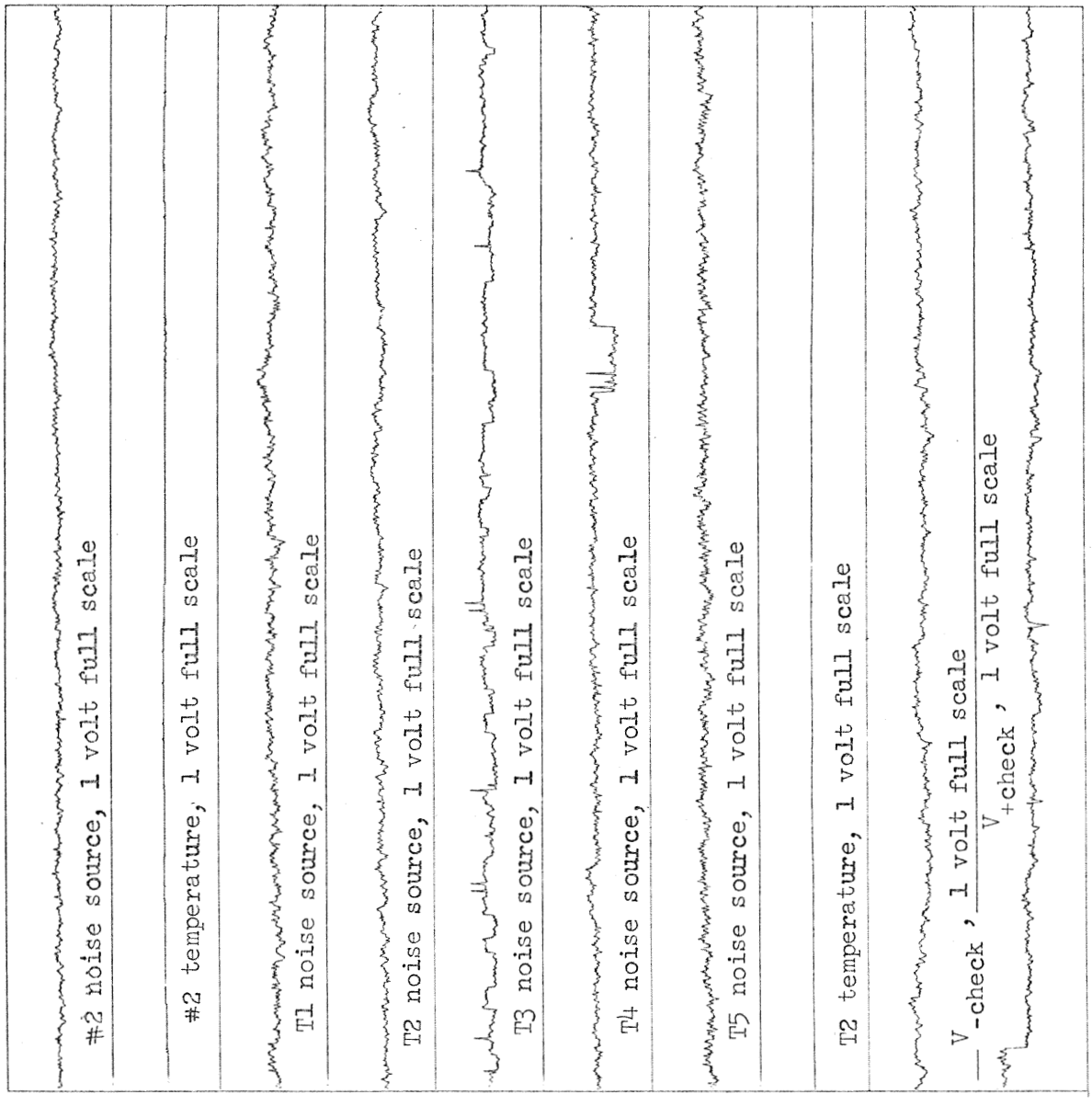


Figure 5.1 Time domain data (1025 points, Δt=10 seconds).

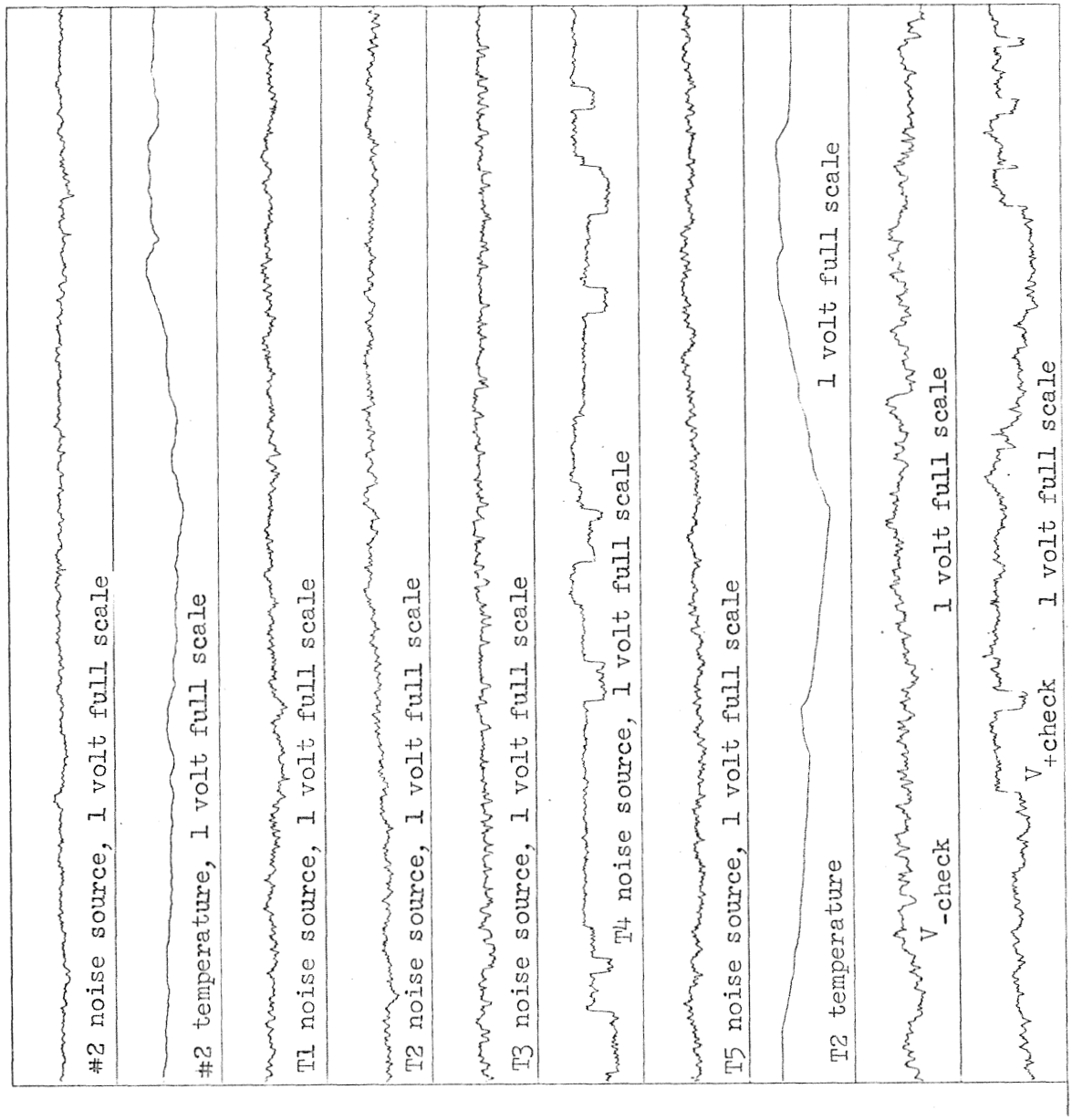


Figure 5.2 Time domain data (1025 points, Δt=100 seconds).

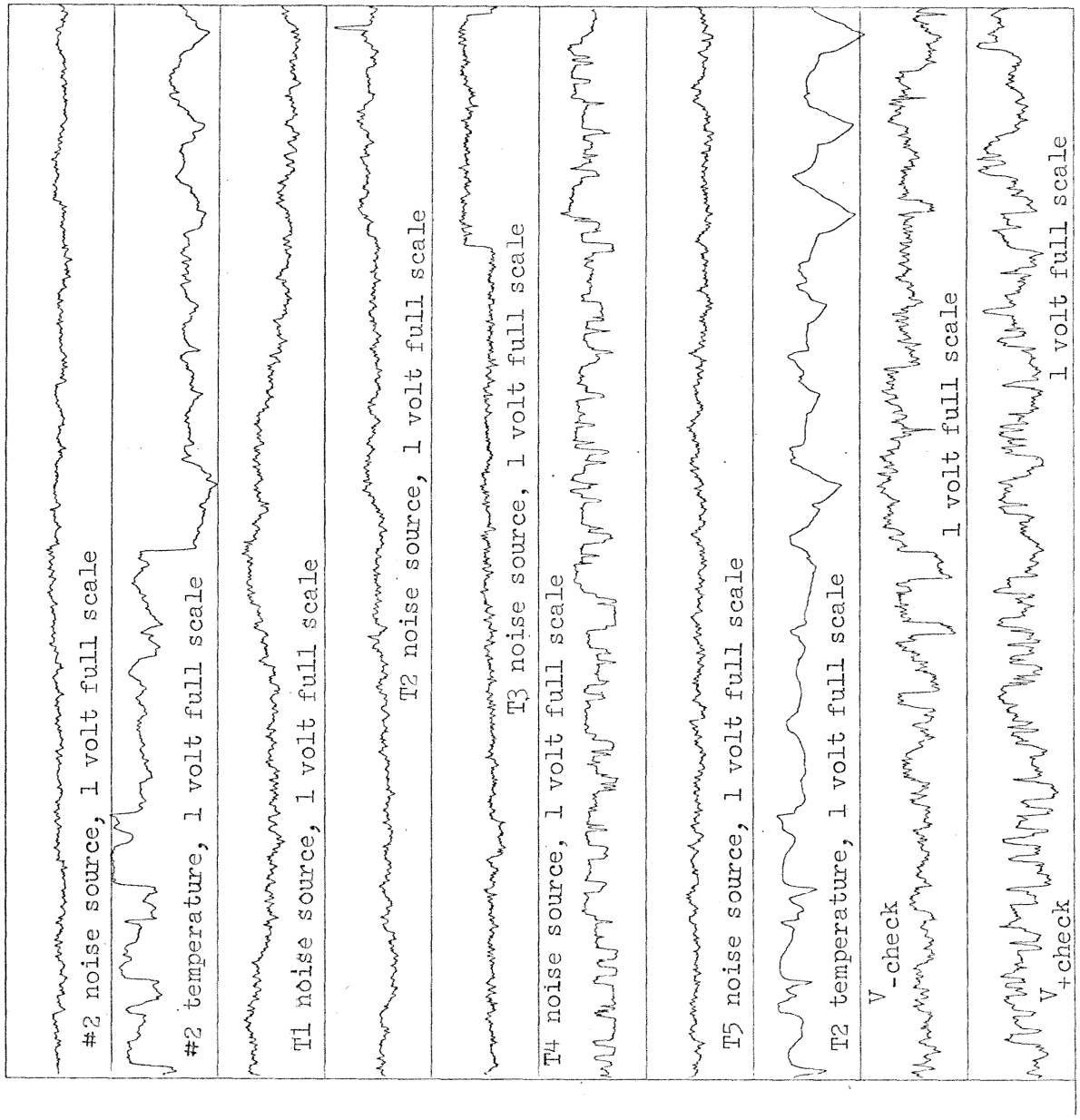


Figure 5.3 Time domain data (1025 points, Δt=1000 seconds).

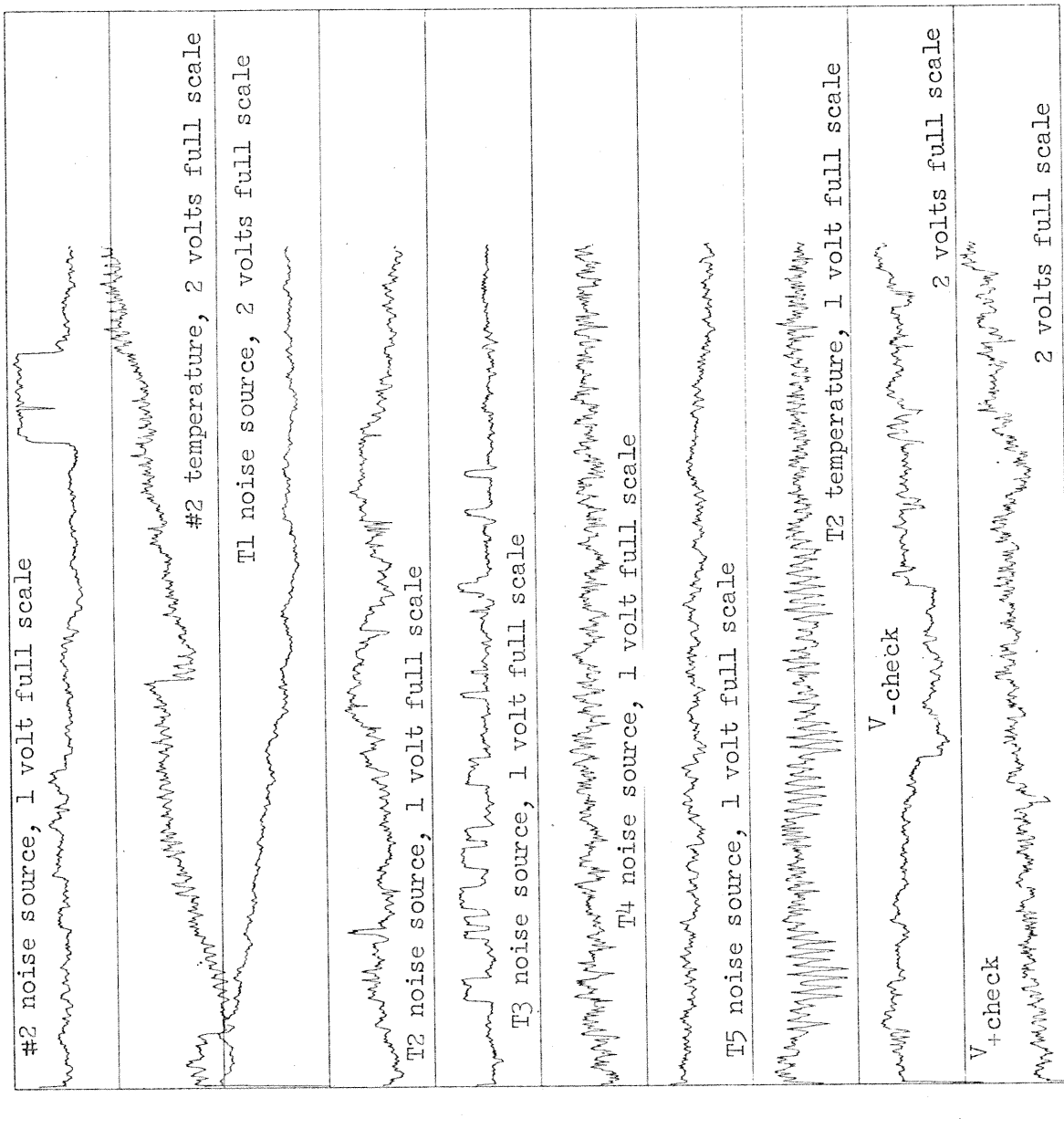


Figure 5.4 Time domain data (1025 points full scale, Δt=10,000 sec.).

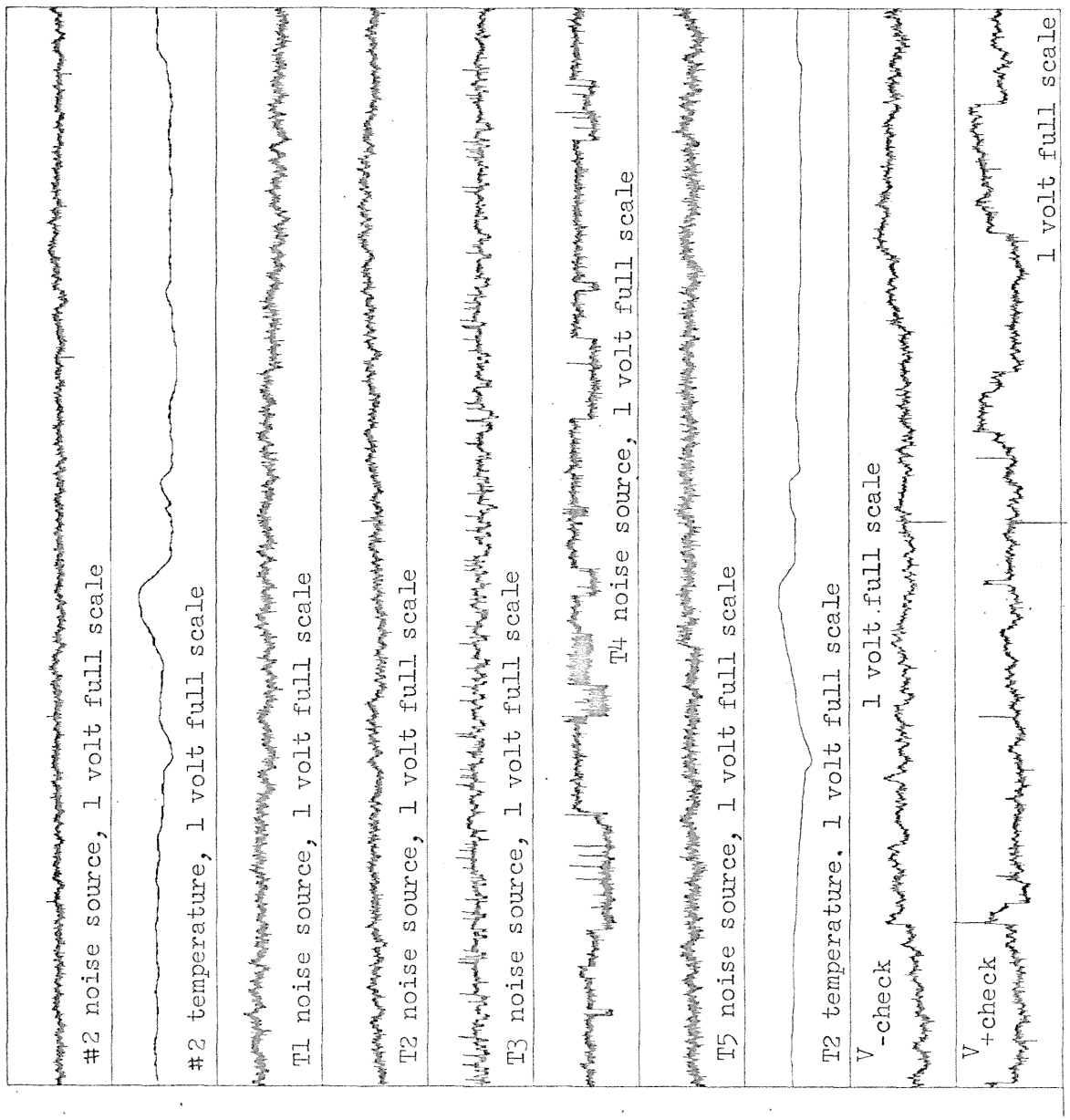


Figure 5.5 Time domain data (10,250 points, Δt=10 seconds).

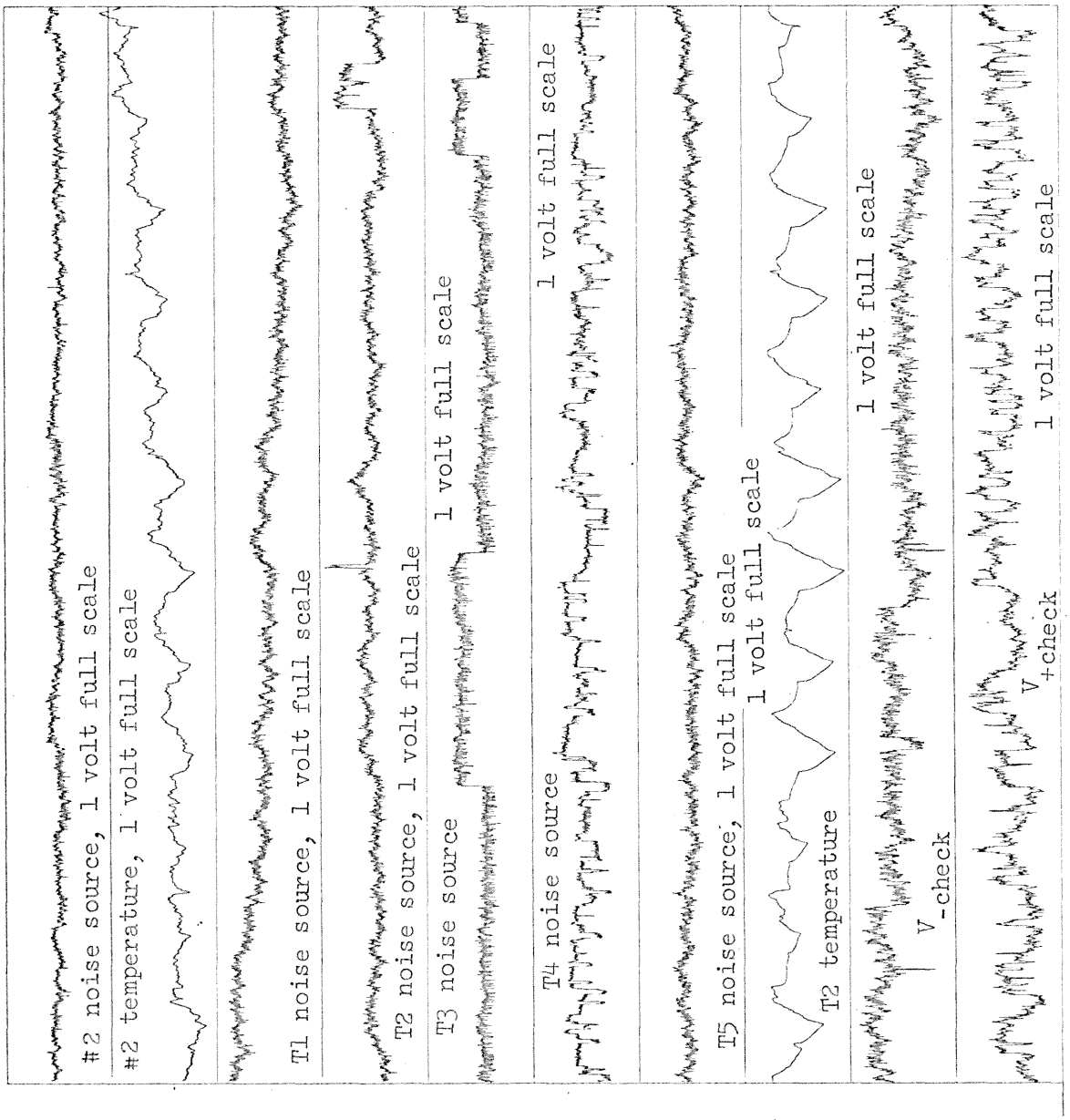


Figure 5.6 Time domain data (10,250 points, Δt=100 seconds).

Figure 5.7 Time domain data (10,250 points full scale, Δt=1000 sec.).

5.4 Power Spectral Density Estimates

The power spectral density estimates were made using the algorithm already discussed (which is the same as used by M. A. Caloyannides[1] and D. S. Blakemore[11]). The power spectral density estimation program was in fact borrowed from Caloyannides (and Blakemore) with slight modifications to allow the program to be run on the Univac 1108, to allow the program to be called from an ALGOL program as a subroutine in order to handle the input data and spectral data estimates more efficiently, and to check for negative power spectral density estimates. Unless otherwise indicated, all estimates were made with M = 100 (number of lags) to yield 2 decades of frequency. Each time the sampling rate is reduced by a factor of ten, therefore, we obtain estimates which overlap the previous estimates over one decade.

Figures 5.8,...,5.22 show the power spectral density estimates for the six noise sources, the temperatures for the six noise sources, room temperature, V_{-check}, and V_{+check}. The effects of the initial under sampling and the non-ideal characteristics of the digital dealiasing filter have not been accounted for although it will be remembered that these appreciably affect only the highest decade of each estimate. The 1/f prewhitening filter was used in all of these estimates although for some of these sources a $1/f^2$ filter might be more appropriate.* The spectral density estimates of the sources represented by Figs. 5.14, ...,5.22 were originally made with no prewhitening filter (but with mean

---

*The 1/f prewhitening filter is closer to f^2 below the first decade anyway,

removal prior to the spectral estimation program). These estimates contained so many negative power spectral density estimates that they were considered unusable and are not presented).

The temperatures of noise sources T1,...,T5 are nearly identical for the lower frequencies as might be expected. At frequencies above 0.001 Hz, however, the temperatures of noise sources T1, T3, and T4 appear to have larger spectral densities than the temperatures of noise sources T2 and T5. This is even more apparent in Figs. 5.23 and 5.24 which show the spectral estimates for the temperatures of T1 and T2 over a larger range of amplitudes. From Table 5.1 we remember that the data for the temperatures of T2 and T5 were taken using the DVM with a resolution of 0.1 millivolt, whereas the other temperature data were taken from the A/D converter with a resolution of approximately 10 millivolts ($\frac{10}{1024}$ volts). Let us assume that the data recorded were the true values of the temperature indicated plus a quantization noise ($n_q(t)$) due to the limits of resolution of the A/D converter. If the normal changes in the true voltage from the temperature indicator during one sampling period were large compared to one quantization level of the A/D converter, then it seems reasonable to assume that $n_q(t)$ is white noise with a probability density of

$$P_{n_q}(x) = \begin{cases} \frac{1}{q} & \text{for } |x| < \frac{1}{2} q \\ 0 & \text{for } |x| \geq \frac{1}{2} q \end{cases}$$

where q is the quantization step size ($\frac{10}{1024}$ volts in this case).

Under these assumptions $\langle q^2 \rangle = \int_{-\infty}^{\infty} x^2 p_{n_q}(x)dx = \frac{2}{q}\int_0^{\frac{1}{2}q} x^2 dx = \frac{1}{3}\left(\frac{q}{2}\right)^2$.

Using the white noise assumption, sampling period of Δt = 10 sec (implying a 1/10 Hz bandwidth), and $q = \frac{10}{1024}$ volts we get

$$S_{n_q}(f) = \frac{1}{12}\left(\frac{10}{1024}\right)^2 \text{ 10 volt}^2 \text{ seconds}$$

$$= 7.947 \times 10^{-5} \text{ volt}^2\text{-sec (-41.0 db)} \quad .$$

Referring again to Fig. 5.23 we see that the spectral density of T1 temperature has a minimum value of about -42.5 ± 0.7 db. Note that subtracting 3 db at the high frequency end due to aliasing is not appropriate here since the total average power of the noise $n_q(t)$ was considered rather than the power spectral density of the quantization noise before sampling. It is interesting that the power spectral density estimate of the T1 noise source temperature actually goes below the power spectral density calculated for the quantization noise. This problem may be resolved (if one considers 1.5 db significant) by noting that the temperature of the noise source actually changed so slowly that the output of the A/D converter often remained fixed for several samples at a time. The quantization noise in this case would not be white or uncorrelated with the temperature and we might expect the sum to yield a lower power spectal density at the higher frequencies (a conclusion made on the basis of the previous comment that the output of the A/D converter often does not change over a period of several samples, which would intuitively imply low high-frequency components).

At this point a slight digression will be made to discuss the thermal transfer function of the oven used to control the temperatures of noise sources T1,...,T5. We now have power spectral density estimates for room temperature and for the temperatures of each of the noise sources. If we assume that all heat sources inside the oven have constant power dissipation then we may estimate the magnitude squared of the thermal transfer function between the room temperature and the temperature of noise source T2 by dividing the spectral estimate of T2 temperature by the spectral estimate for room temperature and multiplying by a constant to account for the different gains of the two temperature measurements.

$$10 \log |\hat{H}(f)|^2 \sim 10 \log \hat{S}_{\substack{T2 \\ temp.}}(f) - 10 \log \hat{S}_{\substack{Room \\ temp.}}(f) - 23.5 \text{ db} .$$

Figure 5.25 shows this estimate of $|H(f)|^2$. Notice that the dimensions of H(f) are in [degrees change of T2 temperature]/[degrees change of room temperature] rather than the usual dimensions involving volts at the input to the measuring channel. As can be seen from Fig. 5.25, the temperature variations at the noise source are reduced by greater than 25 db at the lower frequencies (a factor of 17.8 in temperature amplitudes) as compared to the room temperature. The oven clearly has the characteristics of a low pass filter. The fact that this low pass filter levels off at a frequency of approximately 0.01 Hz is somewhat surprising. The loop gain of the temperature control breaks at a slightly higher frequency than this. Referring back to Fig. 5.24 we see that a white noise with an rms value of .067 mv added

to the T2 temperature measuring channel would be sufficient to cause leveling off of the T2 temperature spectral density to -73.5 db as observed in Fig. 5.24 even though the estimate on quantization noise for the DVM would predict a leveling off at a spectral density of -80.8 db. When one considers that .067 millivolts could result from only 3.3 milliamps of current through 2 feet of #20 copper wire, a noise voltage of this order due to system ground loops would not be surprising.

Referring to Figs. 5.8,...,5.13 we see that if we assume that $S_n(f) = Af^{-\alpha}$ we might estimate the values of α for noise sources #2,T1, ...,T5 to be 1.27, 1.17, 1.3, 1.24, 1.13, 1.08. If we average these six estimates for α we get $\langle\alpha\rangle = 1.20$. It is noted here that this value of α is not far removed from the value of 1.3 found by M. A. Caloyannides when he averaged his ten noise sources together.[1, page 191]

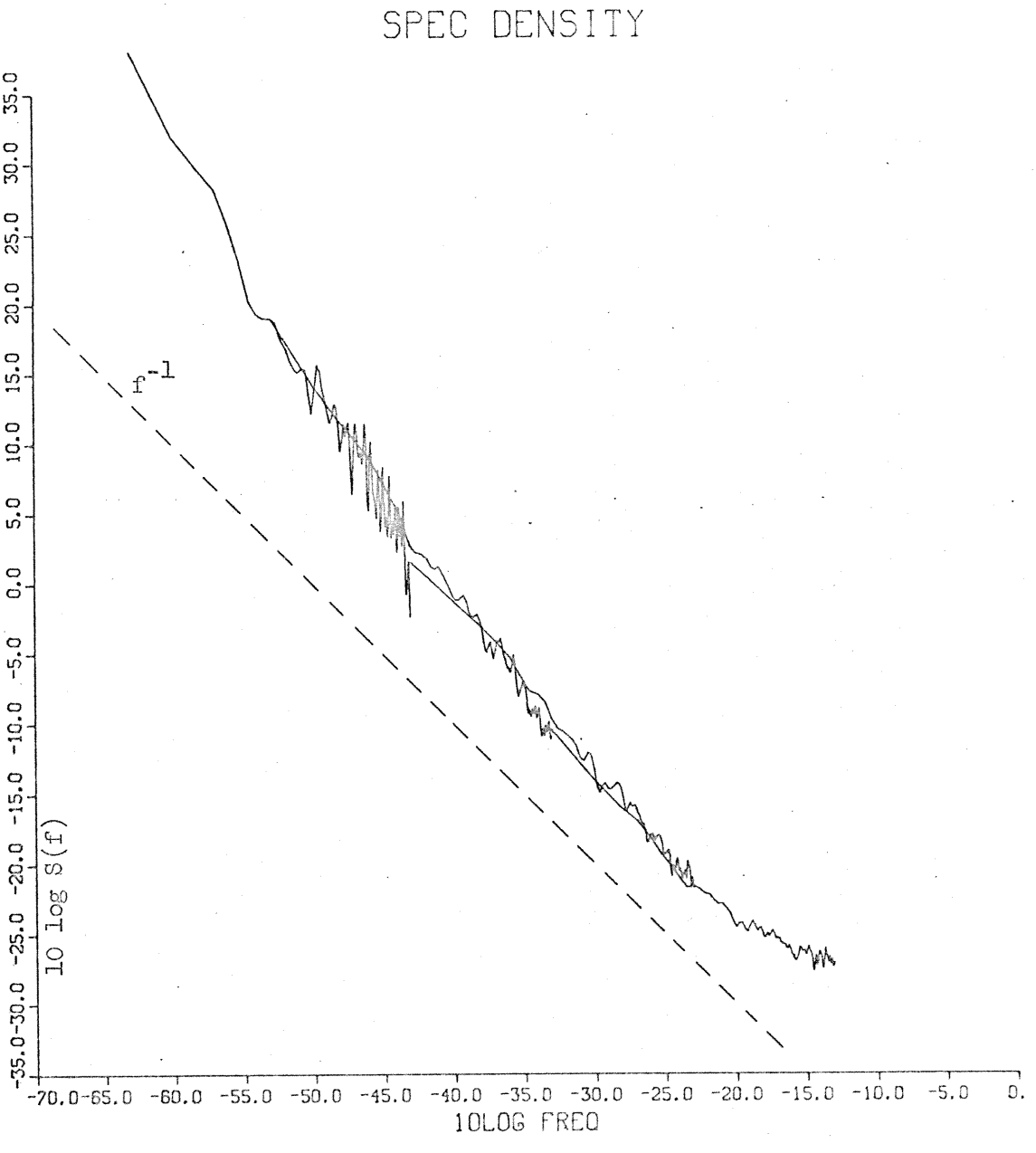


Figure 5.8 Noise source #2 spectral estimates ($-63 \leq 10 \log f \leq -13$).

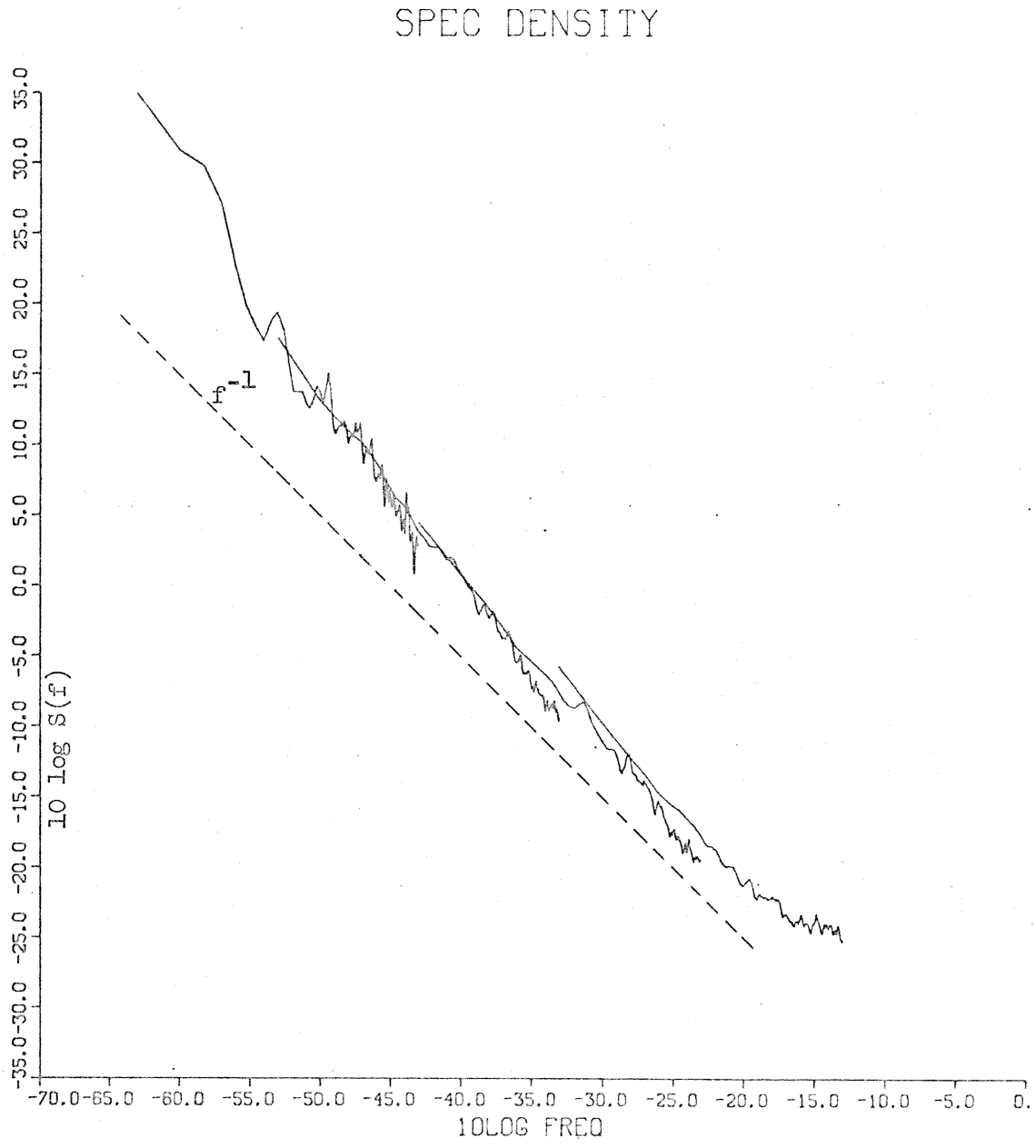


Figure 5.9 Noise source T1 spectral estimates ($-63 \leq 10 \log f \leq -13$).

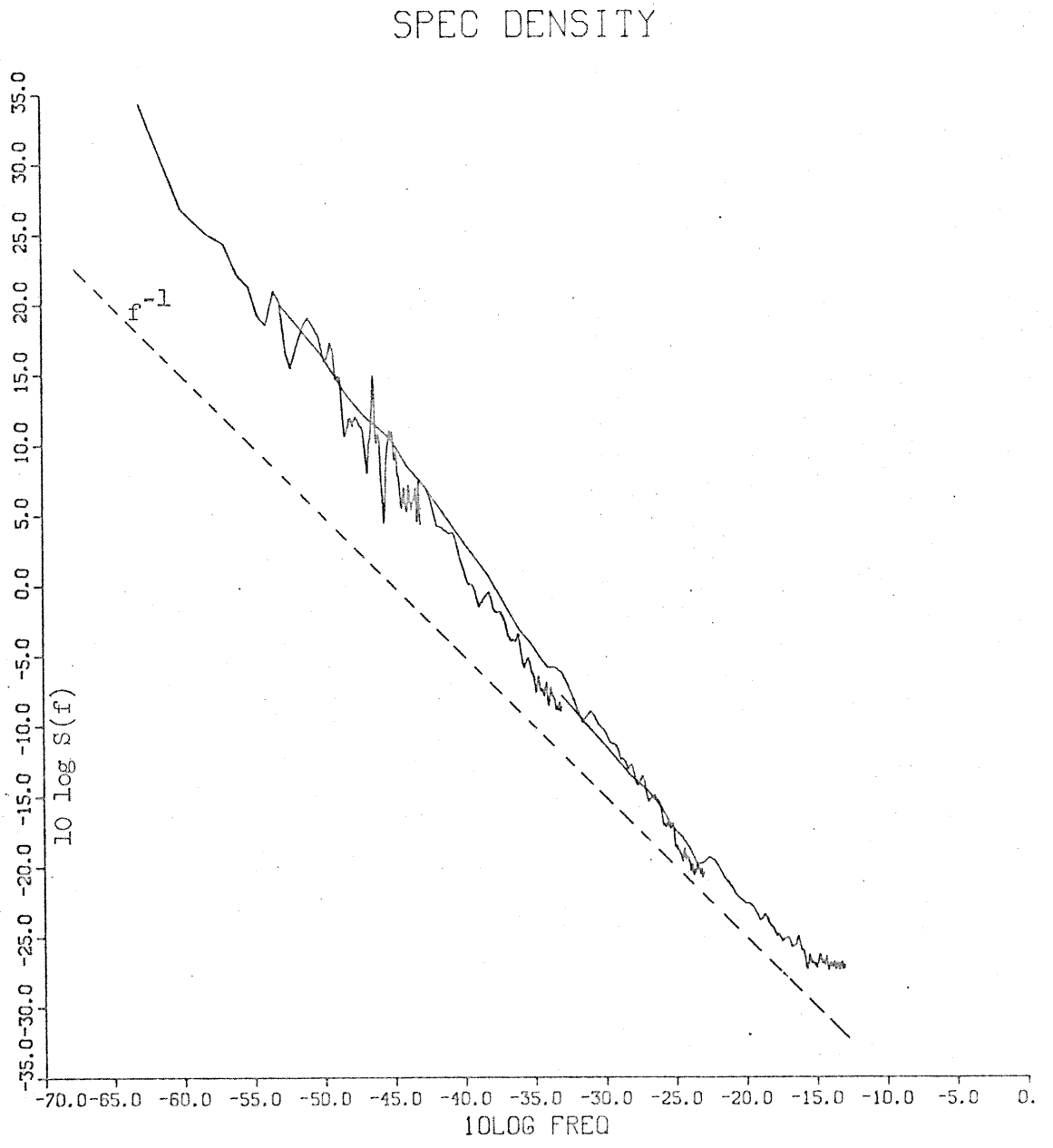


Figure 5.10 Noise source T2 spectral estimates ($-63 \leq 10 \log f \leq -13$).

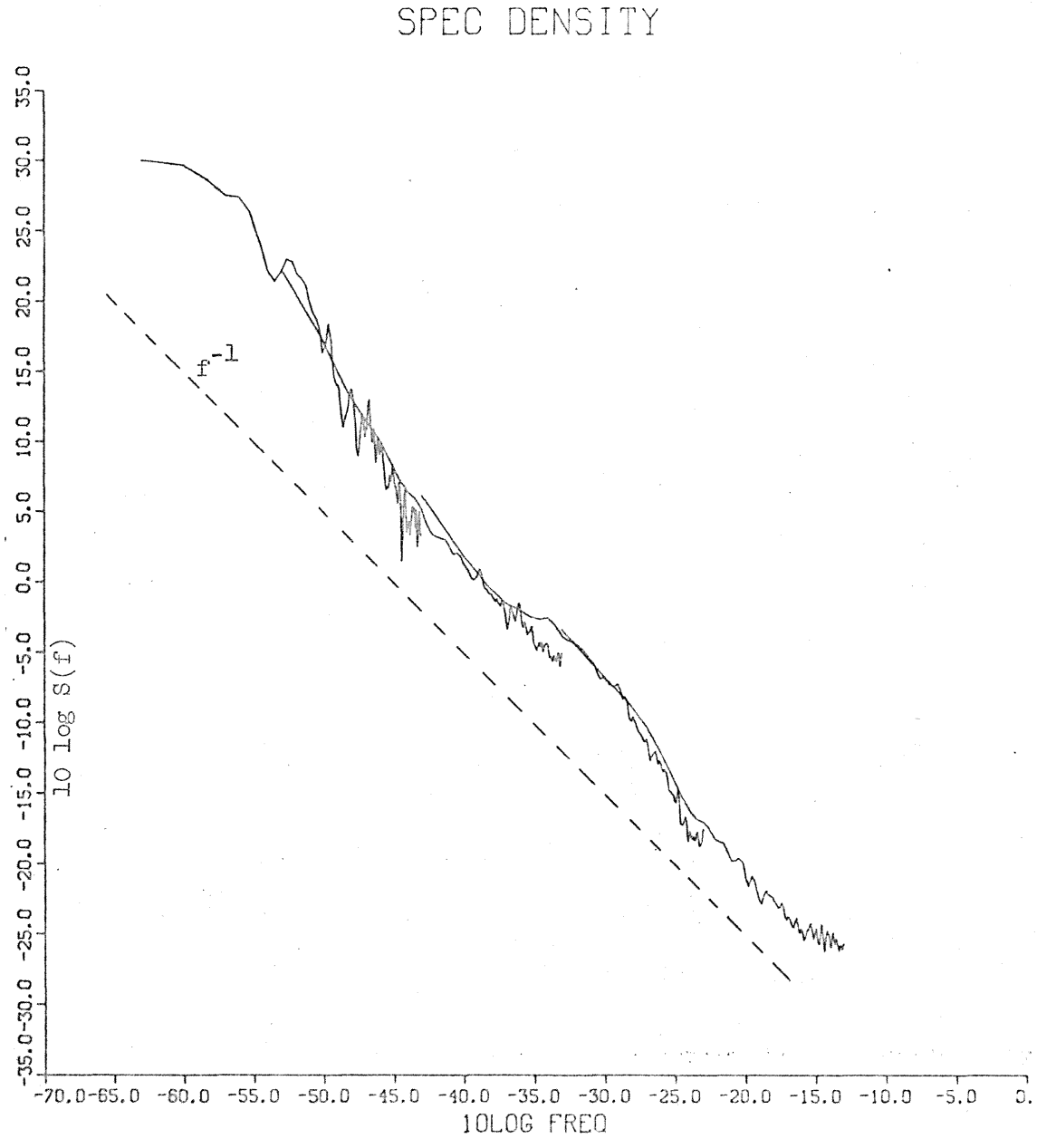


Figure 5.11 Noise source T3 spectral estimates ($-63 \leq 10 \log f \leq -13$).

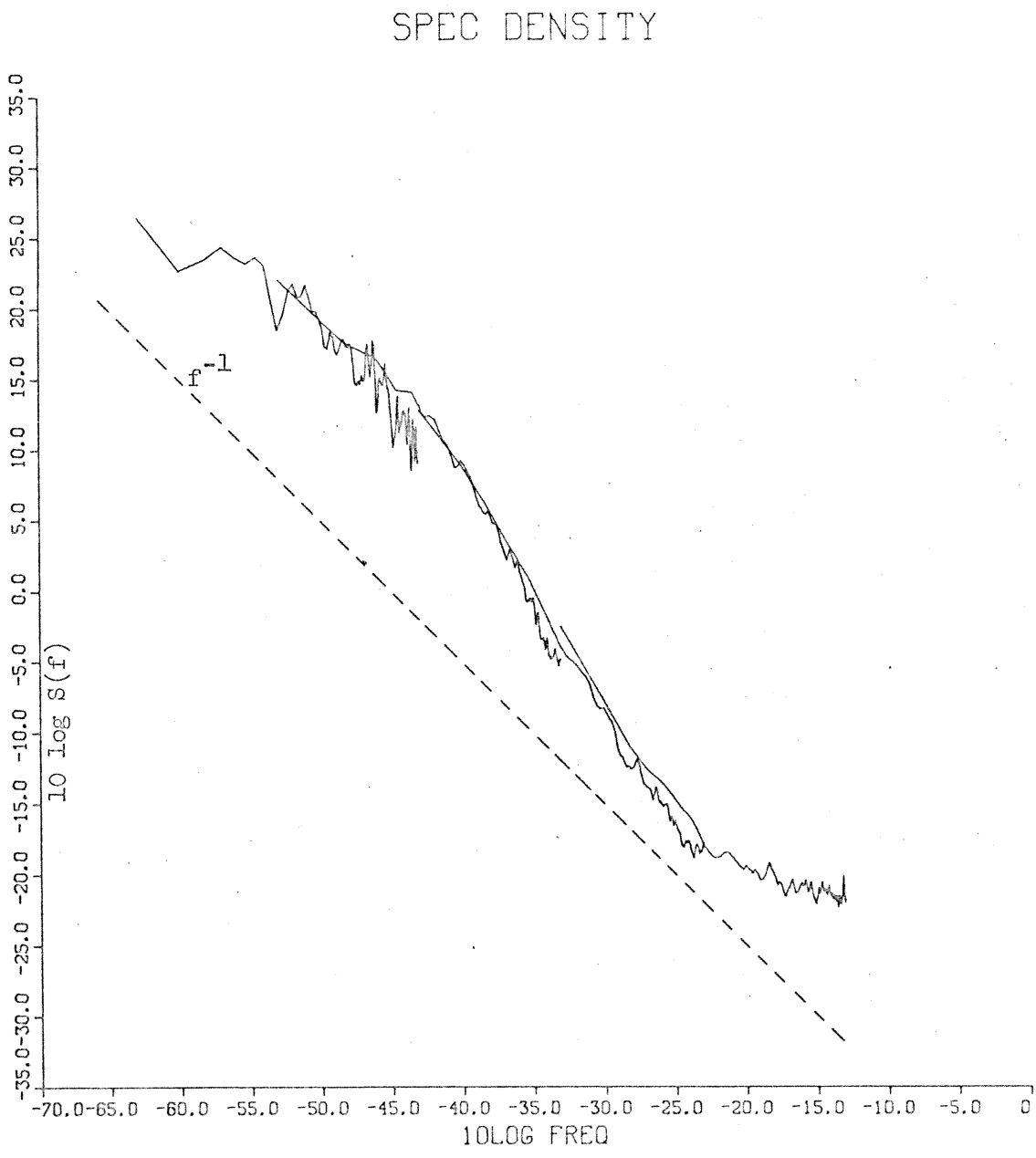


Figure 5.12 Noise source T4 spectral estimates ($-63 \leq 10 \log f \leq -13$).

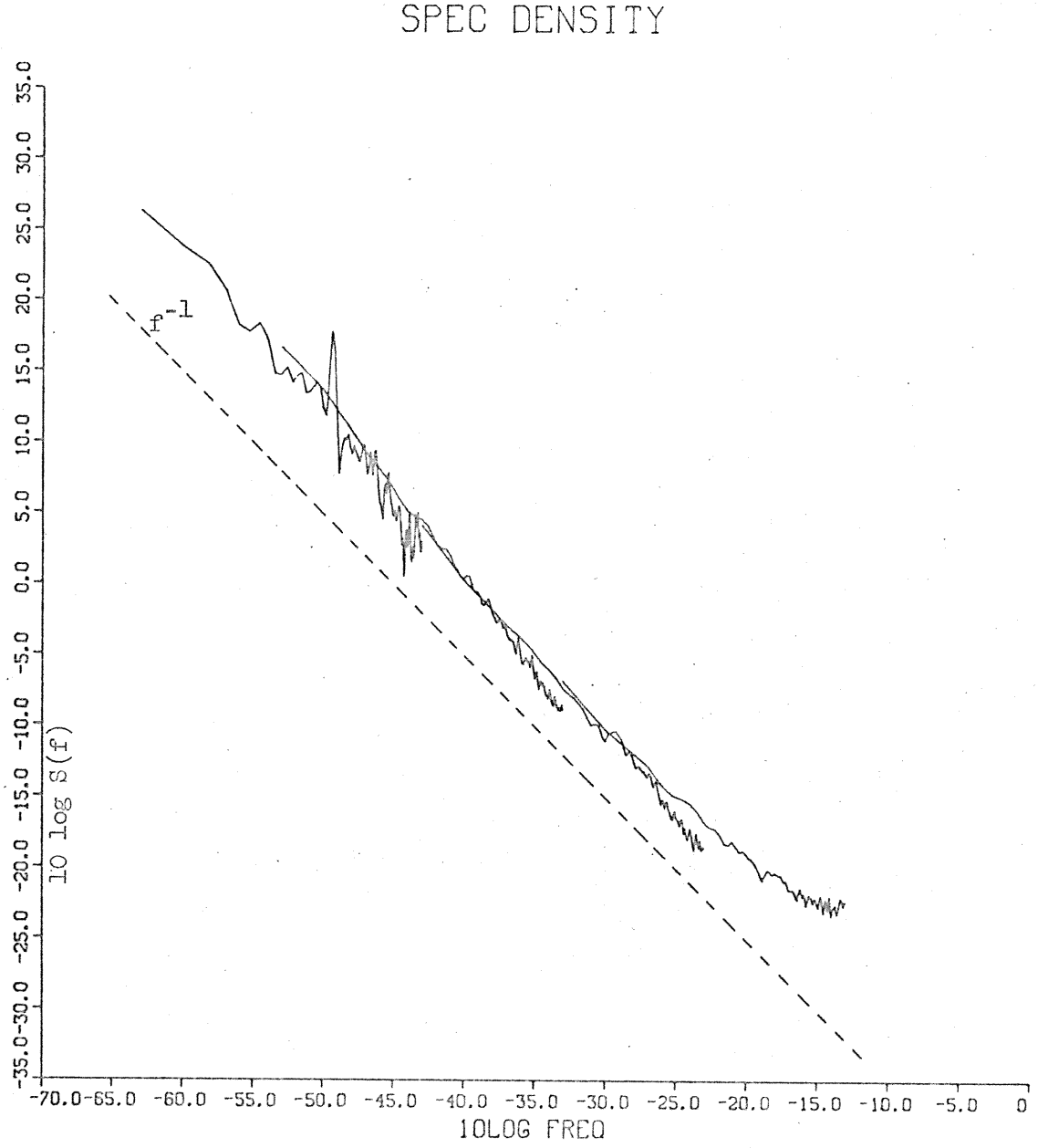


Figure 5.13 Noise source T5 spectral estimates ($-63 \leq 10 \log f \leq -13$).

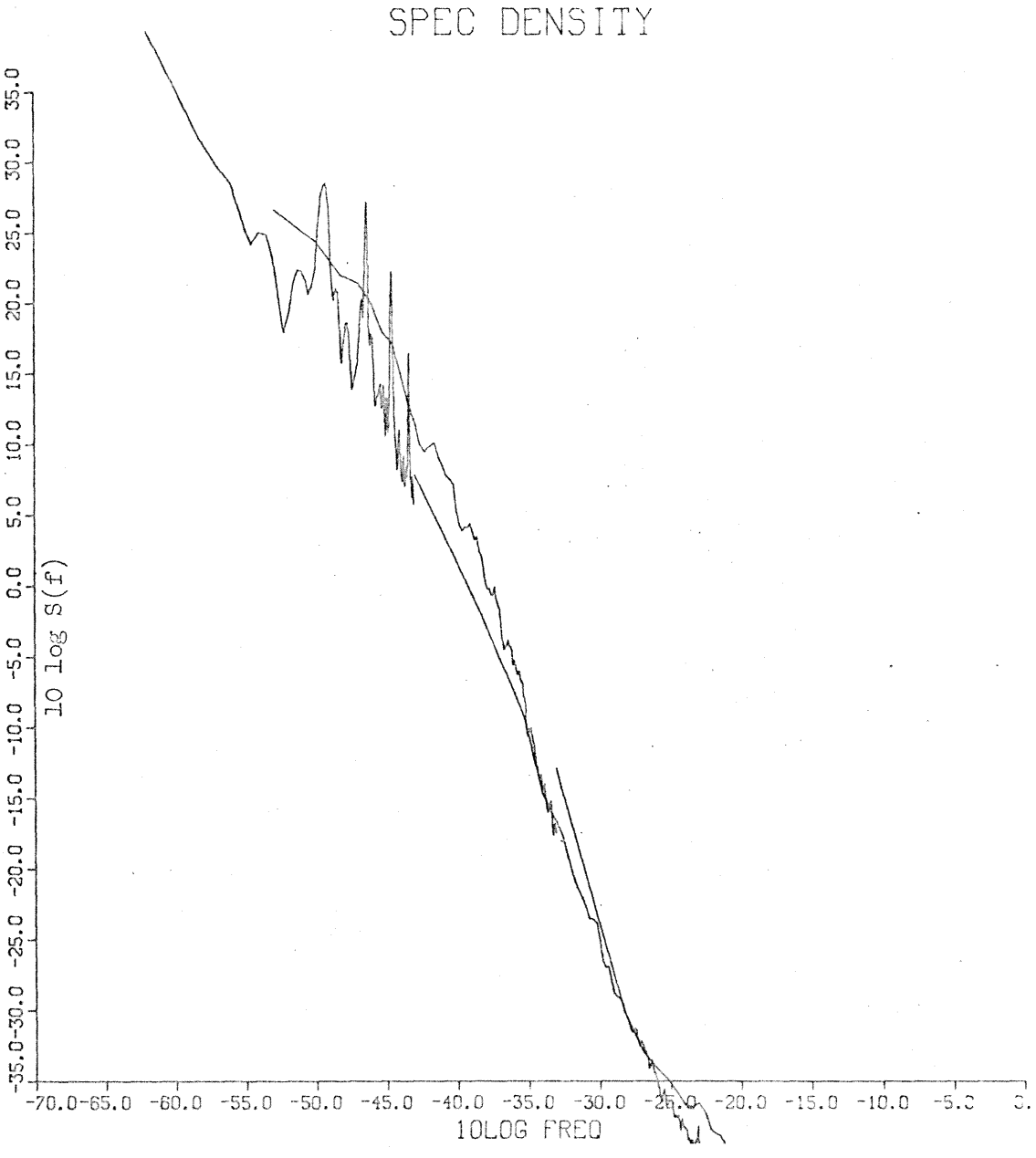


Figure 5.14 #2 temperature spectral estimates ($-63 \leq 10 \log f \leq -13$).

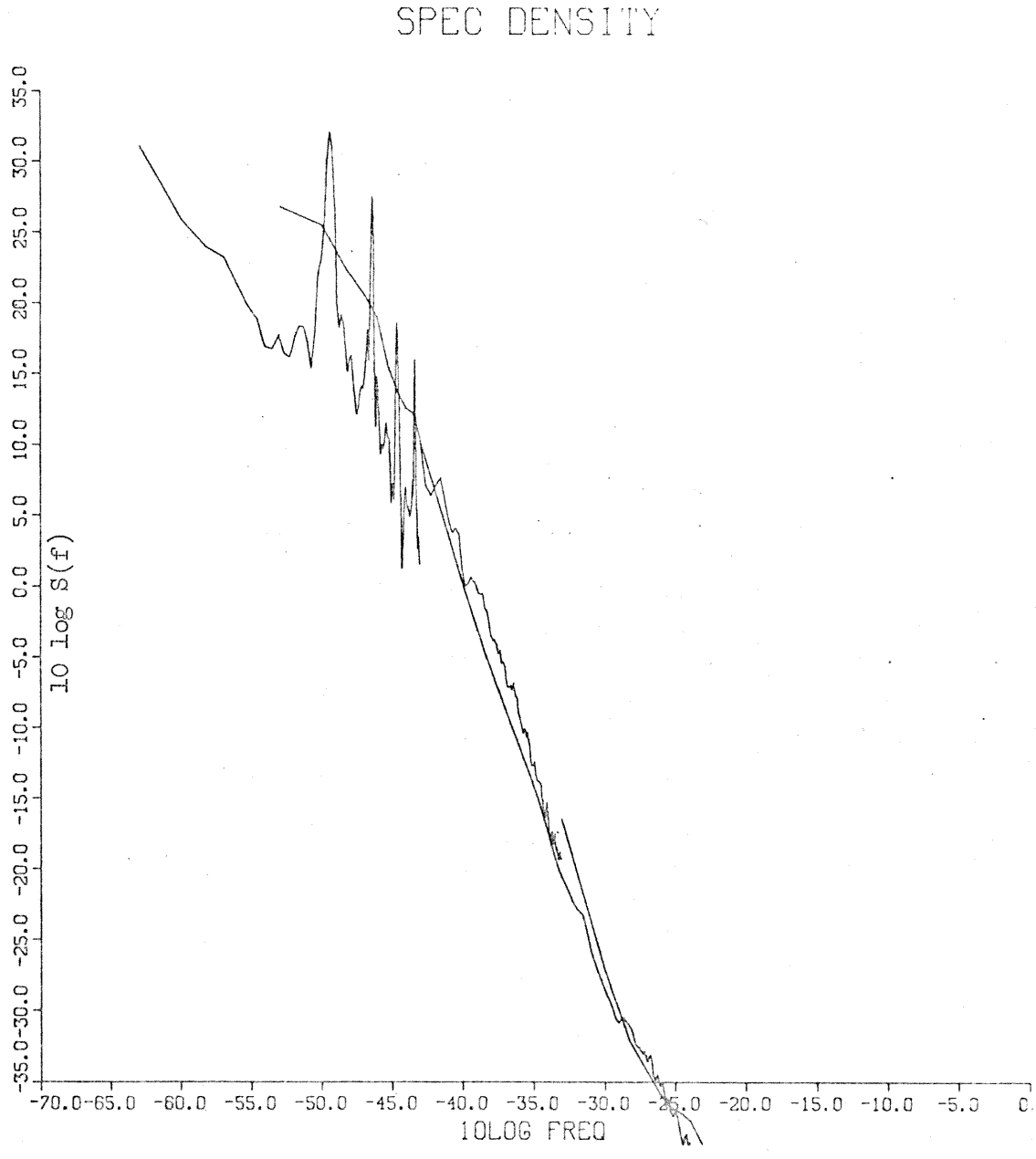


Figure 5.15 T1 temperature spectral estimates ($-63 \leq 10 \log f \leq -13$).

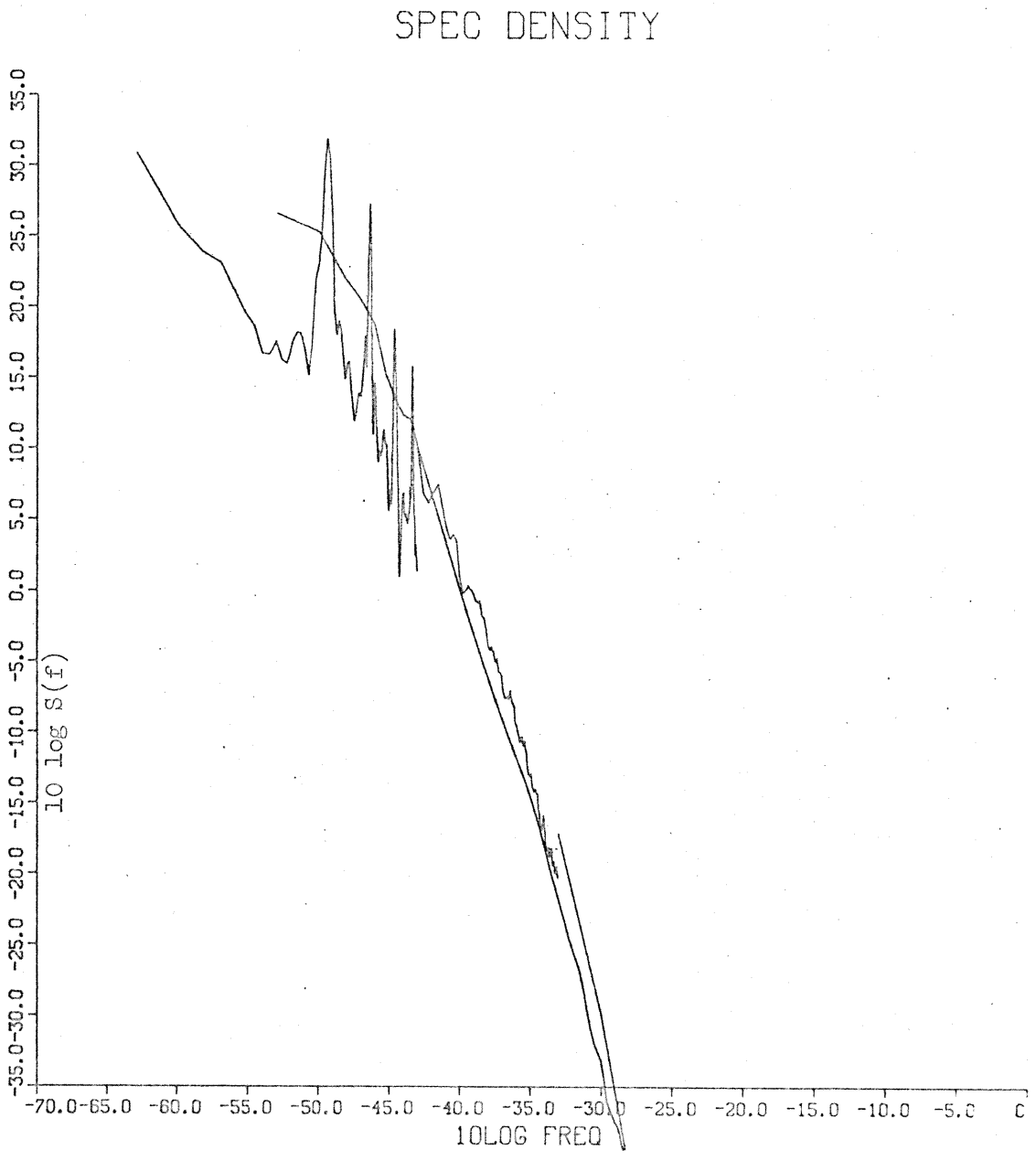


Figure 5.16 T2 temperature spectral estimates ($-63 \leq 10 \log f \leq -13$).

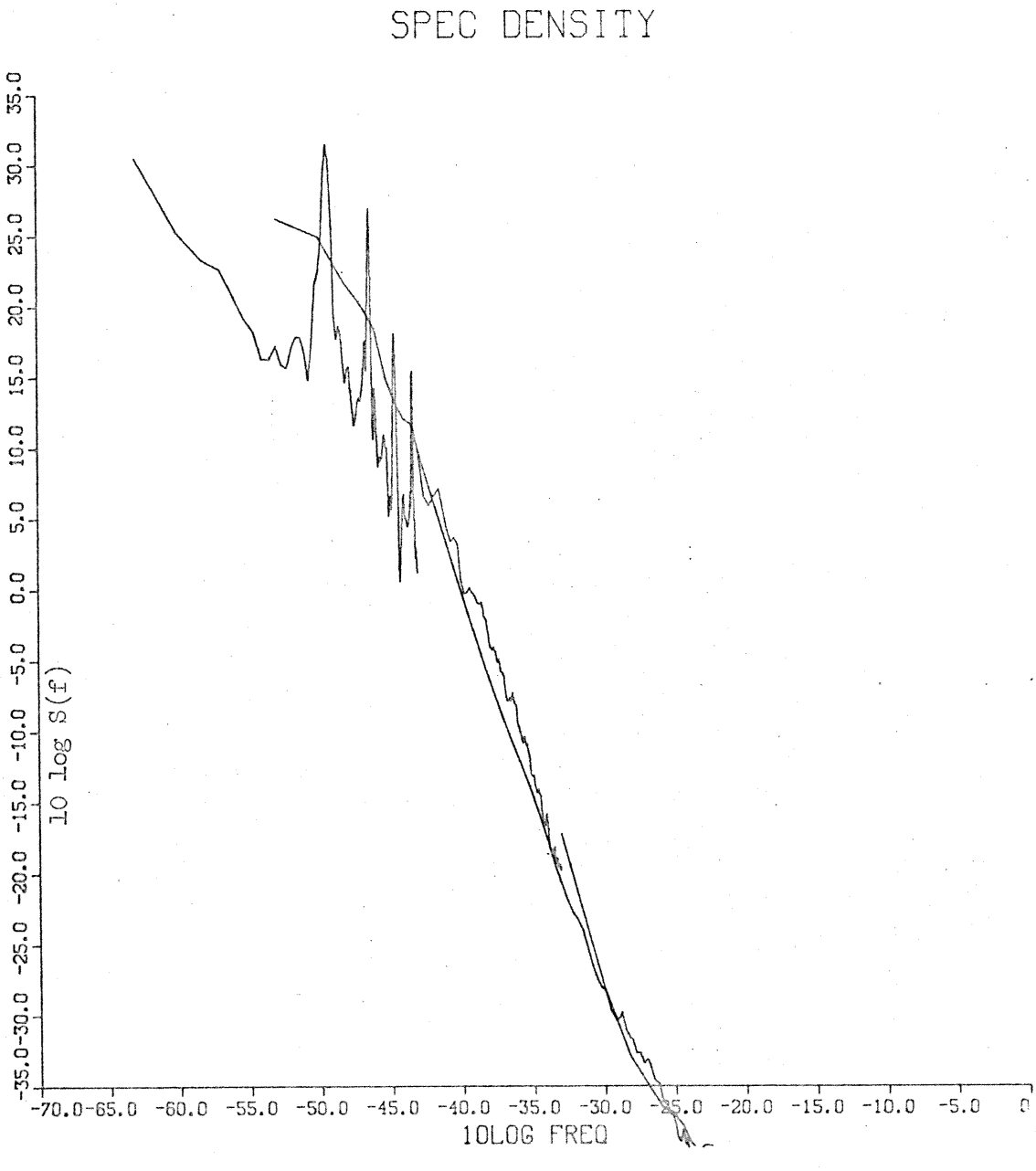


Figure 5.17 T3 temperature spectral estimates ($-63 \leq 10 \log f \leq -13$).

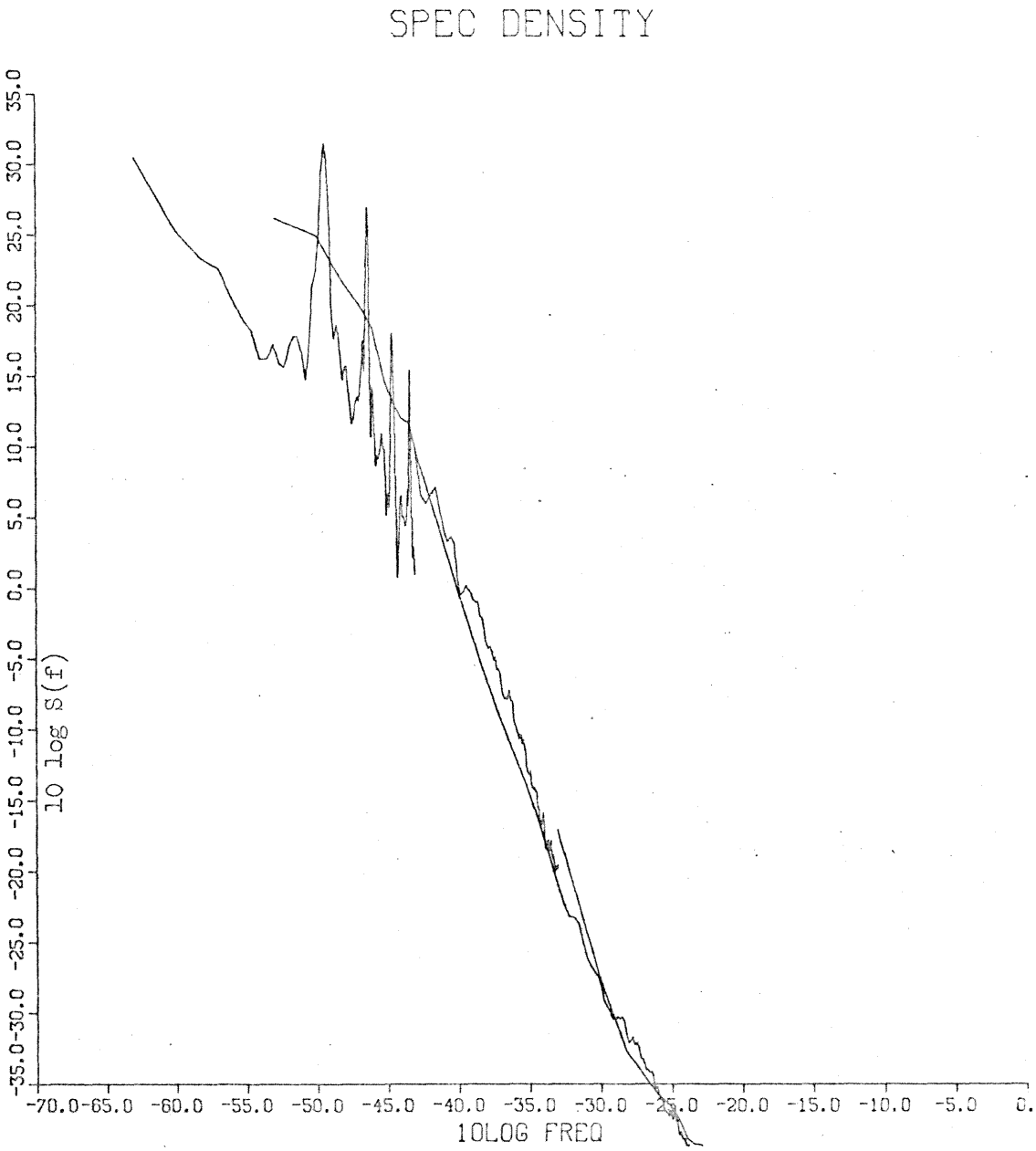


Figure 5.18 T_4 temperature spectral estimates ($-63 \leq 10 \log f \leq -13$).

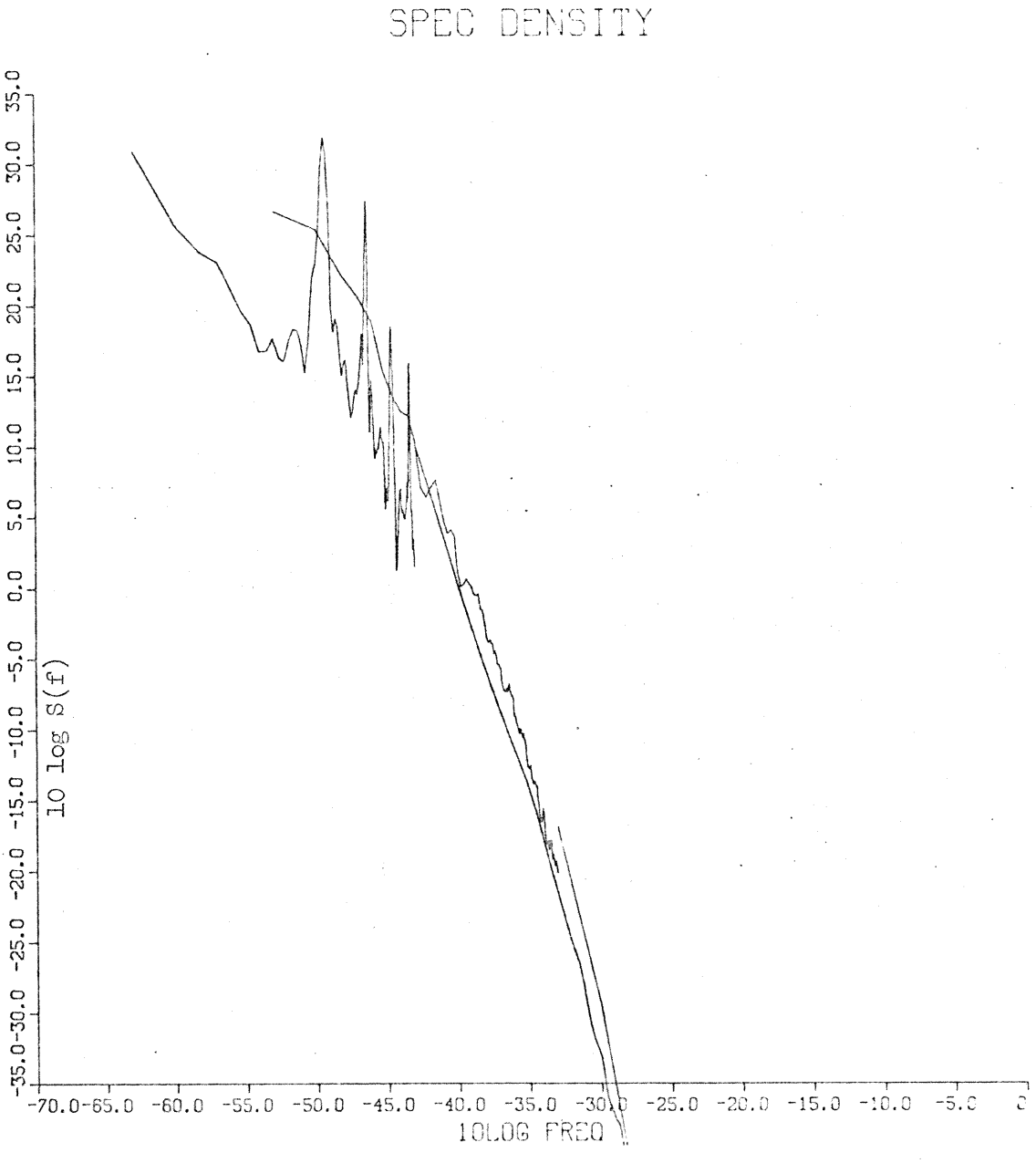


Figure 5.19 T5 temperature spectral estimates ($-63 \leq 10 \log f \leq -13$).

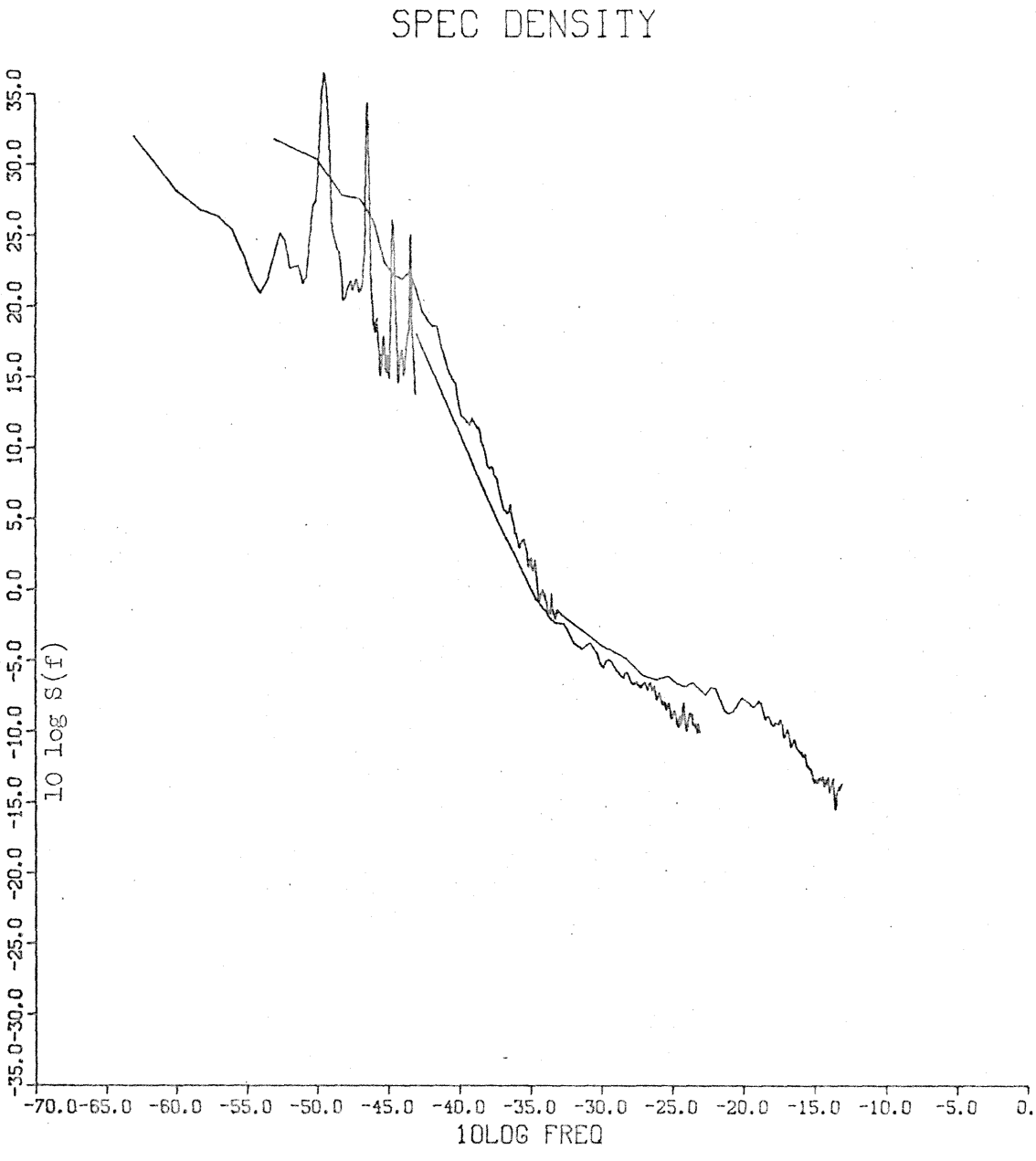


Figure 5.20 Room temperature spectral estimates ($-63 \leq 10 \log f \leq -13$).

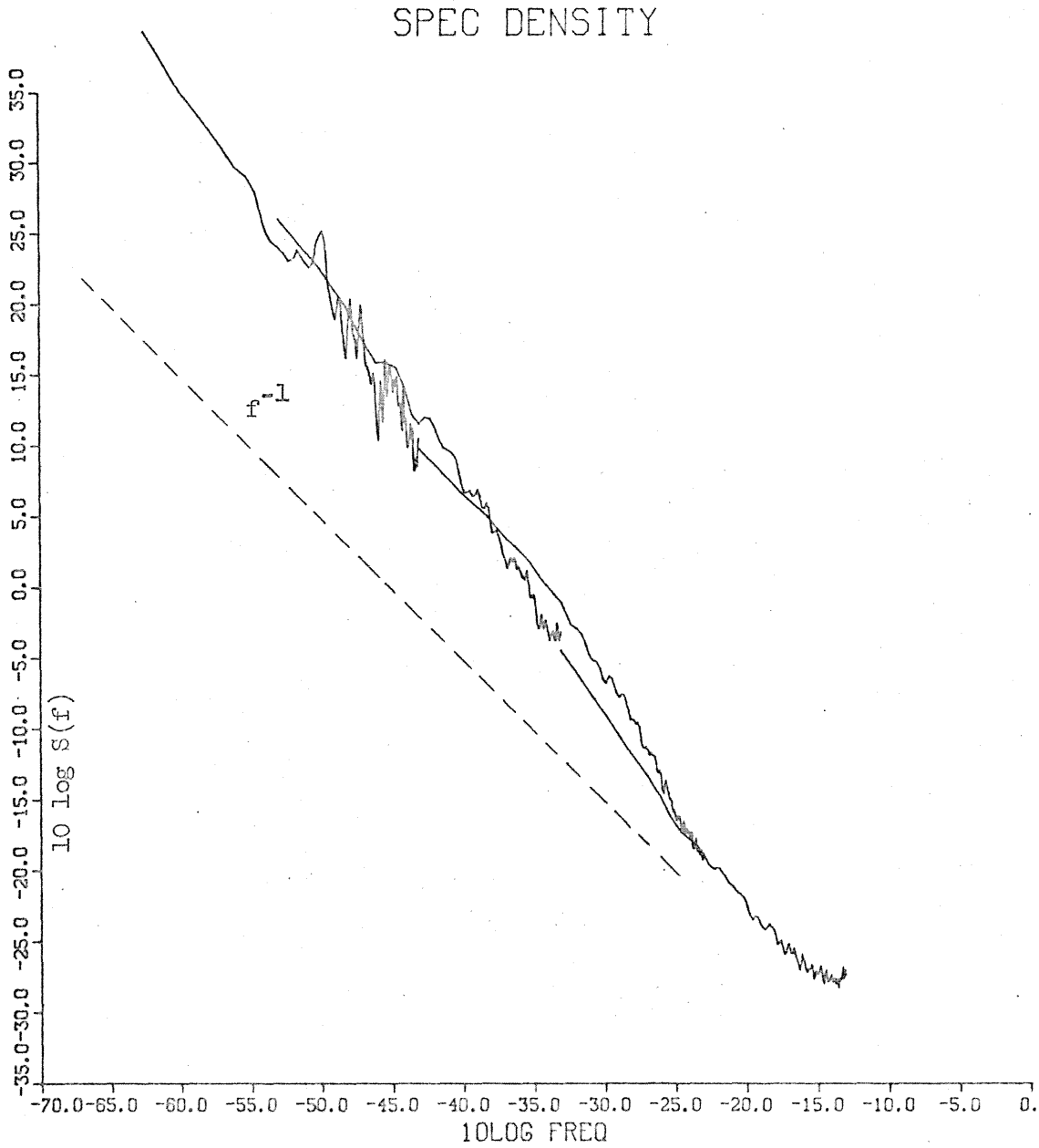


Figure 5.21 V_{-check} spectral estimates ($-63 \leq 10 \log f \leq -13$).

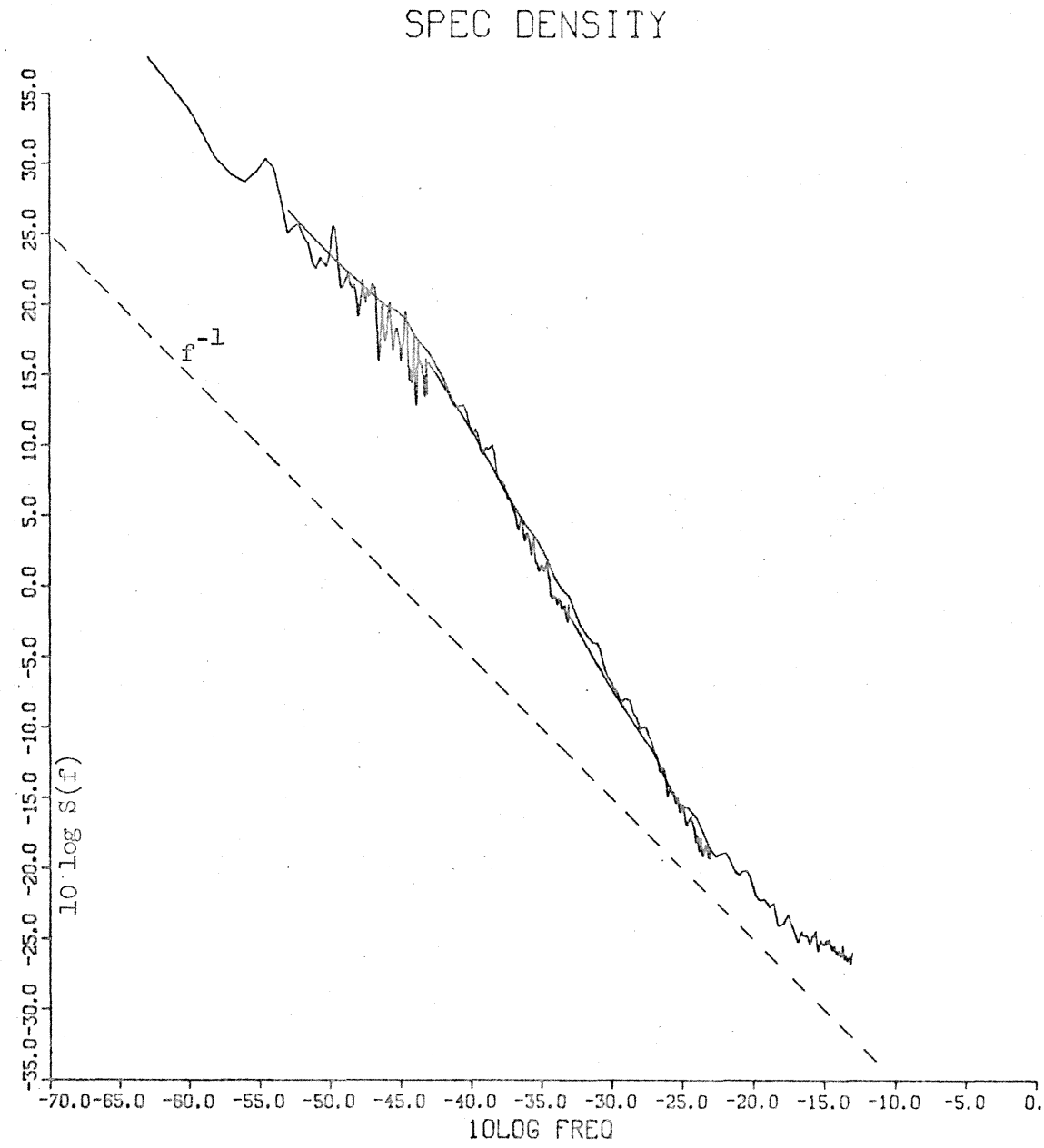


Figure 5.22 V_{+check} spectral estimates (-63= 10 log f =-13).

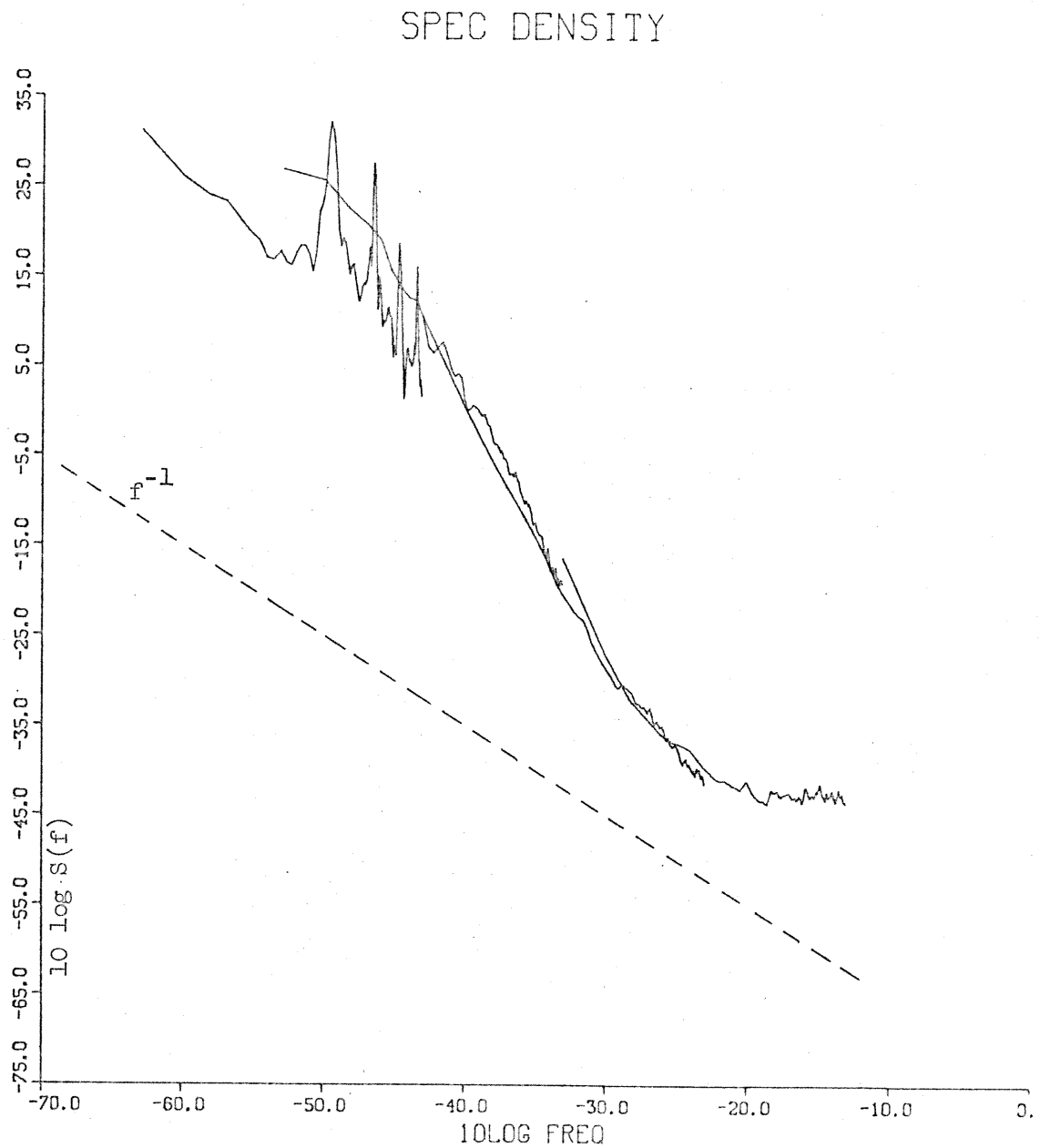


Figure 5.23 T1 temperature spectral estimates ($-63 \leq 10 \log f \leq -13$, compressed vertical scale).

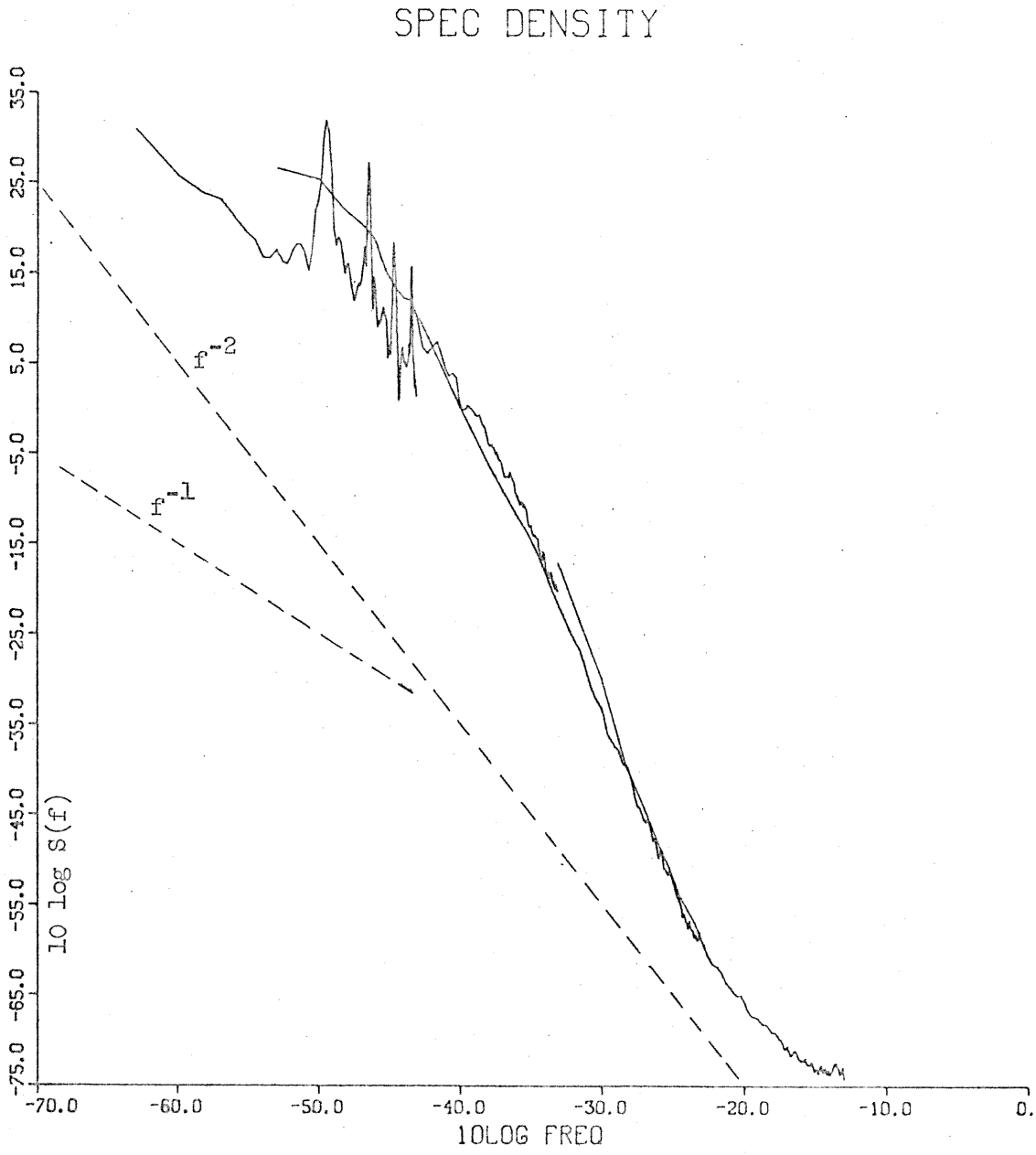


Figure 5.24 T2 temperature spectral estimates ($-63 \leq 10 \log f \leq -13$, compressed vertical scale).

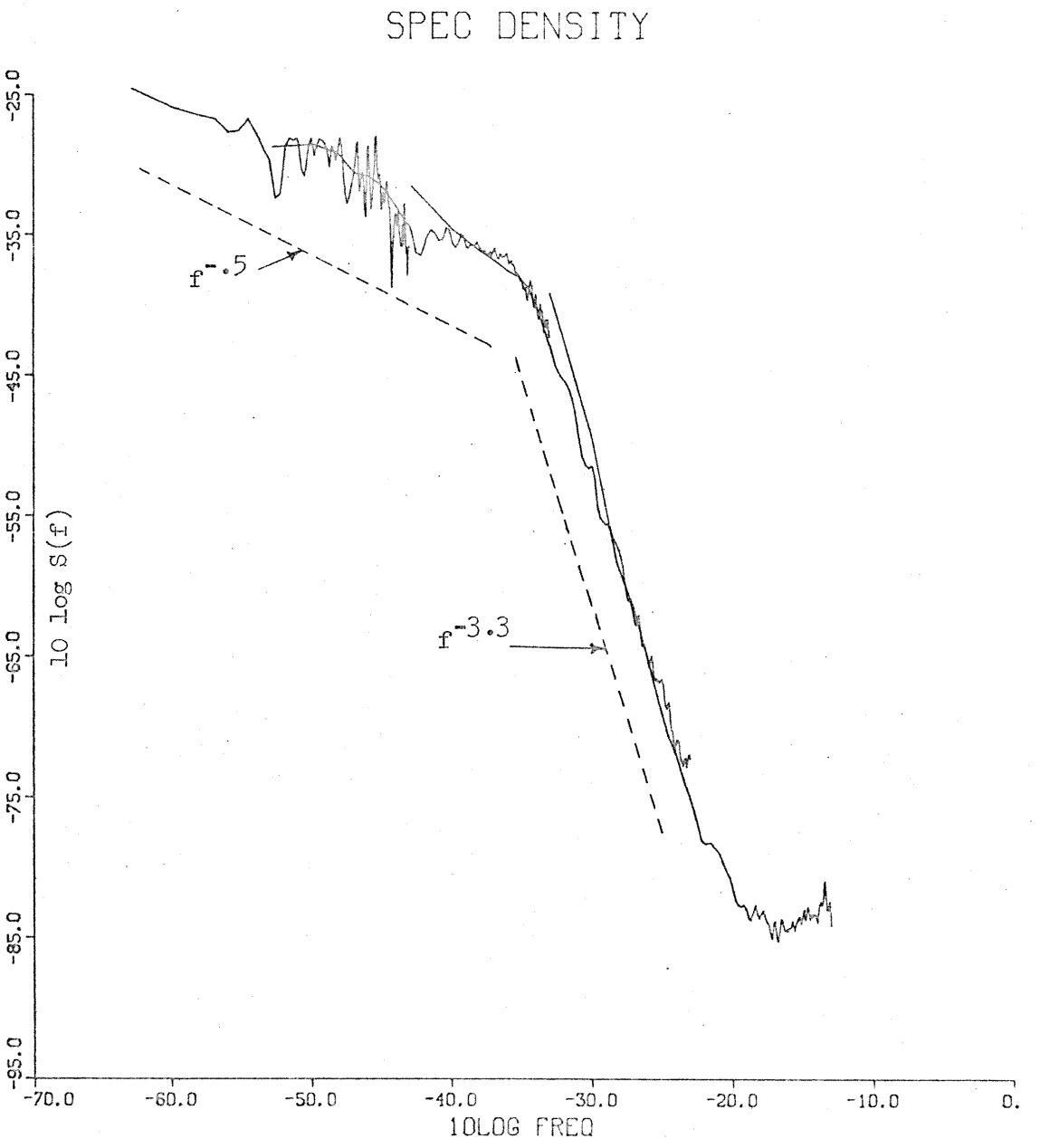


Figure 5.25 Estimate of magnitude squared thermal transfer function from room temperature to noise source T2 (both in units of degrees).

5.5 Temperature Coefficient Estimates

A great deal of effort was expended to reduce the variations of the temperatures of the noise sources. No claim, however, has been made that the temperature of the noise sources has actually been held to a constant. The hope, therefore, is to either subtract any effects of a now known temperature from the data before processing or to show that the temperature does not significantly affect the final results. In order to do this we need a reasonable estimate of the temperature coefficients of each of the amplifiers. Shortly before the start of the long term measurements an attempt was made to measure these coefficients. These measurements were made using the thermistors attached to the noise sources as mentioned before.

Since temperatures were measured by the same thermistor and amplifier circuits in all cases, the coefficients will be referred to in terms of volts change in noise source after amplification divided by volts change out of the thermistor amplifier. For the five T noise sources the temperature was modulated by changing the resistances in the resistor-thermistor bridge of the oven temperature control network. The temperature was actually varied over a range of about 1.7°C (10 volts). Time was allowed for the noise sources to reach an equilibrium temperature before taking the data. The voltage measurements of the noise sources were made by attempting to average by eye the noise over a period of several seconds (See Fig. 5.26).

We are assuming that the actual noise voltage which we can measure is

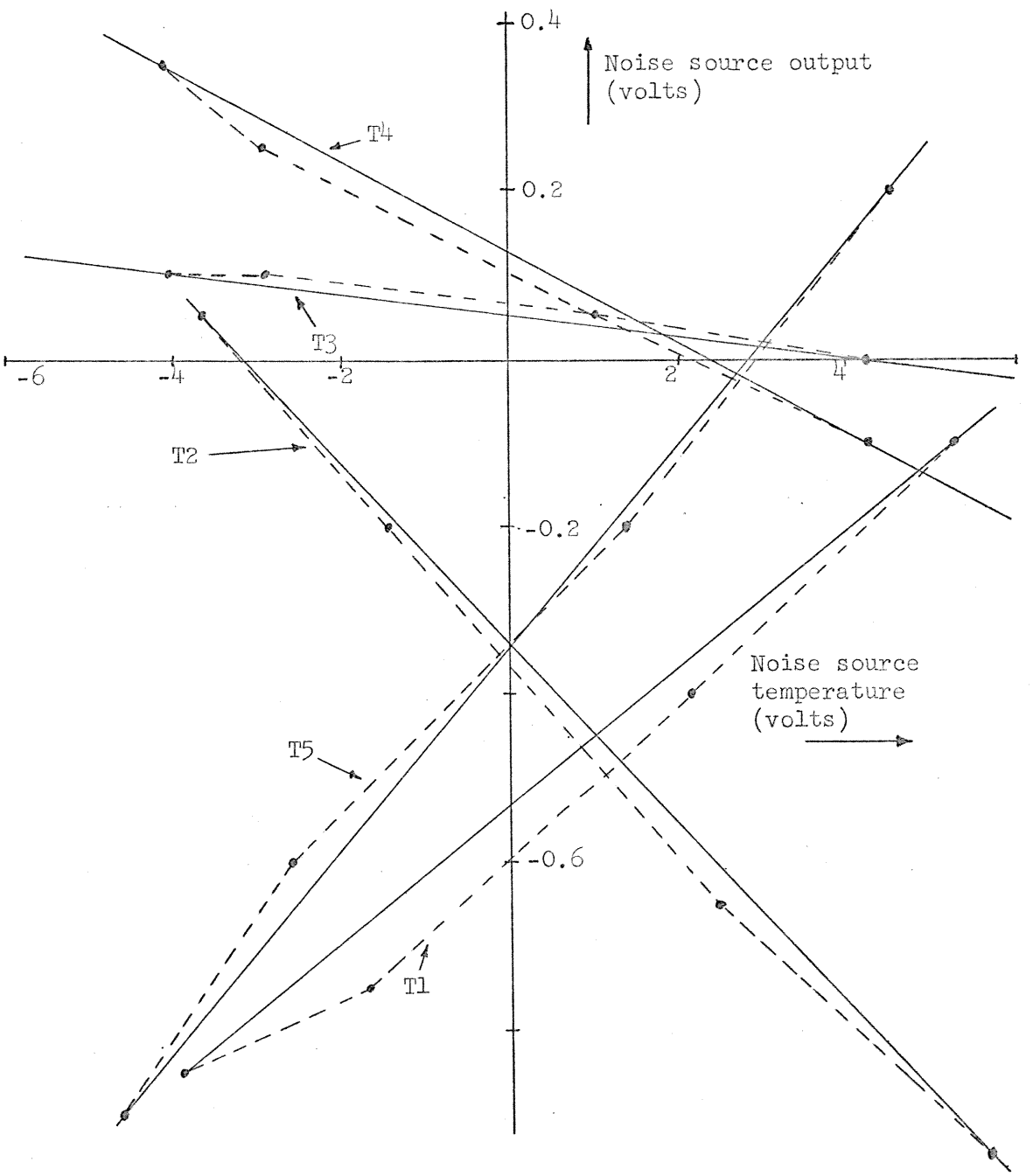


Figure 5.26 Noise source temperature versus noise source output voltage for noise sources T1, T2, T3, T4, and T5.

$$N(t) = n(t) + k\,T(t)$$

where $n(t)$ is the actual noise process which we would like to measure and $T(t)$ is the temperature which we will consider known (in this case $T(t)$ is in units of volts rather than degrees). We wish to measure the temperature coefficient k. The estimate of k will be denoted by $\hat{k}$. Since the total test period was on the order of a couple of hours, the drift of the noise sources must be recognized as an error. In order to estimate the temperature coefficient we are calculating

$$\hat{k} = \frac{(n(t_0) + kT(t_0)) - (n(t_1) - kT(t_1))}{T(t_0) - T(t_1)}$$

$$\hat{k} = k + \frac{n(t_0)-n(t_1)}{T(t_0) - T(t_1)}$$

$$= k + \frac{x(t_0,t_1)}{T_0 - T_1}$$

where $x(t_0,t_1) = n(t_0)-n(t_1)$ is a random variable. If we assume that $n(t)$ is a zero mean stationary process then

$$E\{\hat{k}\} = k + \frac{E\{n(t_0)\} - E\{n(t_1)\}}{T(t_0) - T(t_1)} = k$$

$$\sigma^2_{\hat{k}} = E\{(\hat{k}-k)^2\} = \frac{1}{[T(t_0)-T(t_1)]^2} E\{[n(t_0)-n(t_1)]^2\}$$

$$= \frac{1}{[T(t_0)-T(t_1)]^2} E\{n^2(t_0)-2n(t_0)n(t_1) + n^2(t_1)\}$$

$$= \frac{2}{[T(t_0)-T(t_1)]^2} \left[E\{n^2(t)\} - E\{n(t)n(t+\Delta t)\}\right]$$

where $\Delta t = t_1 - t_0$

$$= \frac{2}{[T(t_0)-T(t_1)]^2} [R_n(0) - R_n(\Delta t)]$$

If we assume that $S_n(f) = \begin{cases} \frac{A}{f} & |f| \leq f_{max} \\ 0 & |f| > f_{max} \end{cases}$ then

$$\sigma^2_{\hat{k}} = \frac{2}{[T(t_0)-T(t_1)]^2} \; 2\int_0^{f_{max}} \frac{A}{f} (1-\cos(2\pi f \Delta t))df \qquad \text{[Eq. 5.5.1]}$$

$$= \frac{4A}{[T(t_0)-T(t_1)]^2} [C + \ell n(2\pi f_{max}\Delta t) - ci(2\pi f_{max}\Delta t)]$$

[by Eq. 8.230.1 of Ref. 17]. f_{max} is determined by the averaging time of the eye and $\frac{1}{f_{max}}$ = 10 sec seems reasonable. Using f_{max} = .1 Hz, and Δt = 1 hour we get

$$\sigma^2_{\hat{k}} \cong \frac{4A}{[T(t_0)-T(t_1)]^2}\,[0.577 + 7.72 + 0] = \frac{4A}{[T(t_0)-T(t_1)]^2}\,[8.297]$$

$$\sigma_{\hat{k}} \sim \frac{5.76\sqrt{A}}{|T(t_0)-T(t_1)|} \qquad \text{where} \quad A = S_n(f = 1\ \text{Hz})$$

Although most of the $\hat{S}_n(f)$ for the 6 noise sources are not very good approximations to A/f, Eq. 5.5.1 implies that $S_n(f)$ for $f << 10^{-3}$ is relatively unimportant for differences taken over the interval of an hour. For the purposes of this estimate, it will be assumed that $S_n(f) = 10^{-4}/f$ volt^2-seconds for all the noise sources. This number will not be justified for all sources except to point out that it looks particularly good for noise sources T4 and T5. This yields

$$\sigma_{\hat{k}} \sim .006 \quad .$$

If n(t) is a gaussian process then we have a 95% probability that k lies in the region $\hat{k} \pm 2\sigma_{\hat{k}} = \hat{k} \pm .012$. Using Cauchy's inequality we have (without the gaussian assumption) a 75% probability of lying in this region.

Another way of looking at this error is the following. If we were somehow able to draw the straight line representing the true temperature coefficient, k, through the first datum point, then changed the temperature and waited an hour to obtain a second datum point, the noise voltage of the second datum point would be a random variable with a standard deviation of $\sigma_{\hat{k}}|T(t_0)-T(t_1)| \sim 0.06$ volts. Noise sources T1 and T5 exhibit drifts slightly larger than predicted. The others appear to be within the standard deviation estimated if one attempts to draw

the optimum straight line. It might be noted that the measurements were not made in the sequence of lowest to highest temperature but rather from next to lowest temperature in sequence to highest temperature and then back to the lowest temperature. This fact could prove useful if one wished to consider the possibility of linear trends in the noise source voltage, or the fact that sufficient time may not have been allowed for the temperature of the noise source to reach equilibrium. The following estimates result from fitting a straight line to the highest and lowest temperature points with an error estimate of $\pm 2\sigma_{\hat{k}}$ predicted above.

Source	Temperature coefficient (volts/volt)
T1	+.082 ± .012
T2	-.108 ± .012
T3	-.012 ± .012
T4	-.054 ± .012
T5	+.121 ± .012

Table 5.5

As a point of interest, a temperature coefficient of .12 volts/volt at the output is equivalent to 7.2 μ volts/°C referred to the input of the operational amplifier as compared to 10 μV/°C and 7 μV/°C typical values published by the manufacturer for the μA709 and μA747 respectively.[18, page 5]

For noise source #2 the temperature coefficient was estimated by recording noise source and temperature voltages during warm up from the intial turn on of the system (Fig. 5.27). Here again the measurements are over a temperature range of about 1.7°C with the data points spread over about $\frac{1}{2}$ hour period of time. Approximating a straight line to these data gives a temperature coefficient of approximately +0.1.

The question now arises whether by performing a cross correlation in the digital computer between the noise data and the temperature data for each source we could measure the temperature coefficient to a greater accuracy (or at least verify these measurements already taken).

The situation, therefore, is that we assume that the actual noise measurements $N(j\Delta t)$ (where Δt is the sampling period) are in fact the sum of the actual noise process we would like to measure and a temperature dependent part where the temperature is known but not the appropriate constant k.

Assume that $N(j\Delta t) = n(j\Delta t) + kT(j\Delta t)$ where $n(t)$ is a zero mean stationary random process and $T(t)$ has zero mean. We define our estimate of k as

$$\hat{k} = \sum_{j=1}^{M} N(j\Delta t)T(j\Delta t) / \sum_{j=1}^{M} T^2(j\Delta t)$$

$$\hat{k} = k + \left(\sum_{j=1}^{M} n(j\Delta t)T(j\Delta t) / \sum_{j=1}^{M} T^2(j\Delta t)\right)$$

If we assume now that $T(j\Delta t)$ is a known function then

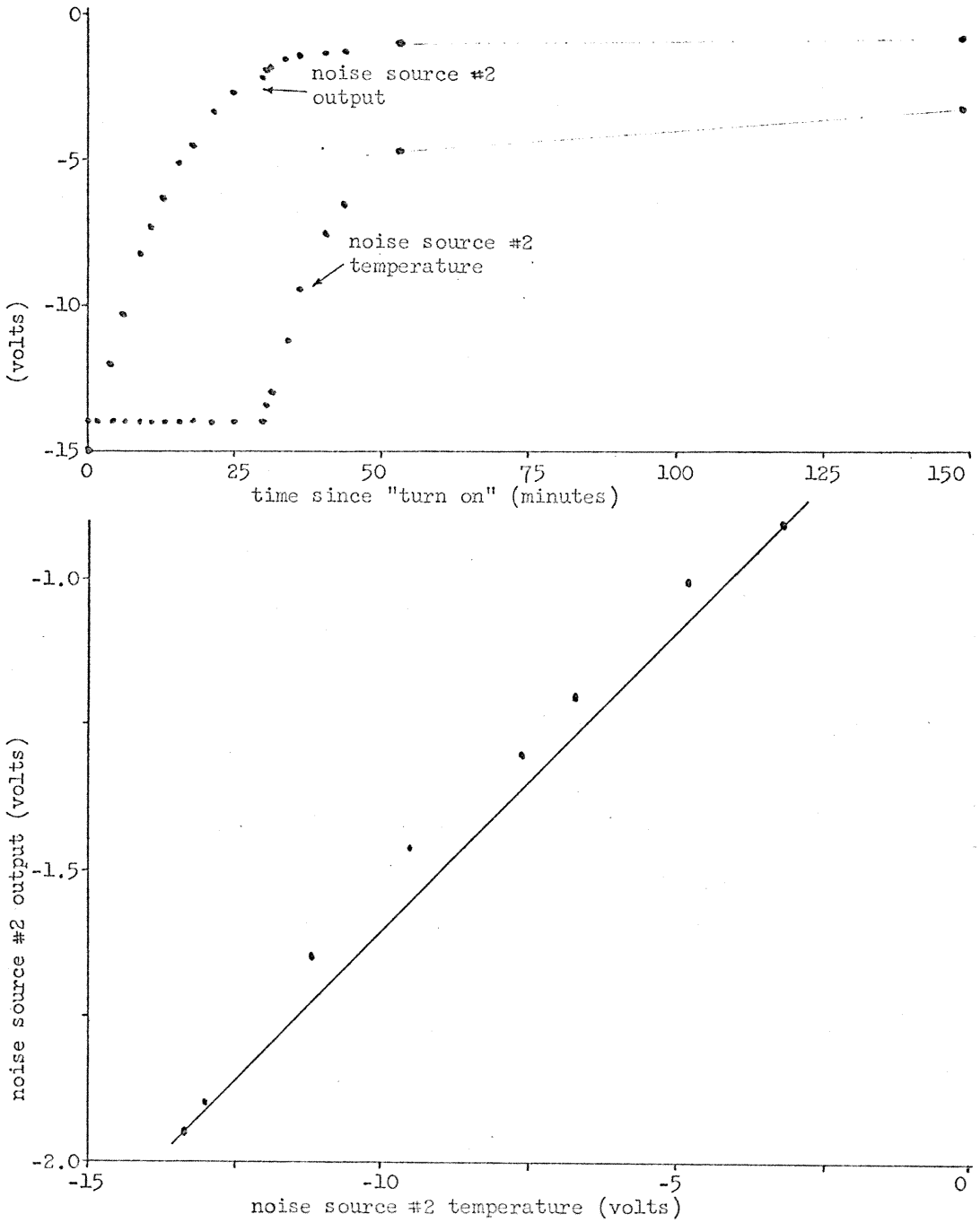


Figure 5.27 Noise source temperature versus noise source output voltage for noise source #2 during oven warm up.

$$E\{\hat{k}\} = k + \sum_{j=1}^{M} E\{n(j\Delta t)\}T(j\Delta t) \Big/ \sum_{j=1}^{M} T^2(j\Delta t)$$

$$= k + 0 \text{ (by our assumption that } n(t) \text{ is a zero mean, stationary process)}$$

$$= k.$$

The estimator, therefore, has a mean value of k and a parameter of interest is the variance of the estimate.

$$\sigma^2_{\hat{k}} = E\{(\hat{k}-k)^2\} = E\left\{ \frac{\left[\sum_{j=1}^{M} n(j\Delta t)T(j\Delta t)\right]\left[\sum_{i=1}^{M} n(i\Delta t)T(i\Delta t)\right]}{\left[\sum_{j=1}^{M} T^2(j\Delta t)\right]^2} \right\}$$

$$= \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} E\{n(i\Delta t)n(j\Delta t)\}\, T(i\Delta t)T(j\Delta t)}{\left[\sum_{j=1}^{M} T^2(j\Delta t)\right]^2}$$

To simplify the calculation at this point, assume that n(t) is band limited white noise (a condition which we may approximate by prewhitening), bandlimited to $f = \frac{1}{2\Delta t}$. In this case

$$E\{n(i\Delta t)n(j\Delta t)\} = \delta_{ij} \langle n(t)^2 \rangle$$

$$\sigma^2_{\hat{k}} = \langle n^2(t)\rangle \sum_{j=1}^{M} T^2(j\Delta t) \Big/ \left[\sum_{j=1}^{M} T^2(j\Delta t)\right]^2$$

$$= \langle n^2(t)\rangle \Big/ \sum_{j=1}^{M} T^2(j\Delta t) = \frac{S_n(f)}{\Delta t \sum_{j=1}^{M} T^2(j\Delta t)} \qquad \text{[Eq. 5.5.2]}$$

For a rough estimate of the sort of accuracy we can expect in measuring k by this method, consider the following. Looking at the estimated power spectral density of the temperatures of the various noise sources (see Fig. 5.14,...,5.19) it appears that most of the power is concentrated near the frequency f_d corresponding to a one day period (i.e., 10 log f_d = -49.4). It seems reasonable that if we prewhiten both N(t) and T(t) by a filter having a transfer characteristic of $|H(f)|^2 \propto f$ with unity gain at frequency f_d we will have n'(t) and T'(t) such that

$$\sum_{j=1}^{M} T'(j\Delta t)^2 \sim \sum_{j=1}^{M} T(j\Delta t)^2 \qquad \text{(if } S_T(f) \text{ is narrow band about } f = f_d \text{)}$$

and $\langle n'(t)^2\rangle \sim S_n(f_d) \times \frac{1}{\Delta t}$ if we prewhiten properly.

Therefore

$$\sigma^2_{\hat{k}} \sim \frac{S_n(f_d)/\Delta t}{M\langle T(t)^2\rangle} \quad .$$

If we use the total length of our record (L seconds) then $M\Delta t = L$ and

$$\sigma^2_{\hat{k}} \sim \frac{S_n(f_d)}{L\langle T(t)^2\rangle} \quad . \qquad \text{[Eq. 5.5.3]}$$

Therefore we do as well estimating k at the lowest sampling rate as at the highest if we use a proper dealiasing filter each time the sample rate is reduced. We shall therefore make estimates of k at $\Delta t = 10^4$ sec using the ~800 samples we have at this rate since it reduces the amount of arithmetic and simplifies the prewhitening filter.

If we estimate $\langle T(t)^2\rangle$ to be ~.012 (the value calculated from data after prewhitening) and $S(f_d) = 15.85$ v^2-sec from the T5 power spectral density estimate we get

$$\sigma^2_{\hat{k}} \sim \frac{15.85}{8 \times 10^6 \times .012} = 1.65 \times 10^{-4}$$

$$\sigma_{\hat{k}} \sim .013 \quad \text{or about a} \quad \pm 13\% \text{ estimate when } k = 0.1$$

Another way of looking at this is as follows. $\hat{S}_T(f)$ has a peak value of $\sim 1.99 \times 10^3$ volt^2-sec at frequency f_d. We have already determined that $\sigma^2_{\hat{k}} \sim \frac{S_n(f_d)}{8 \times 10^6 \times 0.012}$. If we make an estimate for k we should not be surprised to get $\hat{k} = k + \sigma_{\hat{k}}$, giving us an estimate of $n(t)$ as $\hat{n}(t) = N(t) - \hat{k}T(t) = n(t) - \sigma_{\hat{k}}T(t)$. If $n(t)$ and $T(t)$ are uncorrelated we will have

$$\hat{S_n}(f_d) = S_n(f_d) + \sigma^2_{\hat{k}}\, S_T(f_d)$$

$$= S_n(f_d)\left(1 + \frac{1.66 \times 10^3}{8 \times 10^6 \times 0.012}\right)$$

$$10 \log \hat{S_n}(f_d) = 10 \log \hat{S_n}(f_d) + 0.075.$$

That is, if we have an estimate of k which gives less than a 0.1 db "bump" in the noise power spectral density due to temperature, then we have probably done about as well as we can do. In other words, if we cannot see a "bump" in the noise spectral density estimate, we probably cannot make a meaningful estimate of the temperature coefficient.

It would be interesting to find out more about the effects of not prewhitening in this estimator for k. In order to do this I will make the artificial assumption that T(t) is also a zero mean, stationary, random process and that n(t) and N(t) are statistically independent. Although I'm quite sure the reader will not agree that T(t) is likely to be a stationary process, if we assume that it is a very narrow band process it would be difficult to prove or disprove stationarity over ~100 cycles at the center frequency.

Again we have

$$\hat{k} = k + \sum_{j=1}^{M} n(j\Delta t)T(j\Delta t) \Bigg/ \sum_{j=1}^{M} T^2(j\Delta t)$$

$\hat{k} = k + x/y$ where x and y are now random variables

$$E\{x\} = E\left\{\sum_{j=1}^{M} n(j\Delta t)T(j\Delta t)\right\}$$

$$= \sum_{j=1}^{M} E\{n(j\Delta t)\}\ E\{T(j\Delta t)\}$$

by assumption of stationarity and statistical independence

$$E\{x\} = 0$$

$$\sigma^2_x = E\{x^2\} = E\left\{\sum_{i=1}^{M} n(i\Delta t)\ T(i\Delta t) \sum_{j=1}^{M} n(j\Delta t)T(j\Delta t)\right\}$$

$$= \sum_{i=1}^{M}\ \sum_{j=1}^{M} E\{n(i\Delta t)n(j\Delta t)\}\ E\{T(i\Delta t)T(j\Delta t)\}$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{M} R_n\{(i-j)\Delta t\}\ R_T\{(i-j)\Delta t\}$$

$$= \sum_{i=-(M-1)}^{M-1} (M-|i|)R_n(i\Delta t)\ R_T(i\Delta t)$$

Now by using the convolution theorems of the Fourier transform and by denoting convolution as

$$f_1(t)*f_2(t) = \int_{-\infty}^{\infty} f_1(\nu)f_2(t-\nu)d\nu$$

and noting that

$$S_n(f) = \mathcal{F}\left\{R_n(t)\right\}$$

$$S_T(f) = \mathcal{F}\left\{R_T(t)\right\}$$

we get

$$E\left\{x^2\right\} = \sum_{i=-\infty}^{\infty} \mathcal{F}^{-1}\left\{M^2\Delta t \, \frac{\sin^2(\pi f M\Delta t)}{(\pi f M\Delta t)^2} * S_n(f) * S_T(f)\right\}\Bigg|_{t = i\Delta t}$$

$$\sim \left[M^2 \, \frac{\sin^2(\pi f M\Delta t)}{(\pi f M\Delta t)^2} * S_n(f) * S_T(f)\right]_{f = 0} \qquad \text{[Eq. 5.5.4]}$$

where the approximation in Eq. 5.5.4 is due to aliasing if the quantity in brackets is not bandlimited to $f = \frac{1}{2\Delta t}$. The approximation will be good if $M >> 1$ and if $S_n(f) = S_T(f) = 0$ for $|f| > \frac{1}{2\Delta t}$.

$$E\left\{x^2\right\} \sim M^2 \, \frac{1}{M\Delta t}\left[M\Delta t \, \frac{\sin^2(\pi f M\Delta t)}{(\pi f M\Delta t)^2} * S_T(f) * S_n(f)\right]_{f = 0}$$

but as $M \to \infty$, $M\Delta t \, \frac{\sin^2 \pi f M\Delta t}{(\pi f M\Delta t)^2} \to \delta(f)$, so that as $M \to \infty$,

$$E\left\{x^2\right\} \to \frac{M}{\Delta t}\left[S_n(f) * S_T(f)\right]_{f = 0}$$

and for large M

$$E\left\{x^2\right\} \sim \frac{M}{\Delta t} \int_{-\infty}^{\infty} S_n(f)\, S_T(f)\, df \quad .$$

If we assume that $\sum_{i=1}^{M} T^2(j\Delta t) = E\left\{ \sum_{i=1}^{M} T^2(j\Delta t) \right\}$ then it seems reasonable that

$$\sigma^2_{\hat{k}} \sim E\left\{x^2\right\} \Big/ y^2$$

$$\sim \frac{M}{\Delta t} \int_{-\infty}^{\infty} S_n(f)\, S_T(f) df \Big/ M^2 \left(\int_{-\infty}^{\infty} S_T(f) df \right)^2$$

$$\sim \frac{1}{M\Delta t} \frac{\int_{-\infty}^{\infty} S_n(f) S_T(f) df}{\left(\int_{-\infty}^{\infty} S_T(f) df \right)^2}$$

$$\sim \frac{1}{M\Delta t} \frac{\int_{-\infty}^{\infty} S_n(f)\, S_T(f)\, df}{\langle T^2(t) \rangle^2} \qquad \text{[Eq. 5.5.5]}$$

Now it becomes clear that if $S_n(f)$ is "reasonably" smooth and $S_T(f)$ has a narrow bandwidth about $f = f_d$ relative to the smoothness of $S_n(f)$, we will have

$$\sigma^2_{\hat{k}} \sim \frac{1}{M\Delta t} \frac{S_n(f_d) \langle T^2(t) \rangle}{\langle T^2(t) \rangle^2} = \frac{1}{M\Delta t} \frac{S_n(f_d)}{\langle T^2(t) \rangle} \qquad \text{[Eq. 5.5.6]}$$

which is just the answer we got before for white noise. Now, however, using Eq. 5.5.5 we can see what happens if $S_n(f)$ is no longer white. As long as $S_n(f)$ doesn't blow up anywhere (e.g., is not much larger

than $S_n(f_d)$ anywhere) and is reasonably smooth near $f = f_d$, the accuracy of the estimate $\hat{k}$ is dependent only on $S_n(f_d)$. If, however, we have $\int_{-\varepsilon}^{\varepsilon} S_n(f)df \to \infty$ for small ε as in 1/f noise and if $S_T(f = 0) \neq 0$ (no matter how small) we will have $\sigma_{\hat{k}}^2 \to \infty$. In the actual experiment, of course, we have removed the sample mean from N(t) which would keep $\sigma_{\hat{k}}^2$ from actually going to ∞, nevertheless it is clearly desirable to prewhiten in this case.

The estimates were made on the raw data (i.e.,before any attempt was made at removing popcorns and trends). In addition $\frac{1}{M}\sum_{i=1}^{M} T^2(i\Delta t)$ was calculated from the data and an estimate was made for $S_n(f_d)$ from the power spectral density estimates (attempting to ignore any "bump" in the spectral density due to temperature). These were used in estimating $\sigma_{\hat{k}}$ for each source. The estimates are assumed to be $\pm 2\sigma_{\hat{k}}$ to yield at least a 75% probability of being within the specified range even under the (likely) conditions that the noise sources are not gaussian.

The errors in all cases were estimated using Eq. 5.5.6. As was shown, lack of prewhitening can significantly increase the errors above these estimates.

Comparing the original estimates and the estimates made from prewhitened data we see that the error tolerances are not sufficient to explain the differences. In the case of noise source #2 we see (by

noise source	orig. estimate	estimate from data* without prewhitening		estimate from data with prewhitening
#2	+.103 ± .012	+.022	(±.0069)	-.034 ± .03
T1	+.082 ± .012	+.658	(±.027)	+.058 ± .028
T2	-.108 ± .012	-.372	(±.041)	-.163 ± .043
T3	-.012 ± .012	-.058	(±.043)	-.101 ± .045
T4	-.054 ± .012	-.142	(±.053)	-.081 ± .057
T5	+.121 ± .012	+.183	(±.025)	+.133 ± .026

Table 5.6

*The error tolerances in this case ignore the fact that $S_n(f)$ blows up near f = 0. These error tolerances are not, therefore, to be taken seriously. They only serve to demonstrate the evil of using Eq. 5.5.6 blindly.

looking at the raw data) that the temperature has two very large step functions in it. If this is due to some error in the temperature measuring device rather than changes in the temperature of the noise source, the assumption of $N(t) = n(t) + kT(t)$, where T is known, is no longer valid. Rather, $N(t) = n(t) + kT(t)$, $T'(t) = T(t) + s(t)$ where $T'(t)$ rather than $T(t)$ is known and our estimator would yield

$$\hat{k} = \frac{\sum N(t)\, T'(t)}{\sum (T'(t))^2} = \frac{\sum (n(t)T(t) + n(t)s(t) + kT^2(t) + kT(t)s(t))}{\sum (T^2(t) + s^2(t) + s(t)\, T(t))}.$$

Under appropriate assumptions we might expect

$$\hat{k} \sim k \frac{\sum T^2(t)}{\sum [T^2(t) + s^2(t)]}$$

giving us a biased estimate tending to be smaller in magnitude than k with a variance which would be much more difficult to calculate. In addition to this, noise sources #2, T3 and T4 have obvious popcorning which contribute to the power spectral density of each noise source. Removing these step functions should, therefore, decrease the variance of the estimator.

5.6 Effects of Temperature on Power Spectral Density Estimates

We now have power spectral density estimates of both temperature, $T(t)$, and the raw noise, $N(t)$, where we assume that $N(t) = n(t) + kT(t)$. We also have estimates for the temperature coefficients, $\hat{k}$, for each of the noise sources. These temperature coefficients are not as good as we might have hoped for and it seems appropriate to investigate the effects of an imperfect estimate of a temperature coefficient (or the effects of not correcting for temperature for that matter) on the final estimate $\hat{S}_n(f)$. If $n(t)$ and $T(t)$ are uncorrelated random processes, then we know that $S_N(f) = S_n(f) + k^2 S_T(f)$. Therefore $E\{\hat{S}_n(f)\} = E\{\hat{S}_N(f)\} - k^2 E\{\hat{S}_T(f)\}$. If $k^2\hat{S}_T(f) << S_n(f)$ for all but a few narrow intervals of frequency, then we may be able to make reasonable estimates of $S_n(f)$ by extrapolating the values of $\hat{S}_N(f)$ outside these intervals of frequency to the frequencies inside these intervals (this just involves the assumption that $S_n(f)$ is reasonably smooth).

To demonstrate this, we will look at noise source T5 in detail. This source is chosen because it shows the greatest effect of temperature in its spectral estimates (see Figs. 5.8,...,5.13), a consistent set of temperature coefficient estimates was obtained for this source (see Table 5.6), and there is no reason to suspect that the temperature measurements themselves may be inaccurate (as in the case of the #2 noise source temperature measurements).

Figure 5.28 shows the power spectral density estimates of noise source T5 along with the estimates for T5 temperature multiplied by 0.12 (the original $\hat{k}$ estimate from Table 5.6). The small "+" on

the graph shows my attempt to extrapolate $\hat{S}_N(f)$ to obtain $\hat{S}_n(f)$ at frequency f_d (f_d = 1 cycle/day = $\frac{1}{60 \cdot 60 \cdot 24}$ Hz, 10 log f_d = -49.37). The power spectral density estimate $(0.12)^2 S_T(f)$ peaks at about 2 db above the "+". Since spectral densities are supposed to add, we would predict the power spectral density $\hat{S}_N(f)$ to peak at 10 log$(1 + 10^{0.2})$ = 4.1 db above the "+". The value which we get, however, is 6.2 db above the "+". There are several possible reasons for this apparent discrepancy. We may not have extrapolated the true power spectral density $S_n(f)$ sufficiently well, the variances of the $\hat{S}_T(f_d)$ and $\hat{S}_N(f_d)$ estimates are non-zero, and we may not have chosen the proper value for $\hat{k}$. If we assume that only our estimate of k is wrong we may work this problem in reverse and ask what value of $\hat{k}$ does Fig. 5.28 predict. If we assume that k is in fact 0.12α then we solve 10 log$(1 + \alpha^2 10^{-.2})$ = 6.2 db which yields $|\alpha|$ = 1.4 or a new value for $\hat{k}$ of $|\hat{k}|$ = 0.17. This is not in good agreement with the two values of $\hat{k}$ from Table 5.6 which were considered to be valid (0.121 ± 0.12 and 0.133 ± .026). The major error may come from the inaccuracies of the spectral estimates $\hat{S}_T(f)$ and $\hat{S}_N(f)$. The high resolution spectral esimtates which are used here were made using 800 sample points with M = 100 lags. For white noise we remember that this yields a variance of the estimator of $\sigma^2 \sim \frac{100}{800} S^2(f)$ for white noise. If we blindly use this estimate of variance for these sources which have sharp peaks in their spectral densities (no attempt will be made to verify the accuracy of this assumption), a spectral estimate which is one standard deviation away

from the true value would either be 1.3 db high or 1.9 db low. If we assume that the peak in $\hat{S}_T(f)$ at f_d (Fig. 5.28) is σ too low, we would predict the peak in the noise spectral density, $\hat{S}_N(f_d)$, to be $10 \log (1 + 10^{.39}) = 5.4$ db above the "+",requiring (if k = 0.12) that $\hat{S}_N(f_d)$ was actually 0.8 db too high (an error of less than the predicted σ for the estimate $\hat{S}_N(f_d)$).

Figure 5.29 shows the spectral density estimates of noise source T5 before and after 0.12 x T(t) was subtracted (that is $\hat{S}_N(f)$ and $\hat{S}_{\hat{n}}(f)$ where $\hat{n}(t) = N(t) - 0.12\ T(t)$). Notice that the estimates at each end of the frequency scale are nearly identical (higher frequency estimates are not included because of the relatively low values of $S_T(f)$ at higher frequencies). The presence of an additive temperature coefficient appreciably affects the estimate only in those regions where the temperature (multiplied by the temperature coefficient, k) is less than 10 db below the raw noise spectral density as was expected (that is, $\hat{S}_N(f) \sim \hat{S}_{\hat{n}}(f)$ as long as $(\hat{k})^2 \hat{S}_T(f) < \frac{1}{10}\ \hat{S}_N(f)$).

This is an important point because it implies that one may make a meaningful estimate (biased by less than 1 db) without knowing the exact temperature coefficient (or even accurate measurements of the temperature as we suspect may be the case for noise source #2) if it can be shown that $(k - \hat{k})^2\ \hat{S}_T(f) < 0.1\ \hat{S}_N(f)$ over the ranges of frequency for which we wish to make estimtes. In fact, if $k^2\ \hat{S}_T(f) < 0.1\ \hat{S}_N(f)$ we need not bother to correct for temperature and may reasonably assume that $\hat{S}_N(f) = \hat{S}_n(f)$.

Using Fig. 5.29, another example is offered to test the hypothesis

that $\hat{S}_N(f) \sim \hat{S}_{\hat{n}}(f) + (\hat{k})^2\, \hat{S}_T(f)$; this time using the lower resolution estimates at $10 \log f_1 = -50$, $10 \log \hat{S}_N(f_1) + 10 \log[(0.12)^2\, \hat{S}_T(f_1)] + 6.6$. We, therefore, expect that

$$\hat{S}_{\hat{n}}(f) + (0.12)^2\, \hat{S}_T(f) \sim \hat{S}_N(f)$$

$$\hat{S}_{\hat{n}}(f) + 10^{-.66}\, \hat{S}_N(f) \sim \hat{S}_N(f)$$

$$\hat{S}_{\hat{n}}(f) = 0.78\, \hat{S}_N(f)$$

$$10 \log \hat{S}_{\hat{n}}(f) = 10 \log \hat{S}_N(f) - 1.07 \text{ db} \quad .$$

Measuring this value from Fig. 5.29 we get $10 \log \hat{S}_{\hat{n}}(f) \sim 10 \log \hat{S}_N(f) - 1.2$. An error of about 0.1 db. In this case we were making spectral estimates using 8000 data points and M = 100 lags yielding a standard deviation for each of the spectral estimates of $\sim\sqrt{\frac{1}{80}}\, S(f)$ ($\sim \pm 0.5$ db).

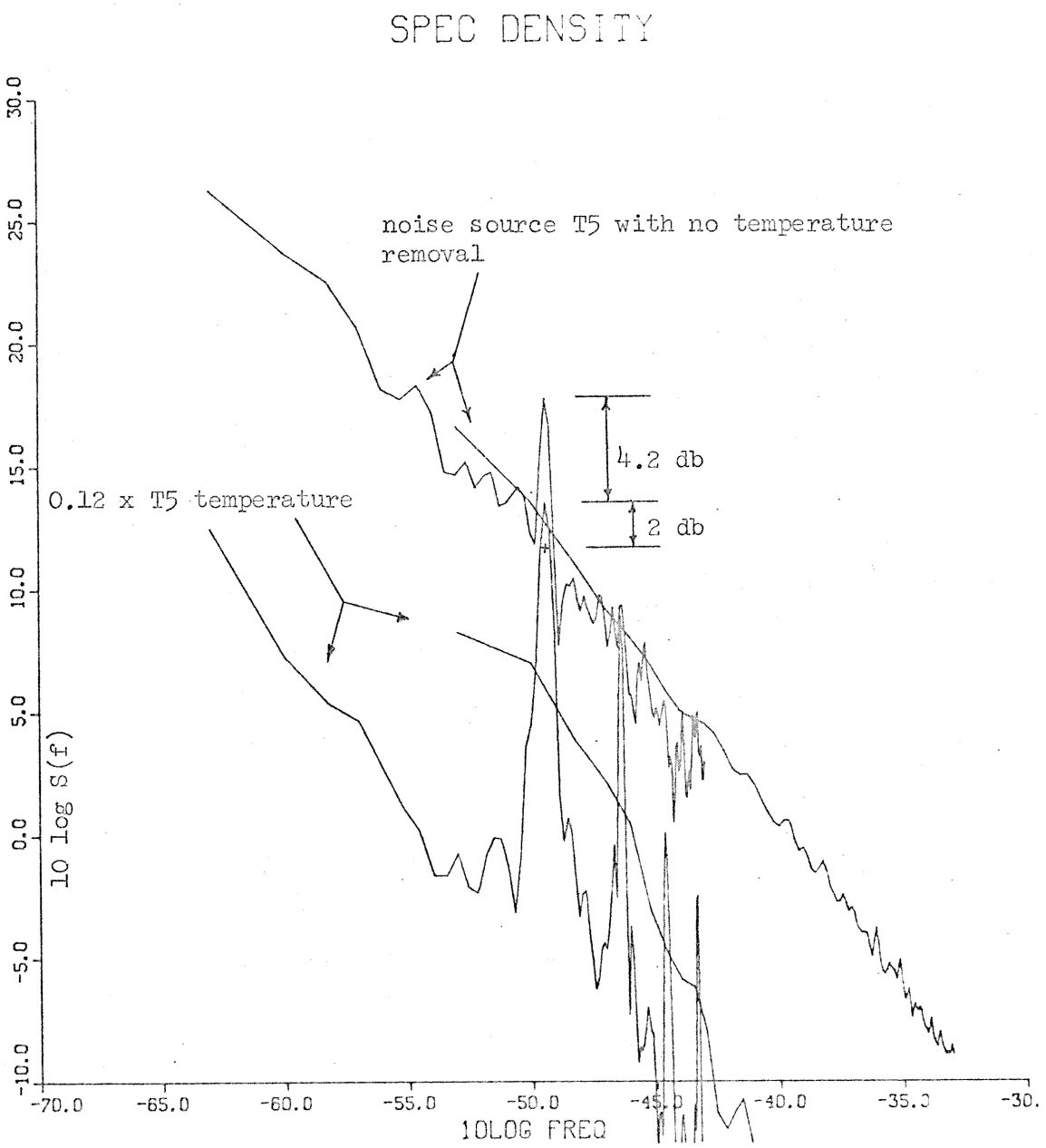


Figure 5.28 Spectral estimates of T5 noise source and 0.12 x T5 temperature ($-63 \leq 10 \log f \leq -33$).

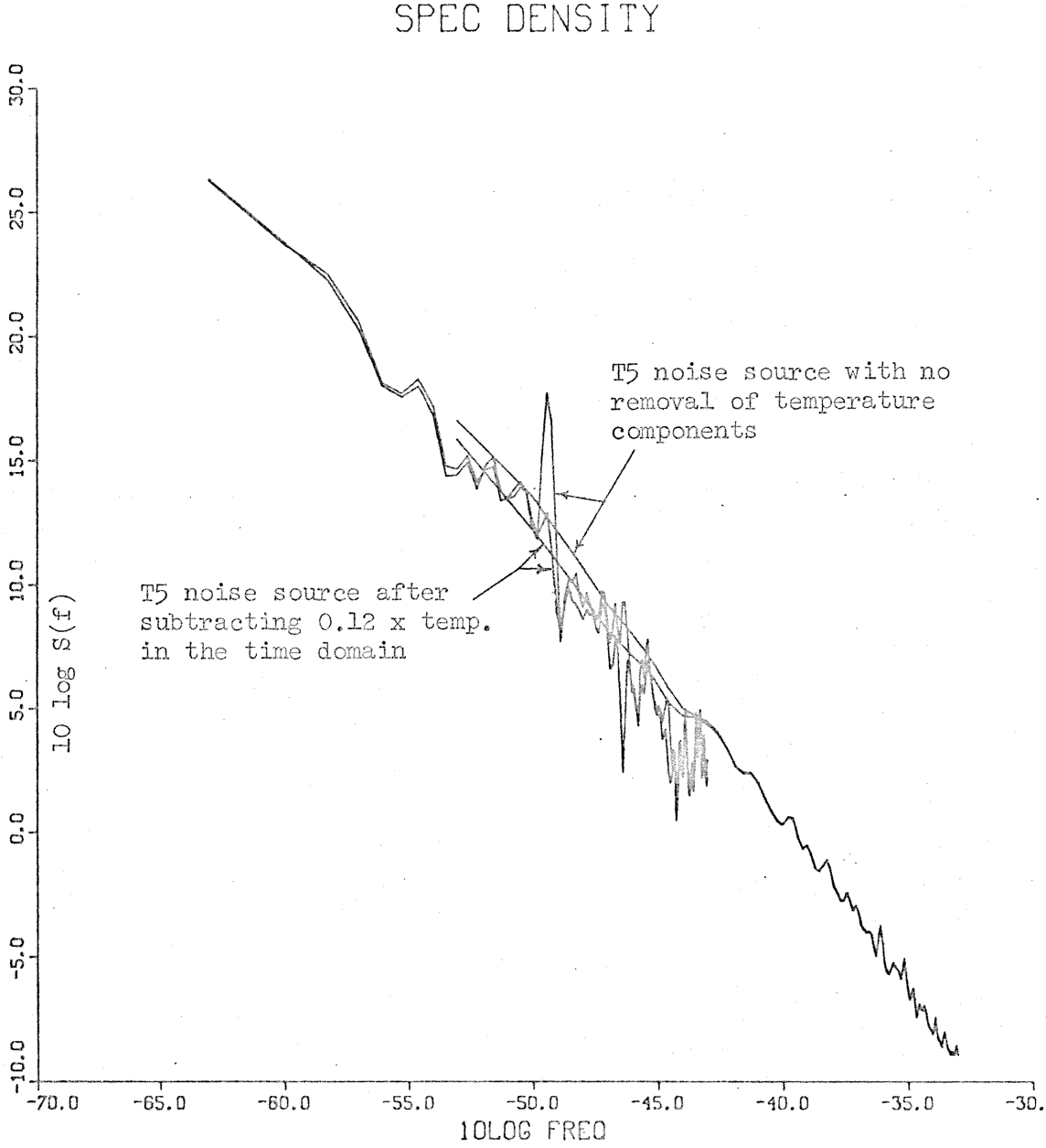


Figure 5.29 Spectral estimates of T5 noise source and T5 noise source minus 0.12 X T5 temperature.

5.7 Effects of Burst ("popcorn") Noise

Referring to Figs. 5.1,...,5.7 it can be seen that noise sources #2, T2, T3, and T4 (as well as V_{-check} and V_{+check}) exhibit a characteristic often referred to as burst noise (an apparent additive series of step functions containing only two distinct levels of amplitude). In some cases (such as noise source T4 in Fig. 5.2) there appear to be multiple burst components within a single noise source. The time periods of the bursts are apparently random variables and can be observed in the data from ~ 10 sec (times shorter than this cannot be resolved with a 10 second sampling period) for noise source T3 in Figs. 5.1 and 5.5 to an "off" time period of ~ 1.61×10^6 sec (18.6 days) for V_{-check} in Fig. 5.7 and an apparent "off" time of greater than (the original step down is presumed to occur before the start of the recording period) 6.1×10^6 sec (70.6 days) for noise source #2 in Fig. 5.7.

This process is not new. It was noted by Caloyannides[1] in his investigation of the 1/f phenomeonon. In his investigation he noted the existence of burst noise and attempted to eliminate it by choosing new noise sources until no burst noise was observed. It is not possible, however, to show the non-existence of burst components having time periods on the order of days with an observation time of a few hours. Caloyannides still had observable popcorning in his final experimental data which he proceeded to ignore with the following comment:

> "Removing them digitally would produce an artificial array of data, representing nothing physical; keeping the noise source as it is would steepen the overall final estimate;

> yet the latter can hardly be called "distortion" or "bias" since "popcorning" does exist in semiconductors and is as natural as flicker noise itself." [1, page 69]

The "naturalness" of popcorn noise does not seem to be the issue here. There is reason to believe that flicker noise may, in fact, be the sum of a large number of popcorn components within a single semiconductor device. It seems that a good step toward understanding the nature of the "pure" 1/f noise (if in fact it exists) is to remove all components contributing to the raw noise which can be identified as separate (even if similar) processes.

A statistical study of burst noise in bipolar semiconductor devices (such as the operational amplifiers used in this investigation) by Puckett[3] indicated that popcorn noise appeared to be a random telegraph wave with a waiting time probability density of "on" time for a given burst of $P_{on}(t) = \frac{1}{\tau_+} e^{-t/\tau_+}$ and a probability density for "off" time duration of $P_{off}(t) = \frac{1}{\tau_-} e^{-t/\tau_-}$ where in general $\tau_+ \neq \tau_-$. Puckett's investigation involved time constants on the order of 0.1 to 1 millisecond because of experimental convenience. He notes, however, that neither upper nor lower limits had been established for τ_+ or τ_-. Puckett shows both theoretically and experimentally that the power spectral density for such a noise source with a peak-to-peak amplitude of h and uncorrelated waiting times is given by[3, page 72]

$$S(f) = \frac{2h^2}{(\tau_+ + \tau_-)[(\frac{1}{\tau_+} + \frac{1}{\tau_-})^2 + (2\pi f)^2]} .$$

If we let $f_0 = \frac{1}{2\pi}\left(\frac{1}{\tau_+} + \frac{1}{\tau_-}\right)$, $\alpha = \tau_-/\tau_+$, and $k = \frac{\alpha h^2}{(1+\alpha)^2\pi}$

we have

$$S(f) = \frac{k}{f_0}\Big/\left[1 + \left(\frac{f}{f_0}\right)^2\right] \qquad \text{[Eq. 5.7.1]}$$

If we have a noise source, $N(t) = n_1(t) + n_2(t)$ which is composed of a flicker component, $S_{n_1}(f) = \frac{A}{|f|}$, and a burst component then the resulting power spectral density of the sum should be

$$S_n(f) = \frac{A}{|f|}\left[1 + \frac{(k/A)\,\frac{f}{f_0}}{\left[1 + \left(\frac{f}{f_0}\right)^2\right]}\right] \qquad \text{[Eq. 5.7.2]}$$

Figure 5.30 shows this function plotted for $f_0 = 10^{-3}$ Hz, A = 1, and various values of k. It is interesting to note that we may predict the peak value of 10 log$[S_N(f)]$ - 10 log $[S_{n_1}(f)]$ by knowing only h and τ_-/τ_+. Therefore, a popcorn component of a given amplitude h produces the same relative effects in the spectral density of $A/|f|$ noise whether the time constants are large or small except that the distortion will occur at different points on the log f axis.

As was mentioned earlier, burst noise characteristics are easily observable in the time domain data plots (Figs. 5.1,...,5.7) for noise sources #2, T2, T3, and T4. For future reference (and for the reader to compare with the spectral estimates in Figs. 5.8,...,5.12) Table 5.7 shows the very crude estimates of τ_+, τ_-, and h (as well as some other values implied by these parameters) which were made directly from these data plots. (Note that if $p(t) = \frac{1}{\tau} e^{-t/\tau}$ then

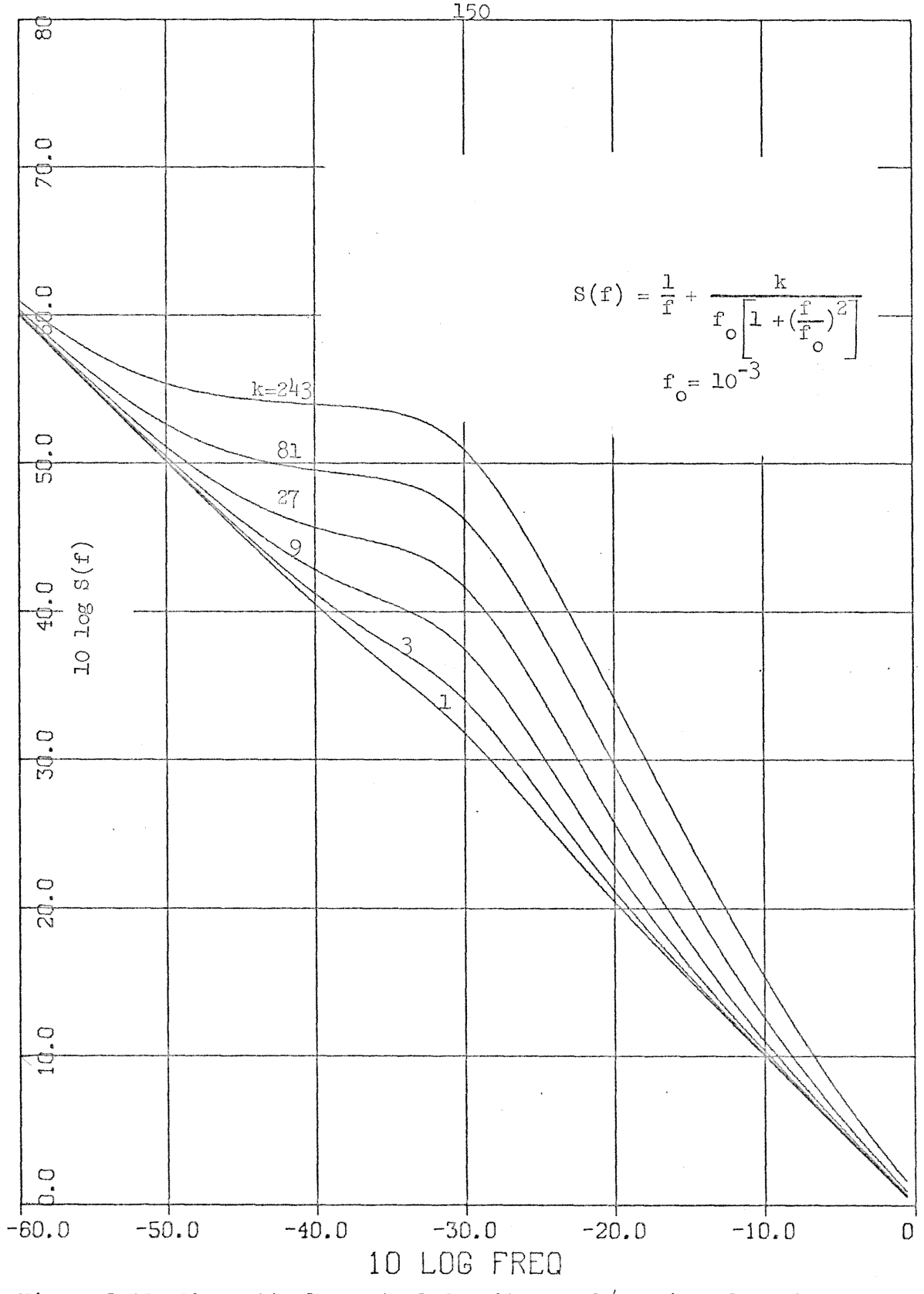


Figure 5.30 Theoretical spectral density for 1/f noise plus single burst component.

$$E\{t\} = \int_0^\infty \frac{t}{\tau} e^{-t/\tau}\, dt = \tau).$$

A reasonable method of estimating f_0 and k from the power spectral density estimate of a popcorn source would be to compare the spectral density to a A/f curve. If $S(f) = \frac{k}{f_0} \Big/ \left[1 + \left(\frac{f}{f_0}\right)^2\right]$ then the peak value of 10 log S(f) - 10 log $(\frac{A}{f})$ occurs at $f = f_0$ and is 0 if $A = \frac{k}{2}$. Estimates of 10 log f_0 and 10 log $\frac{k}{2}$ are included in Table 5.7.

Referring to Figs. 5.8, 5.10, 5.11, and 5.12 we see that noise sources #2 and T3 agree quite well with these estimates but that noise sources T2 and T4 do not. Figure 5.7 reveals that the single burst observed in Fig. 5.6 for source T2 was not representative (τ_ might be better estimated to be on the order of 5×10^7 sec). Figure 5.7 also reveals that in addition to the positive going bursts at the beginning of the record there are some negative going bursts later. These bursts are somewhat suspicious since they appear to limit at a particular value rather than being of uniform height. This suggests that they may actually be missed glitches. Noise source T2 (Fig. 5.7) also has a very large low frequency component observable in the time domain (not characteristic of the other sources) which could account for the fast rise in power spectral density at the lowest frequencies. This low frequency component cannot be easily identified as coming from burst noise although it is possible that it is the result of several small amplitude bursts with long, but approximately equal, time constants.

Figure 5.6 shows that T4 has at least two separate burst components (and probably more). This would account for the discrepancy observed

between Table 5.7 and Fig. 5.12. More will be said about this later.

Noise Source	Figure from which estimates made	ESTIMATES				
		h(volts)	τ_+(sec)	τ_-(sec)	10 log f_0	10 log $\frac{k}{2}$
#2	5.7	0.4	8.5×10^5	7×10^6	-66.8	-26.1
T2	5.6	0.35	2×10^4	5×10^5	-50.8	-31.4
T3	5.1	0.1	6×10^2	5×10^2	-32.3	-34.0
T3	5.7	0.25	1.2×10^5	5×10^5	-57.8	-28.1
T4	5.5	0.2	9×10^3	4×10^3	-42.4	-28.7

Table 5.7

5.8 Removal of Burst Noise Components

Noise source #2 and noise source T3 have burst components with sufficiently few steps that it was considered feasible to determine the amplitude and position of each step by hand. These two sources were plotted at $\Delta t = 10^3$ sec (~ 8000 pts) and the switching times and amplitudes estimated from this plot. For noise source #2 the burst amplitude was estimated to be approximately 0.39 volts with 2 positive going steps and 2 negative going steps. For noise source T3 the amplitude was estimated to be 0.22 volts with 13 positive and 13 negative going steps. The switching times were estimated at $\Delta t = 10^3$ sec because a shorter sample period would have involved a great deal more plotting and effort with little or no advantage.

At this time it was decided to attempt to remove an apparent trend in noise source T1. This source was assumed to have an exponential component of $C\, e^{-t/\tau}$ where C was estimated to be 1.6 volts and τ to be 2.7×10^6 sec. These estimated components were removed and the resulting data low pass filtered and resampled to $\Delta t = 10^4$ sec. These data are shown in Fig. 5.31 and may be compared with those in Figs. 5.7 and 5.4.

The power spectral density estimates (resulting from these modified data) are shown in Figs. 5.32, 5.33, and 5.34 along with the spectral density estimates made before the data were modified. As can be seen from these plots, the power spectral density estimates in all three of these cases were lowered at the low frequency end as had been expected.

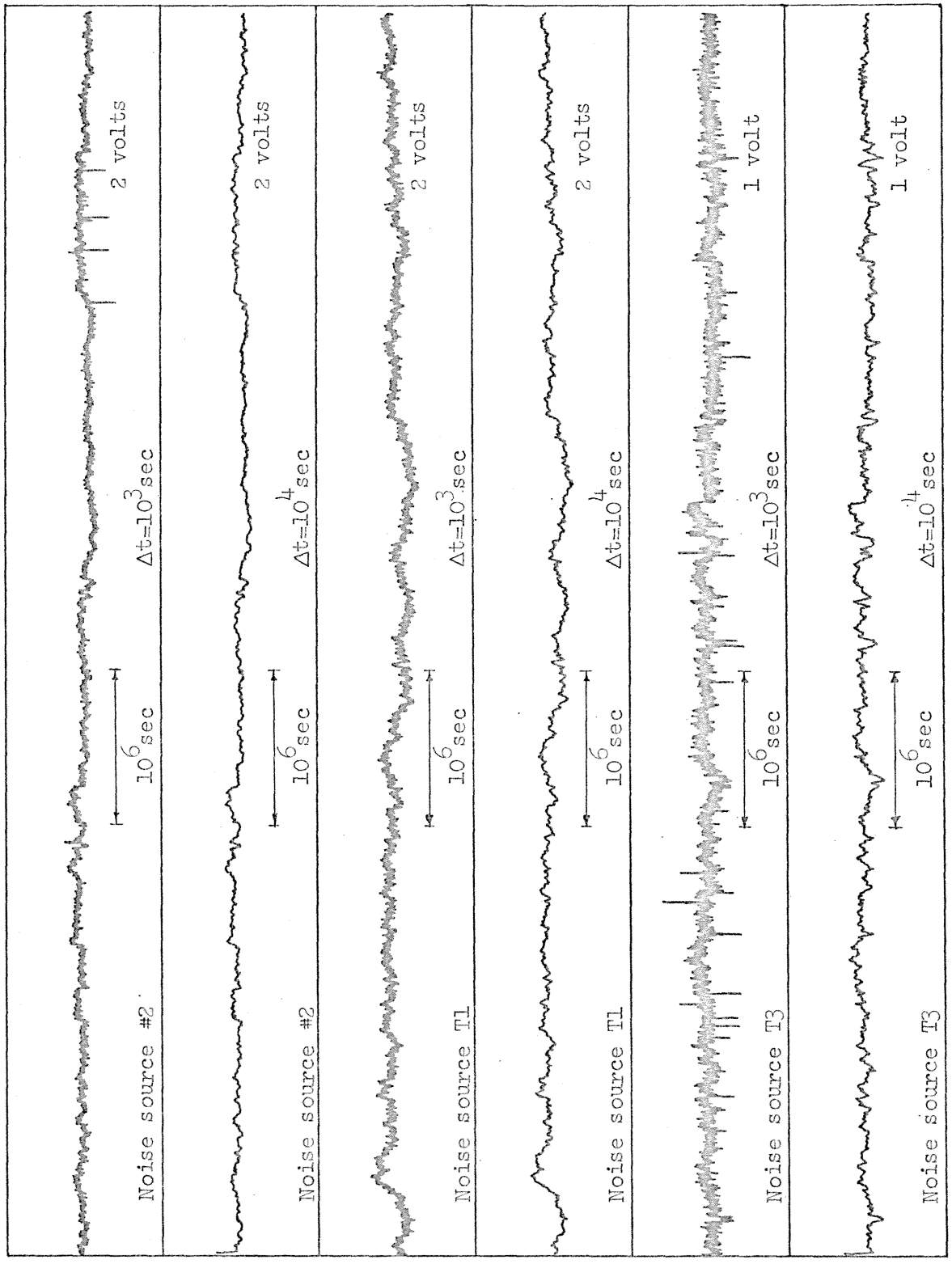


Figure 5.31 Time domain data after removal of estimated burst components from noise sources #2 and T3, and removal of estimated exponential trend from noise source T1.

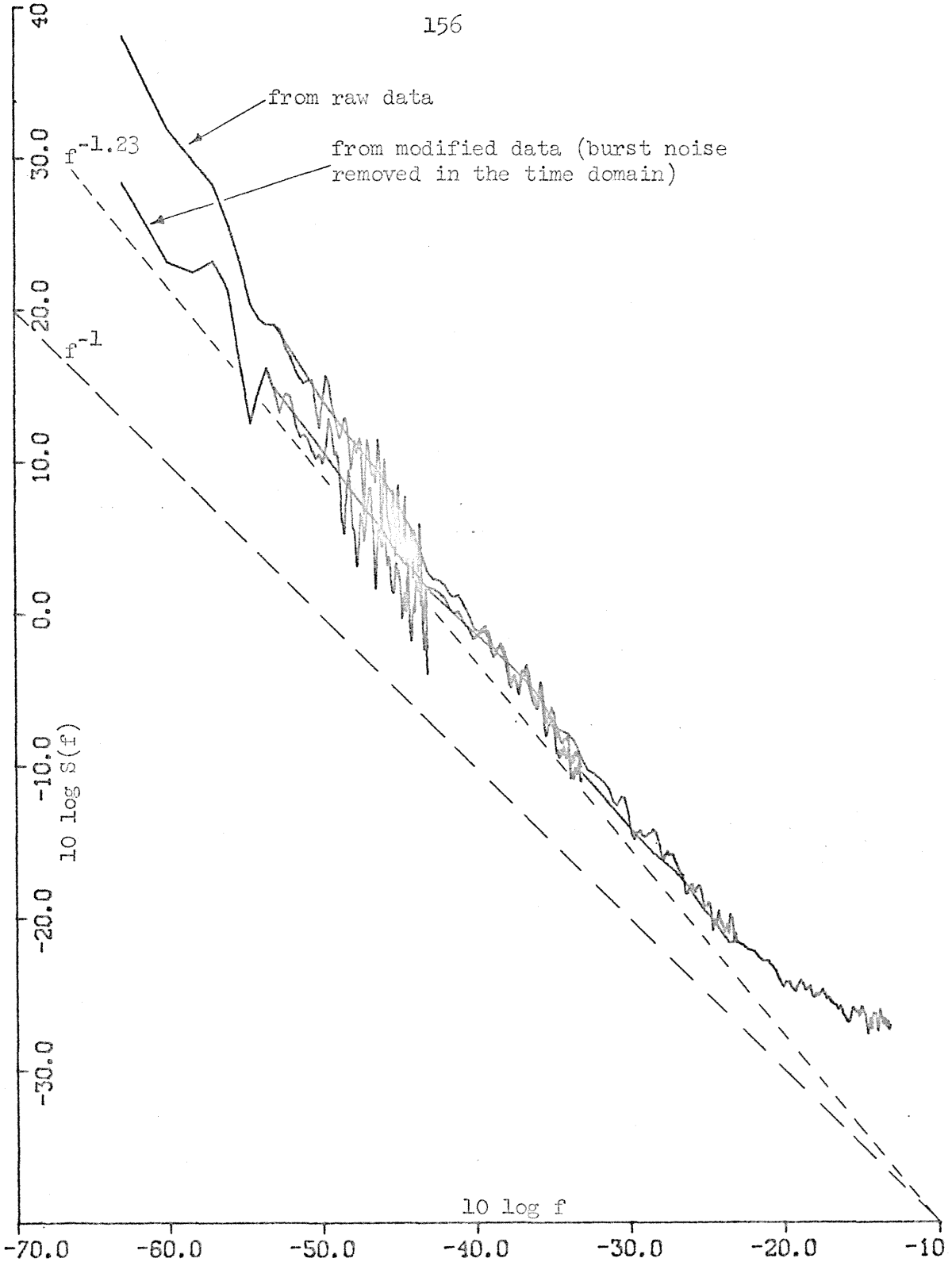


Figure 5.32 Spectral estimates of noise source #2 before and after removal of estimated burst component in the time domain ($-63 \leq 10 \log f \leq -13$).

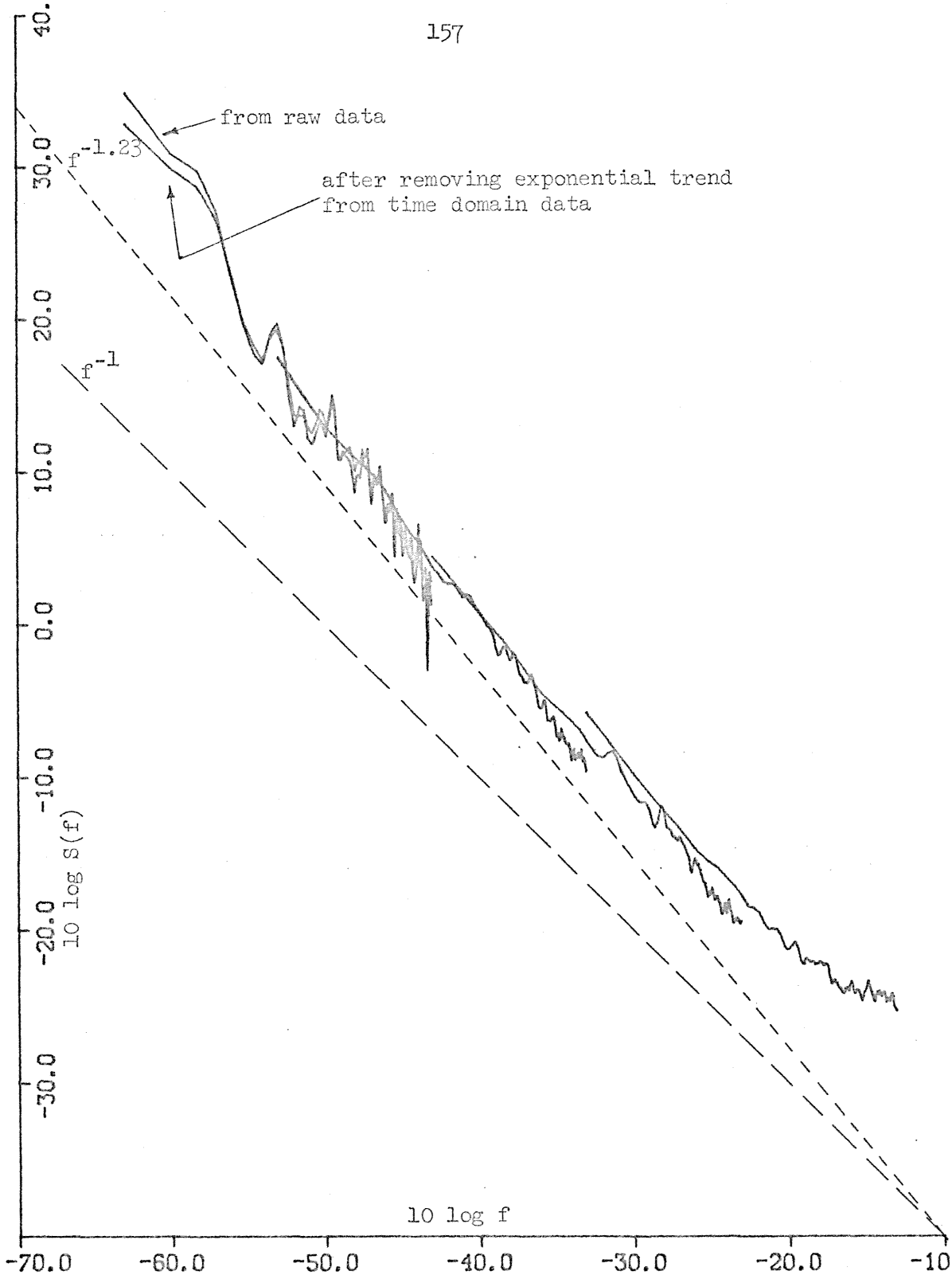


Figure 5.33 Spectral estimates for noise source T1 before and after removal of estimated exponential trend in time domain ($-63 \leq 10 \log f \leq -13$).

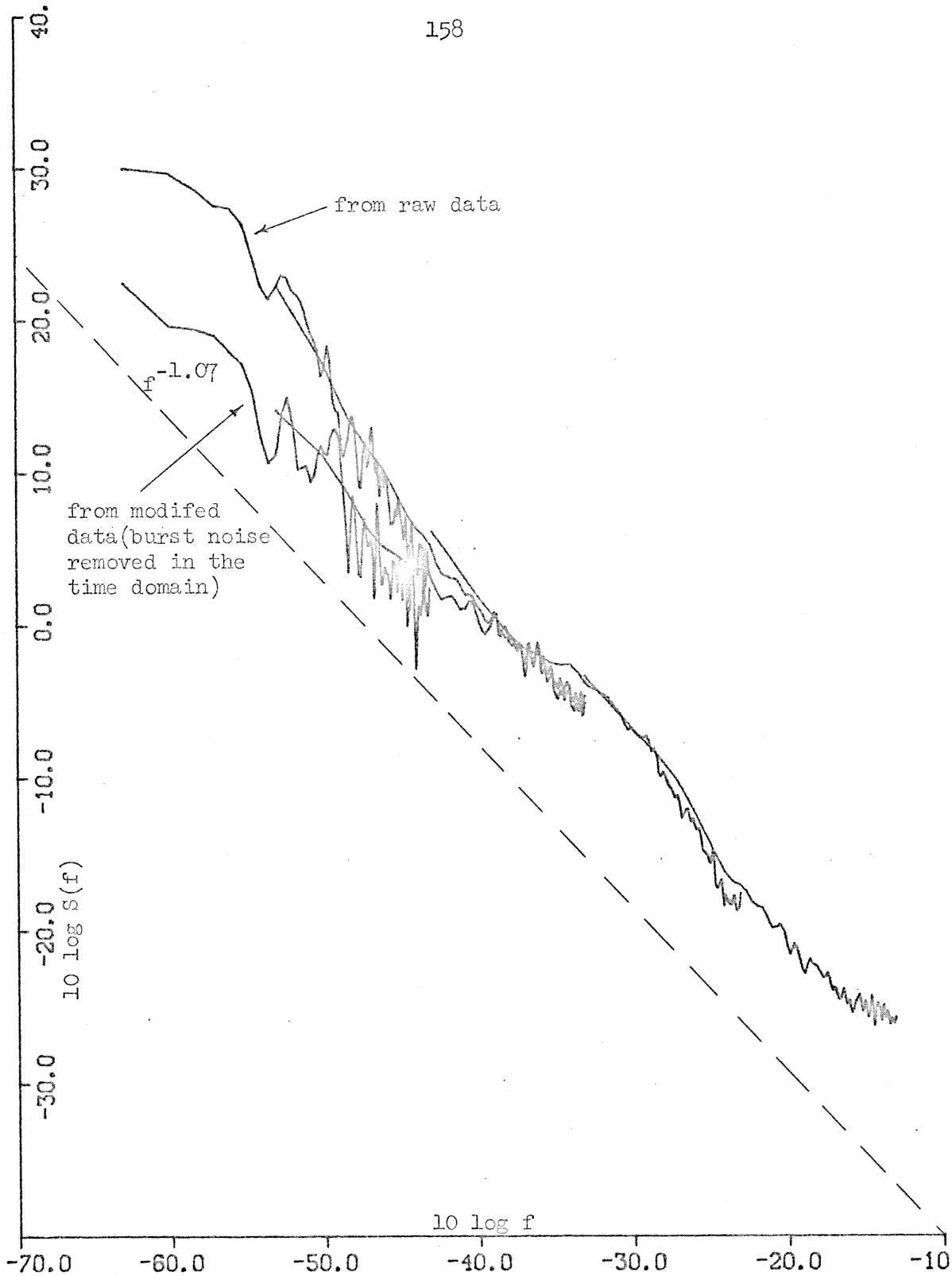


Figure 5.34 Spectral estimates of noise source T3 before and after removal of estimated burst component in time domain ($-63 \leq 10 \log f \leq -13$).

If one attempts to fit these estimates to a $Af^{-\alpha}$ model, the value of α is reduced slightly ($\alpha \sim 1.1$ for noise source T3 and $\alpha \sim 1.23$ for noise source #2 by my estimate) but did not result in any estimates of $\alpha \leq 1$. It should be remembered in estimating α that these spectral estimates have not been corrected for aliasing. Making these corrections would pull the spectral estimates down by about 5 db at 10 log f = -13 if $\alpha \sim 1$ and there was no component due to popcorning at this frequency. No attempt was made to carefully measure the statistics of these burst components because of the small number of data points involved.

Noise source T4 exhibits burst noise with time constants short enough that several hundred bursts may be observed in the data of the first phase of the experiment. The thought of plotting all of the T4 data at Δt = 100 sec (as in Fig. 5.2) and attempting to compile a list of burst times and amplitudes by hand was somewhat overwhelming. A more automated method was required to model the burst components of this source. An obvious method is as follows. For each datum point the local sample mean is calculated over some number of points before and some number after and including this datum point. These two local sample means are subtracted. If there were no bursts in this range we would expect the difference to be close to zero. If there were a single burst at the dividing point between the sample means we would expect the difference to be approximately equal to the burst amplitude. Since the expected differences will be a ramp whenever the sample averages are over more than one point, we must look for peaks in the difference function.

The actual algorithm used was as follows. Given the data points N_i the differences between local means is calculated

$$D_i = \frac{1}{L} \sum_{j=1}^{L} \left\{ N_{i-j+1} - N_{i-j} \right\} \qquad \text{[Eq. 5.8.1]}$$

for some specified L. A threshold, θ, is specified and all α_j and β_j are determined such that for $\alpha_j \leq i \leq \beta_j$:

1) $|D_i| > \theta$

2) all D_i have the same sign

3) for $i = \alpha_j - 1$ and $i = \beta_j + 1$ either $|D_i| < \theta$ or D_i has a different sign than D_{α_j}.

For each j a value k_j, $\alpha_j \leq k_j \leq \beta_j$, is then determined such that $|D_{k_j}| \geq |D_i|$ for all i such that $\alpha_j \leq i \leq \beta_j$. The k_j are then assumed to be the estimated burst times with estimated burst amplitudes of D_{k_j}.

Bursts of shorter duration than L may be missed (or at least assigned an amplitude which is too small). This problem may be minimized by choosing L to be small relative to $\frac{\tau}{\Delta t}$ for both τ_+ and τ_-. If the waiting time distribution is $p(t) = \frac{1}{\tau} e^{-t/\tau}$, then the probability of any given burst lasting for a shorter time than $\gamma\tau$ is $p\{t \leq \gamma\tau\} = 1 - e^{-\gamma}$. Therefore, if we choose $L < \frac{\tau}{100\ \Delta t}$, less than 1% of the bursts will be shorter in duration than L samples.

If we intend to use these data to remove the burst component of a noise source to improve our spectral estimate, it is clear that by removing all but a few isolated short bursts we will significantly

decrease the total average power due to the burst component (particularly at the low frequency end of the spectrum). Multiple short bursts could also, however, cause a burst of long duration to be missed (see Fig. 5.35).

Figure 5.35 shows an example of a contrived burst noise estimated by the above algorithm with L = 4 and $\theta = \frac{5}{8}$ x height of burst noise. This figure demonstrates that isolated short bursts may be missed (as for the bursts starting at i = 2, 14, 40). It also shows a long duration burst which is missed due to nearby short bursts (the burst beginning at i = 44). In this example we could eliminate these problems by decreasing L to 1.

If we have a noise source N(t) = n(t) + b(t) where b(t) is a burst component and we assume for the moment that all of the bursts in b(t) (both plus and minus) are longer than LΔt, then we may still have a problem detecting the bursts. Consider the difference function,

$$D_i = \frac{1}{L}\sum_{j=1}^{L} [N[(i-j+L)\Delta t] - N[(i-j)\Delta t]],$$

as a random process.

If we consider first the case where b(jΔt) = constant for i-L $\leq$ j $\leq$ i+L-1 we have

$$D_i = \frac{1}{L}\sum_{j=1}^{L} n[(i-j+L)\Delta t] - n[(i-j)\Delta t] \quad . \qquad \text{[Eq. 5.8.2]}$$

If n(t) is a stationary process then

$$E\{D_i\} = 0$$

and

$$E\{D_i^2\} = E\left\{ \frac{1}{L^2} \sum_{j=1}^{L} \sum_{k=1}^{L} [n[(i-j+L)\Delta t]-n[(i-j)\Delta t]][n[(i-k+L)\Delta t] - n[(i-k)\Delta t]] \right\}$$

$$= \frac{2}{L^2} \sum_{j=1}^{L} \sum_{k=1}^{L} [R_n\{(j-k)\Delta t\} - R_n\{(j-k-L)\Delta t\}]$$

$$= \frac{2}{L^2} \left\{ LR_n(0)-LR_n(L\Delta t) + \sum_{j=1}^{L-1} (L-j)[2R_n(j\Delta t)-R_n\{(j-L)\Delta t\} - R_n\{(j+L)\Delta t\}] \right\}$$

If we assume that

$$S_n(f) = \begin{cases} \frac{A}{|f|} \text{ for } \varepsilon \leq |f| \leq f_m \\ 0 \text{ for } |f| < \varepsilon \text{ or } |f| > f_m \end{cases}$$

then

$$R_n(\tau) = 2 \int_0^{\infty} S_n(f) \cos(2\pi f\tau)\, df$$

$$= 2A \int_{\varepsilon}^{f_m} \frac{\cos(2\pi f\tau)}{f}\, df$$

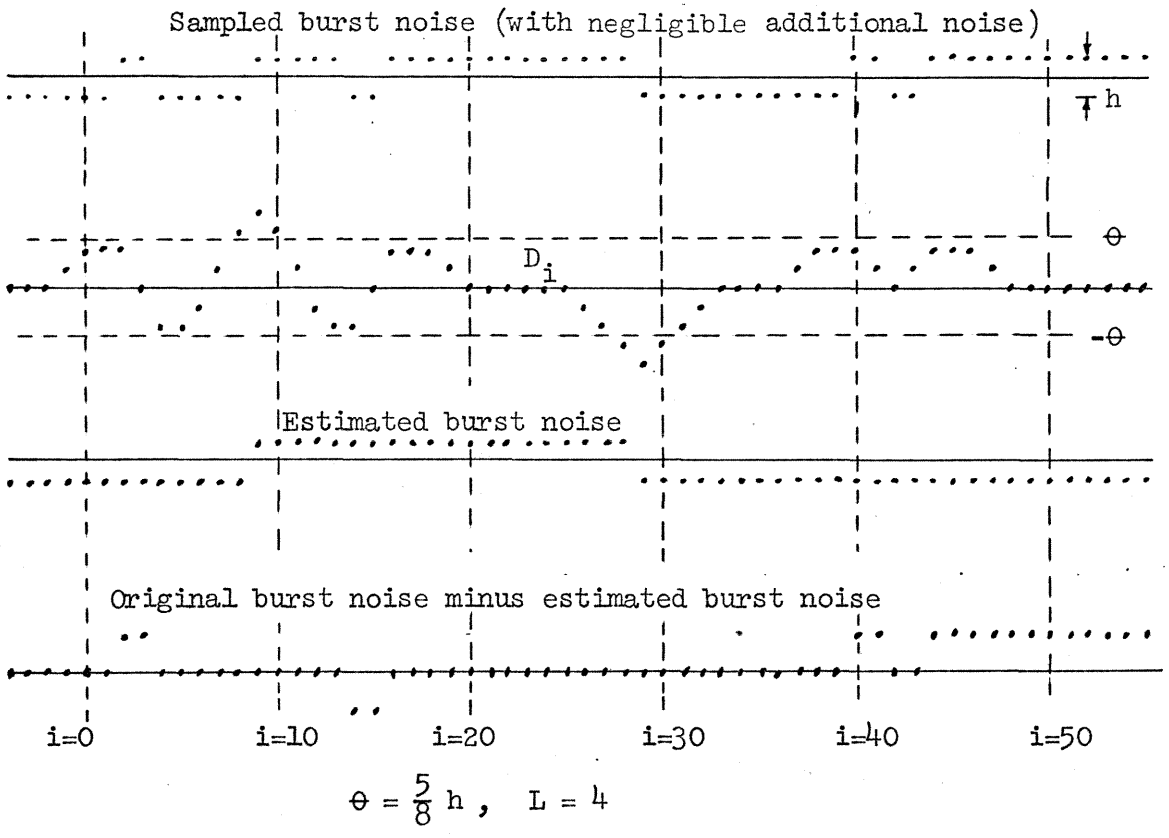


Figure 5.35 Example of burst noise estimation algorithm applied to fictitious burst noise.

$$= \begin{cases} 2A(\ell n\ f_m - \ell n\ \varepsilon), \tau = 0 \\ 2A[ci(2\pi\tau f_m) - ci(2\pi\tau\varepsilon)],\ \tau > 0^* \\ R_n(-\tau),\ \tau<0 \end{cases}$$

$$E\left\{D_i^2\right\} = \frac{4A}{L}\left\{\ell n\ f_m - \ell n\ \varepsilon - ci\ (2\pi L\ \Delta t\ f_m) + ci(2\pi L\ \Delta t\varepsilon)\right.$$

$$+ \sum_{j=1}^{L-1} \frac{L-j}{L}[2\ ci(2\pi j\ \Delta t\ f_m) - 2\ ci(2\pi j\Delta t\ \varepsilon) - ci(2\pi(L-j)\Delta t\ f_m)$$

$$\left. + ci(2\pi(L-j)\Delta t\ \varepsilon) - ci(2\pi(L+j)\Delta t\ f_m) + ci(2\pi(L+j)\Delta t\ \varepsilon)]\right\}$$

$$\lim_{\varepsilon\to 0} E\left\{D_i^2\right\} = \frac{4A}{L}\left\{C + \ell n(2\pi L\ \Delta t\ f_m) - ci(2\pi L\ \Delta t\ f_m) + \right.$$

$$\sum_{j=1}^{L-1} \frac{L-j}{L}\left[2\ ci(2\pi j\ \Delta t\ f_m) - ci(2\pi(L-j)\Delta t\ f_m)\right.$$

$$\left.\left. - ci(2\pi(L+j)\Delta t\ f_m] + \ell n\left(\left(\frac{L}{j}\right)^2 - 1\right)\right]\right\}$$

If we let $f_m = \frac{\beta}{2\Delta t}$ then

$$\lim_{\varepsilon\to 0} E\left\{D_i^2\right\} = \frac{4A}{L}\left\{C + \ell n(\pi L\beta) - ci(\pi L\beta)\right.$$

$$\left. + \sum_{j=1}^{L-1} \frac{L-j}{L}\left[2\ ci(\pi j\beta) - ci(\pi(L-j)\beta) - ci(\pi(L+j)\beta) + \ell n\left(\left(\frac{L}{j}\right)^2 - 1\right)\right]\right\}$$

[Eq. 5.8.3]

---

* $ci(x) = -\int_x^\infty \frac{\cos(t)}{t}\,dt$ (the well known cosine integral)

$ci(x) = C + \ell n(x) + O(x^2)$ where $C = .0577\ \ldots$ (Euler's constant)

The fact that the variance of D_i does not diverge even when $S_n(f) = A/|f|$ even as $f \to 0$ is not surprising when one notes that D_i is the sum of finite differences $N\left((k+L)\Delta t\right) - N(k\Delta t)$. The function $D(i\Delta t) = D_i$ is, therefore, the result of passing $n(i\Delta t)$ through a filter with a zero at $f = 0$. This implies that the variance of D_i will not diverge for $S_n(f) \sim A|f^{-\alpha}|$ $(f \to 0)$ for any $\alpha < 3$ ($S_{n(t)}(f) \sim A|f^{-\alpha}|$ for $|f| << \frac{1}{L\Delta t}$ implies $S_{D(t)}(f) \sim \hat{A}\,|f^{-\alpha+2}|$). Furthermore, if $\alpha \sim 1$ we can see that the value of $E\left\{D_i^2\right\}$ is primarily dependent only on the highest decade or so (of frequency) of $S_n(f)$.

Table 5.8 shows values of $E\left\{D_i^2\right\}/A$ evaluated on the digital computer (using a standard system routine for evaluating ci(x)) for various values of L and β using Eq. 5.8.3.

From Table 5.8 we see that if $S_n(f)$ is a 1/f process and we have a sharp pre-sampling filter with a cutoff frequency of $1/(2\Delta t)$, then the variance of D_i is nearly independent of L. If, however, we are under sampled by a factor of 5 to 10 (as is the case if we use our raw data at $\Delta t = 10$ sec which was filtered with 0.34 Hz low pass filter giving $\beta \sim 7$) we may decrease the variance of D_i significantly by choosing $L \sim 20$ but with little advantage in making L larger than this. $E\left\{D_i^2\right\}$ yields information concerning the probability of having $|D_i| > \theta$ indicating the presence of a burst when, in fact, there is none.

If we now consider the case where $N(t) = n(t) + b(t)$ and

$$b(j\Delta t) = \begin{cases} k + h\,, & i \le j \le i + L - 1 \\ k\,, & i - L \le j \le i - 1 \end{cases}$$

β	L	$E\{D_i^2\}/A$
1	1	6.59
1	10	5.58
1	100	5.55
1	1000	5.55
5	1	13.31
5	10	6.19
5	20	5.86
5	100	5.60
10	1	16.10
10	20	5.99
10	100	5.63
100	1	25.31

Table 5.8

Values for $E\{D_i^2\}/A$ computed using Eq. 5.8.3 assumes

$$S_n(f) = \begin{cases} A/|f| & \text{for } |f| \leq \frac{\beta}{2\Delta t} \\ 0 & \text{for } f > \frac{\beta}{2\Delta t} \end{cases}$$

(where h may be positive or negative depending on the sign of the step) then we have

$$D_i = \frac{1}{L} \sum_{j=1}^{L} \left\{ n\left[(i-j+L)\Delta t\right] - n\left[(i-j)\Delta t\right] \right\} + h.$$

This is the same as Eq. 5.8.2 (when no burst was present) except for the presence of an additive constant h. Therefore

$$E\left\{D_i\right\} = h$$

and

$\sigma^2_{D_i} = E\left\{(D_i-h)^2\right\}$ is the same as $E\left\{D_i^2\right\}$ when there is no burst present in the range $i-L \leq j \leq i+L-1$.

This algorithm was applied to the detection of burst components in noise source T4. Looking at Fig. 5.2, the shortest burst observed is approximately 10^3 seconds in length (100 samples). It was, therefore, decided to use L = 20 in the algorithm for determining the D_i's. (An additional reason for using L > 1 which was not previously mentioned is the finite probability of a sample occurring during the finite rise time of the step due to the dealiasing filter. This would cause an estimate of h which is too low but is at worst $h(1 - \frac{1}{2L})$ if the rise time is less than Δt). The algorithm was performed on the raw data with a threshold of θ = 0.05 volts. Figure 5.36 is a histogram showing the number of D_{k_j} recorded in bins which are 0.01 volts wide.

As we might have expected, there are groupings near ± 0.175 volts indicating a burst component of that amplitude. If we compute the total number of $D_{k_j} \geq 0.12$ volts and the number of $D_{k_j} \leq -0.12$ volts we

find that there are 430 positive and 431 negative values. This is as would be expected (the number of positive steps must equal the number of negative steps in our burst noise model). The positive D_{k_j} seem to be spread over a wider range (have a greater variance) than the negative D_{k_j}. This may be chance, although if true, it does not fit our simple model. Different distributions for positive and negative steps would tend to indicate that either the remaining noise, n(t), is not actually independent of the burst noise or that the burst noise amplitude is not, in fact, a constant and has a distribution which is different for positive and negative steps.

The sample mean and variance were calculated for $D_{k_j} \geq 0.12$ volts and for $D_{k_j} \leq -0.12$ volts. For $D_{k_j} \geq 0.12$ volts the sample mean is 0.1793 volts and the variance is 4.44×10^{-4} volts^2 ($\sigma = 0.02108$ volts). For $D_{k_j} \leq -0.12$ volt the sample mean is -0.1795 volt and the variance is 3.95×10^{-4} volt^2 ($\sigma = 0.01987$ volts). For $|D_{k_j}|$ such that $|D_{k_j}| \geq 0.12$ volts we get a sample mean of 0.1794 and a variance of 4.19×10^{-4} volt^2 ($\sigma = 0.0205$ volts). The difference in standard deviations for positive and negative steps is ~ 6% . No attempt has been made to determine if this difference in sample standard deviations is significant for this number of samples. This would require making assumptions about the nature of the n(t) process, which at this point still contains other burst components.

Referring back to Fig. 5.12, one may attempt to fit an A/|f| curve to the power spectral density of the T4 noise source. Recognizing that the spectral density estimate is too large at the high frequency end due

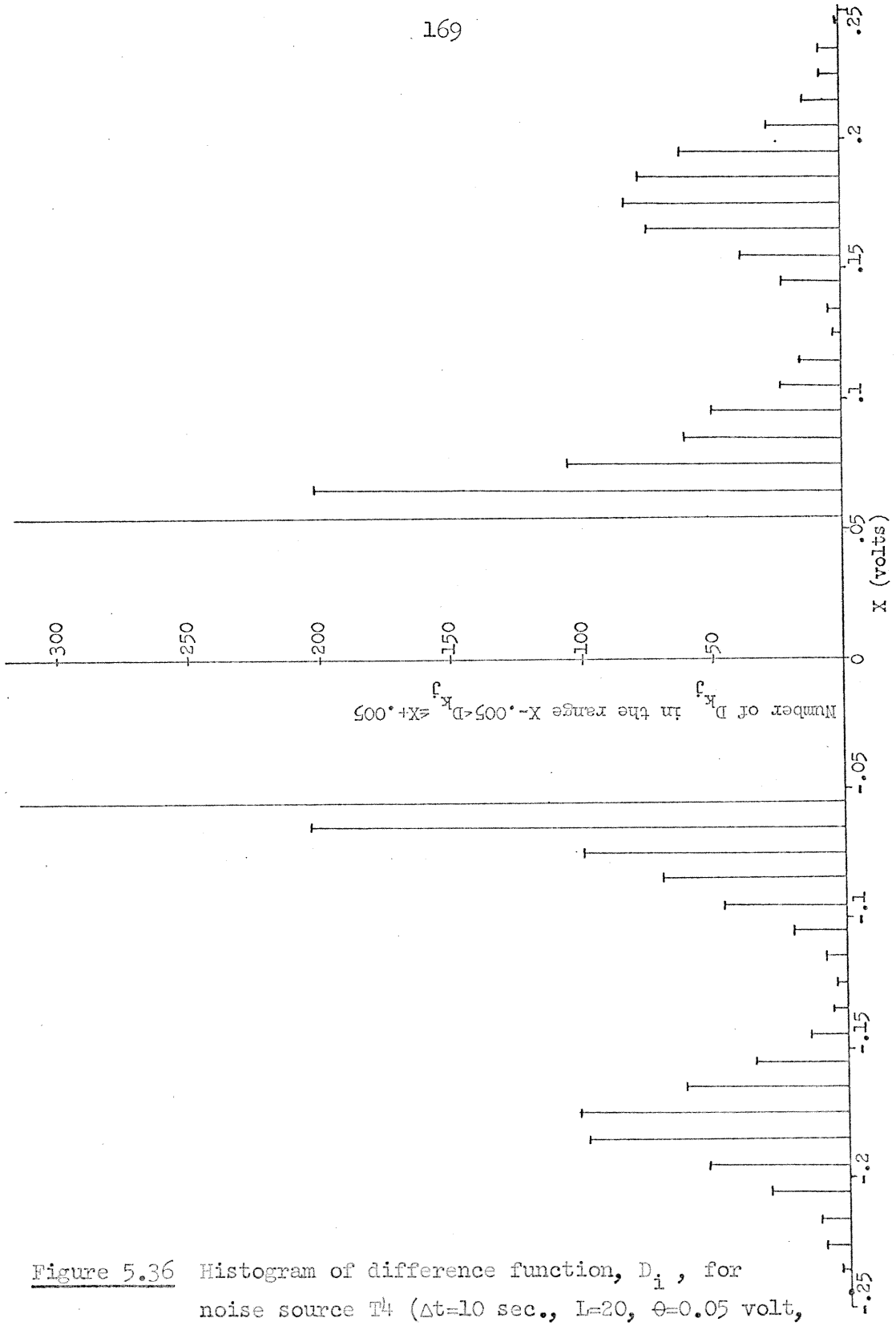


Figure 5.36 Histogram of difference function, D_i, for noise source T4 (Δt=10 sec., L=20, Θ=0.05 volt, and 0.01 volt wide bins).

to aliasing, suspecting that the burst components dominate much below 10 log f = -25, and remembering that the lowest decades are less important than the highest; my estimate for A is 10 log A ~ -44 to -41.

$$A \sim 3.98 \times 10^{-5} \text{ to } 7.94 \times 10^{-5} \text{ volt}^2$$

Using Table 5.8 for L = 20, β = 5 we get a predicted standard deviation in the range $0.015 \leq (E\{D_i^2\})^{\frac{1}{2}} \leq 0.022$ volts as compared to the sample standard deviation of 0.0205 volts for $|D_{k_j}|$ such that $|D_{k_j}| \geq 0.12$ volts. This is quite good agreement considering that we are trying to estimate A from a power spectral density which is highly distorted by burst noise and, more importantly, have no real reason to believe that $S_n(f)$ is in fact $A/|f|$.

Figure 5.2 clearly exhibits at least one other popcorn source with an amplitude of approximately 0.1 volt. Presumably this cannot be observed in Fig. 5.36 because of the non-zero variance of the difference function and the large number of points involved.

Hoping to see some trace of the popcorn component with amplitude of approximately 0.1 volt, another histogram was plotted with a higher resolution (Fig. 5.37). This figure has two interesting features. First, with a little imagination, the effects of another popcorn source with amplitude less than 0.1 volts can be seen (especially on the right hand side representing positive transitions). Second, the groupings about ±.18 volts no longer appear to have such a smooth bell shaped distribution. This may be due to the larger standard deviation caused by the smaller number of data points in each bin (for N = 430, k = 50 we have an

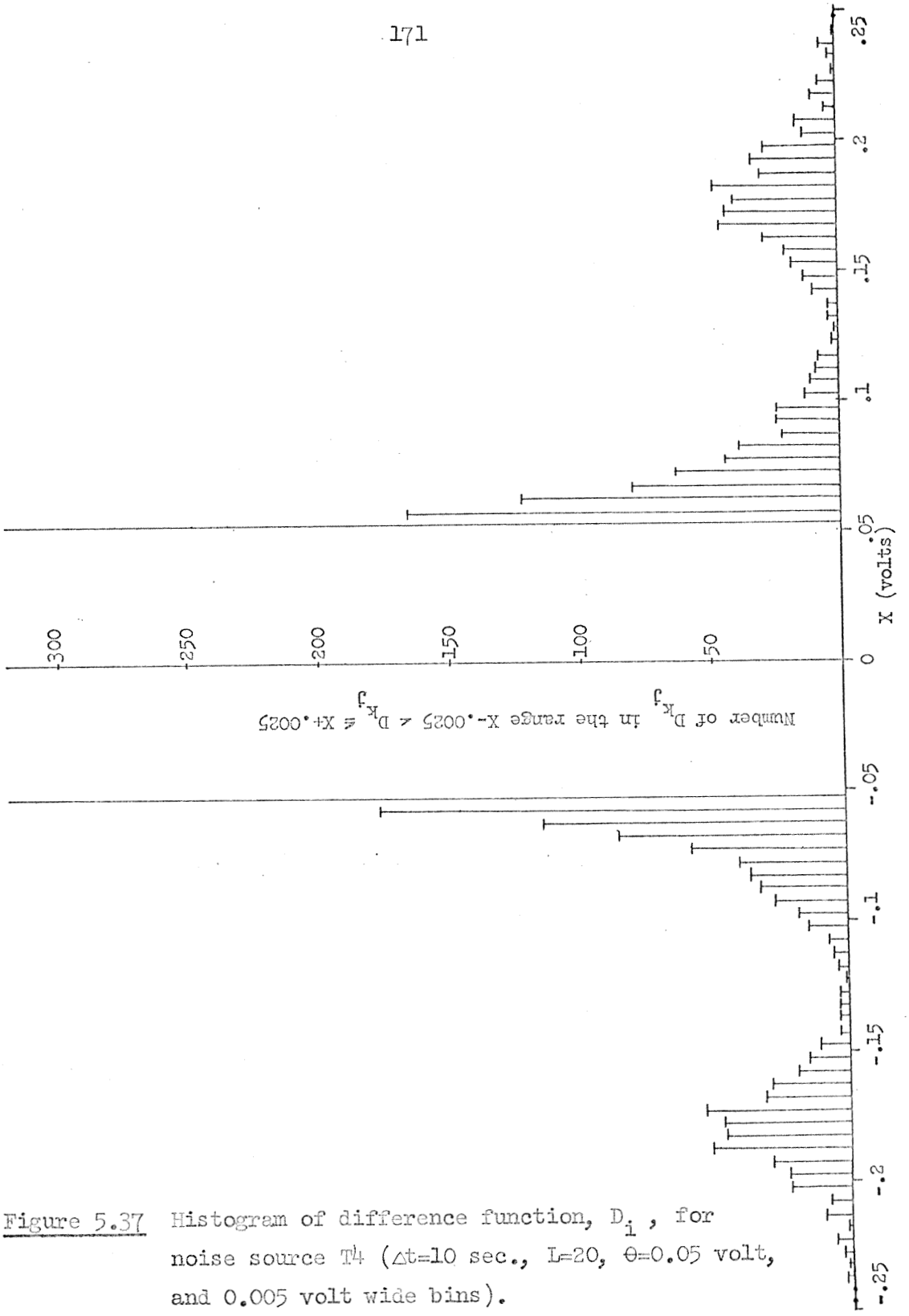


Figure 5.37 Histogram of difference function, D_i, for noise source T4 (Δt=10 sec., L=20, Θ=0.05 volt, and 0.005 volt wide bins).

estimated standard deviation of $\hat{\sigma}_k \sim 7$ from Appendix B). The dual peaks, if they in fact exist, seem to suggest that we are actually dealing with two burst components with approximately the same amplitudes and time constants, or there may be some uncertainty in the amplitude of the single burst process with two preferable states.

The obvious next step would be to assume that all of the D_{k_j} such that $|D_{k_j}| \geq 0.12$ volt represent steps of magnitude ~ 0.18 volts. This procedure, unfortunately, leads to an assumed process with occasional dual + steps or dual - steps. The possibility that this indicated the presence of two burst components of approximately the same amplitude was considered. The T4 noise source was modified by assuming that each time a $|D_{k_j}| \geq 0.12$ volts is encountered there is a step of magnitude 0.18 volts in N(t) and this function is subtracted from the N(t) data. The power spectral density was estimated for these modified data. The spectral density estimates were somewhat smaller at the high frequency end for the modified data but larger at the low frequency end than for the raw data estimates. Since uncorrelated processes added in the time domain produce a spectral density which is the sum of the individual spectral densities, it is clear that a new component with large low frequency spectral density had, in fact, been added.

Figure 5.38 demonstrates the effects of subtracting a series of steps from an ideal burst component when an occasional step in the original function is missed or added. In this example negative, positive, and then negative steps are missed. The resulting function after "removing" our estimate of the burst component looks very much

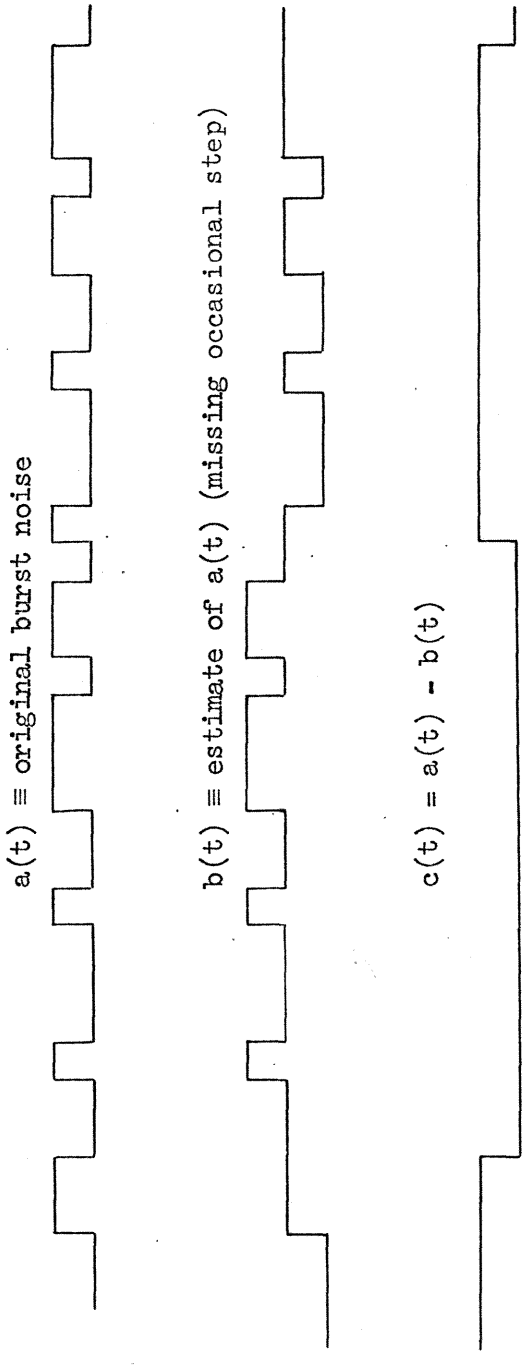


Figure 5.38 Simulated example of subtracting inaccurate estimate of burst component.

like burst noise itself but with longer time constants and the same amplitude as the burst component we were attempting to remove. Looking back at Eq. 5.8.1 we see that if $S_n(f) = A/|f|$, we would expect to find approximately the same relative distortion in the power spectral density due to the burst component except that it would be shifted down in frequency. If we are not so fortunate as to miss (or add) alternate plus and minus steps when attempting to estimate the burst component, then we may add a function with a total average power significantly larger than the burst noise which we are attempting to remove. In fact, if for some reason the estimate favors missing more plus than minus steps (for example) the estimate of the burst component would diverge.

An obvious improvement would be to artificially delete or add steps in the estimated burst component to force it to have only two levels. This insures that we do not add a function to the noise with a diverging average power. Furthermore, if we have missed (or gained) steps infrequently, the result of removing the estimated burst component will be a function with occasional plus and minus bursts with durations on the order of τ_+ or τ_-. If $\tau_+ \sim \tau_-$, for example, and we missed or gained a step approximately once out of every 40 steps then we would expect the average power of the remaining burst component, after subtracting out our estimate of the burst component, to be about $\frac{h^2}{40}$ as compared to $\frac{h^2}{4}$ for the original burst component (or a reduction in the total distortion due to the burst component of 10 db). The situation would be better, of course if we were able to reinsert lost steps at approximately the right places by some form of educated guess.

The burst noise data were, therefore, modified (by hand). Any $|D_{k_j}| \geq 0.12$ volts was considered to be a step in the primary burst component unless this resulted in dual plus or minus steps. If there was a $|D_{k_j}| < .12$ volts (remember that the D_{k_j} were obtained using $\theta = 0.05$ volts) which would correct the duality problem then it would be considered to be a step. At the same time a secondary burst component was considered in the same way for $0.07 \leq |D_{k_j}| < 0.12$ volts. Eighteen values of D_{k_j} had to be modified to arrive at a function with no dual plus or minus steps for $|D_{k_j}| \geq 0.12$ volts. This is about 1 out of 48 steps. Since most of these modified D_{k_j} were close to the ± 0.12 limit we might expect a reduction due to this component of significantly more than 10 db. The attempts to estimate the second burst component ($0.07 \leq |D_{k_j}| < 0.12$ volts) did not yield nearly as nice results (this might be expected from Fig. 5.37). It is assumed, therefore, that we have a reasonably good estimate of the switching times for a burst component in noise source T4 with an amplitude of approximately 0.18 volts and a very poor estimate of a burst component with amplitude less than 0.1 volts.

The probability density of waiting times, t, is expected to be $p(t) = \frac{1}{\tau} e^{-t/\tau}$ which implies $E\{t\} = \tau$ and $P\{kT \leq t < (k+1)T\} = (1 - e^{-T/\tau}) e^{-kT/\tau}$, $k = 0,1,\ldots$

$\text{Ln}(P\{kT \leq t < (k+1)T\}) = \text{const} - \frac{T}{\tau} k$. If we add the total time that a burst component is in a given state and divide it by the number of bursts, N, we will have an estimate of τ. For this value of τ, if

$p(t) = \frac{1}{\tau} e^{-t/\tau}$ and successive bursts are indeed independent, then $E\left\{\frac{n_k}{N}\right\} = (1 - e^{-T/\tau})\, e^{-kT/\tau}$ where n_k is the actual number of observed bursts of width $kT \leq t < (k+1)T$, and $\sigma_{\left(\frac{n_k}{N}\right)}$ is

$$\left[\frac{1}{N}\left(1 - e^{-T/\tau}\right)e^{-kT/\tau}\left\{1 - \left(1 - e^{-T/\tau}\right)e^{-kT/\tau}\right\}\right]^{1/2}$$

(see Appendix B). Figures 5.39,...,5.42 show the estimated τ for each of the states for each of the two burst components (with amplitudes of .18 volt and ~ .1 volt). The function $E\left\{\frac{n_k}{N}\right\} = \left(1 - e^{-T/\tau}\right)e^{-kT/\tau}$ is then plotted versus k for this value of τ and T along with error bars to show $\sigma_{\frac{n_k}{N}}$ and points representing the observed $\frac{n_k}{N}$ on a semilog coordinate system.

The estimated 0.18 volt burst component appears to match an exponential waiting time distribution. The 0.1 volt burst is not as convincing but this is understandable in light of the difficulties encountered in estimating this component. This supports the conclusions of Puckett[3] for burst noise with time constants of more than an hour (Puckett made no statistical measurements of bursts with time constants greater than a few milliseconds). In addition to extending Puckett's experimental measurements of waiting time distributions to bursts of significantly greater time constants, we are now in a position to test the hypothesis that the waiting time for each burst is statistically independent of all of the other bursts. Puckett was not able to do this because of experimental limitations. He recorded a set of burst waiting times which were separated in time by some unknown number of bursts.

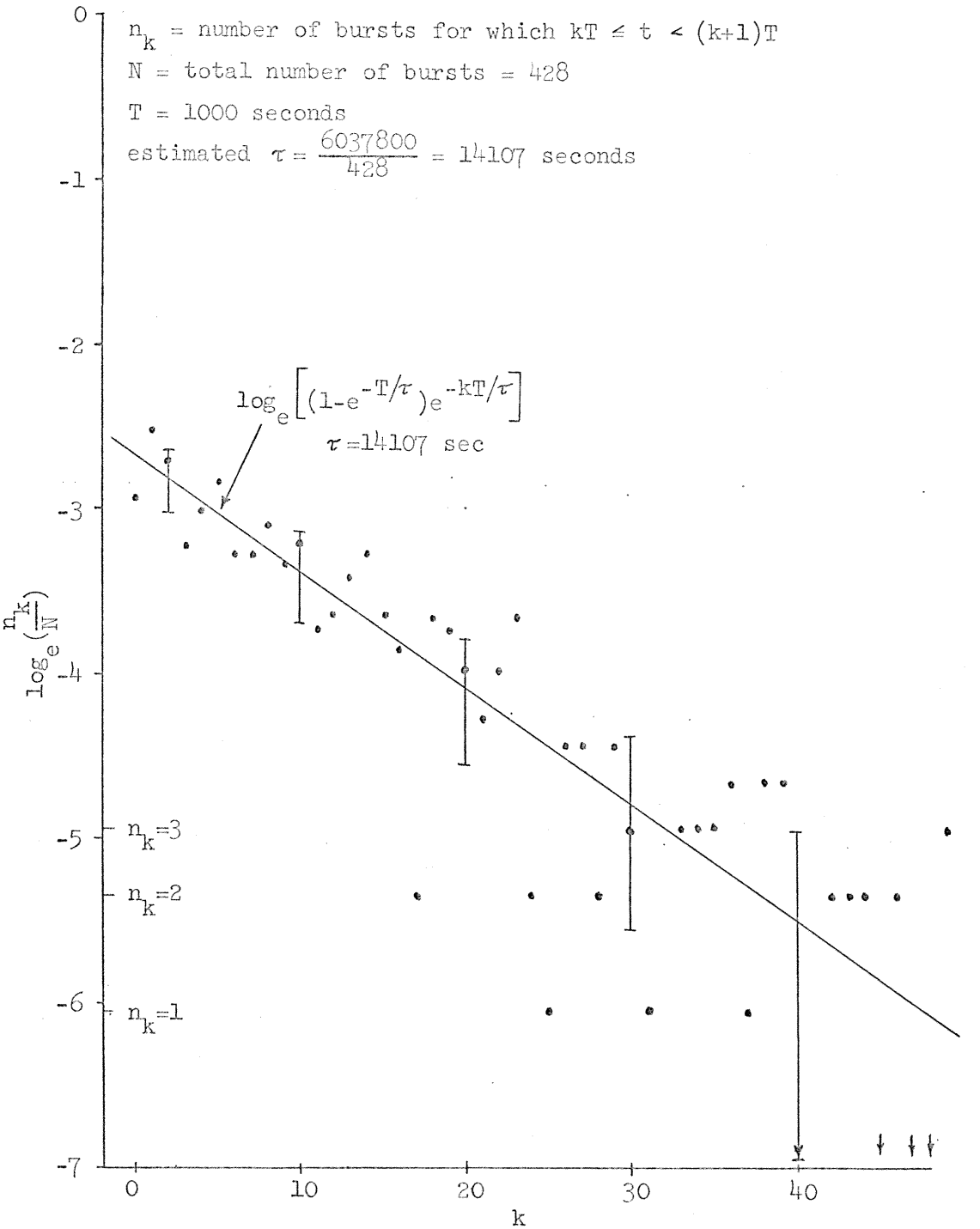


Figure 5.39 Histogram of estimated +0.18 volt burst waiting times for noise source T4 data.

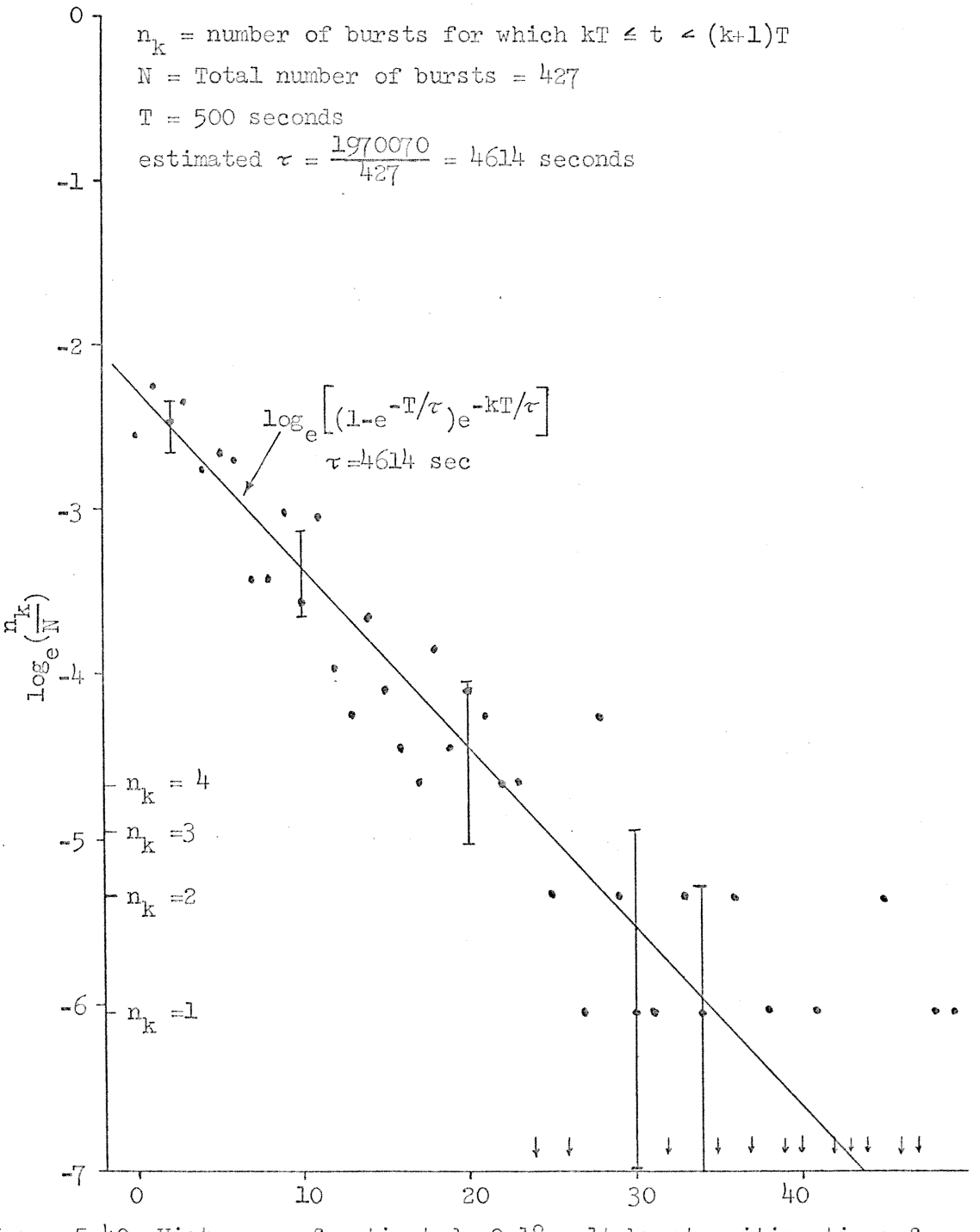


Figure 5.40 Histogram of estimated -0.18 volt burst waiting times for noise source T4 data.

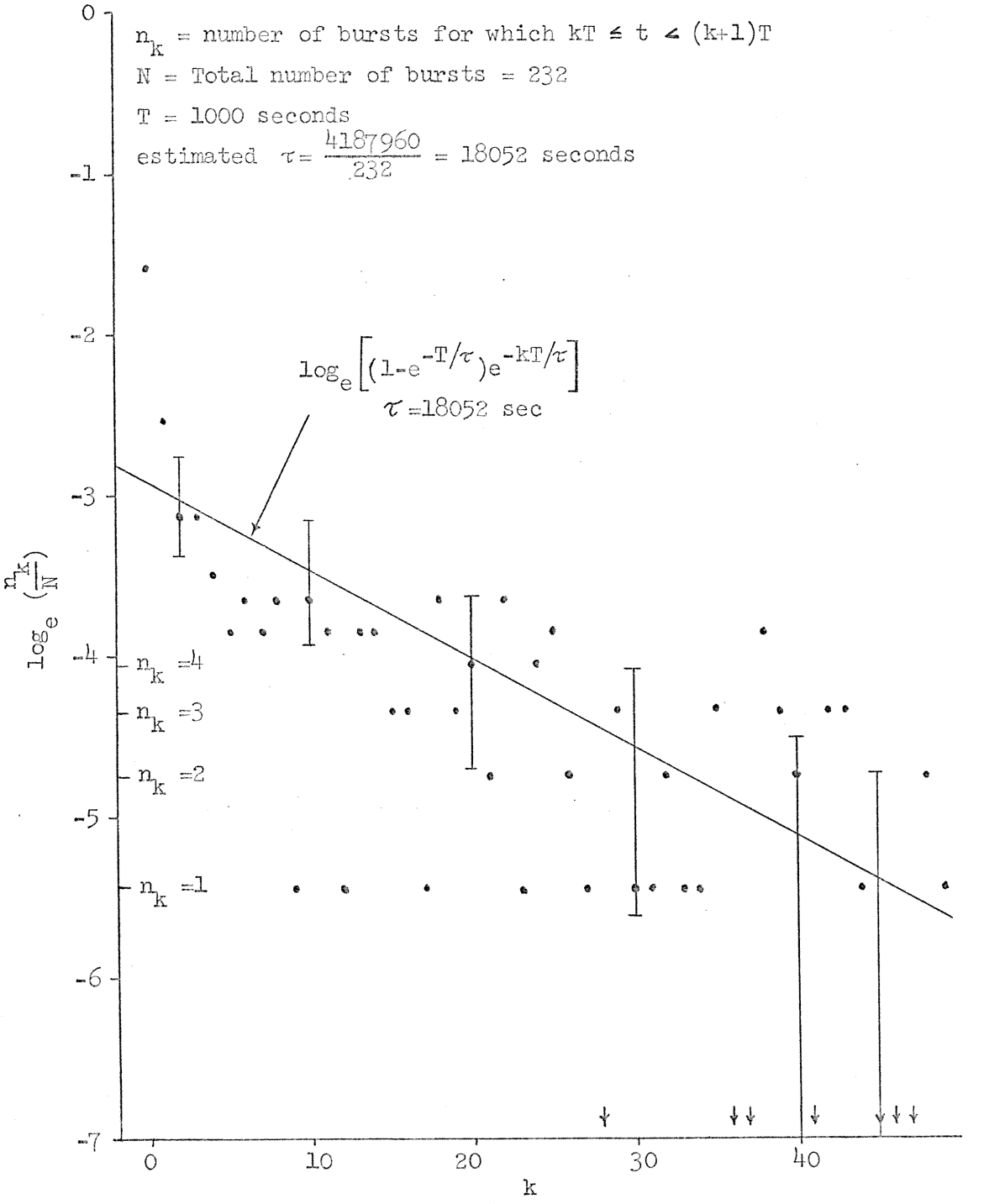


Figure 5.41 Histogram of estimated +0.1 volt burst waiting times for noise source T4 data.

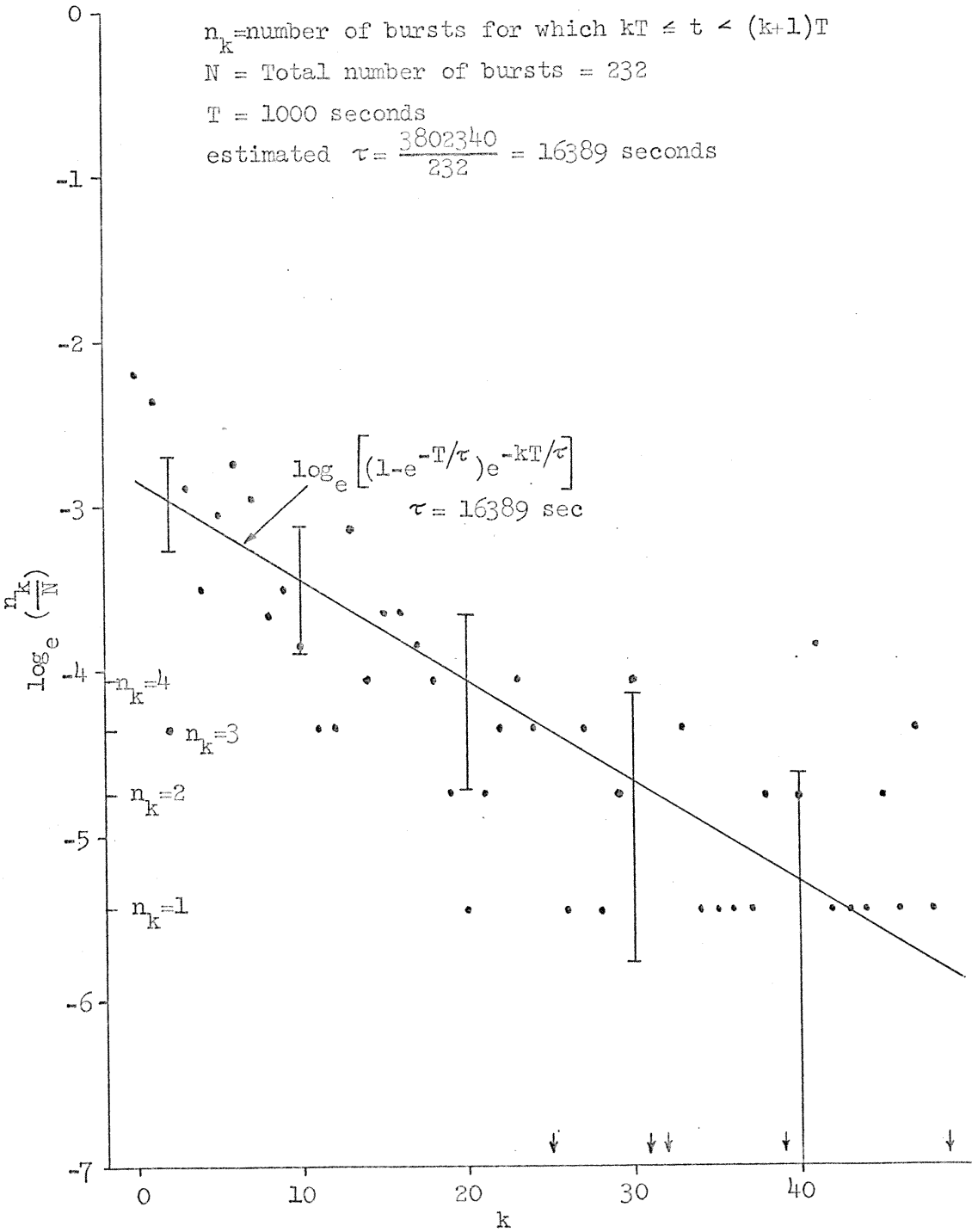


Figure 5.42 Histogram of estimated -0.1 volt burst waiting times for noise source T4 data.

The method chosen for testing the hypothesis that the waiting times are statistically independent is as follows. Let $t_+(j)$ $\left(1 \leq j \leq n_+\right)$ and $t_-(j)\left(1 \leq j \leq n_-\right)$ be the ordered sequence of estimated waiting times for the positive and negative bursts respectively. Let

$$c_{\pm}(\ell) = \frac{1}{n_{\pm}-\ell} \sum_{i=1}^{n_{\pm}-\ell} \left[t_{\pm}(i) - \hat{\tau}_{\pm}\right] \left[t_{\pm}(i + \ell) - \hat{\tau}_{\pm}\right], \quad 0 \leq \ell < n_{\pm}$$

and

$$c(\ell) = \frac{1}{M_1 - M_2 + 1} \sum_{i=M_1}^{M_2} (t_+(i) - \hat{\tau}_+)(t_-(i + \ell) - \hat{\tau}_-), \quad -n_+ < \ell < n_-$$

where

$$M_1 = \begin{cases} 1, & \ell \geq 0 \\ 1-\ell, & \ell < 0 \end{cases} \qquad \text{and } M_2 = \min(n_+, n_- - \ell).$$

If the waiting times are statistically independent and

$$p_{t_{\pm}}(t) = \frac{1}{\hat{\tau}_{\pm}} e^{-t/\hat{\tau}_{\pm}}$$

then

$$E\left\{c_{\pm}(\ell)\right\} = \begin{cases} 0, & \ell \neq 0 \\ \hat{\tau}_{\pm}^2, & \ell = 0 \end{cases} \qquad \text{and } E\left\{c(\ell)\right\} = 0 \quad .$$

If there is some dependency of a waiting time on previous bursts of either sign, then we might hope to see values of $c_{\pm}(\ell)$ $(\ell \neq 0)$ or $c(\ell)$ significantly different from zero. In order to determine the meaning of "significantly" we must consider two points. First of all, we do not know $\tau_{\pm}$ exactly; we have only estimates

$$\hat{\tau}_{\pm} = \frac{1}{n_{\pm}} \sum_{i=1}^{n_{\pm}} [t_{\pm}(i)] \quad ,$$

whose values are given for the burst components we are now considering in Figs. 5.39,...,5.42. If $p_{t_{\pm}}(t) = \frac{1}{\tau_{\pm}} e^{-t/\tau_{\pm}}$, if we have estimated the burst time periods accurately over the time span available to us, and if the bursts are statistically independent then. $E\left\{\hat{\tau}_{\pm}\right\} = \tau_{\pm}$.

$$E\left\{\hat{\tau}_{\pm}^{2}\right\} = E\left\{\frac{1}{n_{\pm}^{2}} \sum_{i=1}^{n_{\pm}} t_{\pm}(i) \sum_{j=1}^{n_{\pm}} t_{\pm}(j)\right\}$$

$$= \frac{1}{n_{\pm}^{2}} \sum_{i=1}^{n_{\pm}} \sum_{j=1}^{n_{\pm}} E\left\{t_{\pm}(i) t_{\pm}(j)\right\}$$

$$= \frac{1}{n_{\pm}^{2}} \left[n_{\pm} E\left\{t_{\pm}^{2}\right\} + \left(n_{\pm}^{2} - n_{\pm}\right) E\left\{t_{\pm}\right\}^{2}\right]$$

$$= \frac{2}{n_{\pm}} \tau_{\pm}^{2} + \left(1 - \frac{1}{n_{\pm}}\right) \tau_{\pm}^{2}$$

$$\sigma_{\hat{\tau}_{\pm}} = E\left\{\hat{\tau}_{\pm}\right\} \Big/ \sqrt{n_{\pm}} \quad .$$

$$E\left\{c_{\pm}(\ell)\right\} = \frac{1}{n_{\pm}-\ell} \sum_{i=1}^{n_{\pm}-\ell} E\left\{\left(t_{\pm}(i) - \hat{\tau}_{\pm}\right)\left(t_{\pm}(i+\ell) - \hat{\tau}_{\pm}\right)\right\}$$

$$= \frac{1}{n_{\pm}-\ell} \begin{cases} \sum_{1=1}^{n_{\pm}-\ell} E\left\{t_{\pm}^{2}\right\} - 2\,\hat{\tau}_{\pm} E\left\{t_{\pm}\right\} + \hat{\tau}^{2}_{\pm}, & \ell = 0 \\ \sum_{i=1}^{n_{\pm}-\ell} E\left\{t_{\pm}\right\}^{2} - 2\,E\left\{t_{\pm}\right\}\hat{\tau}_{\pm} + \hat{\tau}^{2}_{\pm}, & \ell \neq 0 \end{cases}$$

$$= \begin{cases} 2\tau_{\pm}^{2} - 2\tau_{\pm}\hat{\tau}_{\pm} + \hat{\tau}_{\pm}^{2}, & \ell = 0 \\ \tau_{\pm}^{2} - 2\tau_{\pm}\hat{\tau}_{\pm} + \hat{\tau}_{\pm}^{2}, & \ell \neq 0 \end{cases}$$

$$= \begin{cases} \tau_{\pm}^{2} + (\tau_{\pm} - \hat{\tau}_{\pm})^{2}, & \ell = 0 \\ (\tau_{\pm} - \hat{\tau}_{\pm})^{2}, & \ell \neq 0 \end{cases}$$

$$E\left\{c_{\pm}(\ell)\right\} = \begin{cases} \tau_{\pm}^{2}, & \ell = 0 \\ 0, & \ell \neq 0 \end{cases} \qquad \text{if } \hat{\tau}_{\pm} = \tau_{\pm}\,.$$

If we let $y_{\pm}(i) = t_{\pm}(i) - \hat{\tau}_{\pm}$ then the $y_{\pm}(i)$ are statistically independent (by our assumption that the $t_{\pm}(i)$ are statistically independent) and (dropping the $\pm$)

$$E\{y(i)\} = \tau - \hat{\tau}$$

$$E\{y(i)^2\} = 2\tau^2 - 2\hat{\tau}\tau + \hat{\tau}^2$$

$$E\{y(i)^3\} = 6\tau^3 - 6\hat{\tau}\tau + 3\hat{\tau}^2\tau - \hat{\tau}^3$$

$$E\{y(i)^4\} = 24\tau^4 - 24\hat{\tau}\tau^3 + 12\hat{\tau}^2\tau^2 - 4\hat{\tau}^3\tau + \hat{\tau}^4.$$

Letting $N = n_{\pm} - \ell$ we have

$$E\{c_{\pm}(\ell)^2\} = \frac{1}{N^2} E\left\{\sum_{i=1}^{N}\sum_{j=1}^{N} y(i)y(i+\ell)y(j)y(j+\ell)\right\}$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} E\{y(i)y(i+\ell)y(j)y(j+\ell)\}$$

$$E\{c_{\pm}(\ell)^2\} = \frac{1}{N^2}\begin{cases} NE\{y^4\} + (N^2 - N)\,E\{y^2\}^2 & , \ell = 0 \\ NE\{y^2\}^2 + 2(N-\ell)E\{y^2\}E\{y\}^2 + (N^2-3N+2\ell)E\{y\}^4 & , \ell \neq 0 \end{cases}$$

$$\sigma^2_{c_{\pm}(\ell)} = \begin{cases} \frac{1}{N}(20\tau_{\pm}^4 - 16\hat{\tau}_{\pm}\tau_{\pm}^3 + 4\hat{\tau}_{\pm}^2\tau_{\pm}^2) & , \ell = 0 \\ \frac{1}{N}(5\tau_{\pm}^4 - 8\hat{\tau}_{\pm}\tau_{\pm}^3 + 4\hat{\tau}_{\pm}^2\tau_{\pm}^2) + \frac{\ell}{N^2}(-2\tau_{\pm}^4 + 4\hat{\tau}_{\pm}\tau_{\pm}^3 - 2\hat{\tau}_{\pm}^2\tau_{\pm}^2) & , \ell \neq 0 \end{cases}$$

$$\sigma^2_{c_\pm(\ell)} = \begin{cases} \frac{8}{N}\ \tau_\pm^4 & \ell = 0 \\ \\ \frac{1}{N}\ \tau_\pm^4 & \ell \neq 0 \end{cases} \qquad \text{if} \quad \hat{\tau}_\pm = \tau_\pm \ .$$

For $c(\ell)$ we find

$$E\{c(\ell)\} = (\tau_+ - \hat{\tau}_+)(\tau_- - \hat{\tau}_-)$$

and

$$E\{c(\ell)^2\} = \frac{1}{(M_1-M_2+1)^2} \sum_{i=M_1}^{M_2} \sum_{j=M_1}^{M_2} E\{(y_+(i)y_-(i+\ell)y_+(j)y_-(j+\ell)\}$$

$$= \frac{1}{(M_1-M_2+1)^2} \left[(M_1-M_2+1)\ E\{y_+^2\}\ E\{y_-^2\} + \Big((M_1-M_2+1)^2 - (M_1-M_2+1)\Big)\ E\{y_+\}^2\ E\{y_-\}^2\right]$$

$$\sigma^2_{c(\ell)} = \frac{1}{M_1-M_2+1}\left[3\tau_+^2\tau_-^2 - 2\tau_+^2\hat{\tau}_-\tau_- + \hat{\tau}_+^2\hat{\tau}_-^2 - 2\hat{\tau}_+\tau_+\tau_-^2 + \hat{\tau}_+^2\tau_-^2\right]$$

$$\sigma^2_{c(\ell)} = \frac{1}{M_1-M_2+1}(\tau_+\tau_-)^2 \quad \text{if} \quad \hat{\tau}_\pm = \tau_\pm \ .$$

Figures 5.43,...,5.54 show $c_\pm(\ell)$ and $c(\ell)$ for both the 0.18 volt and 0.1 volt burst components along with the $\pm\ \sigma$ estimates. I can see no evidence of a dependency relationship between bursts. It is still possible that some weak dependency does exist which could be detected with a larger number of data points (thus decreasing σ). The results

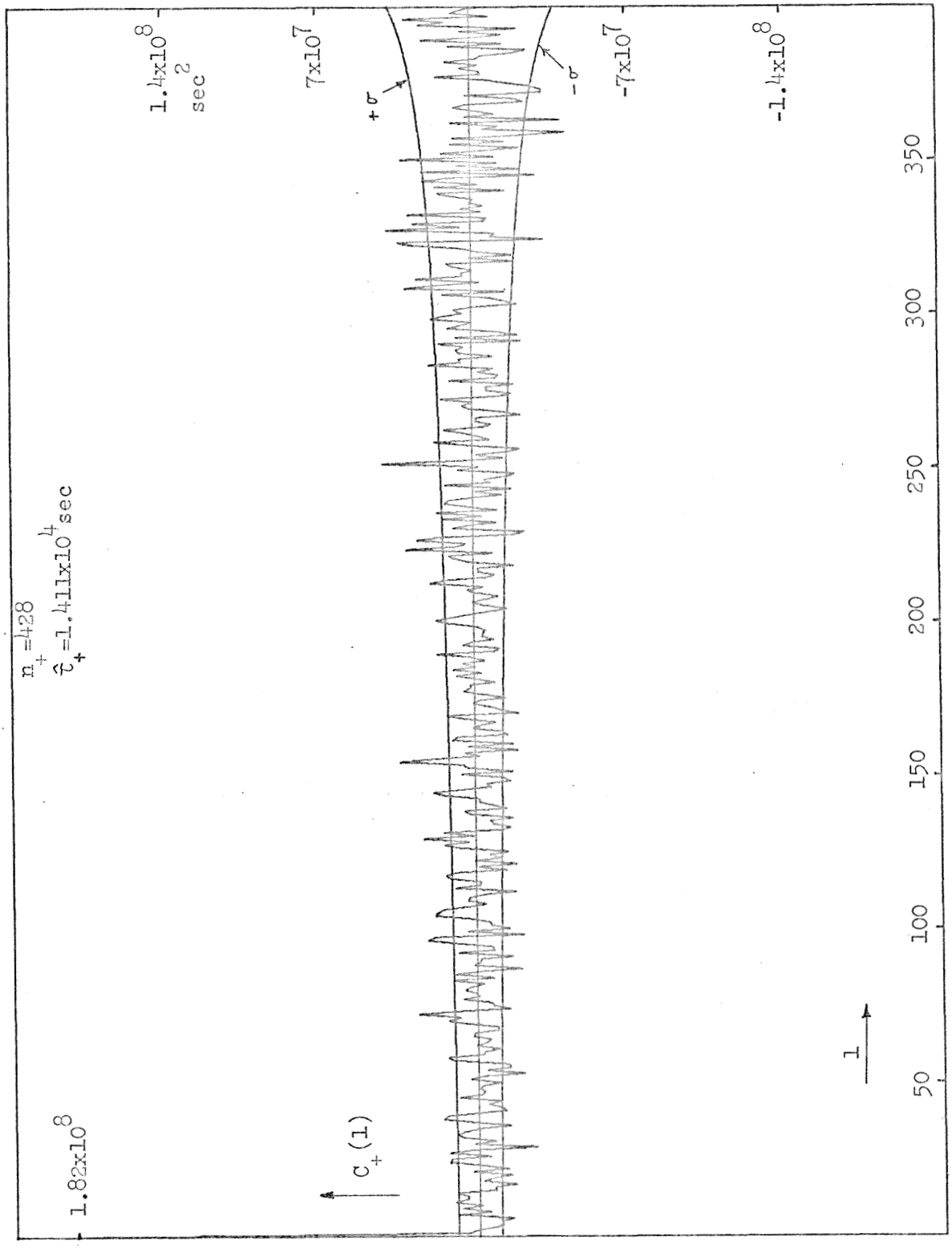


Figure 5.43 Mean lag products of estimated +0.18 volt burst waiting times for T4 noise source.

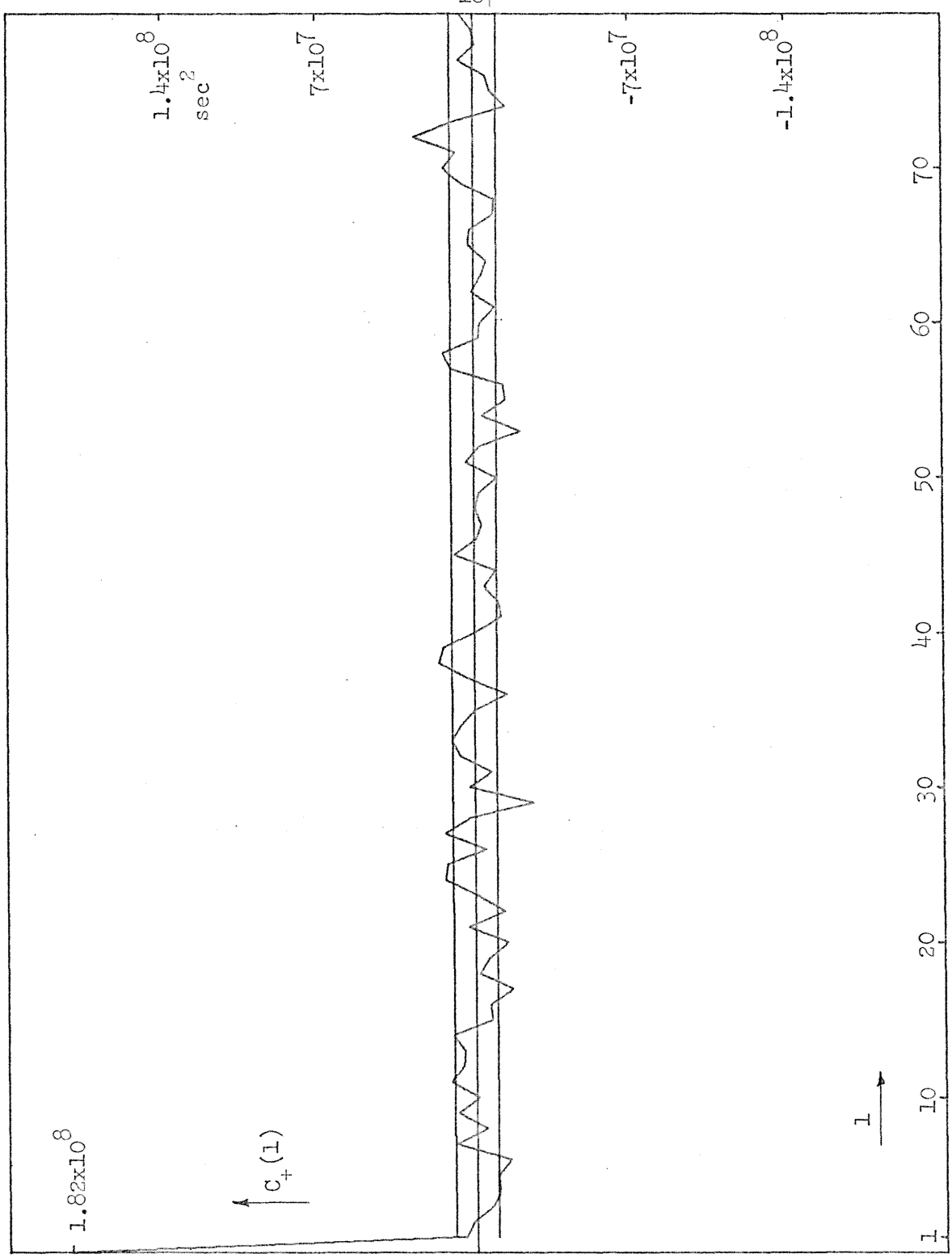


Figure 5.44 Mean lag products of estimated +0.18 volt burst waiting times for T4 noise source (expanded scale).

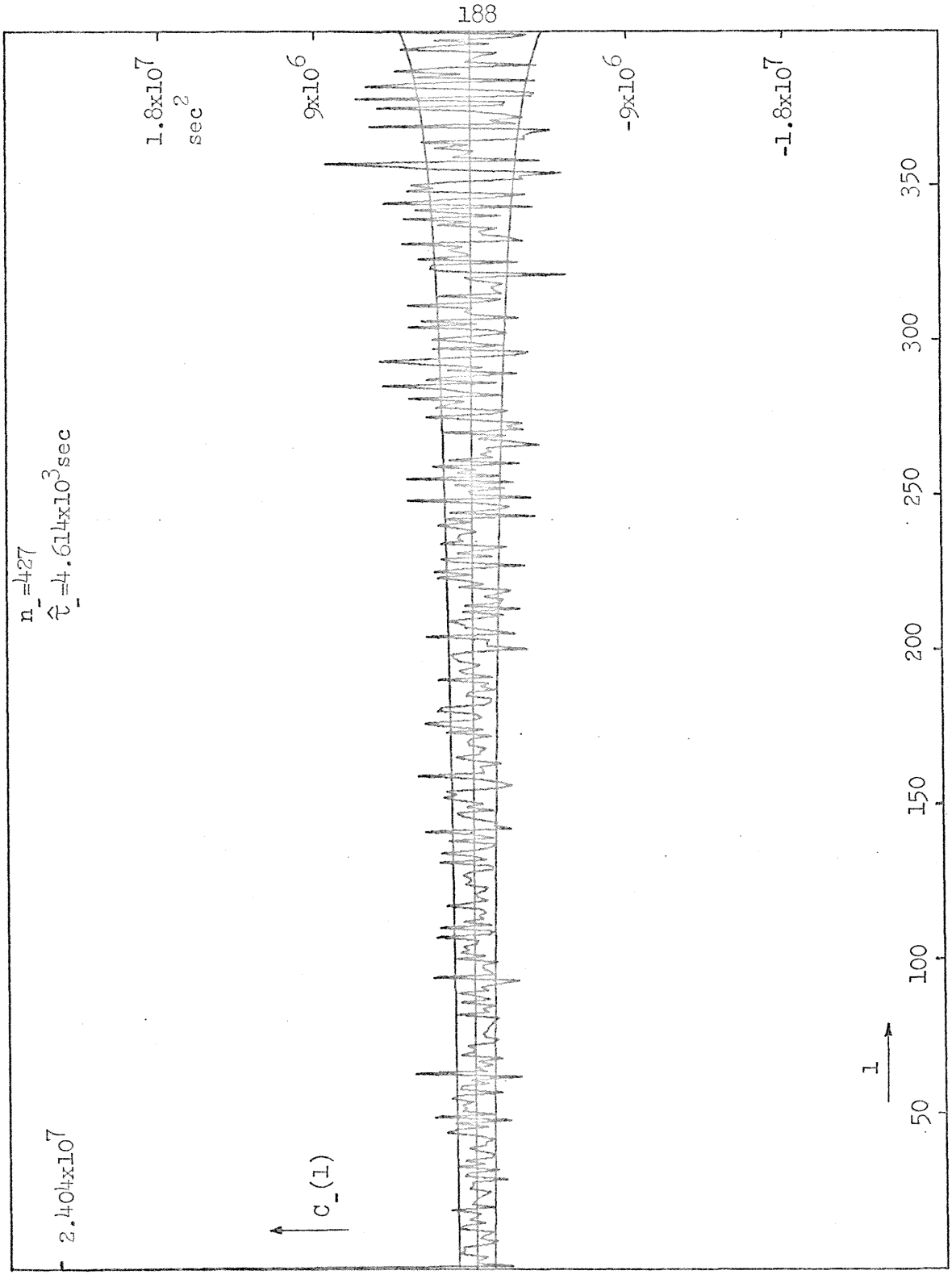


Figure 5.45 Mean lag products of estimated -0.18 volt burst waiting times for T4 noise source.

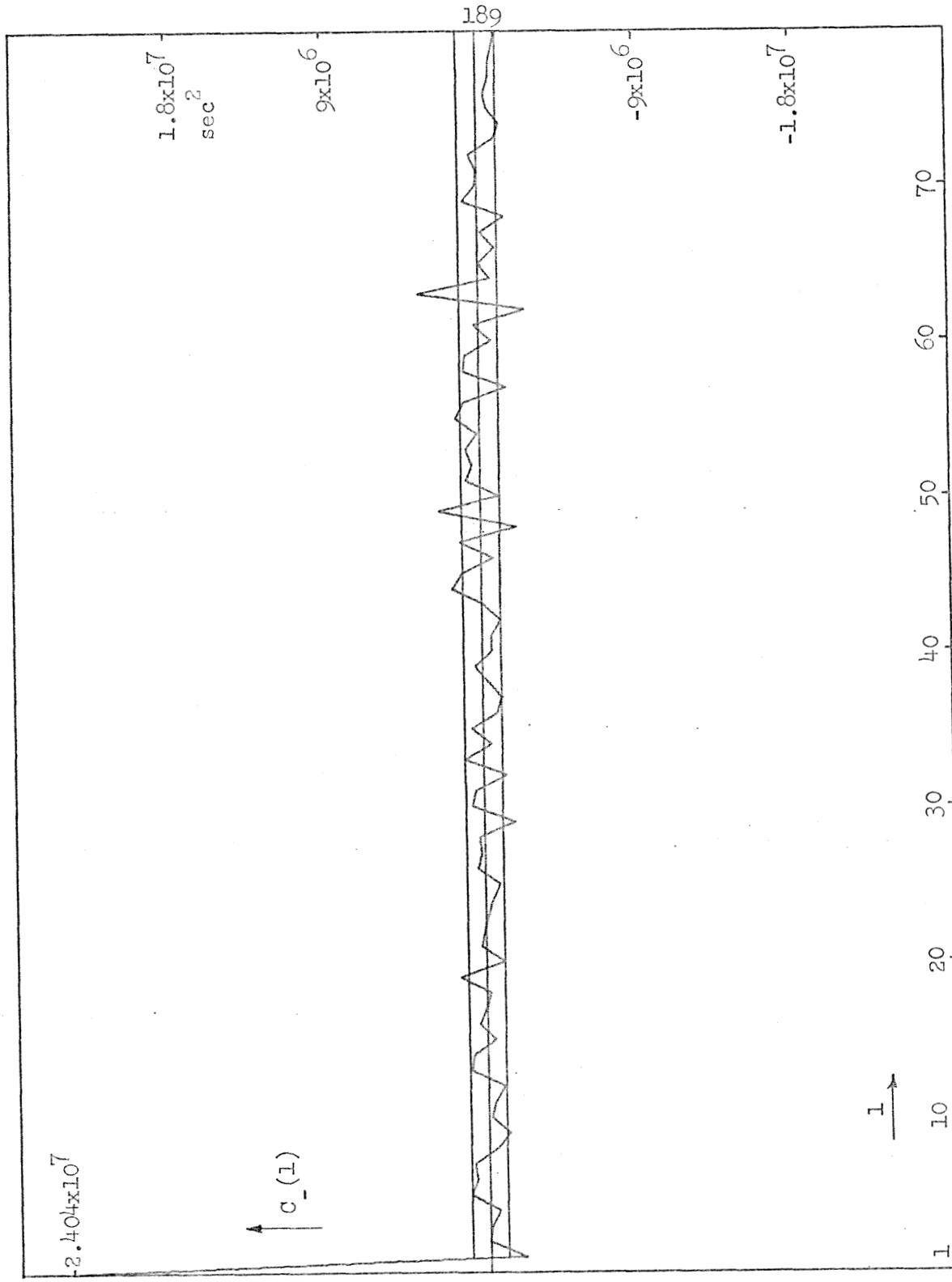


Figure 5.46 Mean lag products of estimated -0.18 volt burst waiting times for T4 noise source (expanded scale).

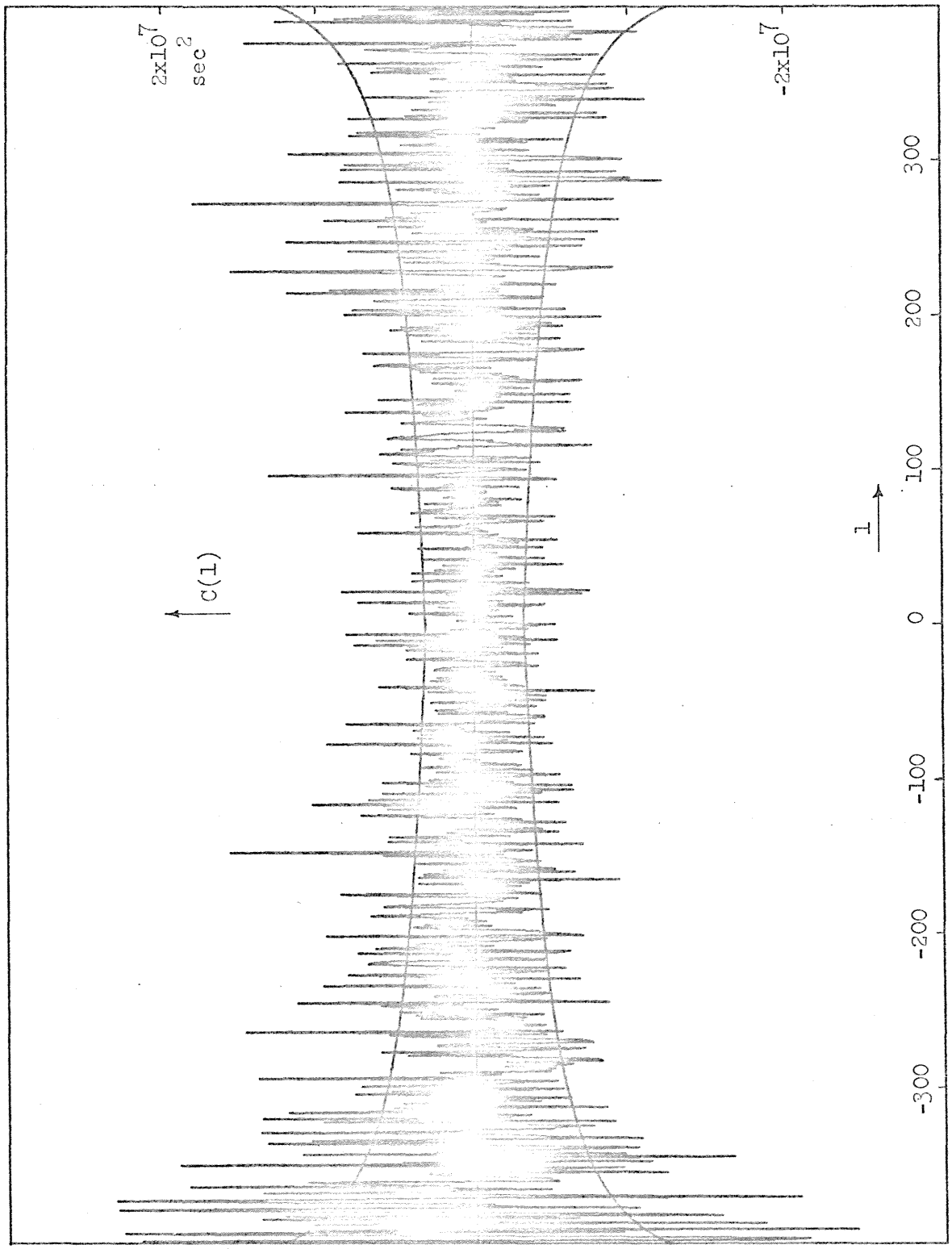


Figure 5.47 Mean lag cross products of estimated ±0.18 volt burst waiting times for T4 noise source.

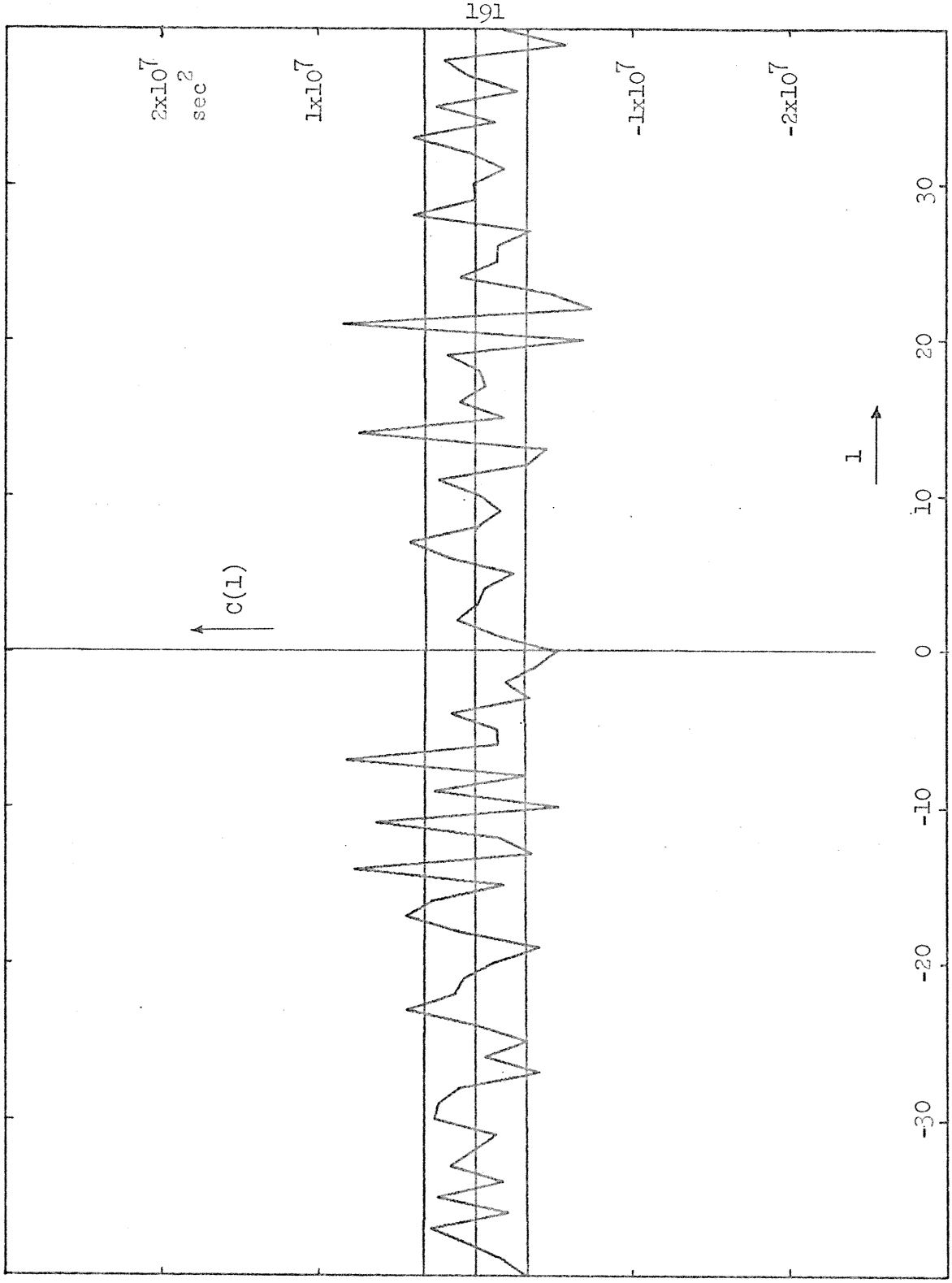


Figure 5.48 Mean lag cross products of estimated ±0.18 volt burst waiting times for T4 noise source (expanded scale).

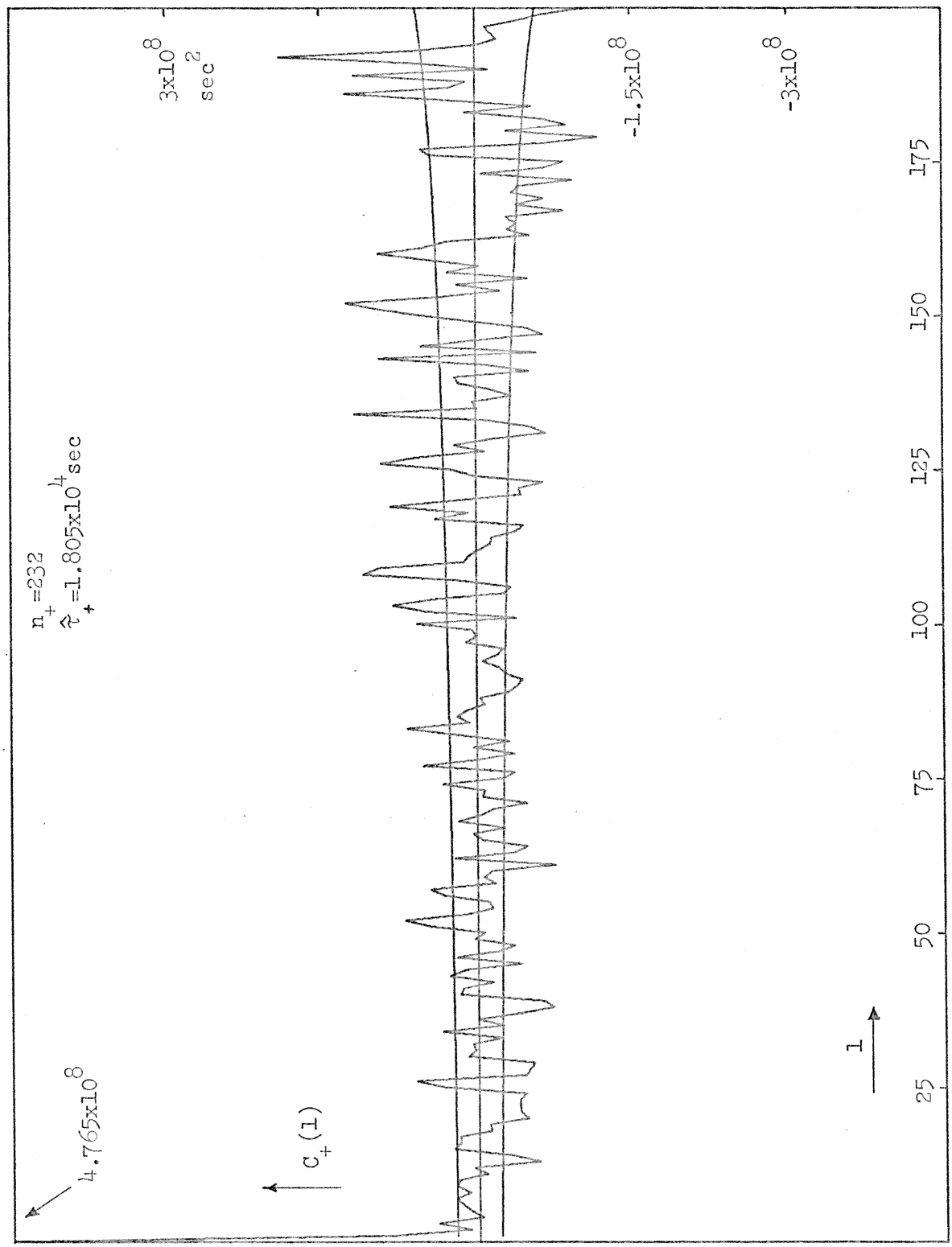


Figure 5.49 Mean lag products of estimated +0.1 volt burst waiting times for T4 noise source.

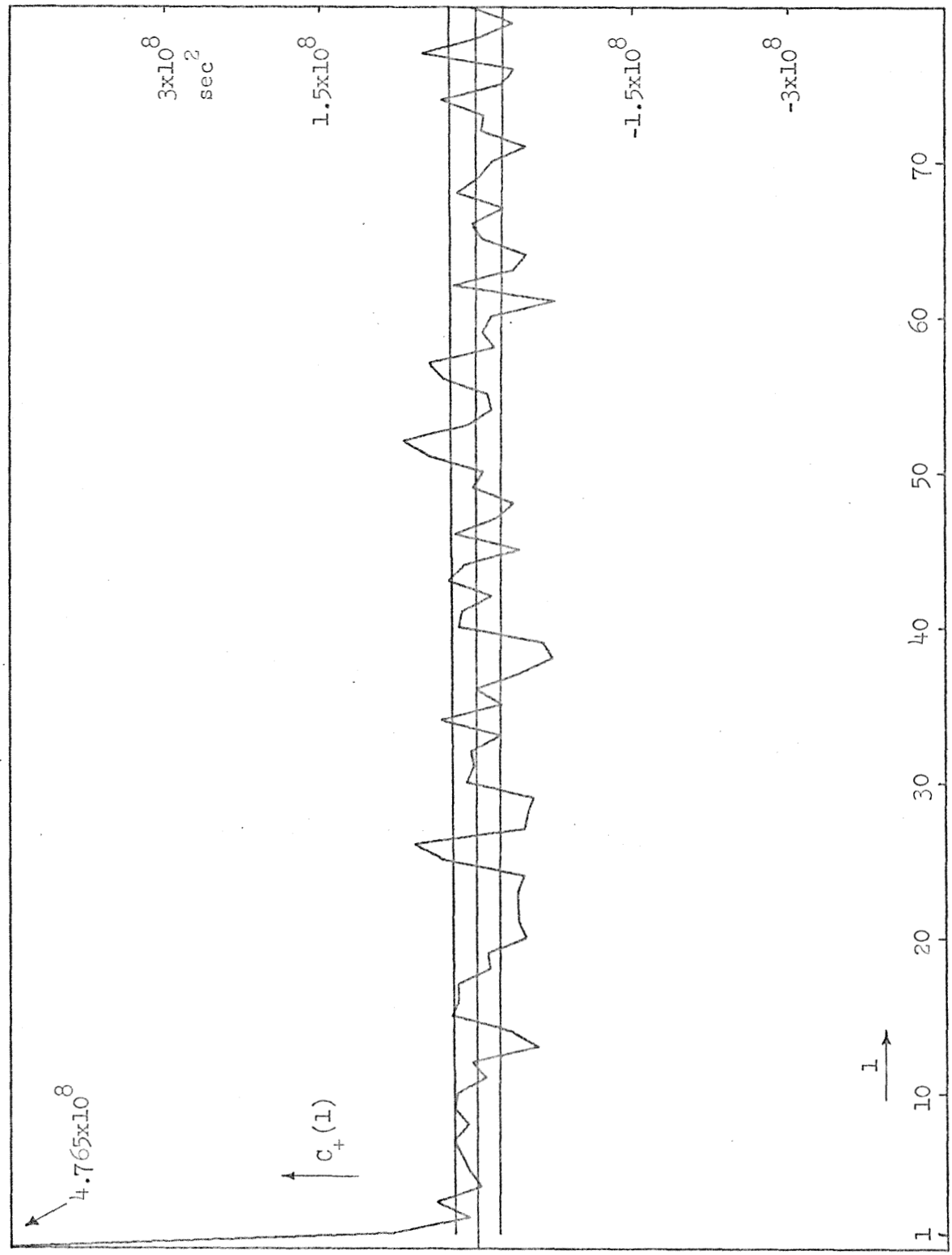


Figure 5.50 Mean lag products of estimated +0.1 volt burst waiting times for T4 noise source (expanded scale).

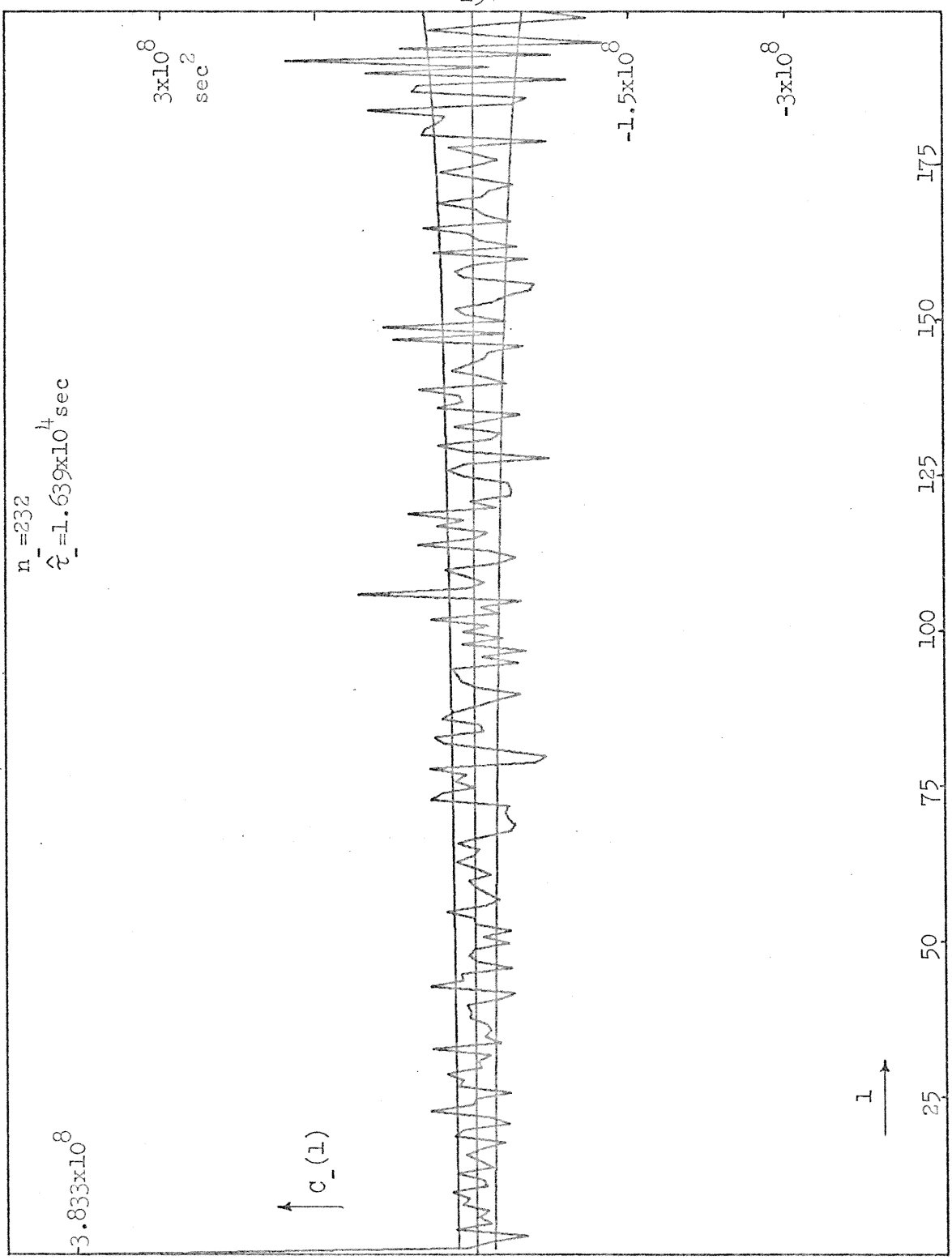


Figure 5.51 Mean lag products of estimated -0.1 volt burst waiting times for T4 noise source.

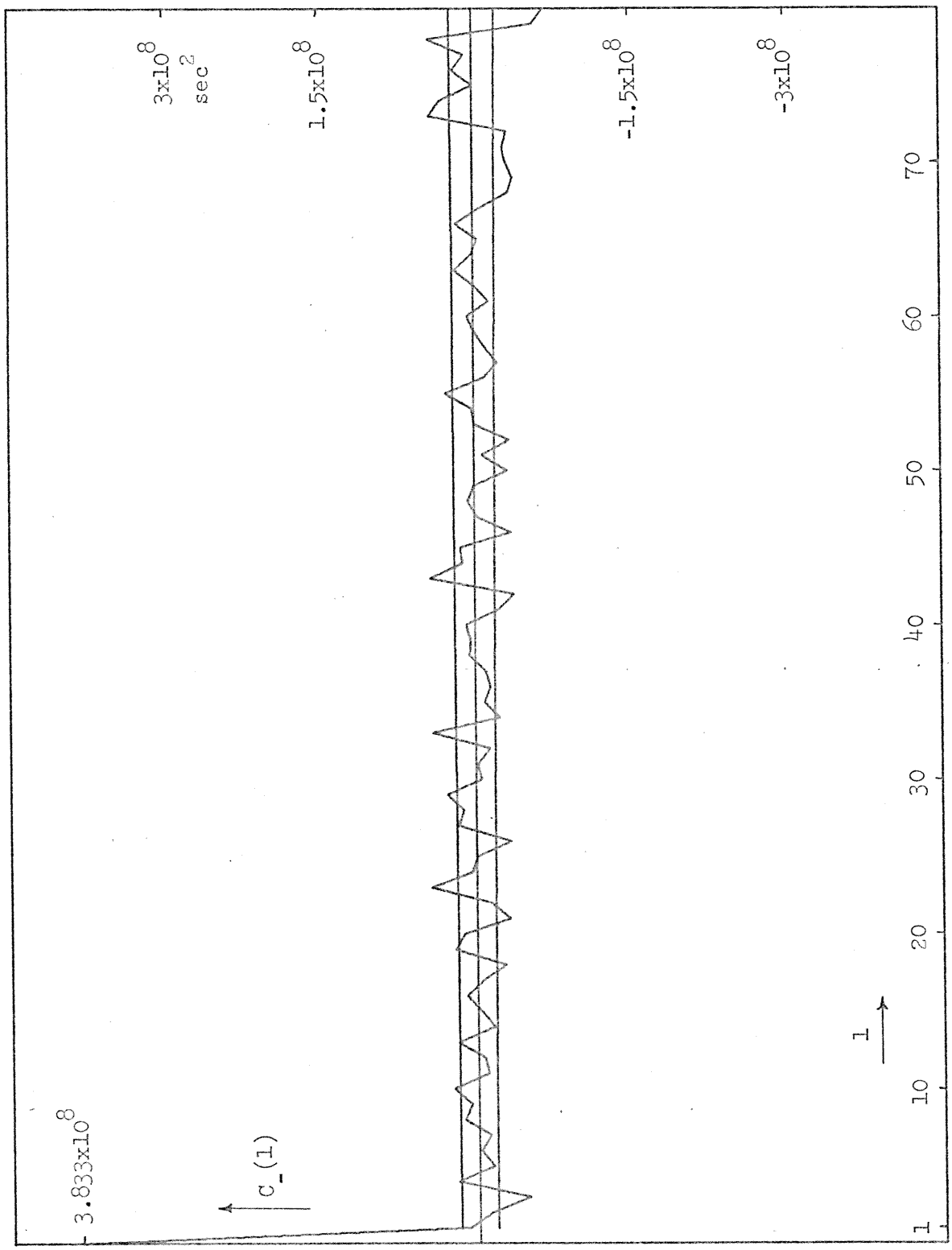


Figure 5.52 Mean lag products of estimated -0.1 volt burst waiting times for T4 noise source (expanded scale).

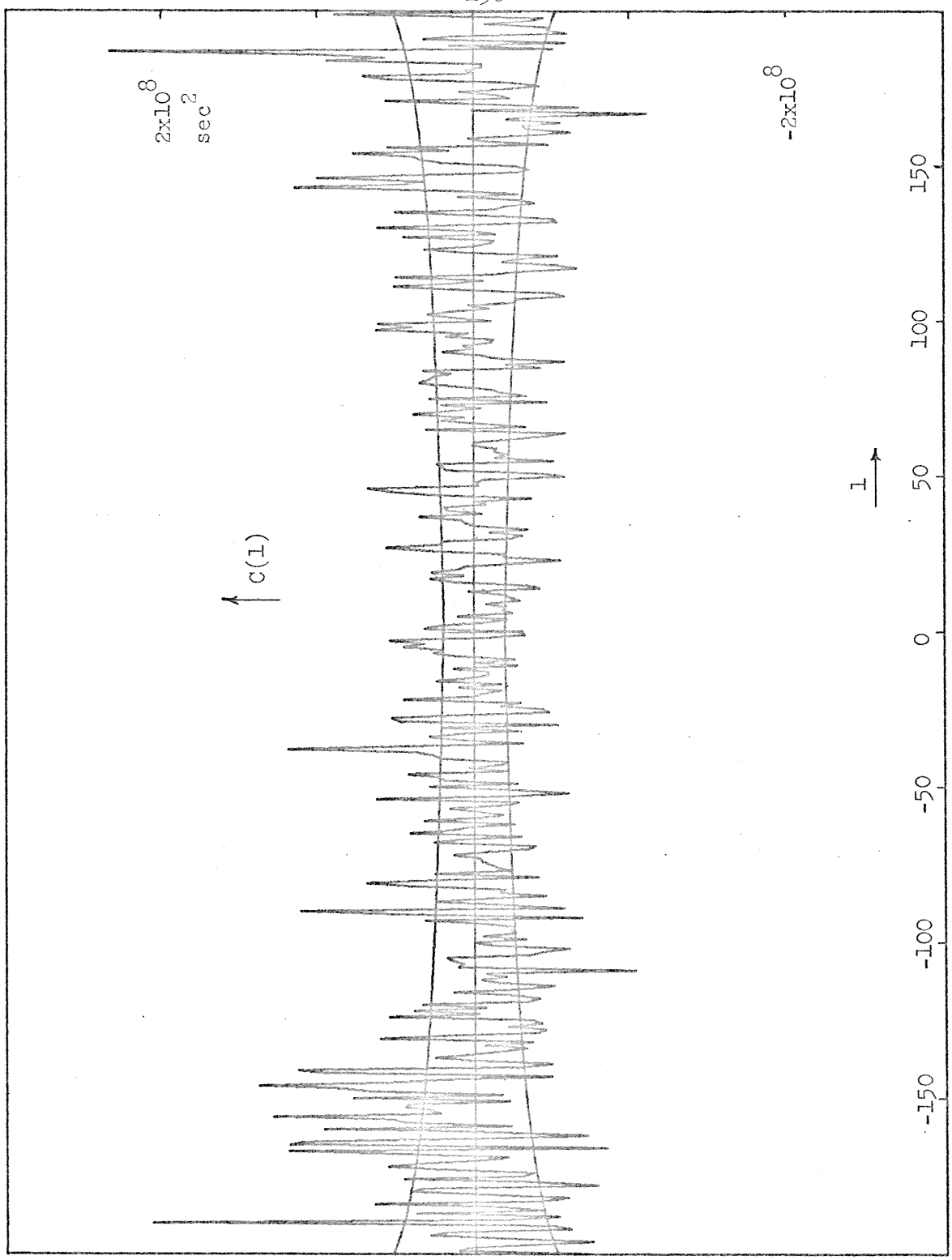


Figure 5.53 Mean lag cross products of estimated ±0.1 volt burst waiting times for T4 noise source.

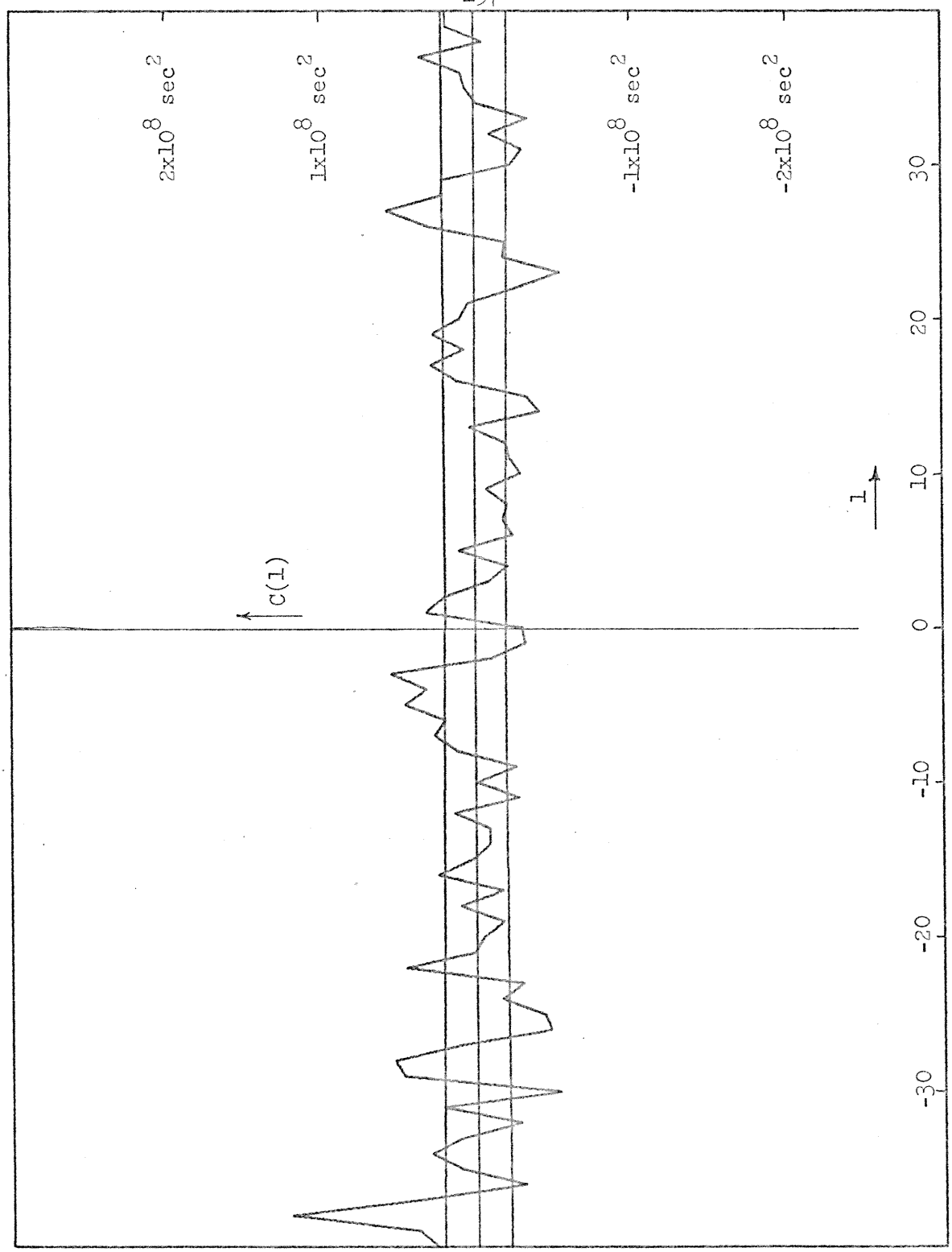


Figure 5.54 Mean lag cross products of estimated ±0.1 volt burst waiting times for T4 noise source (expanded scale).

are, therefore, inconclusive but tend to support the hypothesis of independence.

The three estimates which have been discussed for the two obvious burst components of noise source T4 will be referred to as:

1) $b_1(t)$ - The estimate with a step amplitude of 0.18 volts with occasional dual + or - steps resulting in more than two levels of amplitude.
2) $b_2(t)$ - The estimate with a step amplitude of 0.18 volts which is constrained to have only two amplitude levels.
3) $b_3(t)$ - The estimate with an amplitude of 0.09 volts which is constrained to have only two levels.

Letting the raw noise data be N(t), Figs. 5.55, 5.56, and 5.57 show the spectral density estimates for $N(t) - b_1(t)$, $N(t) - b_2(t)$, and $N(t) - b_2(t) - b_3(t)$ along with the spectral estimates for N(t) in each case.

From these estimates (and those in Figs. 5.32 and 5.34) it appears that removing the burst noise in the time domain does in fact allow the underlying noise process to be observed when the burst amplitude is large enough to estimate the transition times with high certainty. It is apparently useless, however, to attempt to remove the effects of a burst component which is difficult to distinguish by eye. The cancellation of burst noise effects may apparently be improved at the low frequency end of the spectrum by adding (or deleting) steps to insure that the estimate of the burst noise has only two levels.

No evidence has been found to suggest that the spectral density

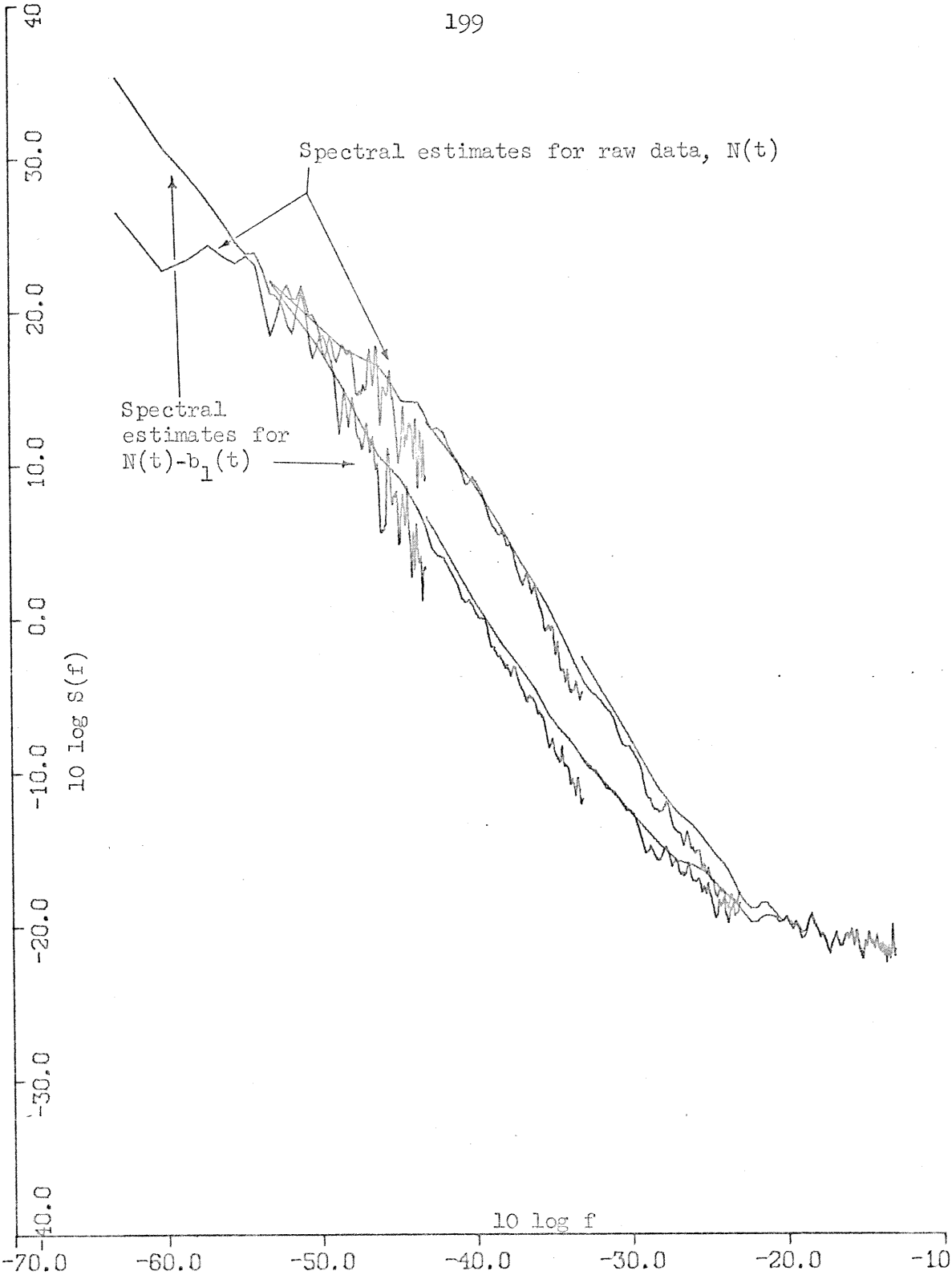


Figure 5.55 Spectral estimates of T4 noise source before and after subtracting $b_1(t)$ burst component estimate.

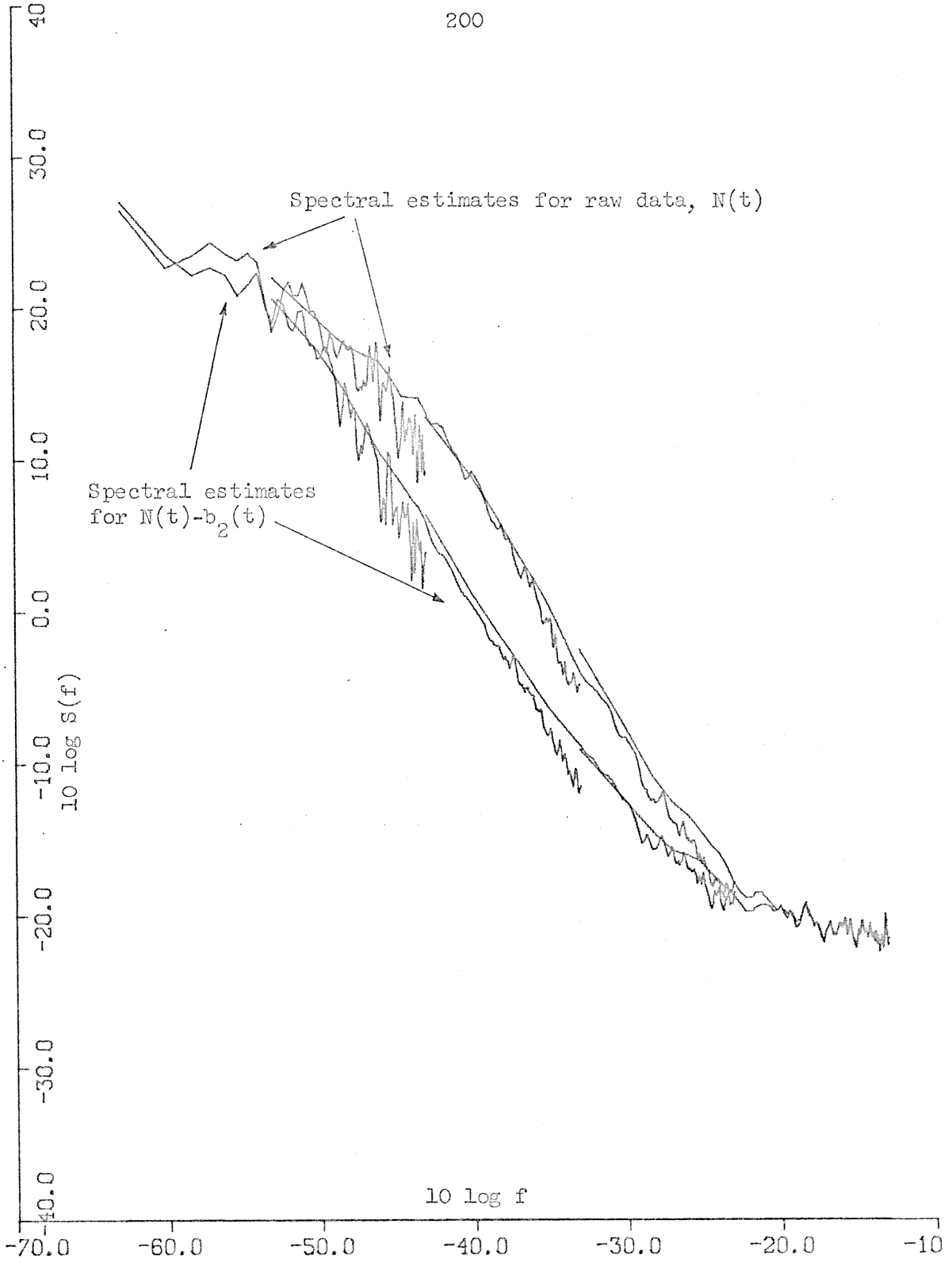


Figure 5.56 Spectral estimates of T4 noise source before and after subtracting $b_2(t)$ burst component estimate.

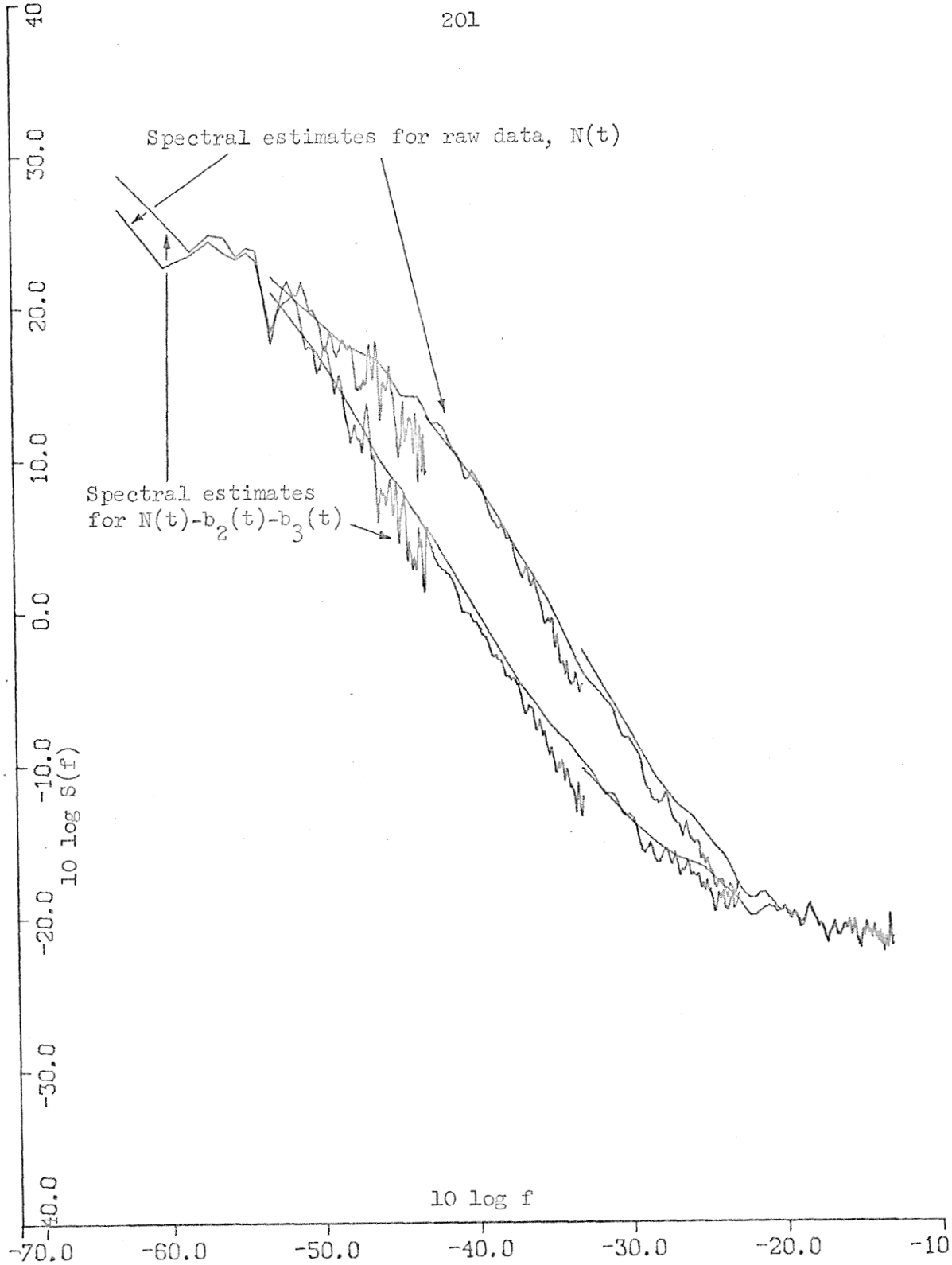


Figure 5.57 Spectral estimates of T4 noise source before and after subtracting $b_2(t)$ and $b_3(t)$ burst component estimates.

of a burst component is other than that derived by Puckett[3] or that the spectral density estimates of the underlying process plus a burst component are not approximately the sum of the spectral density estimates of the components. If one assumes that the underlying process has a spectral density of $S(f) = |f^{-\alpha}|$ with $\alpha \sim 1$ it should be sufficient to remove burst components only with very long and very short time constants to establish density estimates at either end of the frequency spectrum in order to measure the value of α.

Chapter VI

HIGH SAMPLE RATE DATA

6.1 Time Domain Data

As was mentioned previously, the digital voltmeter (DVM) is not capable of taking data with a sampling period much shorter than Δt = 1 second. This limits the sample period for each source to Δt = 10 seconds when multiplexing ten channels. During the first phase of the experiment the sources were sampled at this rate with a presampling filter cutoff frequency (-3 db) of 0.34 Hz. This results in aliasing errors for the highest frequency estimates of an undetermined magnitude.

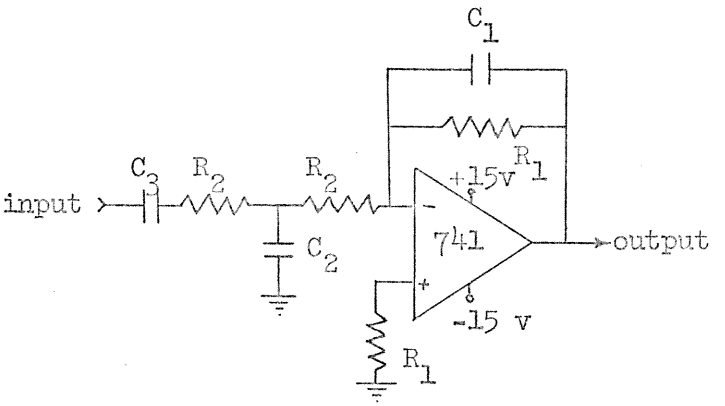
To obtain estimates in the region up to 0.5 Hz, the sources could have been separately sampled at Δt = 1 second by the DVM either before or after the multiplexed data. For sampling rates faster than 1 sample per second, however, the necessary record length is less than a few hours (to obtain 10% estimates down to .005 Hz we need 10^4 seconds = 2.78 hours). The requirements for multiplexing and reliability having been removed, it was a simple task to assemble the necessary equipment to make these measurements in parallel with the lower frequency measurements. The equipment consisted of an integrated circuit amplifier with high input impedance (so as not to appreciably affect the loading of the noise source channels when attached) with a two pole RC presampling filter. For the higher sample rate measurements a series capacitor was inserted so that the gain of the amplifier could be increased to 10 (this decreases the quantization noise by 20 db). An

oscillator set to about 1% accuracy controlled the sampling rate of a 10 bit analog to digital (A/D) converter with a range of ±5 volts. The output of the A/D converter was divided into two parts, each of which was recorded on one frame of a 7 track, IBM compatible, write only magnetic tape transport. Since the maximum recording rate of the tape transport was 10^3 frames/second, the maximum sampling rate was 500 samples/second.

The data were recorded during the period of August 10, 1973 to August 15, 1973. 10^4 samples were recorded for each noise source channel at sample rates of 1.0, 10.0, and 500 samples/second. The buffer amplifier/pre-sampling filter used for each of the sample rates is shown in Fig. 6.1. Figure 6.2 illustrates the transfer function of this filter as implemented at the various sample rates.

Figures 6.3, 6.4, and 6.5 show the first one tenth of each data record at a density of 125 points per inch. Figures 6.6, 6.7, and 6.8 show all of the 10^4 data points in each sample record at a density of 1250 data points per inch.

An interesting feature is observed in the T4 noise source plot of Fig. 6.6. There appears to be a relatively high frequency popcorn component which is active only during widely spaced intervals. This component is not observed in Fig. 6.7 (apparently this record was made during a quiet interval), but is again observed in Fig. 6.8 as a classical popcorn noise with an apparent time constant of several seconds (the slope in the popcorn wave form is accounted for by the high pass presampling filter which has a time constant of 1.6 seconds).



sample rate (Hz)	R_1 (KΩ)	R_2 (KΩ)	C_1 (μf)	C_2 (μf)	C_3 (μf)
1.0	162	82.5	1.27	5.10	shorted
10.0	1665.0	82.5	.0133	0.481	9.784
500.0	1665.0	82.5	249 pf	0.0099	9.784

Figure 6.1 High sample rate buffer amplifier / presampling filter.

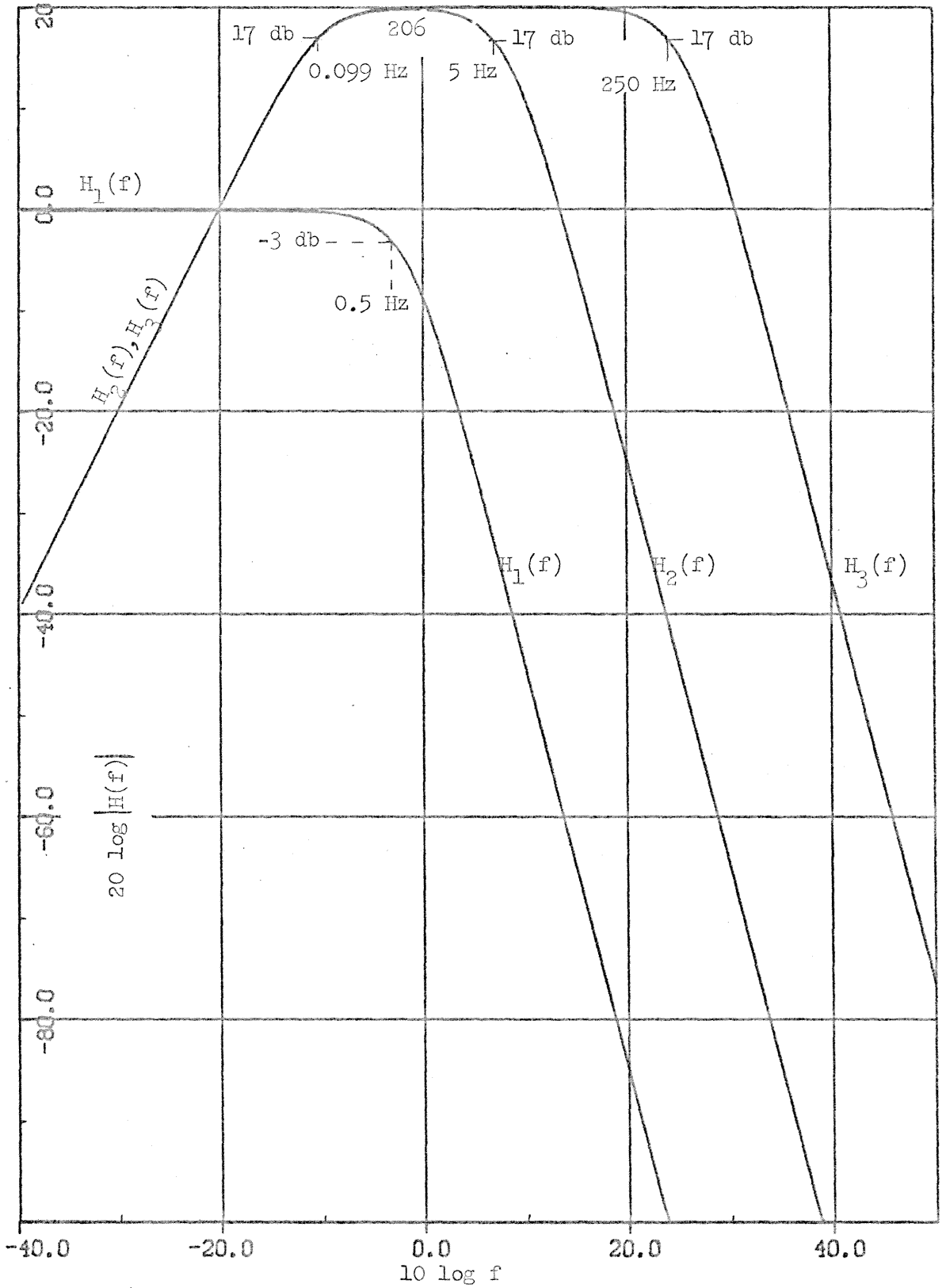


Figure 6.2 Magnitude of presampling filter transfer functions used for Δt=1.0, 0.1, and 0.002 second sampling ($H_1(f)$, $H_2(f)$, and $H_3(f)$).

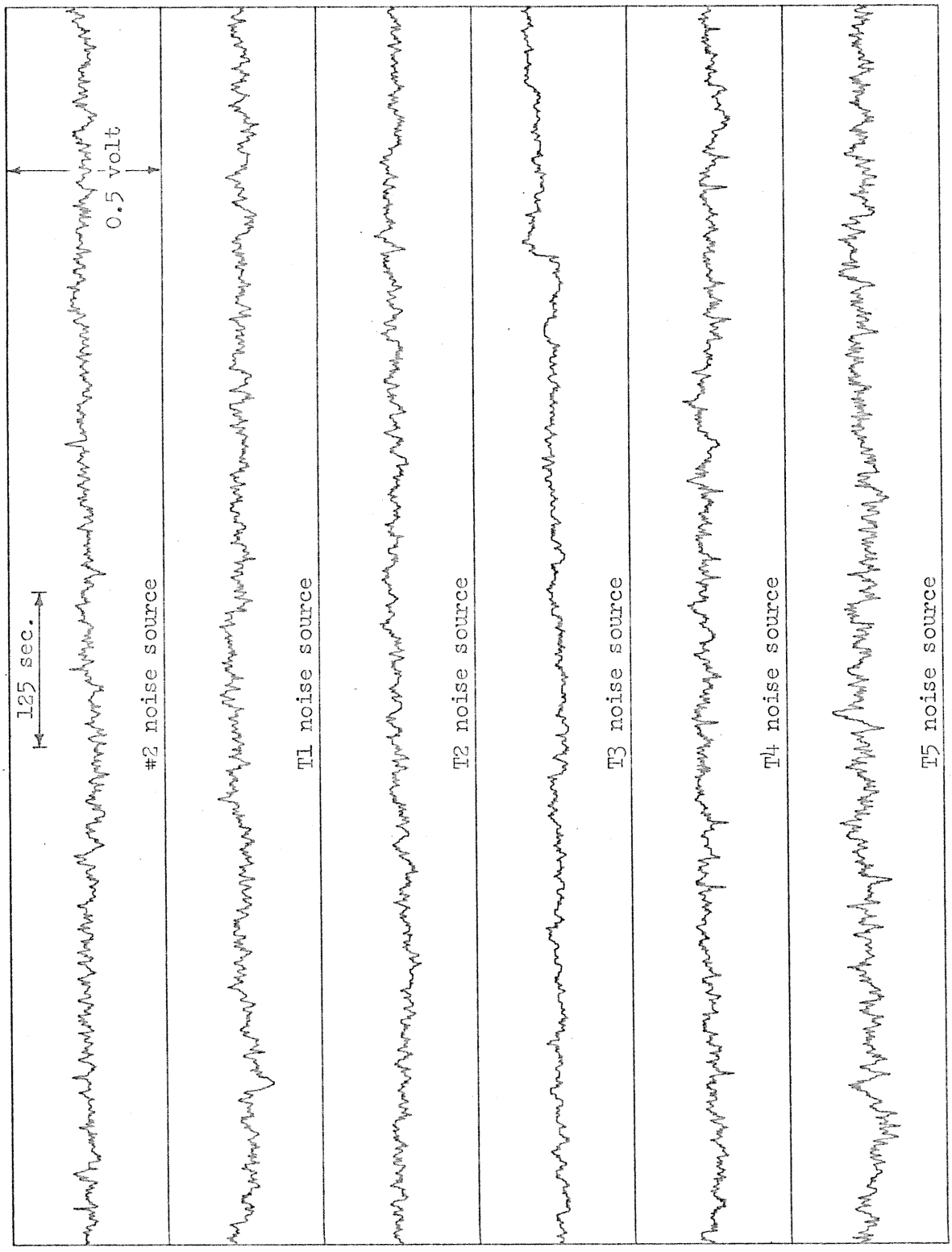


Figure 6.3 Time domain data (1,000 points, Δt=1 sec., 0.5 volt full scale vertical).

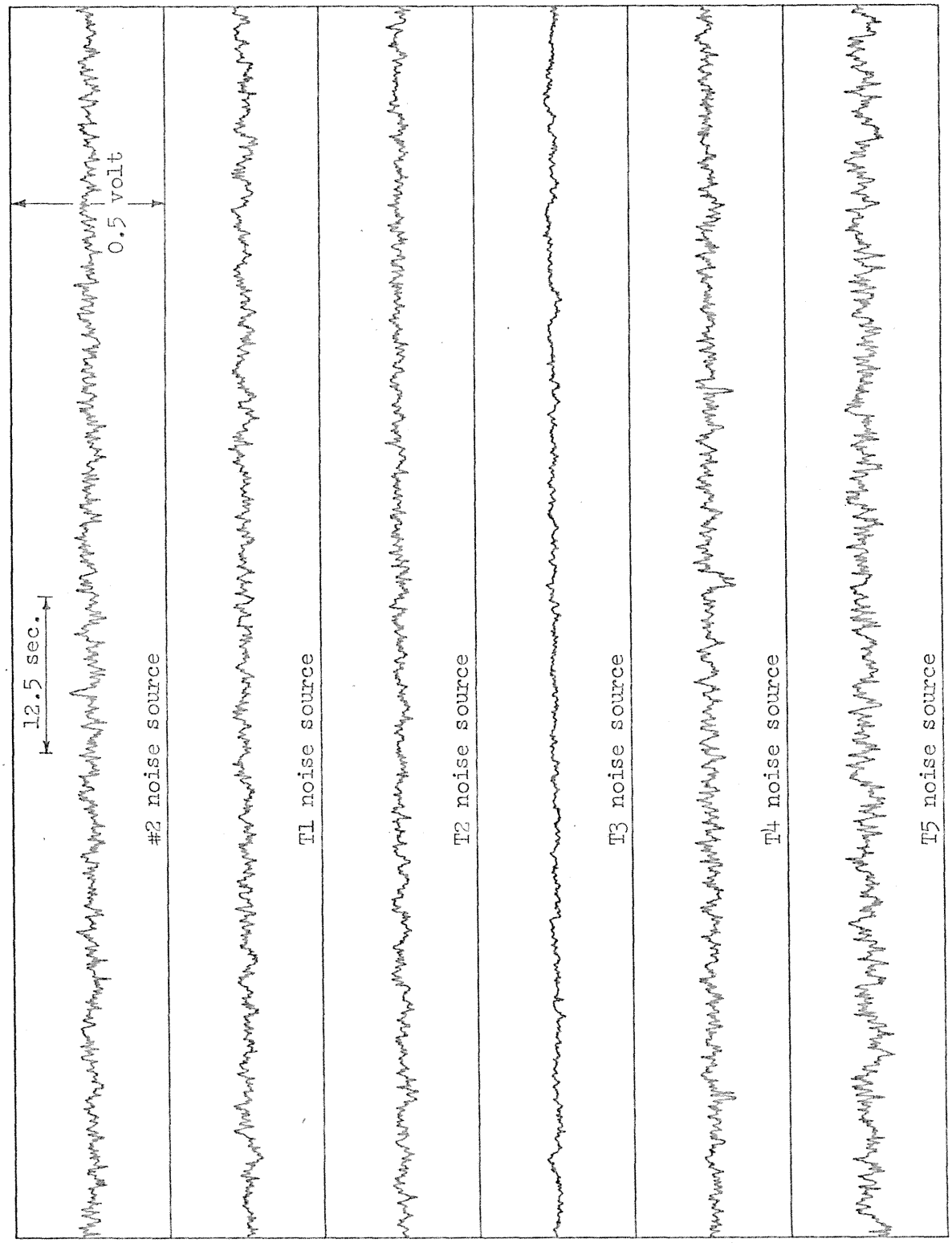


Figure 6.4 Time domain data (1,000 points, Δt=0.1 sec., 0.5 volt full scale vertical).

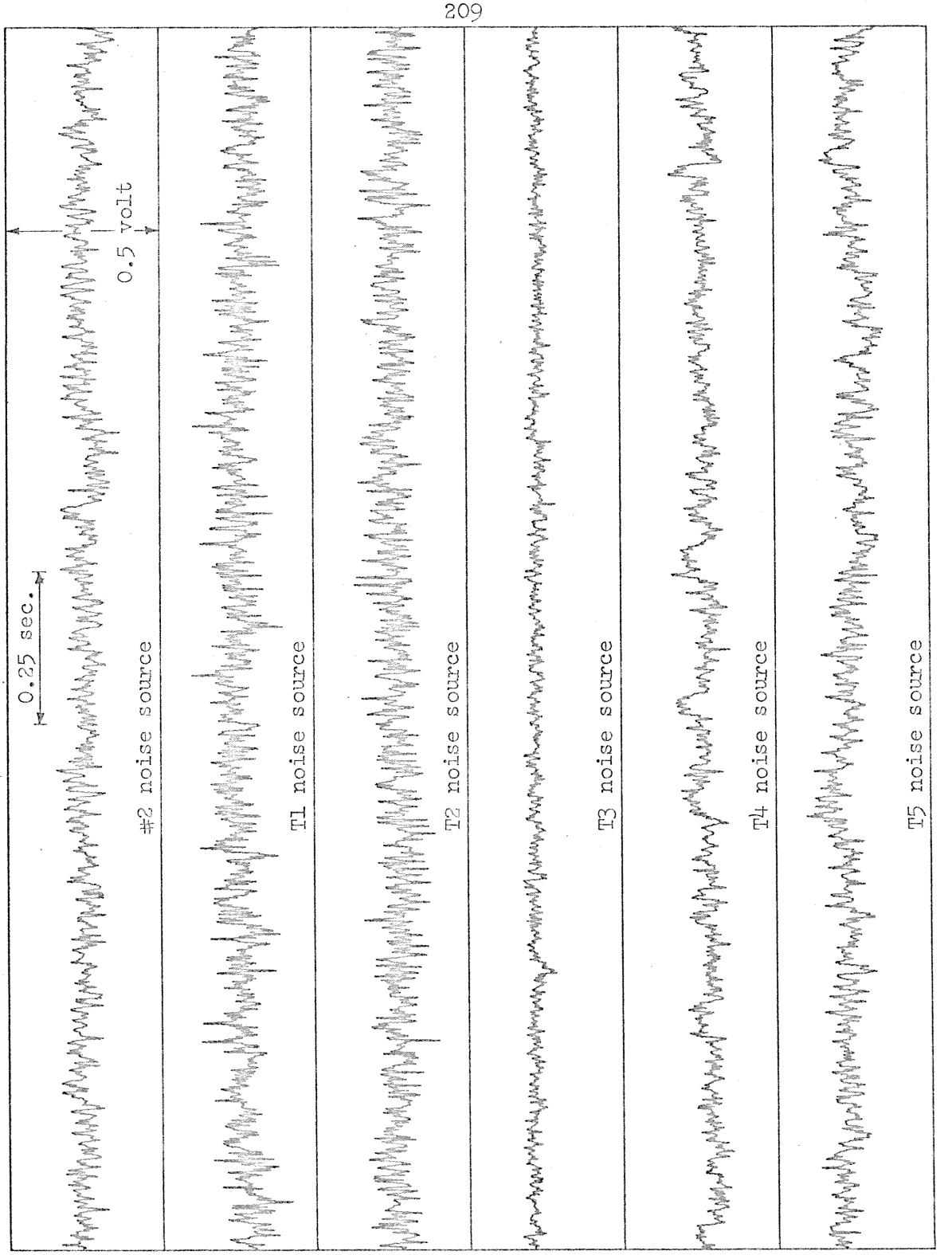


Figure 6.5 Time domain data (1,000 points, Δt=0.002 sec., 0.5 volt full scale vertical).

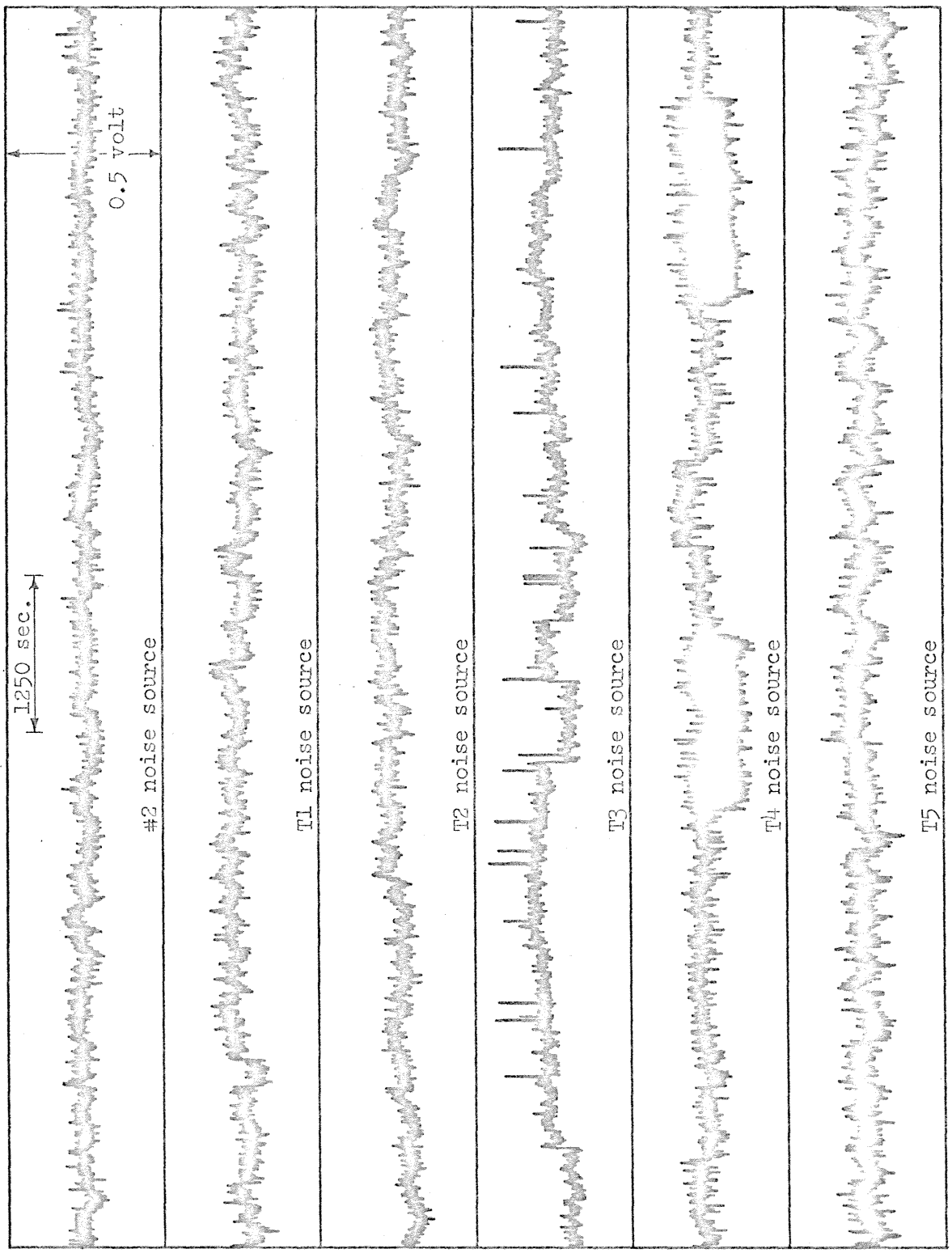


Figure 6.6 Time domain data (10,000 points, Δt=1 sec., 0.5 volt full scale vertical).

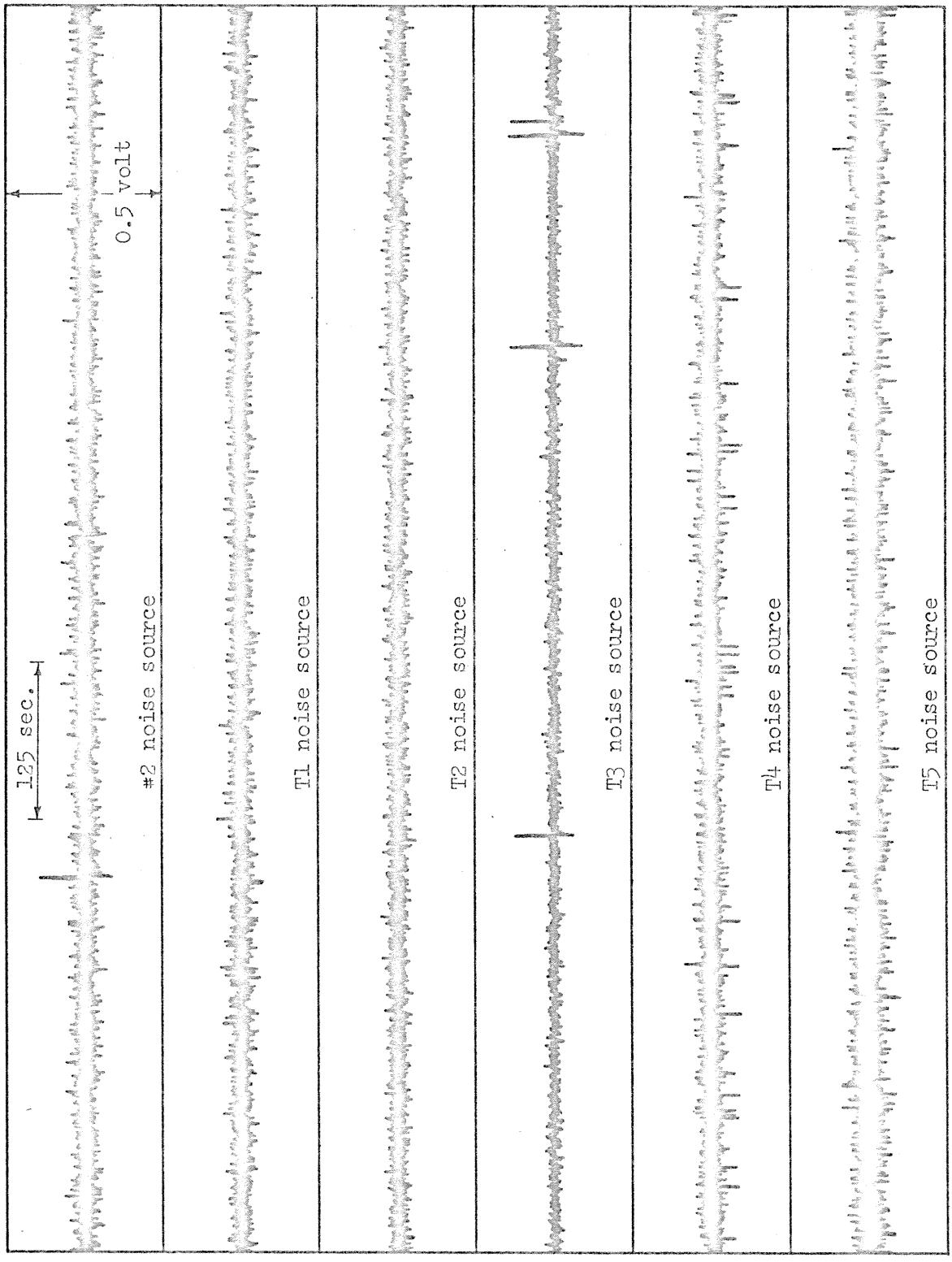


Figure 6.7 Time domain data (10,000 points, Δt=0.1 sec., 0.5 volt full scale vertical).

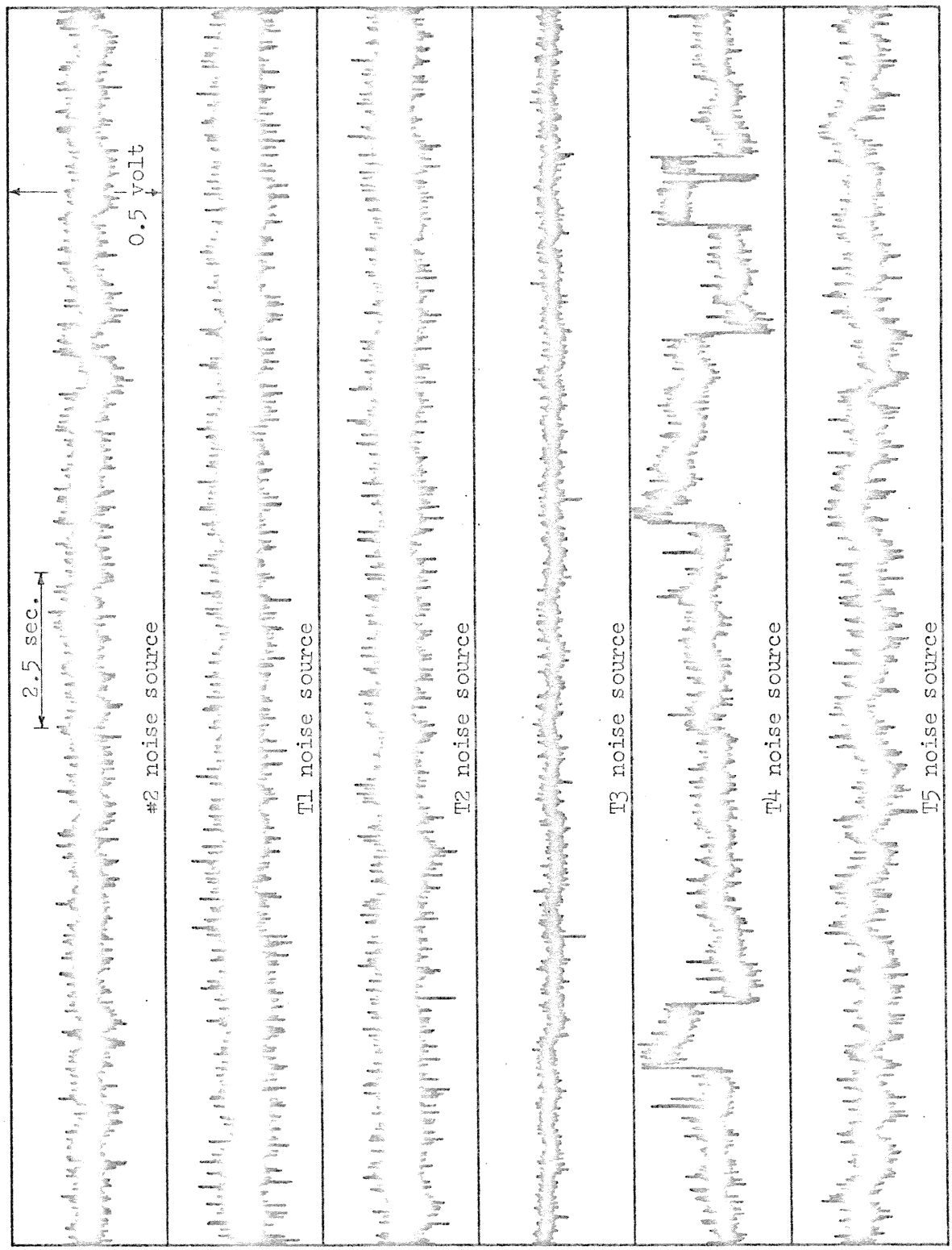


Figure 6.8 Time domain data (10,000 points, Δt=0.002 sec., 0.5 volt full scale vertical).

One possible explanation for this phenomenon is that some of the very rare bursts with waiting times several orders of magnitude longer than the time constant of the burst component are being observed. If a time constant of 5 seconds for both the "on" and "off" states is assumed (see Fig. 6.8) then the exponential distribution of waiting times predicts that the probability of a burst with a waiting time of greater than 2000 seconds is $e^{-\frac{2000}{5}} = 2 \times 10^{-174}$. Since the average burst rate is assumed to be 0.2/second, we would expect to find a burst of this length once each 2.5×10^{174} seconds (8×10^{166} years).

Since we observe at least two of these quiet periods in Fig. 6.6, it is quite reasonable to assume that this burst component does not fit the simple model. In addition, there appears to be a level shift between the noisy and quiet periods (i.e., neither the "on" nor "off" state during the noisy regions lies at the level one might expect by extrapolating from the quiet regions).

Having made this observation, it seems appropriate to diverge a bit more to see if there is any other evidence of burst components not fitting the simple model. Figure 5.7 reveals that noise source #2 has a questionable wave form. Plotting this source at $\Delta t = 100$ seconds (not shown here) clearly demonstrates the following series of burst waiting times:

$$t_- \geq 6.1 \times 10^6 \text{ seconds}$$
$$t_+ = 3.19 \times 10^5$$
$$t_- = 1.82 \times 10^4$$
$$t_+ = 2.1 \times 10^5$$

$t_- = 1.3 \times 10^3$

$t_+ = 1.3 \times 10^5$

$t_- = 8 \times 10^2$

$t_+ = 3.12 \times 10^4$

$t_- = 2.0 \times 10^2$

$t_+ = 1.377 \times 10^5$

$t_- = 6.74 \times 10^6$

The following question arises concerning the series of negative burst waiting times. How likely is a sequence of waiting times with this wide a distribution if the waiting times satisfy an exponential distribution and are independent? Let $P(\tau)$ be the probability of finding i waiting times $\leq$ x seconds and j waiting times $\geq$ y seconds out of a sequence of i + j bursts given that the waiting time probability distribution is $p(t) = \frac{1}{\tau} e^{-t/\tau}$.

$$P(\tau) = \left(1 - e^{-x/\tau}\right)^i \left(e^{-y/\tau}\right)^j \frac{(i+j)!}{i!j!} \quad .$$

Since the value of τ is not known, we will maximize $P(\tau)$ over all τ .

$$\frac{dp(\tau)}{d\tau} = P(\tau) \frac{-ix\, e^{-x/\tau} + jy(1 - e^{-x/\tau})}{\tau^2 (1 - e^{-x/\tau})}$$

$$\frac{dP(\tau)}{d\tau} = 0 \text{ implies } \frac{1}{\tau} = \frac{1}{x} \ln \left(1 + \frac{ix}{jy}\right)$$

$$\max_{\tau} P(\tau) = \left(\frac{ix}{ix+jy}\right)^i \left(\frac{jy}{ix+jy}\right)^{j\frac{y}{x}} \frac{(i+j)!}{i!\, j!}$$

If we let, for the above negative bursts, $x = 1.82 \times 10^4$ seoncds, $i = 4$, $y = 6.1 \times 10^6$ seconds, $j = 2$ we get

$$\max_{\tau} P(\tau) = 3.4 \times 10^{-10} \text{ for } \tau = 3.06 \times 10^6 \text{ seconds}$$

This, therefore, is a second example of a popcorn noise component having a very low probability of fitting the simple model.

One hypothesis which could explain this type of abnormal behavior is that burst components may modulate each other as well as add. There are undoubtedly other possible explanations for this form of abnormal behavior. This should also serve as a warning that there may be other forms of behavior (which have been overlooked in the current observation of only 6 noise sources) which do not fit the simple model which has been presented.

6.2 High Frequency Spectral Estimates

The power spectral density estimates which were made using these data are shown in Figs. 6.9,...,6.14. In Fig. 6.9 one will note the 60 Hz spectral line which is apparently due to pick up in the long unshielded wire used in measuring this signal. For all six sources the 0.5 to 5 Hz estimate levels off at about .1 Hz. This is due to the high-pass filter used when recording the Δt = 0.1 and 0.002 second data. The apparently low estimates for T4 in Fig. 6.13 for the .05 to 5 Hz range are due to the already discussed non-constant burst component. This component was later subtracted out (an estimated burst component was first modified to that which would result from passing an ideal burst component through the high pass filter used in recording the original data) for the Δt = 0.002 sec data. This results in matching the 2.5 Hz to 250 Hz and .05 to 5 Hz estimates as can be seen in Fig. 7.7. No attempt was made to remove this component from the Δt = 1 second data since the burst component becomes too distorted by the low pass presampling filter to easily estimate and remove it.

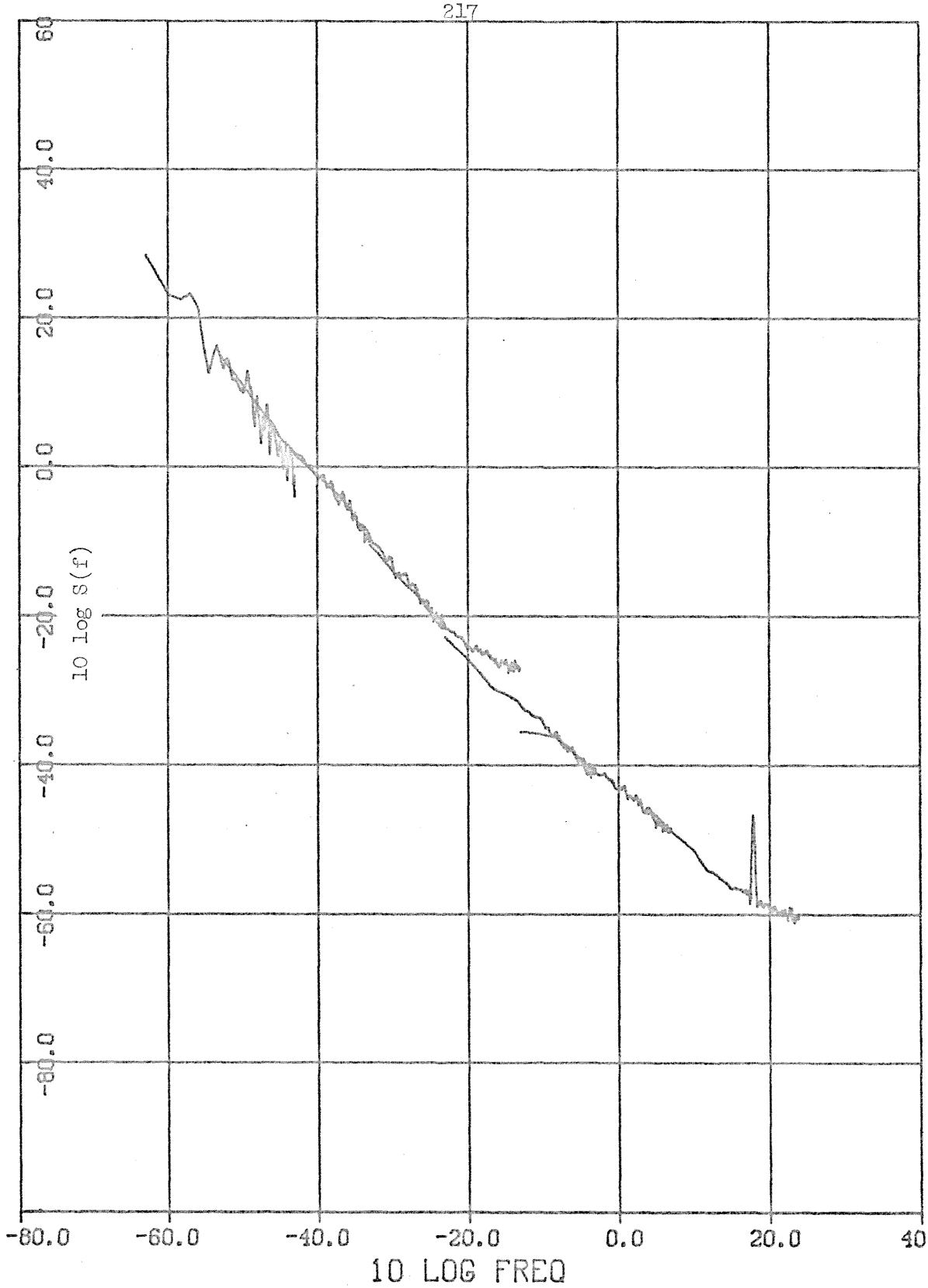


Figure 6.9 Spectral estimates for #2 noise source ($-63 \leq 10 \log f \leq 24$).

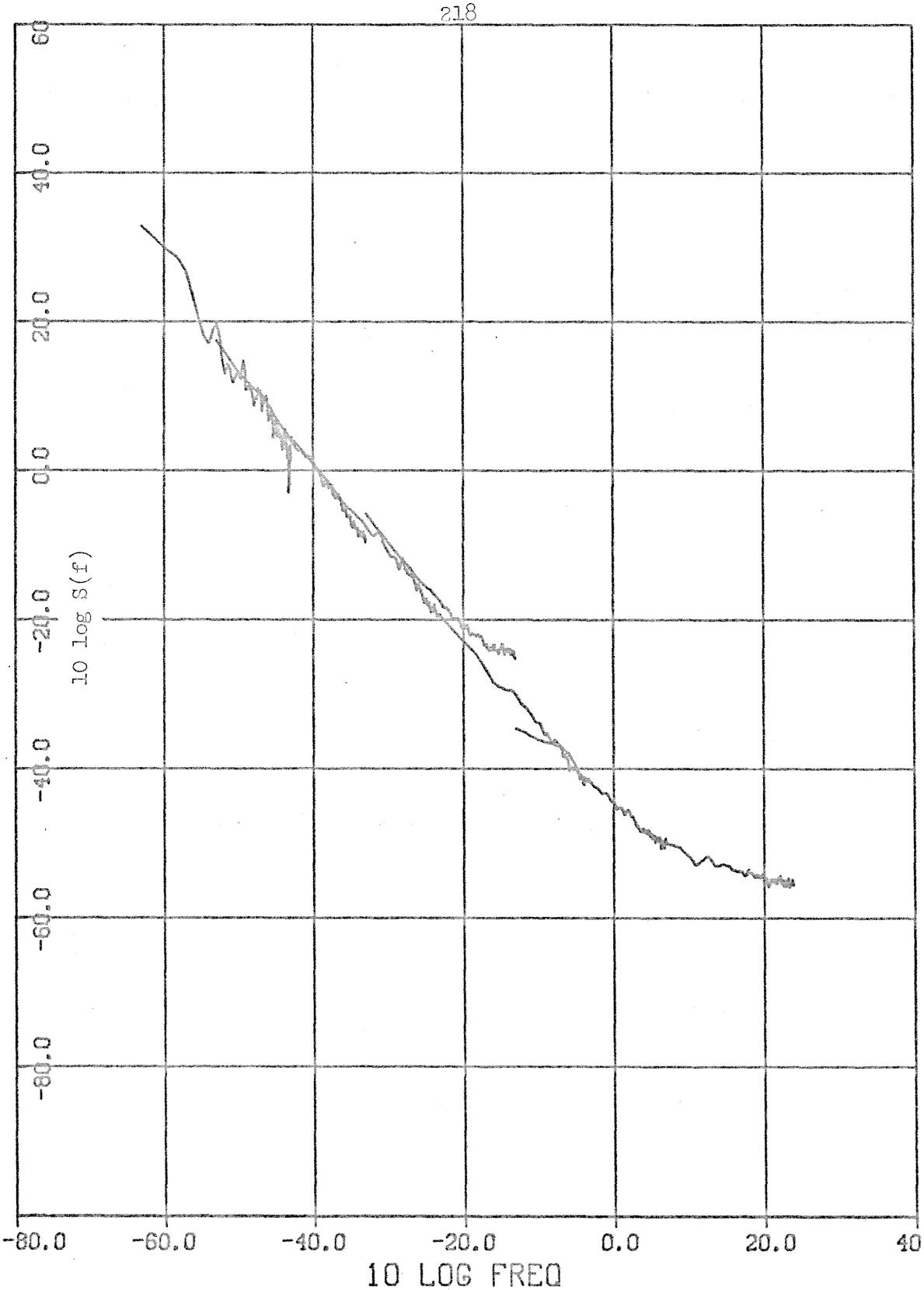


Figure 6.10 Spectral estimates for T1 noise source ($-63 \leq 10 \log f \leq 24$).

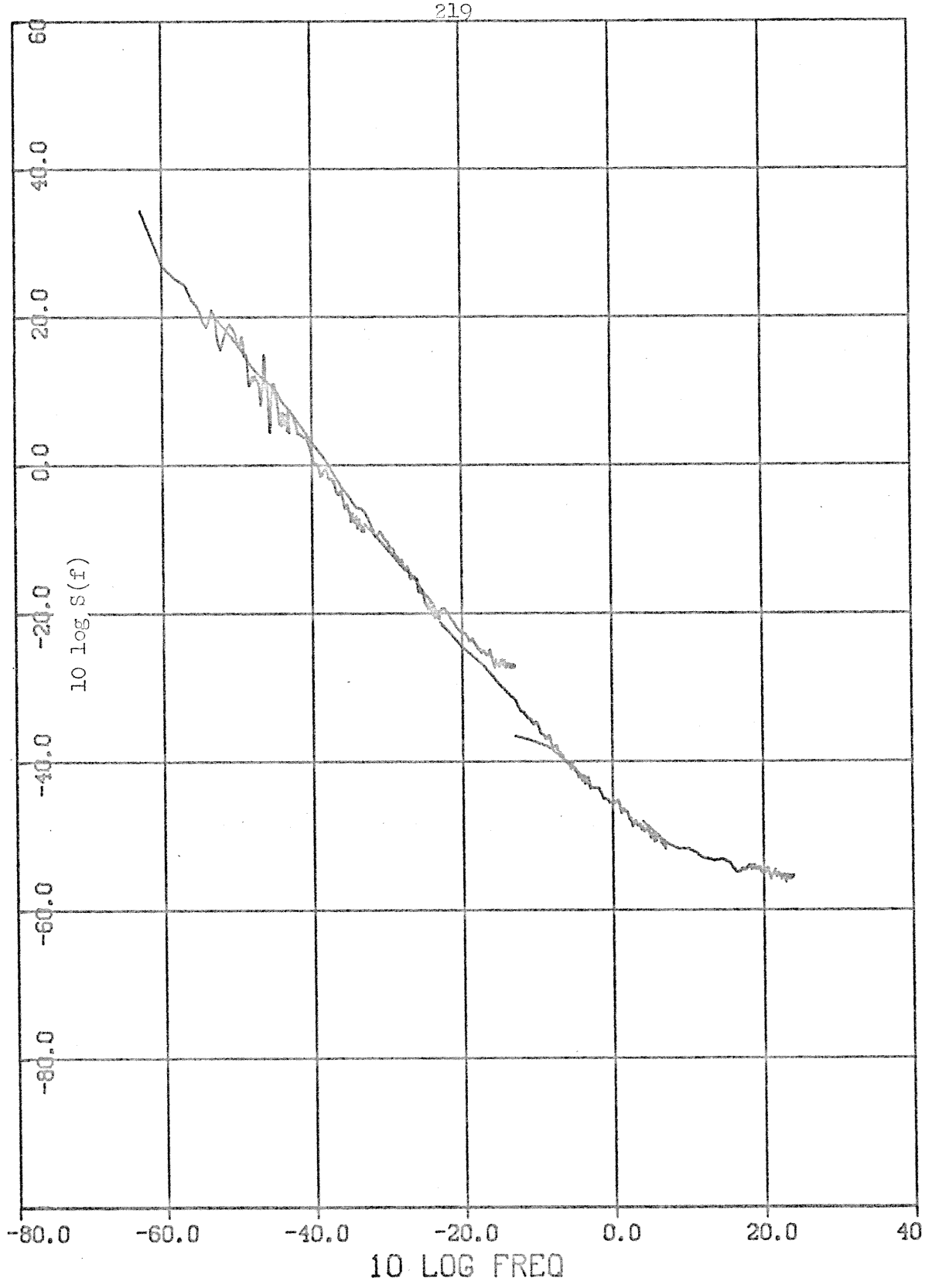


Figure 6.11 Spectral estimates for T2 noise source ($-63 \leq 10 \log f \leq 24$).

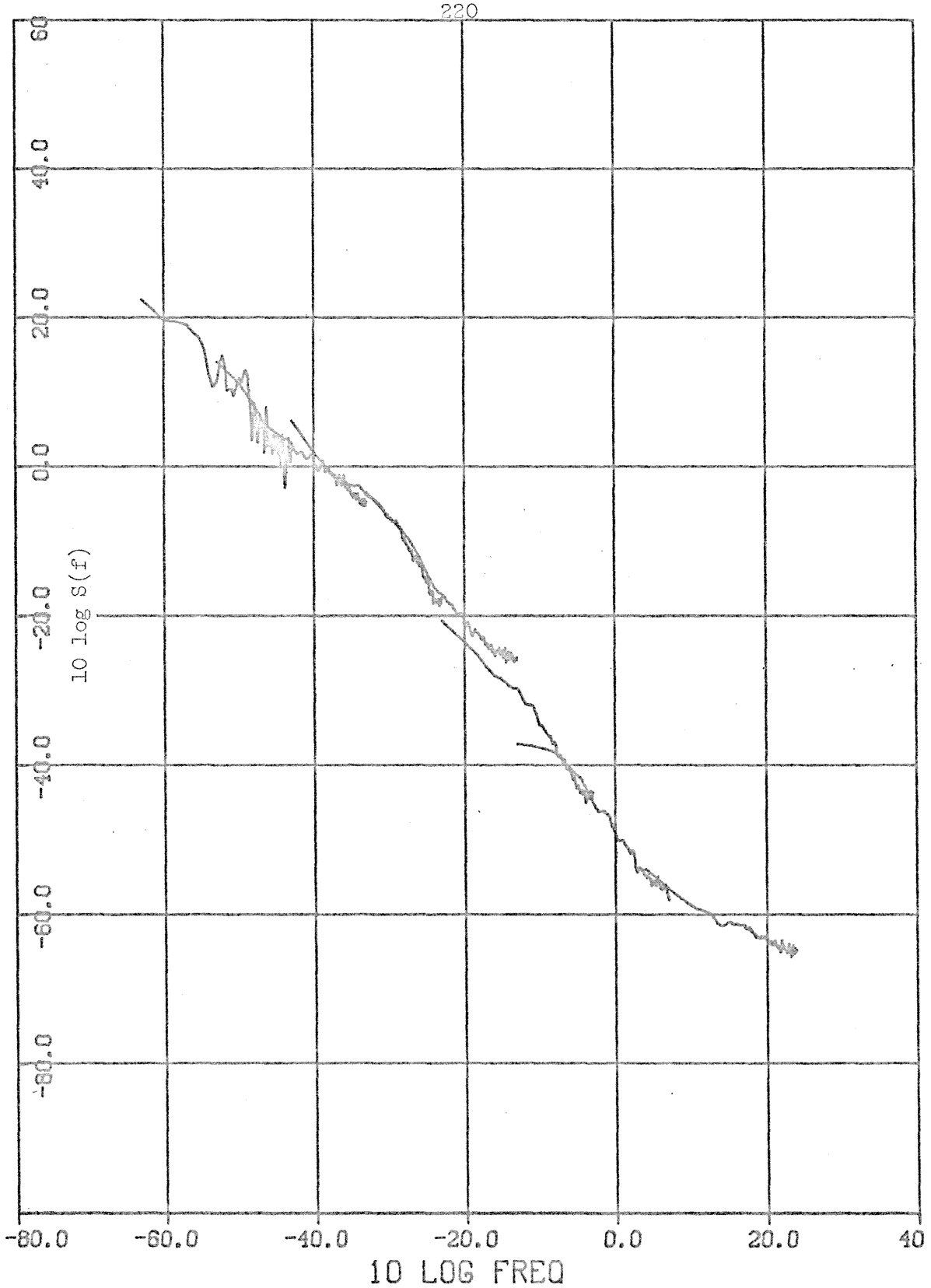


Figure 6.12 Spectral estimates for T3 noise source ($-63 \leq 10 \log f \leq 24$).

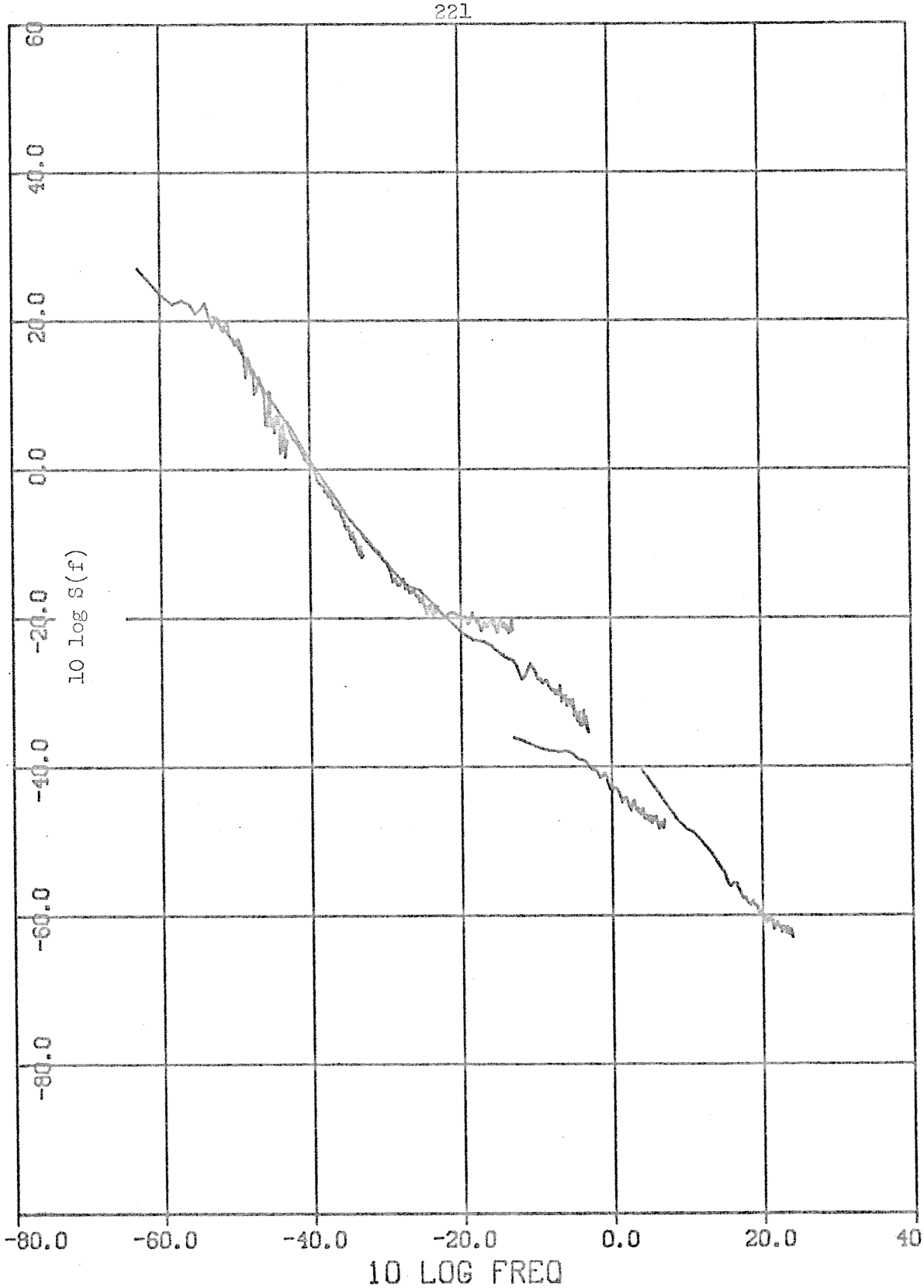


Figure 6.13 Spectral estimates for T4 noise source ($-63 \leq 10 \log f \leq 24$).

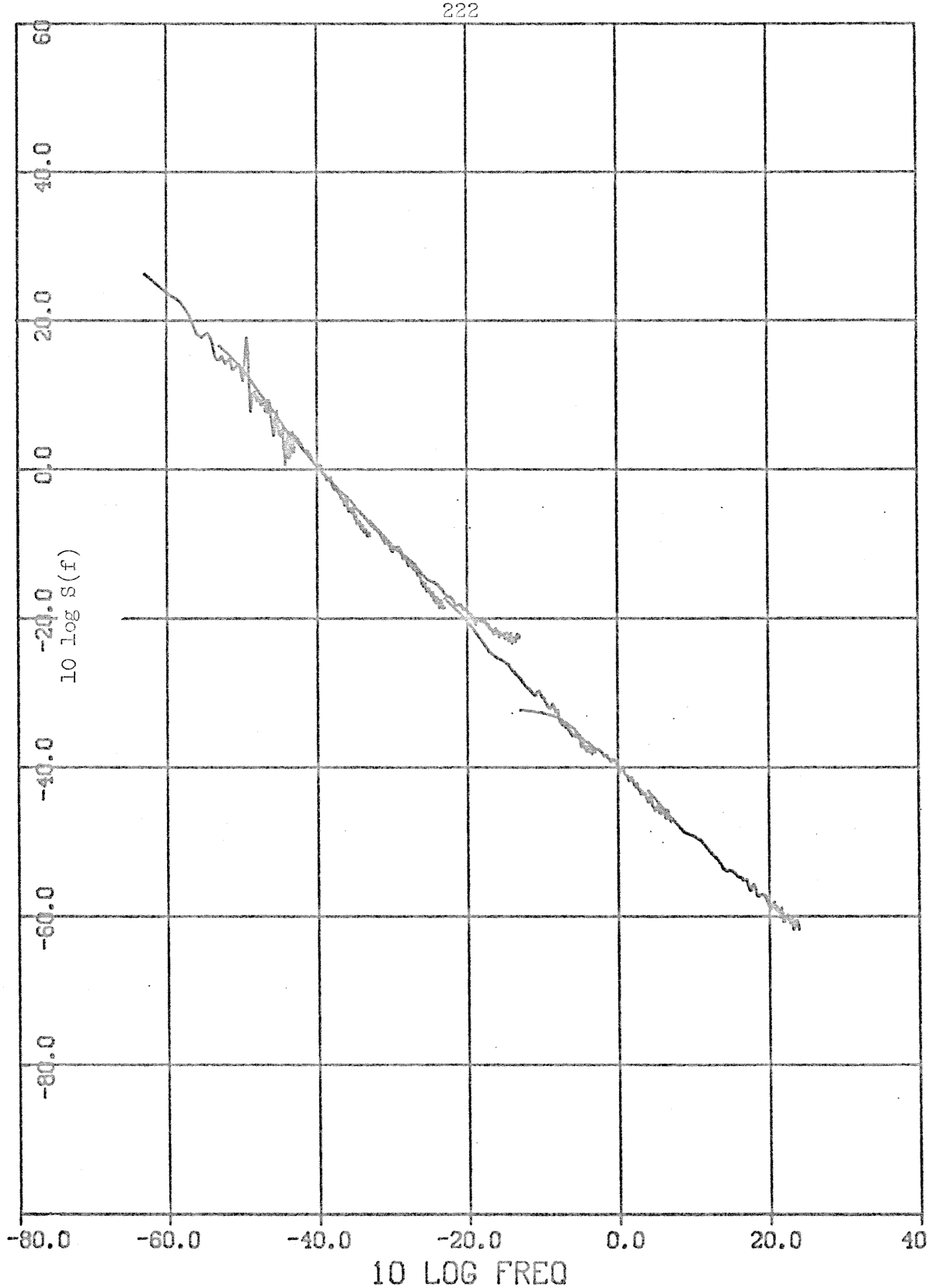


Figure 6.14 Spectral estimates for T5 noise source ($-63 \leq 10 \log f \leq 24$).

Chapter VII

EXPERIMENTAL DATA PHASE II

7.1 Time Domain Data

The sample period of Δt = 10 seconds (for each channel) was found to be rather expensive to maintain. Therefore, at 11:11 PST March 2, 1973 the sample period was increased to Δt = 100 seconds as the sample times and DVM values were recorded on paper. These values were then compared with the data stored on magnetic tape. The phase I (Δt = 10 seconds) data were then resampled to Δt = 100 seconds in such a way as to match the phase II data within a sampling period discontinuity of less than or equal to ±5 seconds. For convenience in programming, some data at the beginning of phase I have not been retained and the effective start of the data recording is moved to approximately 18:20 PST November 29, 1972. The recording was continued until 307500 samples of each channel at Δt = 100 seconds were available. This goal was reached at about 16:03 PST November 20, 1973. The final sample time predicted by 18:20 PST November 29, 1972 + 30,750,000 seconds is 16:00 PST November 20, 1973 indicating an average sample rate error of about 0.0006% when averaged over the 355.9 day experiment.

The data are stored on tapes labeled Spectral Data-7 and Spectral Data-8 having 131 and 91 files respectively. Table 7.1 lists the "glitches" which were removed from the "raw" data and is an extension of Table 5.4 (see Section 5.3 for detailed explanation).

The data were filtered (digitally) and resampled to $\Delta t = 10^3$ seconds

DATAFILE 7	OF TAPE 3703	(SPECTRAL DATA-7)		3 CARDS
A 3(885)THRU	886 CHANGE TO -12300		FROM -14890	
A 3(890)	CHANGE TO -12400		FROM -14590	
A 4(885)	CHANGE TO - 3900		FROM -15670	
DATAFILE 42	OF TAPE 3703	(SPECTRAL DATA-7)		1 CARDS
A 3(288)	CHANGE TO -13700		FROM -15060	
DATAFILE 49	OF TAPE 3703	(SPECTRAL DATA-7)		2 CARDS
A 3(940)THRU	966 CHANGE TO -13000		FROM -18580	
A 4(940)THRU	978 CHANGE TO - 2800		FROM - 3084	
DATAFILE 56	OF TAPE 3703	(SPECTRAL DATA-7)		3 CARDS
A 3(334)	CHANGE TO -13000		FROM -16070	
A 3(337)	CHANGE TO -13000		FROM -17320	
A 4(334)	CHANGE TO - 5400		FROM -11640	
DATAFILE 68	OF TAPE 3703	(SPECTRAL DATA-7)		3 CARDS
A 3(597)	CHANGE TO -11000		FROM -15460	
A 4(590)	CHANGE TO - 7600		FROM -10656	
A 4(597)	CHANGE TO - 7600		FROM -13920	
DATAFILE 71	OF TAPE 3703	(SPECTRAL DATA-7)		3 CARDS
A 3(623)	CHANGE TO -11200		FROM -12720	
A 4(623)	CHANGE TO - 9000		FROM -12800	
A 6(516)THRU	517 CHANGE TO - 2400		FROM - 8248	
DATAFILE 103	OF TAPE 3703	(SPECTRAL DATA-7)		3 CARDS
A 3(231)	CHANGE TO - 8700		FROM -10785	
A 4(174)THRU	175 CHANGE TO -13700		FROM -14290	
A 4(227)	CHANGE TO -13500		FROM -15870	
DATAFILE 109	OF TAPE 3703	(SPECTRAL DATA-7)		2 CARDS
A 3(900)THRU	911 CHANGE TO - 8750		FROM -19740	
A 4(900)THRU	911 CHANGE TO -15300		FROM -36610	
DATAFILE 115	OF TAPE 3703	(SPECTRAL DATA-7)		4 CARDS
A 3(481)	CHANGE TO - 7600		FROM - 5520	
A 3(798)THRU	812 CHANGE TO - 8000		FROM -16090	
A 4(481)	CHANGE TO -15700		FROM -11769	
A 4(798)THRU	812 CHANGE TO -15600		FROM -30960	
DATAFILE 116	OF TAPE 3703	(SPECTRAL DATA-7)		2 CARDS
A 3(967)	CHANGE TO - 7700		FROM - 9200	
A 3(971)	CHANGE TO - 7700		FROM -13840	
DATAFILE 117	OF TAPE 3703	(SPECTRAL DATA-7)		1 CARDS
A 4(3)	CHANGE TO -16500		FROM -11772	
DATAFILE 124	OF TAPE 3703	(SPECTRAL DATA-7)		2 CARDS
A 3(534)THRU	559 CHANGE TO - 7500		FROM - 8092	
A 4(534)THRU	560 CHANGE TO -17000		FROM -18320	
DATAFILE 2	OF TAPE 1419	(SPECTRAL DATA-8)		4 CARDS
A 3(488)THRU	497 CHANGE TO - 7250		FROM -17720	
A 3(506)	CHANGE TO - 7250		FROM - 8269	
A 4(491)THRU	495 CHANGE TO -17600		FROM -19070	
A 4(504)THRU	507 CHANGE TO -17500		FROM -20420	
DATAFILE 8	OF TAPE 1419	(SPECTRAL DATA-8)		2 CARDS
A 3(393)THRU	419 CHANGE TO - 6600		FROM - 7146	
A 4(393)THRU	417 CHANGE TO -18000		FROM -19850	
DATAFILE 14	OF TAPE 1419	(SPECTRAL DATA-8)		2 CARDS
A 3(311)THRU	325 CHANGE TO - 6750		FROM - 7587	
A 4(311)THRU	323 CHANGE TO -17500		FROM -18820	
DATAFILE 20	OF TAPE 1419	(SPECTRAL DATA-8)		3 CARDS
A 3(31)	CHANGE TO - 6900		FROM -10393	
A 3(406)	CHANGE TO - 6750		FROM -12010	
A 4(31)	CHANGE TO -19500		FROM -28110	

Table 7.1 a Correction data (phase II).

```
DATAFILE  29   OF TAPE  1419 (SPECTRAL DATA-8)    1 CARDS
A 6( 348)          CHANGE TO  -3300    FROM -12237
DATAFILE  31   OF TAPE  1419 (SPECTRAL DATA-8)    2 CARDS
A 3(1009)THRU 1025 CHANGE TO  -6250    FROM  -5745
A 4(1017)THRU 1025 CHANGE TO -17500    FROM -22830
DATAFILE  32   OF TAPE  1419 (SPECTRAL DATA-8)    3 CARDS
A 3(   1)THRU   16 CHANGE TO  -6500    FROM  -8099
A 4(   1)THRU   17 CHANGE TO -17500    FROM -23590
A 4( 242)          CHANGE TO -18000    FROM -21150
DATAFILE  38   OF TAPE  1419 (SPECTRAL DATA-8)    2 CARDS
A 3( 975)          CHANGE TO  -6000    FROM -22060
A 4( 975)          CHANGE TO -17100    FROM -42510
DATAFILE  53   OF TAPE  1419 (SPECTRAL DATA-8)    1 CARDS
A 6( 858)THRU  859 CHANGE TO  -4800    FROM -10570
DATAFILE  55   OF TAPE  1419 (SPECTRAL DATA-8)    1 CARDS
A 4( 732)          CHANGE TO -16200    FROM -20970
DATAFILE  56   OF TAPE  1419 (SPECTRAL DATA-8)    1 CARDS
A 4( 424)          CHANGE TO -15900    FROM -11761
DATAFILE  59   OF TAPE  1419 (SPECTRAL DATA-8)    1 CARDS
A 4( 210)          CHANGE TO -15000    FROM -11720
DATAFILE  61   OF TAPE  1419 (SPECTRAL DATA-8)    1 CARDS
A 4( 591)          CHANGE TO -15500    FROM -11769
```

Table 7.1 b Correction data (phase II).

where burst components of .39 volt amplitude for noise source #2 and .22 volt and .35 volt amplitudes for source T3 were estimated from plots and removed. These data were then filtered and resampled to $\Delta t = 10^4$ seconds and $\Delta t = 5 \times 10^4$ seconds using the filters discussed in Appendix C. Figure 7.1 shows representative data at $\Delta t = 10^4$ seconds but where noise sources #2 and T3 have not had burst components removed. Noise source #2 and T5 do not extend over the full period of the experiment. This is due to a failure in the power supply regulator for the #2 noise source and, apparently, an intermittent poor switch contact in the analog channel of the T5 noise source (the alternate channel was also not working properly). Spectral estimates for these sources are, therefore, based on 1.547×10^7 seconds for #2 and 1.763×10^7 seconds for T5.

It is now obvious that noise sources T1 and T2 have a common low frequency component with an apparent period of very nearly one year. This suggests that these amplifiers (which are contained in a common, plastic package, 747 dual operational amplifier) are influenced by some external parameter other than voltage supply and temperature (such as barometric pressure or humidity). Since this alleged parameter is not among those recorded, no attempt has been made to remove this component. The effects of this component will be discussed in later paragraphs.

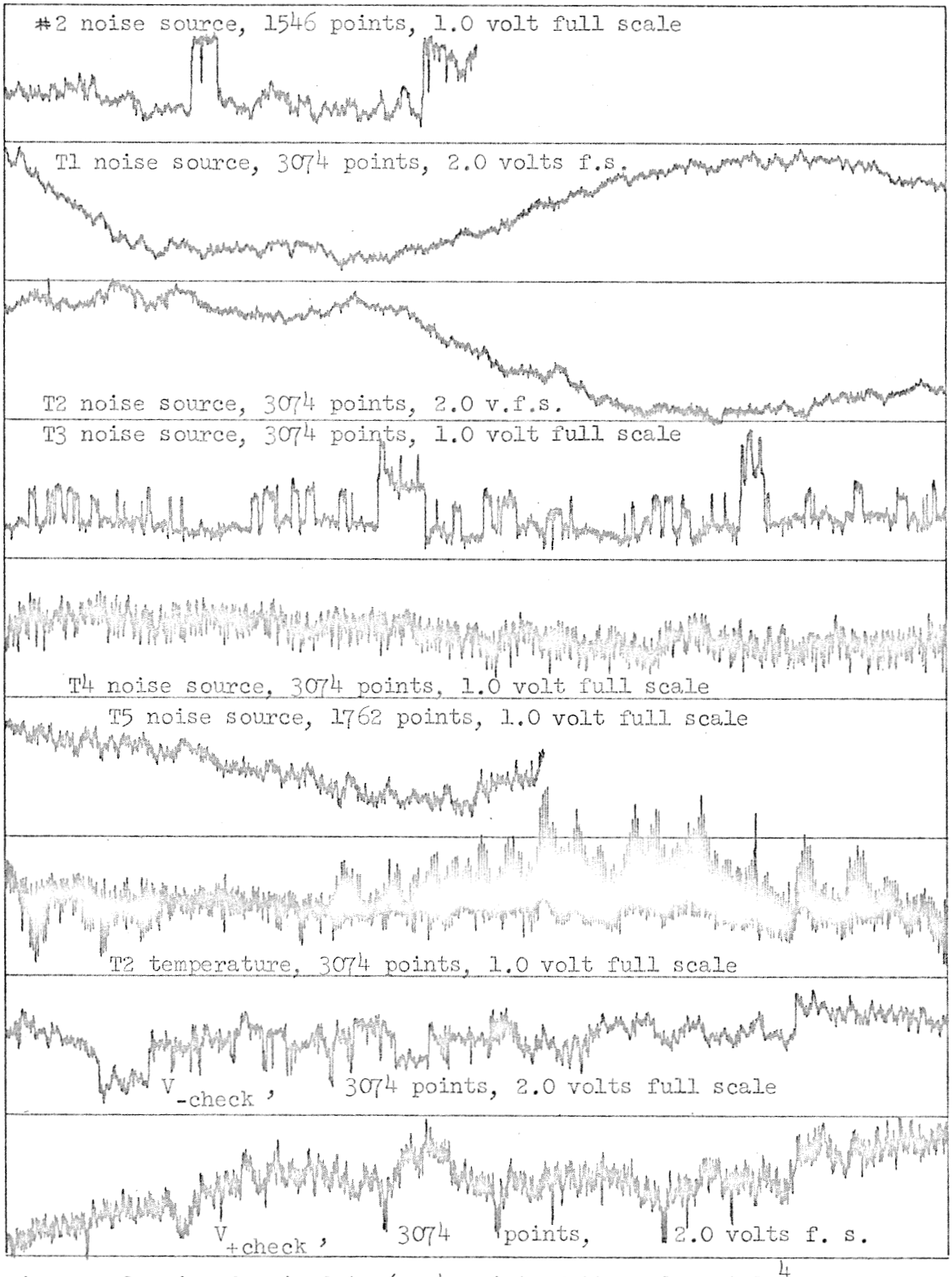


Figure 7.1 Time domain data (3074 points full scale, $\Delta t = 10^{4}$ sec., no components removed).

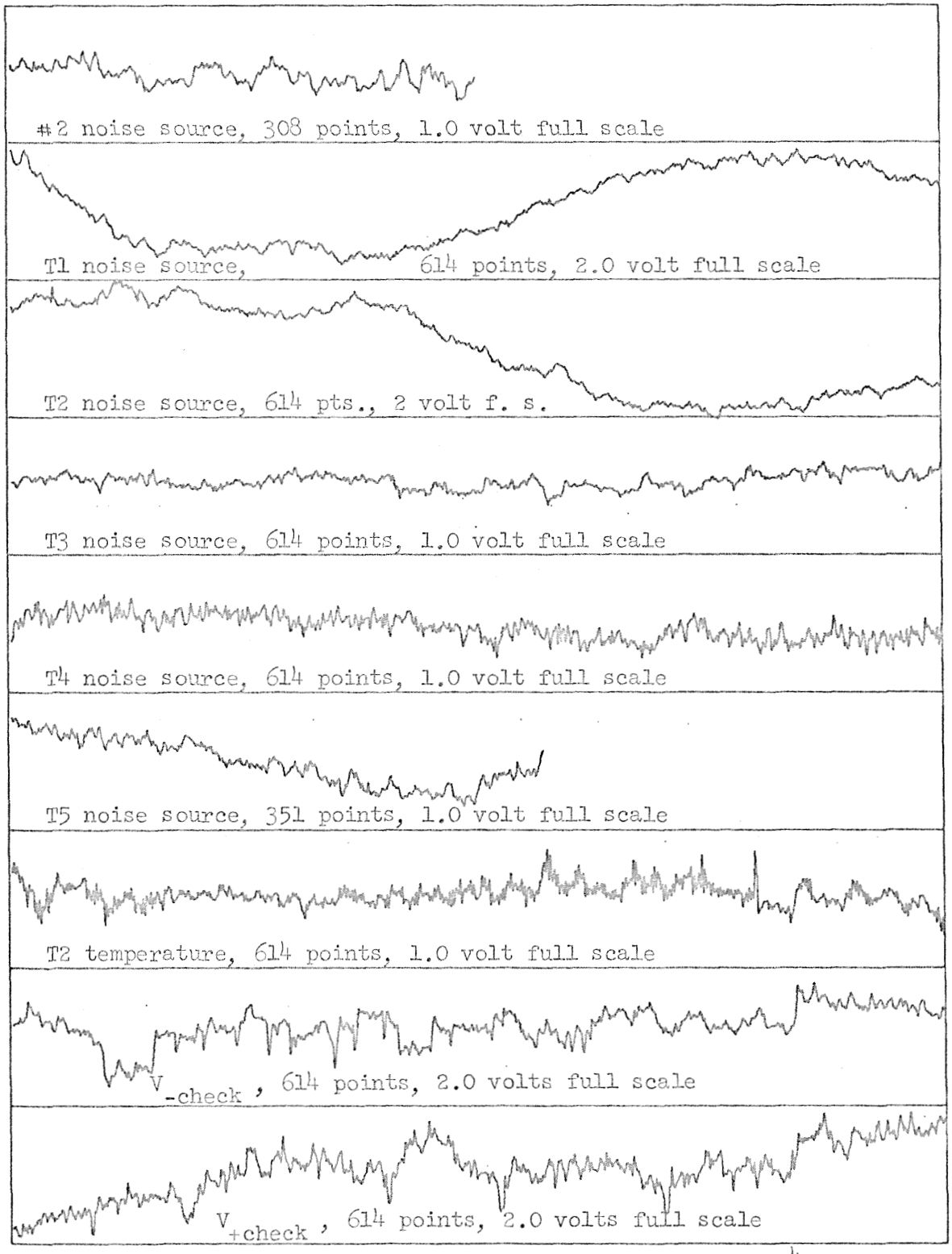


Figure 7.2 Time domain data (614 points full scale, $\Delta t = 5 \times 10^{4}$ sec., burst components have been subtracted from #2 and T5 noise sources at the 0.001 Hz sample rate).

7.2 Power Spectral Density Estimates

Power spectral density estimates were made using these data at $\Delta t = 10^4$ and $\Delta t = 5 \times 10^4$ seconds (using 100 time lags) after removing temperature components for noise sources T1,...,T5 (assuming temperature coefficients of 0.047, -0.146, -.098, -.103, and 0.114) and using the versions of #2 and T3 for which burst components had been removed. The resulting spectral estimates are shown in Figs. 7.3,...,7.8 along with the "best" estimates made for each of the other sample rates discussed earlier.

Figures 7.3,...,7.8 have been corrected for aliasing. This was done by "eyeballing" a straight line (on this log-log scale) to the estimates made at the next higher sample rate and subtracting, on the linear scale, the aliasing which would occur due to this estimated spectral density (taking into account the effects of the appropriate dealiasing filters). Where the higher frequency estimates are "bumpy" the slope and amplitude are estimated near the folding frequency since the digital dealiasing filters insure that only the first octave beyond the $\frac{1}{2\Delta t}$ folding frequency contributes significantly to the aliasing (except for the $\Delta t = 10$ seconds estimates). The $\Delta t = 10^4$ seconds and 5×10^4 seconds T4 noise source estimates were made with no attempt to remove the burst noise as was done for the higher frequency estimates. The aliasing was, therefore, estimated by extrapolation from the lower frequency estimates. It is clear in all of these figures that not only is the relative variance increased in the regions where aliasing was significant, but that a significant bias may be introduced by inaccurate

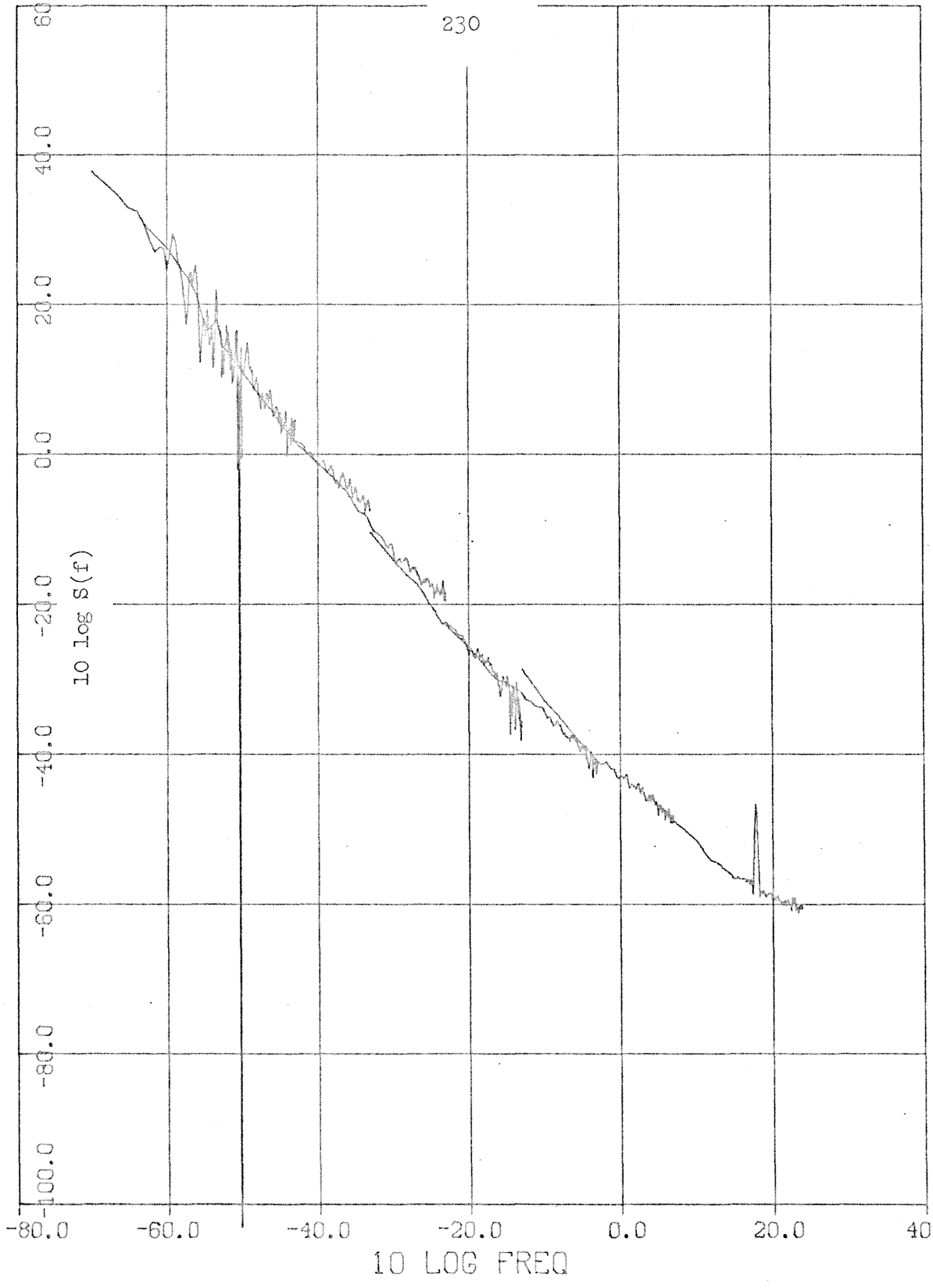


Figure 7.3 Spectral estimates for #2 noise source ($-70 \leq 10 \log f \leq 24$, corrected for aliasing and response of filters).

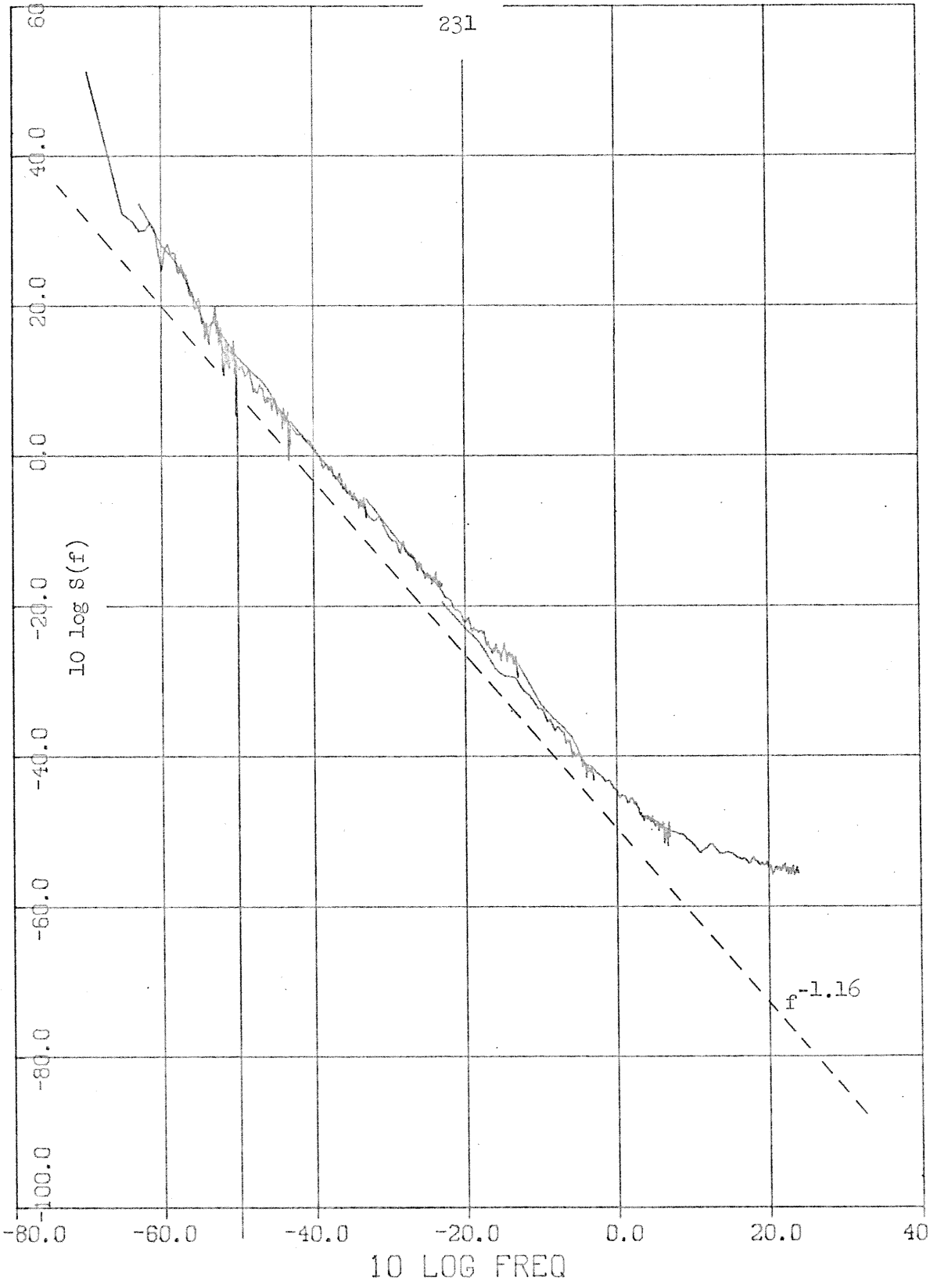


Figure 7.4 Spectral estimates for T1 noise source ($-70 \leq 10 \log f \leq 24$, corrected for aliasing and response of filters).

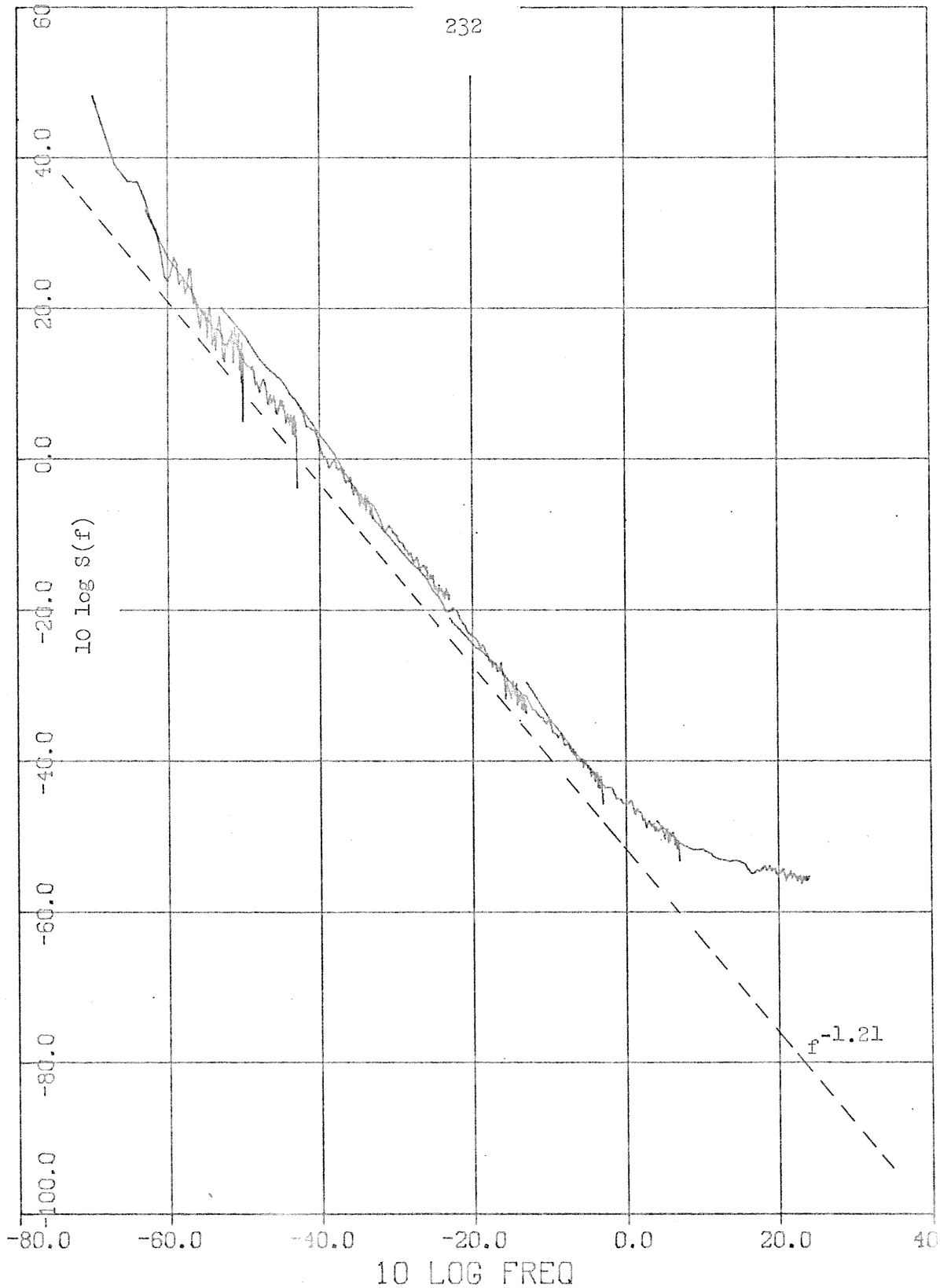


Figure 7.5 Spectral estimates for T2 noise source ($-70 \leq 10 \log f \leq 24$, corrected for aliasing and response of filters).

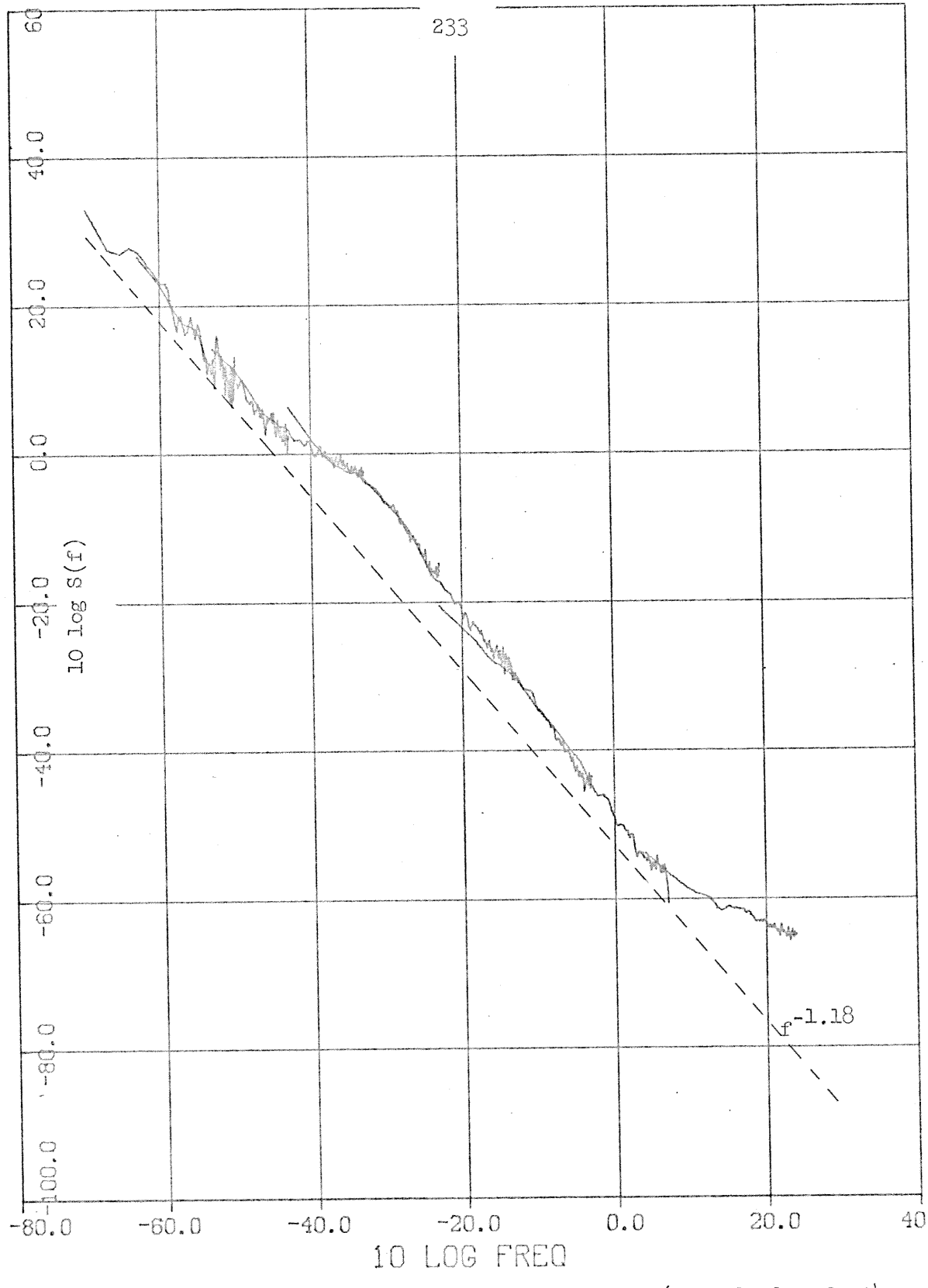


Figure 7.6 Spectral estimates for T3 noise source ($-70 \leq 10 \log f \leq 24$, corrected for aliasing and response of filters).

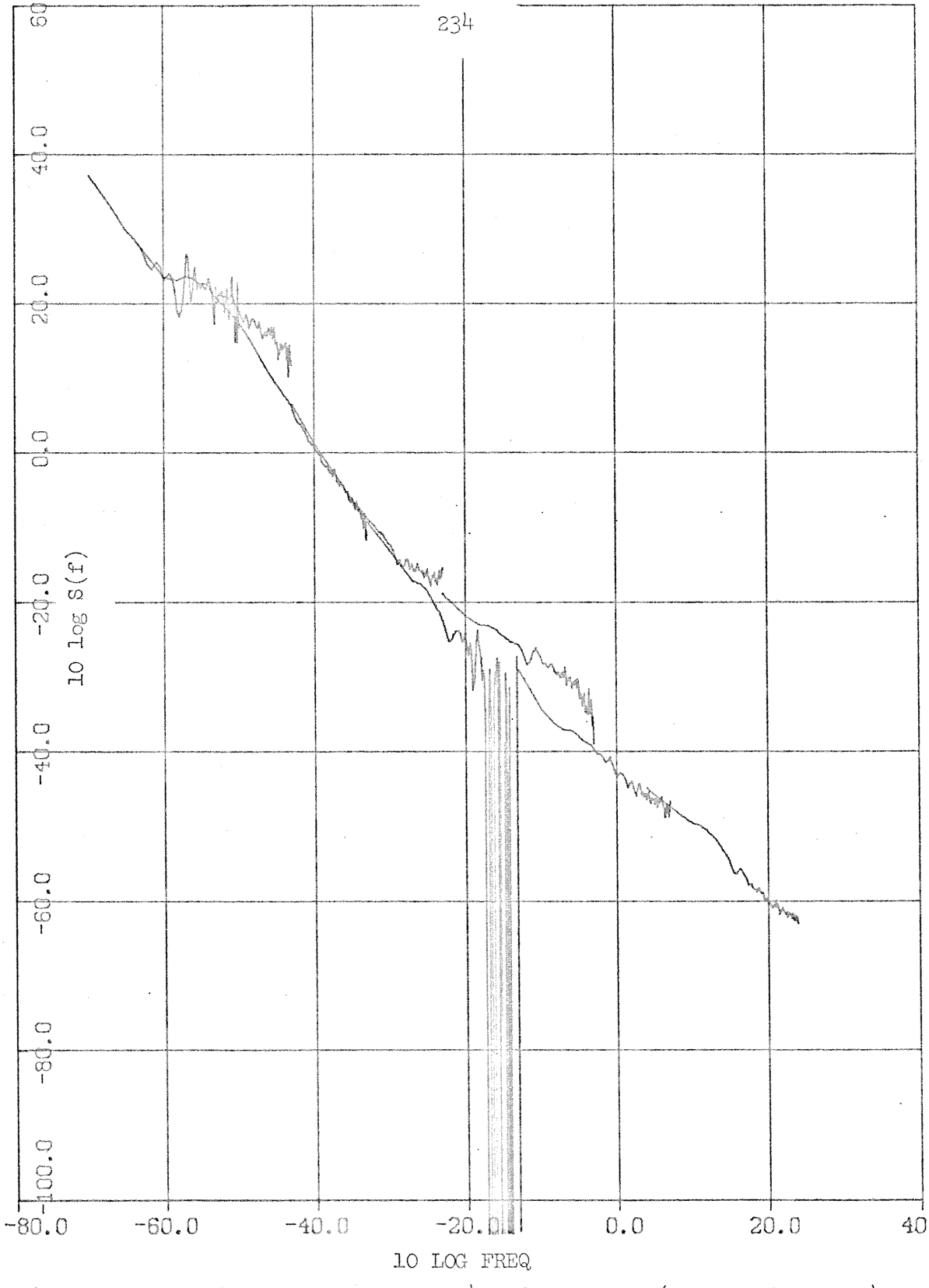


Figure 7.7 Spectral estimates for T4 noise source ($-70 \leq 10 \log f \leq 24$, corrected for aliasing and response of filters).

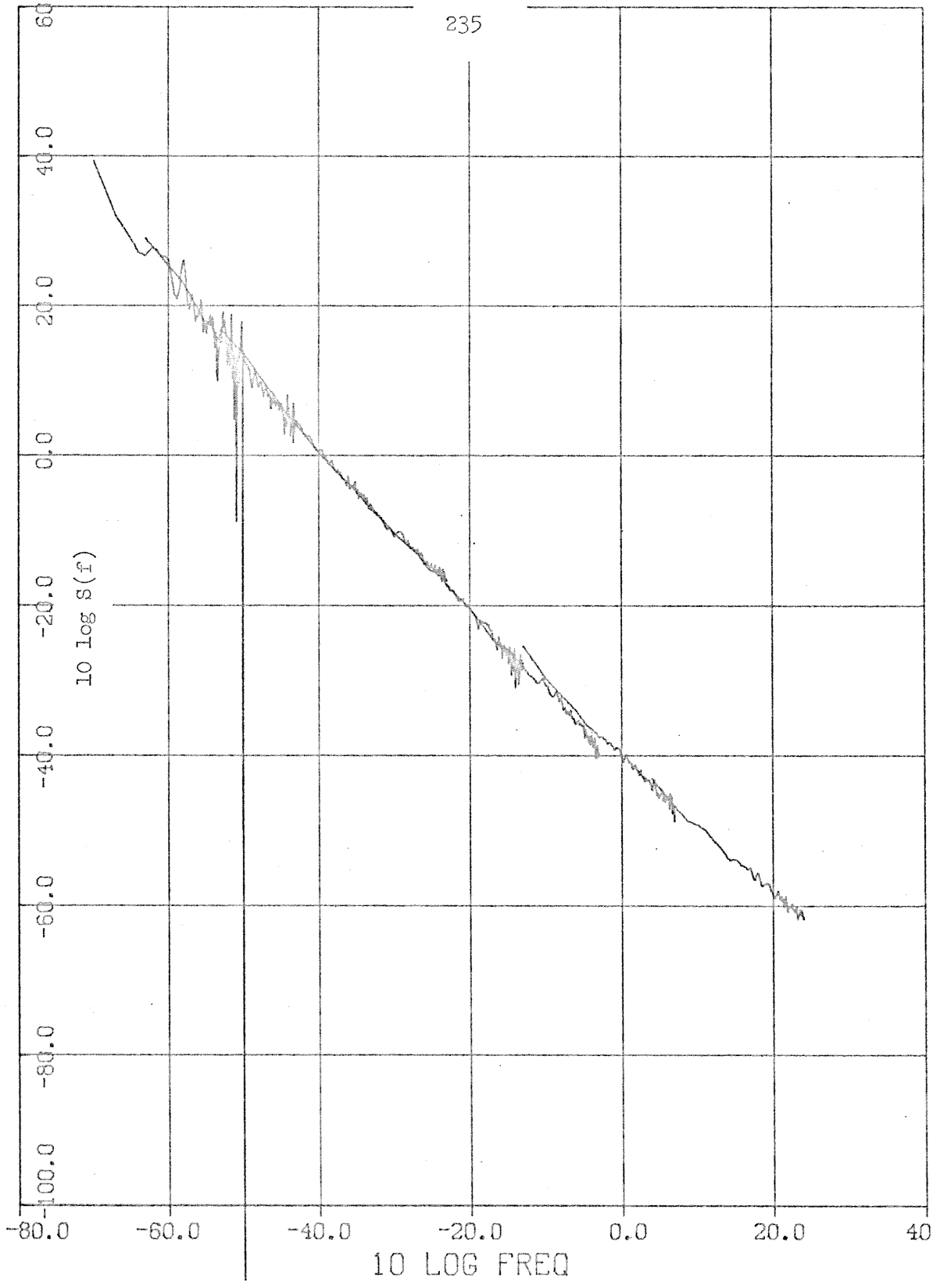


Figure 7.8 Spectral estimates for T5 noise source ($-70 \leq 10 \log f \leq 24$, corrected for aliasing and response of filters).

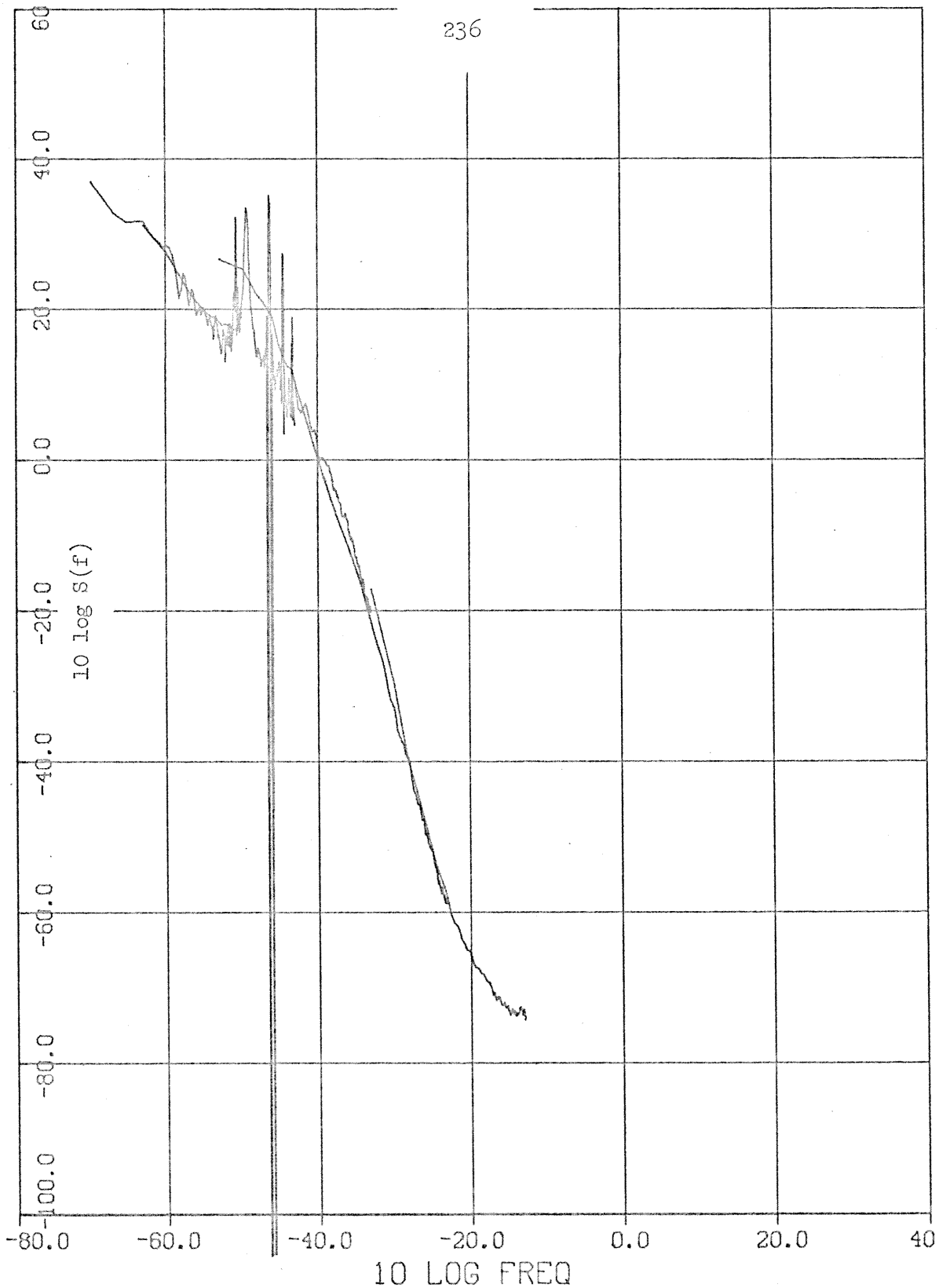


Figure 7.9 Spectral estimates for T2 temperature ($-70 \leq 10 \log f \leq -13$).

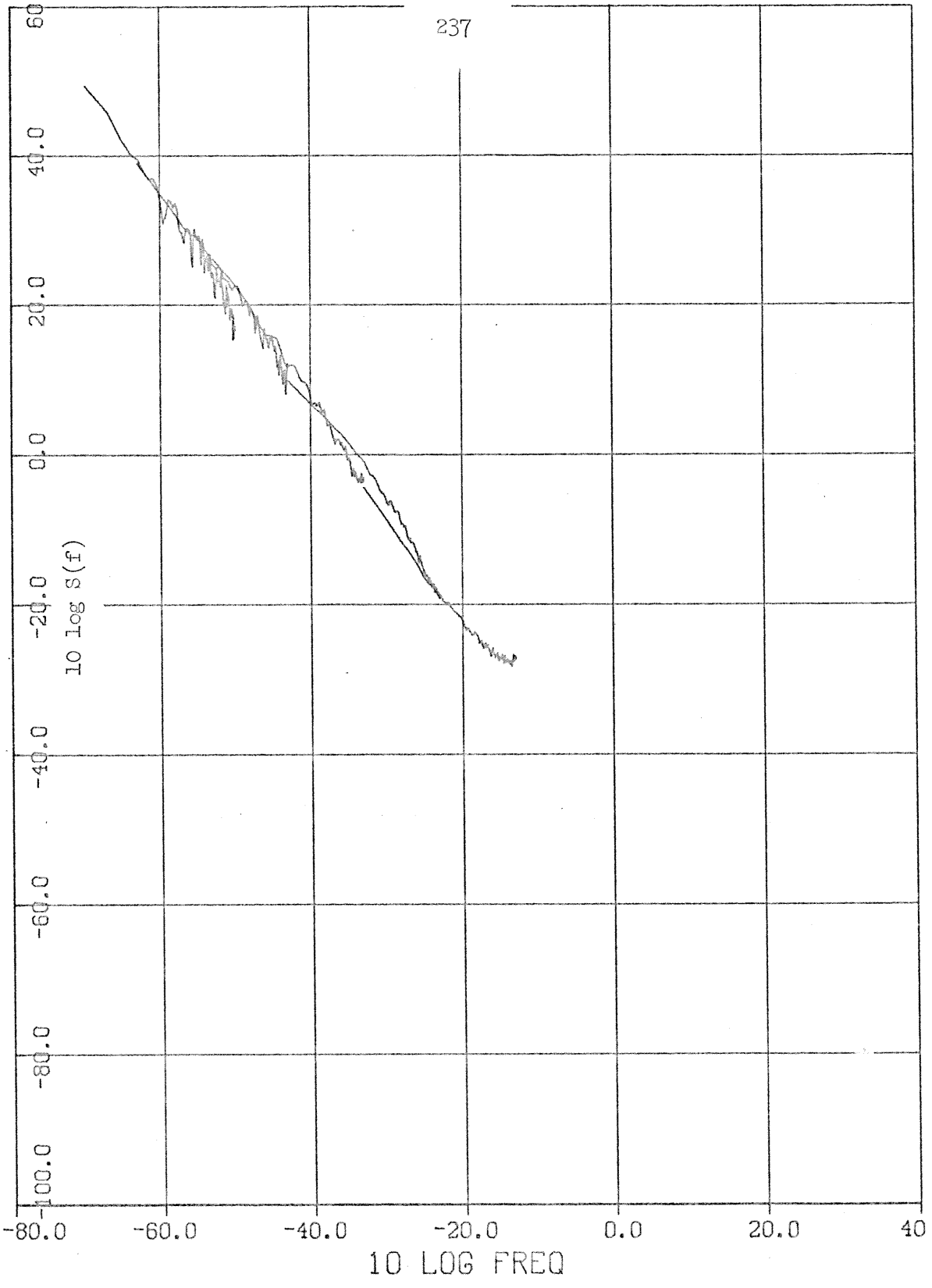


Figure 7.10 Spectral estimates for V_{-check} ($-70 \leq 10 \log f \leq -13$).

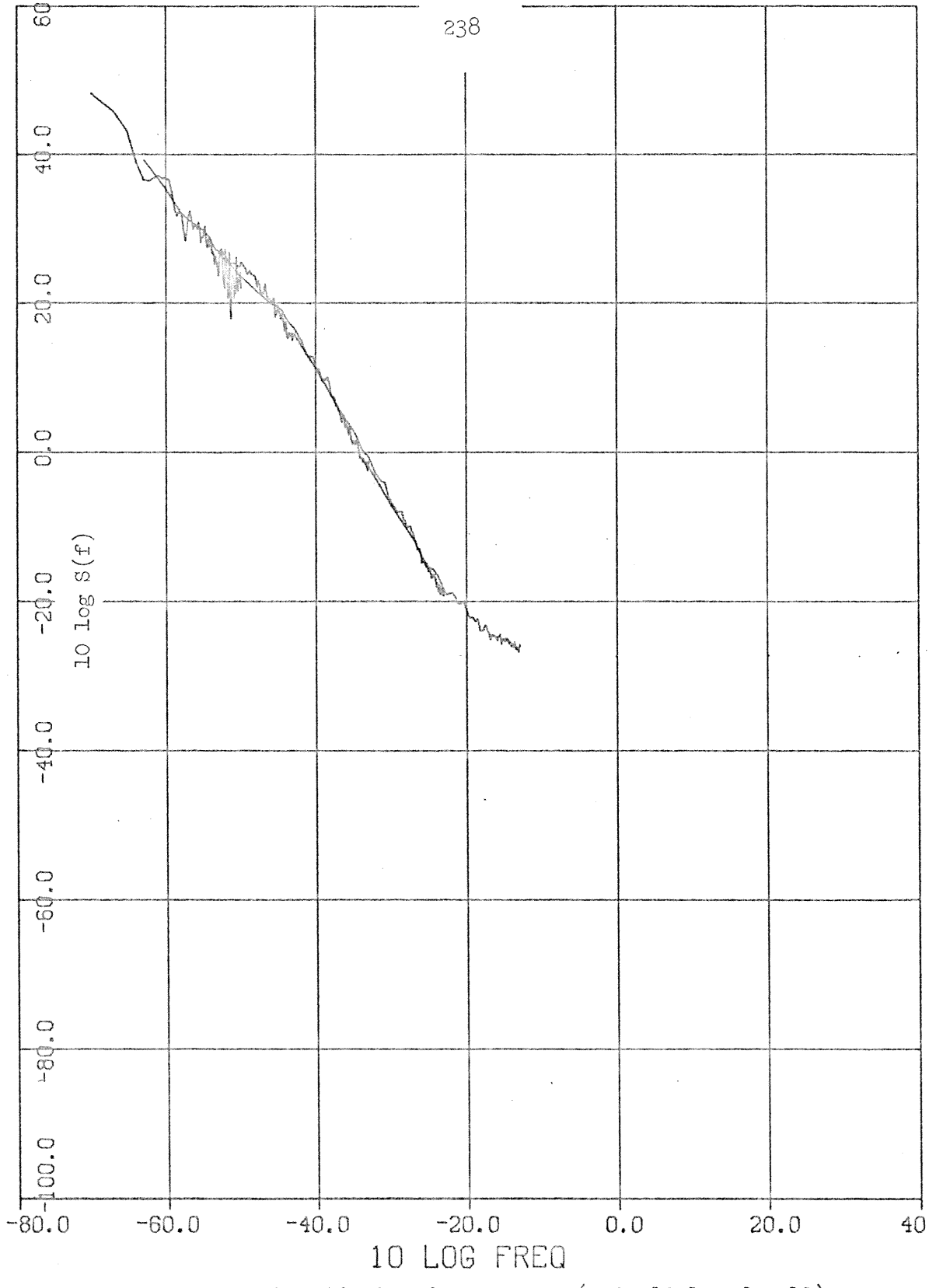


Figure 7.11 Spectral estimates for V_{+check} ($-70 \leq 10 \log f \leq -13$).

estimation of the higher frequency spectral density. Note that the high pass filter used in the $\Delta t = 0.1$ second estimates has also been accounted for.

Using Eq. 2.5.3:

$$\text{relative standard deviation} \cong \sqrt{\frac{M}{N'}} \times 100\%$$

(with $N' = N - 0.3\,M$),

the relative standard deviation (in the regions where aliasing was not significant) should be approximately 10% for estimates based on sample periods less than $\Delta t = 10^3$ seconds, 11% for $\Delta t = 10^3$ seconds, 18% (26% for #2 and 24% for T5) for $\Delta t = 10^4$ seconds, and 41% (60% for #2, and 56% for T5) for $\Delta t = 5 \times 10^4$ seconds.

For the purpose of looking for a low frequency break in the 1/f trend, a 50% standard deviation proves not to be a great liability for any of the six noise sources investigated. An error of up to two standard deviations would result in an estimate between $-\infty$ and 3 db above the true value. It is, therefore, not particularly unlikely that we should find estimates indicating a breakdown of the 1/f trend when there is none, but very unlikely that the estimates would fall along the $f^{-\alpha}$ trend when, in fact, the true (windowed) spectral density is 3 db or more below the trend. None of the Figs. 7.3,...,7.8 indicate a breakdown in the $f^{-\alpha}$ trend even down to 10^{-7} Hz, making such a breakdown seem highly unlikely (unless there is some bias, see Section 2.8 for example).

Noise sources T1 and T2,as mentioned previously, appear to have a

large component (see Figs. 7.1, 7.2) with average power of (very roughly) 0.5 volt^2. Most of this power appears (from the time domain plot) to be concentrated near a frequency corresponding to a 1 year period ($f = 3.17 \times 10^{-8}$ Hz). Assuming the true power spectral density of this component to be a delta function at 3.17×10^{-8} Hz yields a bias of 51 db at $f = 10^{-7}$ Hz using the Q_2 spectral window on which all the estimates presented here are based. The spectral density estimates of T1 and T2 at 10^{-7} Hz are 51 db and 48 db. These estimates are, therefore, probably dominated by this component. At 2×10^{-7} Hz (the next higher frequency estimate) a sine wave of this amplitude and frequency would produce only a 33 db bias so that estimates above this frequency may be meaningful depending on the exact spectral components of this bias component).

The power spectral density estimates for the power supply voltage monitors (v_{-check} and v_{+check}), and the temperature channel of the T2 noise source are shown in Figs. 7.9, 7.10, and 7.11 for completeness.

7.3 Self Noise Estimation for Thermistors and Data Channels

The thermistors used to measure the noise source temperatures may themselves exhibit appreciable noise. The five thermistors corresponding to the noise sources T1,...,T5 should be at approximately the same temperature. The extent to which this approximation is valid, however, is not known. The noise source amplifiers themselves would be expected to dissipate varying amounts of power according to their output voltage levels, for example, resulting in temperature gradients within the oven. The component of temperature for each source due to room temperature fluctuations might differ in amplitude and phase due to the geometry of the oven.

A worst case estimate for the noise power spectral density of any thermistor may be obtained by subtracting the output of any other thermistor from the output of the thermistor of interest (in the time domain) and estimating the power spectral density of the difference. This follows from the fact that the power spectral densities of uncorrelated components add. This process yields an estimate which is greater than or equal to the sum of the noise spectral densities of each of the two thermistors, the equality holding if there is no other channel noise and the thermistors are at the same temperature.

The temperature coefficients of the five thermistor outputs may differ by a few percent since the amplifier gains are determined by standard 1% resistors and the thermistors were not matched exactly.

Let the output voltages of the thermistors attached to noise sources T1,...,T5 be referred to as TT1,...,TT5 (notice that TT2 and TT5 were

measured using the digital voltmeter while TT1, TT3, and TT4 were sampled by the A/D converter giving us information regarding the accuracy of the A/D converter). K_1, K_3, K_4, and K_5 were determined such that TT1 - K_1 x TT2, TT3 - K_3 x TT2, TT4 - K_4 x TT2, and TT5 - K_5 x TT2 are approximately constant over 2000 samples at $\Delta t = 10^4$ sec. These values were arrived at by estimating, for example, K_1 as

$$K_1 = \frac{\langle TT1 \times TT2 \rangle}{\langle TT2^2 \rangle}$$

after first removing the sample means from TT1 and TT2. TT1 - $\hat{K}_1$TT2 and TT2 were then plotted (as in Fig. 7.12). A correction was then made to obtain K_1 by observing the magnitude and polarity of the component of TT1 - $\hat{K}_1$TT2 which resembles TT2. The values arrived at were $K_1 = 1.023$, $K_3 = .960$, $K_4 = .957$ and $K_5 = 1.02$. Figure 7.12 shows TT1-K_1TT2, TT2, TT3-K_3TT2, TT4-K_4TT2, and TT5-K_5TT2. As can be seen, it was possible to cancel the large narrow band component for TT1-K_1TT2 almost completely. This is reassuring since these thermistors were mounted touching each other, encased together in a high thermally conductive compound. The remaining narrowband component may be explained as the quantization error of the A/D converter. The step size of the A/D is ~ 10 millivolts. Since the derivative of the thermistor output is usually less than 1 millivolt per sample, for $\Delta t = 100$ seconds, it is not surprising to find that the quantization noise (with a peak to peak voltage of .01 volts) is highly correlated with the signal being quantized. The cancellation was less successful for the other thermistors which were mounted some distance away from the T2 thermistor. In all cases the self noise of

the thermistors (band limited to f = 5 x 10^{-5} Hz) is less than .012°C over this 7.6 month period with the exception of a possible trend common to all five thermistors.

Figures 7.13,...,7.16 show the power spectral density estimates of these temperature channel differences. Comparing these figures with Figs. 5.15,...,5.19 we see that the thermistor noise is at least 25 db below the power spectral density of the temperatures over these ranges. This is quite sufficient since all estimates of the noise source temperature coefficients are less than 0.2 and the effect of temperature on the noise power spectral density estimates has already been shown to be small.

The possibility still exists, however, that the large low frequency component of temperature (10 log f ~ -60) is in fact common to all of the channels (due to a ground loop for example). The low frequency power spectral density of temperature is assumed not to be due to power supply variations since $V_{\pm check}$ are, by theoretical calculation, 54 db more sensitive to power supply variations than the temperature measuring channels. The power spectral density estimates of $V_{\pm\ check}$ are, however, the same order of magnitude as those for the temperature data (see Figs. 5.21 and 5.22).

The possibility that there is low frequency noise added to all of the channels is rejected for the following reason. The noise source channels are identical to the temperature channels (the final amplifiers before leaving the oven are even contained in the same integrated circuit for the noise and temperature channel of each source), and should exhibit

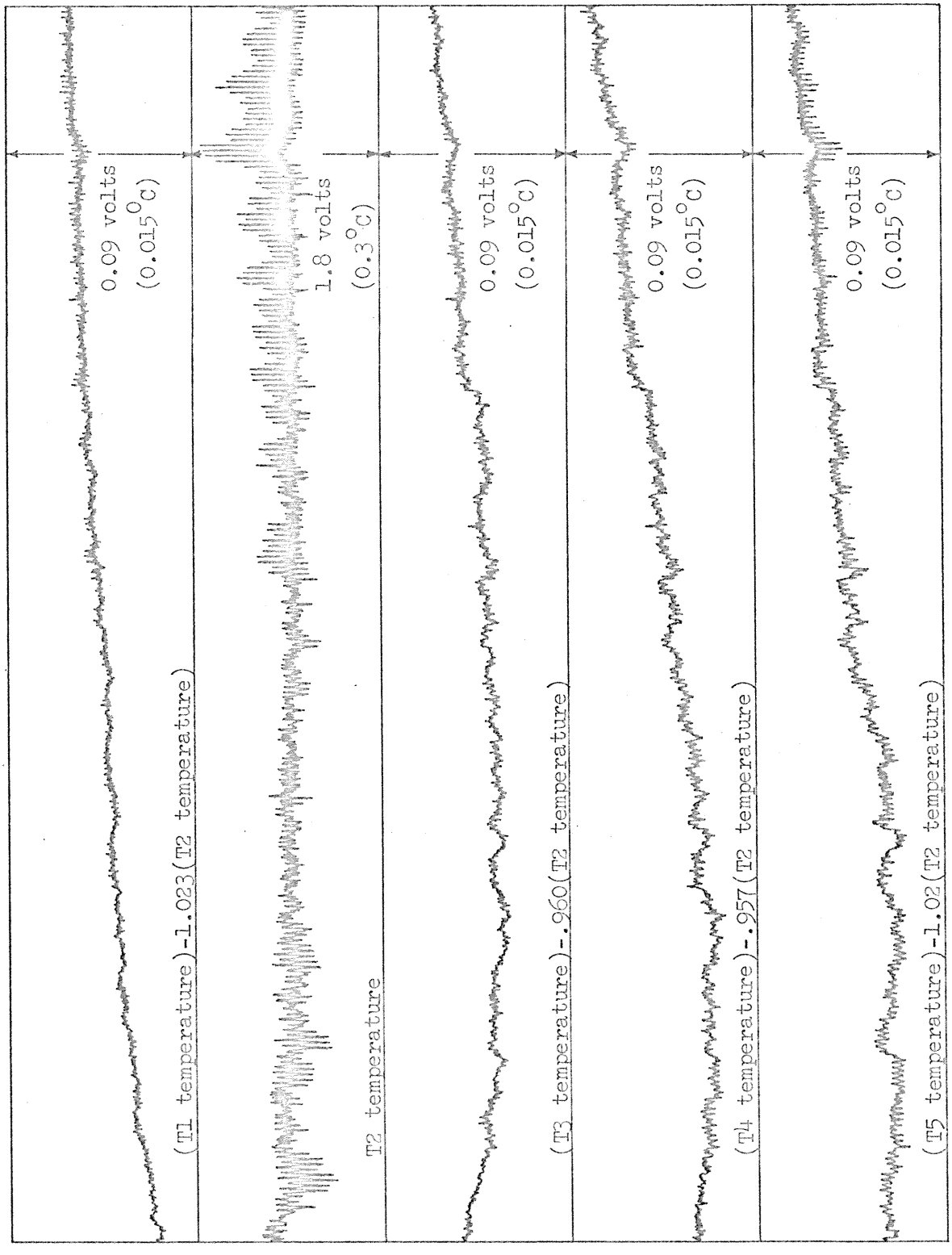


Figure 7.12 Time domain plots of normalized temperature differences (2000 points, Δt=10,000 seconds).

the same noise if it exists. Noise sources T4 and T5, however, have power spectral density estimates which are lower (at the low frequency end) than that of the temperature channels.

Figures 7.13,...,7.16, therefore, establish upper bounds on both the low frequency thermistor noise and the measuring channel noise for 5 typical channels.

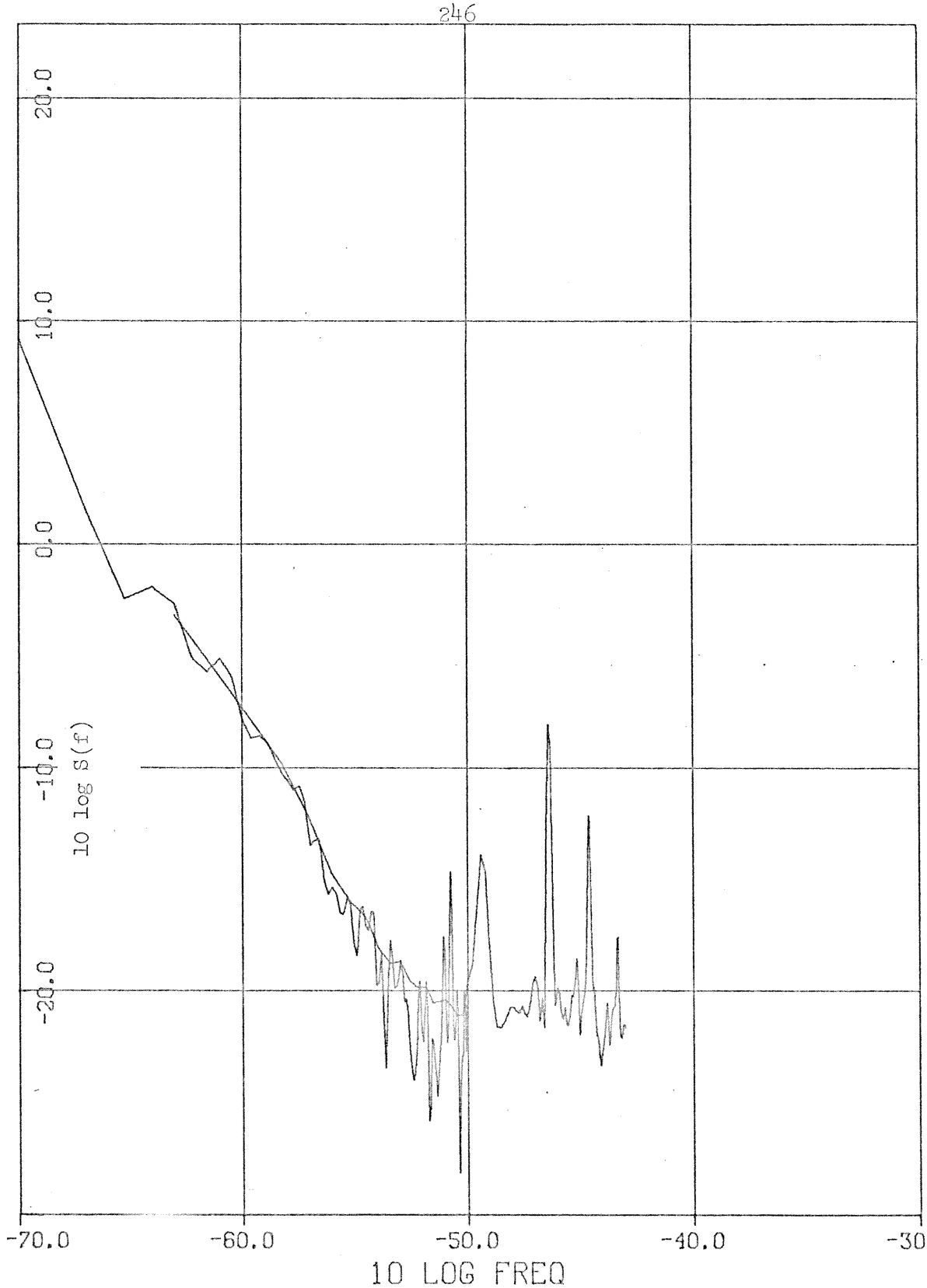


Figure 7.13 Spectral density estimates of T1 temperature minus normalized T2 temperature ($-70 \leq 10 \log f \leq -43$).

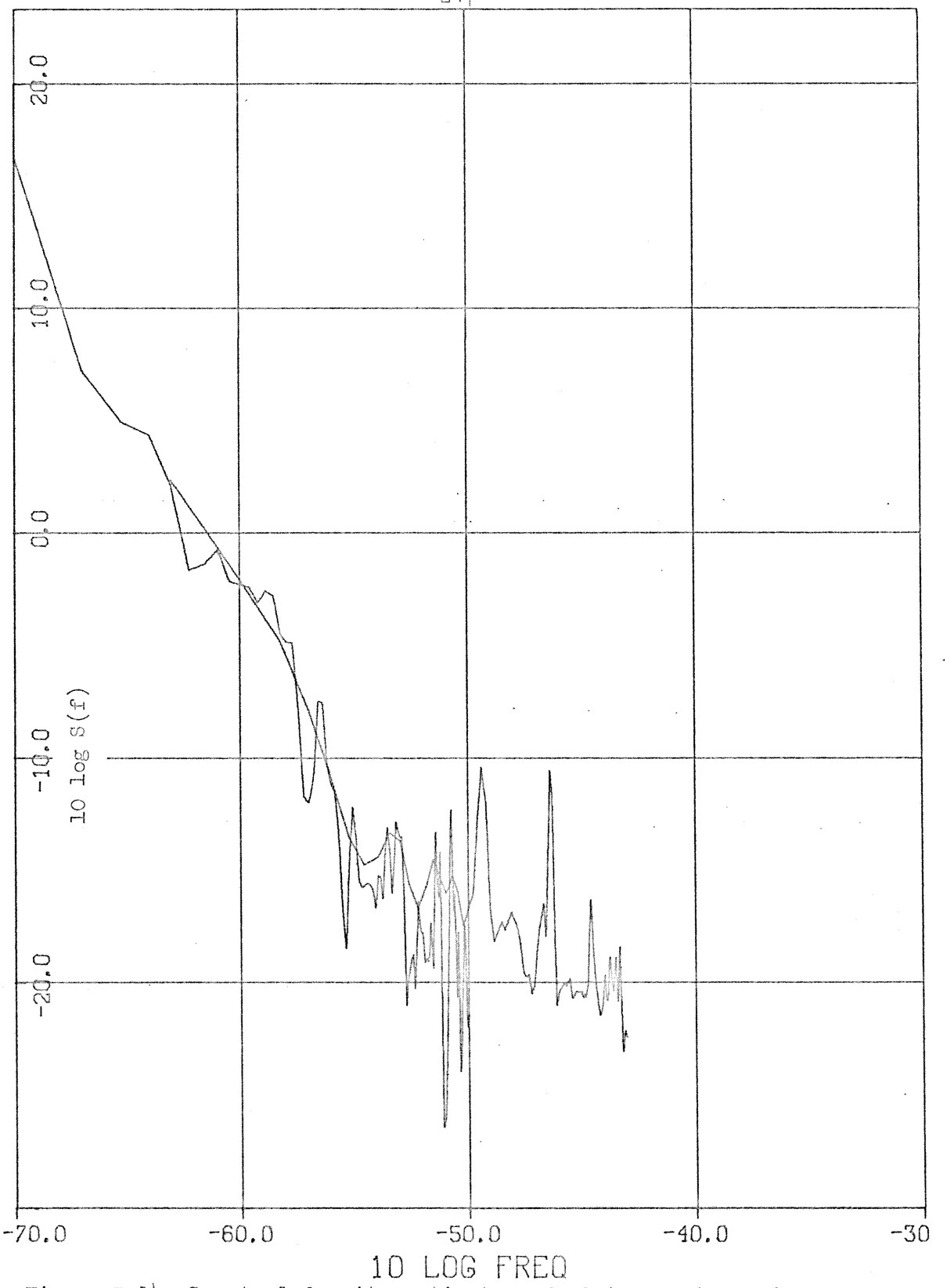


Figure 7.14 Spectral density estimates of T3 temperature minus normalized T2 temperature ($-70 \leq 10 \log f \leq -43$).

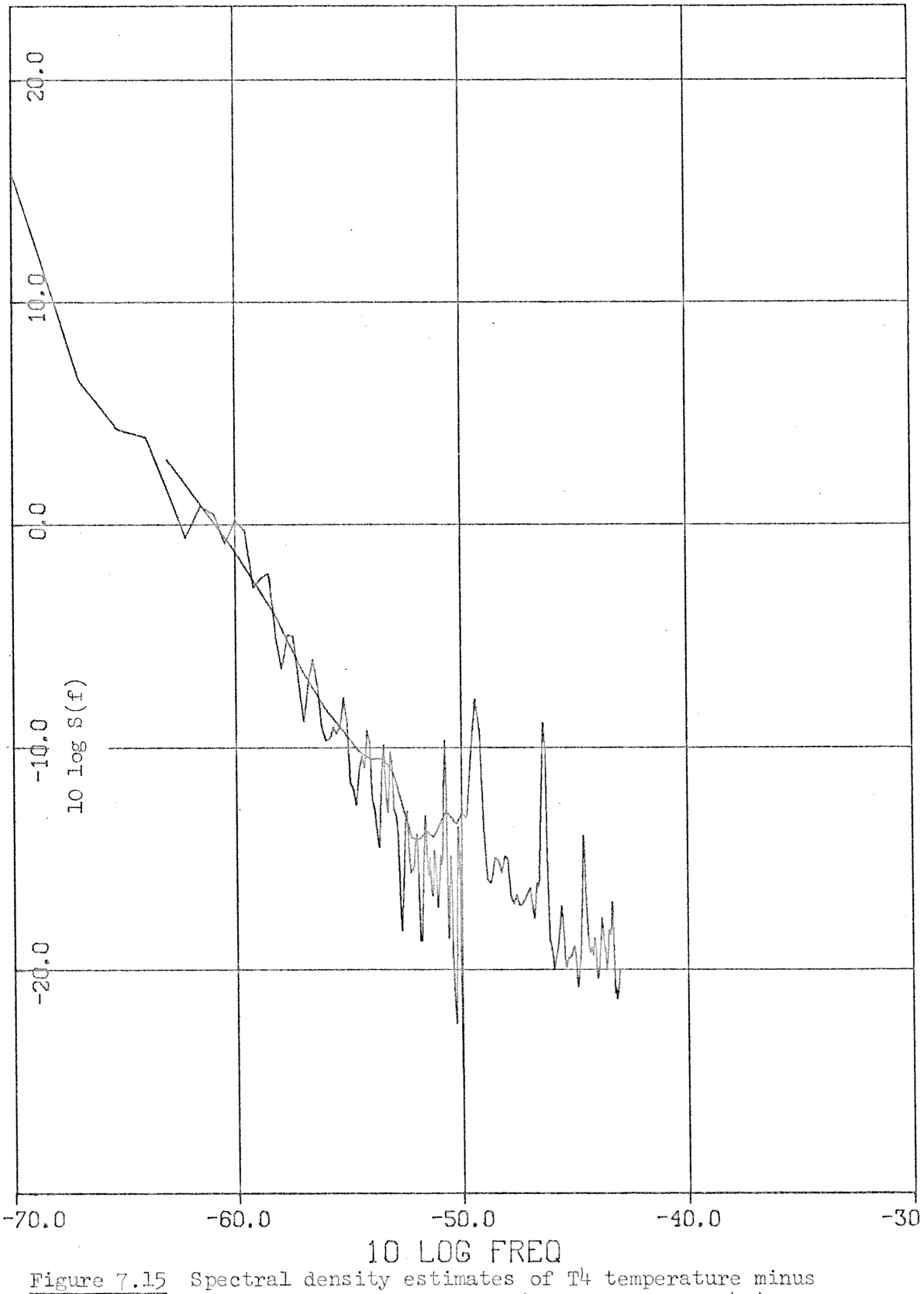


Figure 7.15 Spectral density estimates of T4 temperature minus normalized T2 temperature ($-70 \leq 10 \log f \leq -43$).

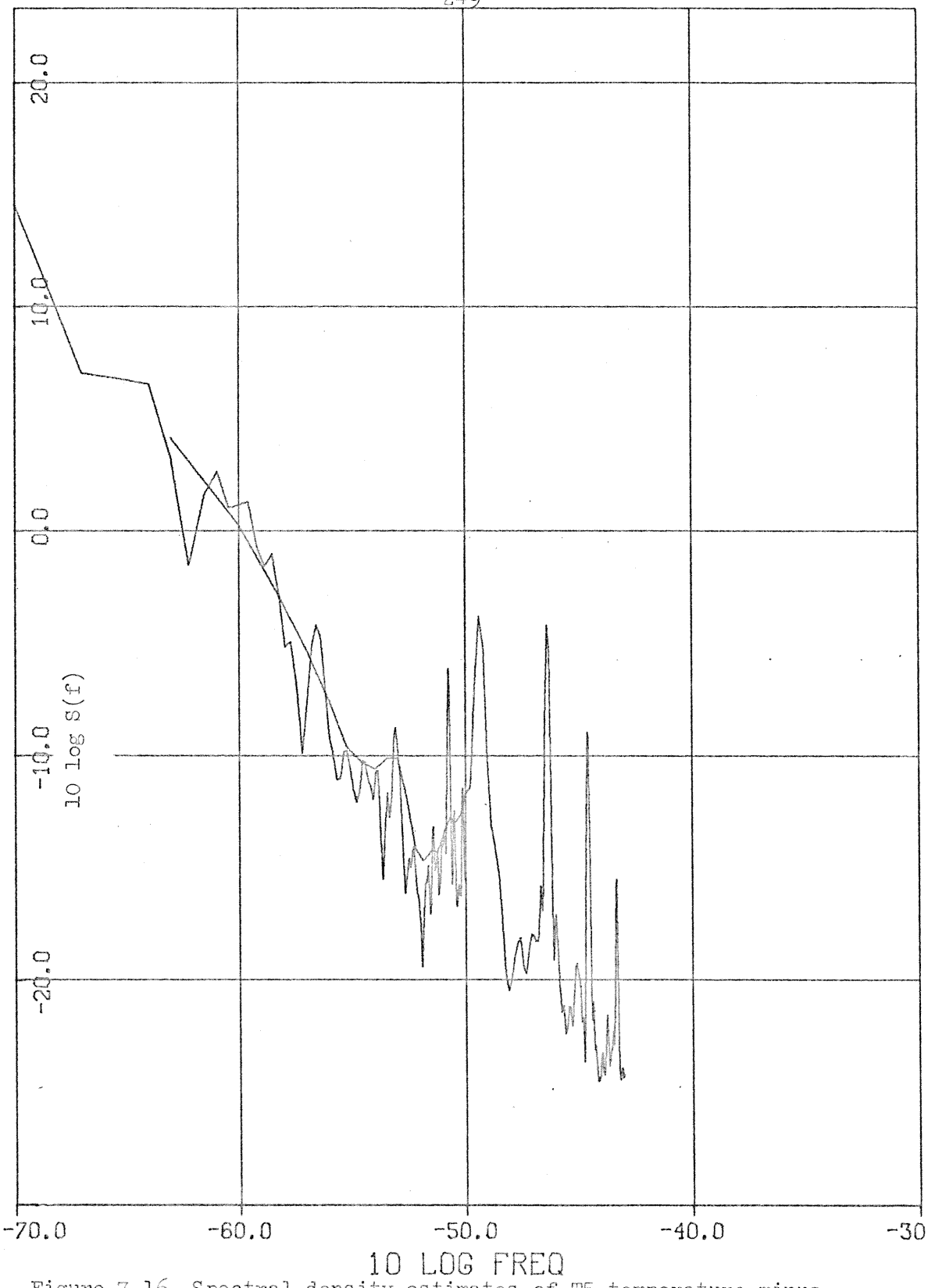


Figure 7.16 Spectral density estimates of T5 temperature minus normalized T2 temperature ($-70 \leq 10 \log f \leq -43$).

7.4 Power Supply Coefficients

The output of the noise source amplifiers is not expected to be entirely independent of supply voltage, although the component of output noise due to supply voltage noise should be negligible based on typical specifications given by the operational amplifier manufacturer. The high degree of regulation which is performed prior to the supply lines entering the oven and the high loop gain of these regulators assures that the supply voltage fluctuations (see Fig. 7.1) are dominated by the internal noise of the regulator supply circuitry.

For noise source #2 there are no data concerning the power supply noise. This fact is not disastrous since the power supply noise is itself just a form of semiconductor noise resulting in an observed noise which originates from a more complex semiconductor source than originally intended.

In the case of noise sources T1,...,T5, however, one would like to assure that the contribution due to the power supply (which is common to all five noise sources) is in fact negligible to assure that the noise data represent 5 statistically independent noise sources. It is for this purpose that V_{-check} and V_{+check} were measured during the course of the experiment.

Assume that $N(t) = n(t) + ks(t)$ and $S(t) = s(t) + r(t)$ (where $N(t)$ is the observed noise, $s(t)$ is either the + or - supply voltage and $r(t)$ is the reference supply voltage), and that $n(t)$, $s(t)$ and $r(t)$ are zero mean, stationary, statistically independent processes. All of the values will be expressed in the units of volts at the appropriate

measuring channel. We will also assume that the power spectral densities of the processes are similar, $S_r(f) \sim k_1 S_s(f) \sim k_2 S_n(f)$, since they are all semiconductor noise and that $k_1 \sim 1$ since s(t) and r(t) are generated by regulators which employ similar components (see Fig. 4.4).

Under these circumstances an obvious estimator for k (given that S(t) and N(t) are known for $0 \leq t \leq T$) is:

$$\hat{k} = 2 \frac{\frac{1}{T}\int_0^T N(t)S(t)\,dt}{\frac{1}{T}\int_0^T S^2(t)\,dt} \qquad \text{[Eq. 7.4.1]}$$

$$\hat{k} = \frac{2\,\frac{1}{T}\int_0^T k\,s^2(t) + n(t)s(t) + n(t)r(t) + ks(t)r(t)\,dt}{\frac{1}{T}\int_0^T s^2(t) + r(t)s(t) + r^2(t)\,dt} \qquad \text{[Eq. 7.4.2]}$$

Taking the expectations of the numerator and denominator individually yields:

$$\hat{k} \to k \frac{2}{1 + k_1} \cong k$$

under the assumption that $k_1 \cong 1$.

The variance of this estimator would be extremely difficult to calculate although a reasonable estimate of the accuracy may be obtained by calculating the variances for the various terms in the numerator and

denominator of Eq. 7.4.2 individually.

Let

$$x = \frac{1}{T}\int_0^T s(t)r(t)\,dt \qquad \text{and} \qquad y = \frac{1}{T}\int_0^T s^2(t)\,dt \quad .$$

An accurate estimate for k requires that $\frac{\sigma_x}{E\{y\}}$ be made as small as possible where prewhitening might be considered as a tool to achieve an improved estimate.

$$\sigma_x^2 = E\left\{\frac{1}{T^2}\int_0^T\int_0^T s(t)s(\tau)r(t)r(\tau)\,dt\,d\tau\right\}$$

$$= \frac{1}{T^2}\int_0^T\int_0^T R_s(t-\tau)\,R_r(t-\tau)\,dt\,d\tau$$

$$= \frac{1}{T^2}\int_{-T}^T (T-|\tau|)\,R_s(\tau)R_r(\tau)\,d\tau$$

$$\sigma_x^2 = \frac{1}{T^2}\left\{W(f) * S_s(f) * S_r(f)\right\}\Big|_{f=0}$$

where * denotes convolution and

$$W(f) = \frac{\sin^2(\pi fT)}{\pi^2 f^2} \quad .$$

Since both s(t) and r(t) are semiconductor noise, a reasonable approximation to their apparent spectral densities might be:

$$S_r(f) \tilde{=} \begin{cases} \frac{A}{|f|}, & \frac{1}{T} \le |f| \le \frac{1}{2\Delta t} \\ 0, & \text{otherwise} \end{cases}$$

$$S_s(f) \tilde{=} \begin{cases} \frac{B}{|f|}, \frac{1}{T} \leq |f| \leq \frac{1}{2\Delta t} \\ 0, \quad \text{otherwise} \end{cases}$$

The low frequency cutoff would occur due to the removal of sample means from the original data.

These assumptions imply:

$$S_r(f)*S_s(f) \sim \begin{cases} 2\int_{f+\frac{1}{T}}^{1/2\Delta t} AB \frac{1}{\nu}\frac{1}{\nu-f}\, d\nu, & 0 \leq f \leq \frac{2}{T} \\ \int_{f+\frac{1}{T}}^{1/2\Delta t} AB \frac{1}{\nu}\frac{1}{\nu-f}\, d\nu + \int_{\frac{1}{T}}^{f-\frac{1}{T}} AB \frac{1}{\nu}\frac{1}{|-f+\nu|}\, d\nu, & \frac{2}{T} \leq f \leq \frac{1}{2\Delta t} - \frac{1}{T} \end{cases}$$

$$= \frac{2AB}{f} \begin{cases} \ell n[(1-2\Delta t f)(1+Tf)], & 0 \leq f \leq \frac{2}{T} \\ \ell n[(1-2\Delta t f)(1+Tf)(Tf-1)], & \frac{2}{T} \leq f \leq \frac{1}{2\Delta t} - \frac{1}{T} \end{cases}$$

$$\sim \frac{2AB}{f} \begin{cases} \ell n(1+Tf), & 0 \leq f \leq \frac{2}{T} \\ \ell n([Tf]^2-1), & \frac{2}{T} \leq f << \frac{1}{2\Delta t} \end{cases}$$

if $T >> 2\Delta t$. This yields (very roughly):

$$\sigma_x^2 = \frac{1}{T^2}\left\{ W(f) * S_s(f) * S_r(f) \right\}\Big|_{f=0} \sim AB$$

$$\frac{\sigma_x}{E\{y\}} \sim \frac{AB}{\frac{1}{2\Delta t}} = \frac{1}{2}\sqrt{\frac{A}{B}} \quad \frac{1}{\ell n\left(\frac{T}{2\Delta t}\right)} \cdot$$

$$2\int_{\frac{1}{T}} (B/f)\, df$$

The same calculation, however, assuming prewhitening

$$\left(S_r(f) \sim \begin{cases} A, |f| \leq \frac{1}{2\Delta t} \\ 0, |f| > \frac{1}{2\Delta t} \end{cases}, \quad S_s(f) = \begin{cases} B, |f| \leq \frac{1}{2\Delta t} \\ 0, |f| > \frac{1}{2\Delta t} \end{cases} \right) \text{ yields:}$$

$$\sigma_x^2 = \frac{2}{T^2} \int_0^{\frac{1}{\Delta t}} \frac{\sin^2(\pi f T)}{\pi^2 f^2} AB \left(\frac{1}{\Delta t} - f\right) df$$

$$\sim \pi \sqrt{\frac{AB}{T\Delta t}}$$

and

$$\frac{\sigma_x}{E\{y\}} \sim \pi \sqrt{\frac{A}{B}} \sqrt{\frac{\Delta t}{T}} \cdot$$

It appears, therefore, that prewhitening holds a clear advantage for large values of $T/\Delta t$.

The variances for terms such as $y = \frac{1}{T}\int_0^T s^2(t)dt$ require more information about the statistics of the process. If one assumes that $s(t)$ is a gaussian process then:

$$\sigma_y^2 = \frac{4}{T^2}\int_0^T (T-\alpha)\, R_s^2(\alpha)\, d\alpha$$ which for band-limited white gaussian noise

$$\left(S_s(f) = \begin{cases} B, |f| \le \frac{1}{2\Delta t} \\ 0, |f| > \frac{1}{2\Delta t} \end{cases} \right)$$

implies

$$\sigma_y^2 \sim \frac{2B^2}{T\Delta t} \text{ for } T >> \Delta t \text{ or } \frac{\sigma_y}{E\{y\}} \sim \sqrt{\frac{2\Delta t}{T}} \text{ for } T >> \Delta t \ .$$

We are now prepared to place tolerance estimates on each of the terms of Eq. 7.4.2 assuming prewhitening and that

$$S_r(f) \sim S_s(f) \sim k_2 S_n(f) \sim \begin{cases} A, |f| \le \frac{1}{2\Delta t} \\ 0, |f| > \frac{1}{2\Delta t} \end{cases} .$$

$$\hat{k} \sim 2\, \frac{k \pm k\sqrt{\frac{2\Delta t}{T}} \pm 2\pi\sqrt{\frac{1}{k_2}}\sqrt{\frac{\Delta t}{T}} \pm k\pi\sqrt{\frac{\Delta t}{T}}}{2 \pm 2\sqrt{\frac{2\Delta t}{T}} \pm \pi\sqrt{\frac{\Delta t}{T}}}$$

For $T = 2000\ \Delta t$ (as will be used in the estimates made later).

$$\hat{k} \sim 2\, \frac{k(1 \pm 0.1) \pm 0.14\sqrt{1/k_2}}{2 \pm 0.133}$$

and if $k_2 > 5$ as is suggested by careful comparison of Figs. 7.4,..., 7.8, 7.10 and 7.11 then we have $\hat{k} \sim k(1 \pm 0.1) \pm 0.06$.

The experimental data were prewhitened at $\Delta t = 10^4$ seconds using the standard prewhitening filter. 2039 data points were used for all noise sources except T5 for which only 1752 points were available. The resulting estimates (replacing summations for integrals in Eq. 7.4.1) are:

Noise Source	$\hat{k}$ (+ supply)	$\hat{k}$ (-supply)
T1	0.014	-0.021
T2	0.002	-0.001
T3	0.013	0.005
T4	-0.032	0.029
T5	0.033	0.042

Since all of these estimates are below the estimated error of $\pm$ 0.06 it seems safe to say (under all of the assumptions made in this section) that the total spectral component due to both power supply voltages lies below the curve resulting from shifting the power spectral density estimates for either V_{+check} or V_{-check} down 24 db (the spectral density estimate for either V_{-check} or V_{+check} represents the power spectral density for a sum of two power supply voltages).

Comparison of Figs. 7.11, 7.4,...,7.8 indicates that it is very unlikely that power supply voltage contributes more than 10% to the power spectral density at any given frequency. As a point of interest, the values of k indicated (ranging from $|k| = 0.001$ to 0.042) represent power supply rejection ratios of -112 db to -80.0 db, and $|k| = 0.06$

represents a supply rejection ratio of -76 db referenced to the input of each operational amplifier. The manufacturer specifies[18] worst case supply rejection ratios of -76 db and -74 db, and typical rejection ratios of -90 db and -92 db for the 747 and 709 operational amplifiers respectively. The above estimate of k = 0.06 corresponds almost exactly with the manufacturer's worst case specifications, giving an additional degree of confidence that the power supply noise contributes less than 10% to the noise power spectral density at any frequency.

Chapter VIII

CONCLUSIONS

How well does the $f^{-\alpha}$ model fit the observed phenomenon and what is the "correct" value of α? Having estimated the power spectral densities of six noise sources over more than 9 decades of frequency, one might hope to shed some light on this question. As can be seen in Figs. 7.4, 7.5, and 7.6, a straight line fits the spectral density estimates for T1, T2, and T3 quite well over many decades (the leveling out into white noise is to be expected at high frequencies and the low frequency bias for T1 and T2 has already been discussed). These straight line approximations yield α's of 1.16, 1.21, and 1.18. Noise sources #2, T4, and T5 (in Figs. 7.3, 7.7, and 7.8), however, seem to defy the straight line model. If one approximates a straight line over a few decades at a time, the slope, α, seems to increase monotonically with decreasing frequency (just the opposite of what is needed to get finite power) giving local values of α between 0.87 and 1.36 for #2, 0.87 and 1.24 for T4, and 0.91 and 1.24 for T5.

Another method of estimating α, which has the virtue of giving a single value representing the overall trend, is to draw a line through the end points. To avoid possible bias at the high and low frequency ends, points were chosen at 10 log f = -60 and 17. This method yields values for α of 1.09, 1.05, and 1.05 for #2, T4, and T5.

The "best" estimates for noise sources #2, T1, T2, T3, T4, and T5 will be assumed to be α = 1.09, 1.16, 1.21, 1.18, 1.05, and 1.05. It is

this author's prejudice that a value of $\alpha = 1.0$ is a more "reasonable" constant for a physical process than 1.12 (the average of the above estimates). The fact that none of the above estimates fall below the magic value of 1.0, however, lends little support to this hypothesis. It is, of course, possible that some unforeseen bias in the experiment is responsible for an inflated value for α. It is also possible, on the other hand, that the apparent increasing value of α with decreasing frequency (observed in three of the noise source estimates) reveals the true character of this process.

Perhaps more important than establishing a value for α, however, is establishing bounds on the low frequency break in the 1/f characteristic (if such a break exists). None of the noise sources (Figs. 7.3,...,7.8) give any indication of a break even down to 10^{-7} Hz (note, however, that spectral estimates for noise sources T1 and T2 are certainly not good below 3×10^{-7} Hz and questionable for higher frequencies).

A more careful statement which takes into account the variance of the estimates, the low frequency bias due to the spectral window used (see Section 2.8), and a particular form of low frequency breakdown seems in order. It has already been mentioned that the standard deviation, σ, of the low frequency estimates is nearly 50% in some cases. Assume that none of the estimates for noise sources #2, T3, T4, and T5 (Figs. 7.3, 7.6, 7.7, and 7.8) fall more than 3 db below the 1/f trend line (the placement of this line is highly subjective, although I feel that this assumption holds for rather pessimistic placement for the noise sources indicated), then the probability is very low that the

expected value for the estimate could lie more than 6db below the trend line (this would require an estimate more than 2σ above the mean value). Referring to Table 2.1, if the true spectral density is of the form $\frac{1}{|f| + f_c}$ then the probability that $f_c \geq 5 \times 10^7$ Hz is negligible. We may conclude, therefore, that the true power spectral density (for noise sources #2, T3, T4, T5) cannot (with reasonable probability) closely resemble $\frac{1}{|f| + f_c}$ at low frequencies with 10 log $f_c \geq -63$.

Spectral estimates for noise sources T1 and T2, however, are thought to be dominated by some unknown parameter at the lowest frequencies. If we assume that these estimates are good for $f \geq 10^{-6}$ Hz then the above argument implies that the true spectral density is not likely to resemble $\frac{1}{|f| + f_c}$ with 10 log fc ≥ -55.23 ($f_c \geq 3 \times 10^{-6}$ Hz).

I consider both of these estimates to be rather conservative. The $f = 10^{-7}$ Hz estimate of T5 (Fig. 7.8) is nearly 3 db above the trend line, which implies (for this case) that a break of the form $\frac{1}{|f| + f_c}$ with $f \geq 10^{-7}$ Hz is very improbable.

Popcorn noise has been observed with waiting times greater than 10^6 seconds. This fact lends support to the hypothesis that 1/f noise may result from a sum of many such components (see Appendix A for example). Waiting time distributions for two popcorn components (having estimated mean waiting times of 1.4×10^4, 4.6×10^3, 1.8×10^4, and 1.6×10^4 seconds) and spectral estimates of noise containing these components show good agreement with previous estimates by Puckett[3] at higher frequencies. Other popcorn components have been observed, however, which indicate the possibility of more complex behavior.

Appendix A

SIMULATED 1/f NOISE

One hypothesis as to the origin of 1/f noise is that it is the sum of a large number of burst components. As has already been discussed, the power spectral density of a random telegraph wave (which appears to accurately describe popcorn noise) is

$$S(f) = \frac{k_1}{f_o} \Bigg/ \left[1 + \left(\frac{f}{f_o}\right)^2\right]$$

where

$$k_1 = \frac{\alpha h^2}{(1+\alpha)^2 \pi} \quad ,$$

$$\alpha = \frac{\tau_-}{\tau_+} \quad , \text{ and}$$

$$f_o = \frac{1}{2\pi}\left(\frac{1}{\tau_+} + \frac{1}{\tau_-}\right)$$

(see Eq. 5.7.1).

For simplicity, assume $\tau = \tau_+ = \tau_-$ (it is clear that the following argument is easily extended for $\tau_+ \neq \tau_-$ if the α's are similarly distributed for all values of f_o). We have, therefore, for each burst component

$$S(f) = \frac{h^2}{2\pi} \frac{1}{f_o} \frac{1}{1 + \left(\frac{f}{f_o}\right)^2}$$

where

$$f_o = \frac{1}{\pi\tau} \quad .$$

Burst components have been observed with apparent time constants large enough to account for 1/f noise down to the lowest frequencies measured. This fact makes the hypothesis appealing since physical processes with time constants on the order of 10^6 seconds are rather hard to come by within a small integrated circuit. The question is, how does one choose a physically realistic distribution for the burst amplitudes and time constants to create a 1/f spectral density.

Bernamont[24] was probably the first of many[25,26] to recognize that a 1/f spectral density could be obtained by the following method. The physical model for burst noise which will be used is apparently due to C.A. Mead and R.J. Whittier[3, page 38]. The theory supposes the existence of trapping centers near the base emitter junction of the transistor (they could be located in the oxide layer above the junction for example) which are positioned in such a way as to modulate the flow of current in the junction as an electron is trapped or released. The detailed physics of the model is beyond the scope of this discussion. The only point which is important to the following discussion is to note that the probability (per unit time) of an electron tunneling through to a trapping center (or leaving it) is proportional to e^{-2kx} where x is the thickness of the non-conducting medium which the electron must tunnel through and k is a constant of the materials involved. This gives $f_o \propto e^{-2kx}$ where f_o is the characteristic frequency of the burst component created by this trapping center. If we assume that these trapping centers are located in the oxide layer (we are ignoring the question of how (or whether) a single electron trapped in the oxide layer might cause such a macroscopic effect as popcorn noise[3]) an

obvious choice for the distribution of trapping centers as a function of x (the distance into the oxide layer) is a uniform distribution over most of the width of the layer.

The relationship $f_o \propto e^{-2kx}$ indicates that to produce an n decade range of characteristic frequencies we require a region of thickness $\frac{n}{2k \log e}$ in the oxide layer. Since 1/k is normally less than a few angstroms[19], characteristic frequencies ranging from 10^{-7} Hz to 10^{3} Hz might be explained by the trapping centers in a ~ 20 Å thick region in the oxide layer. It seems, therefore, not unreasonable to assume that the amplitudes of the burst components, over the region of frequency with which we are concerned, are essentially independent of x.

Using again the relationship $f_o \propto e^{-2kx}$ we find $d(\log f_o) = -(2k \log e)\, dx$. Therefore, a uniform density of trapping centers in x implies a uniform distribution in log f of burst components. Since the spectral density of a burst component at its characteristic frequency, f_o, is proportional to $1/f_o$ it appears that a 1/f spectrum is likely to result from the summation.

Assume N trapping centers uniformly distributed in the range $x_o \leq x \leq x_1$ which each produce a burst component of amplitude $h = \sqrt{\Delta x}\ h_o$ (the Δx has been included to prevent the spectral density from diverging as we let $N \to \infty$) and $f_o = c\, e^{-2kx}$. The expected noise spectral density (assuming that the components add and are uncorrelated) is:

$$S_\Sigma(f) = \sum_{j=0}^{N} \frac{h_o^2 \, \Delta x}{4\pi} \; \frac{1}{ce^{-2k(x_o + j\Delta x)}} \; \frac{1}{1 + \frac{f^2}{c^2} \, e^{4k(x_o + j\Delta x)}}$$

where

$$\Delta x = \frac{x_1 - x_0}{N} \quad .$$

$$\lim_{N\to\infty} S_\Sigma(f) = \int_{x_0}^{x_1} \frac{h_o^2}{4\pi} \; \frac{1}{ce^{-2kx}} \; \frac{1}{1 + \frac{f^2}{c^2} \, e^{4kx}} \, dx$$

$$= \frac{h_o^2}{8\pi} \; \frac{1}{f} \int_{x_0}^{x_1} \frac{1}{\cosh(2kx + \ell n \, \frac{f}{c})} \, dx$$

$$= \frac{h_o^2}{8\pi} \; \frac{1}{kf} \; \tan^{-1} \left[\frac{\frac{f}{c} \, (e^{2kx_1} - e^{2kx_0})}{1 + \frac{f^2}{c^2} \, e^{2kx_1} \, e^{2kx_0}} \right] \quad \text{[Ref. 17, Eq. 1.490.1]}$$

Let

$f_{max} = ce^{-2kx_0}$, $f_{min} = ce^{-2kx_1}$, then

$$\lim_{N\to\infty} S_\Sigma(f) = \frac{h_o^2}{8\pi k} \; \frac{1}{f} \; \tan^{-1} \left[\frac{f_{max} - f_{min}}{\frac{f_{min} f_{max}}{f} + f} \right]$$

$$\simeq \frac{h_o^2}{8\pi k} \; \frac{1}{f} \; \tan^{-1} \left[\frac{1}{\frac{f_{min}}{f} + \frac{f}{f_{max}}} \right] = \frac{1}{f} \, A(f) \, \frac{h_o^2}{8\pi k}$$

for $f_{min} << f_{max}$.

In the range $f_{min} \leq f \leq f_{max}$, $A(f) \sim \frac{\pi}{2}$ and $10 \log[A(f)/(\frac{\pi}{2})] = -3$ db occurs at $f = f_{min}$ and $f = f_{max}$ so that

$$\lim_{N\to\infty} S_{\Sigma}(f) \propto 1/f \quad \text{is}$$

is a good approximation for all $f_{min} < f < f_{max}$.

A low density of trapping centers might also be expected to produce a 1/f trend but with some ripple. If we assume equal amplitude burst components spaced by factors of β in frequency (that is, spaced by $10 \log \beta$ on the $10 \log f$ axis) then the expected spectral density is:

$$S_{\Sigma}(f) = \sum_{j=0}^{N-1} \frac{k}{\beta^j f_{min}} \Bigg/ \left[1 + \left(\frac{f}{\beta^j f_{min}}\right)^2\right].$$

We are interested in the deviation of this spectral density from 1/f. Let

$$D(f) = S_{\Sigma}(f) \Big/ (1/f) \quad \text{then}$$

$$D(f) = \sum_{j=0}^{N-1} \frac{kf}{\beta^j f_{min}} \Bigg/ \left[1 + \left(\frac{f}{\beta^j f_{min}}\right)^2\right] .$$

For large β ($\beta \geq 2$ for example) and large N it is clear that only a few terms need be evaluated for middle range frequencies.

Evaluating D(f) for $\beta = 10$ (one burst component per decade of frequency) reveals a peak to peak ripple in the middle frequency ranges of only about 4.5 db. For $\beta = 2$ (one component per octave of frequency) the peak to peak ripple of D(f) becomes less than 0.0001 db. It is, therefore, possible to explain a spectral density, which deviates from

1/f by much less than the accuracy with which we might ever hope to measure, over 20 decades (10^{-10} to 10^{10} Hz for example) as a sum of only ~ 70 burst components (to estimate the spectral density of a data array over only one decade to a standard deviation of .0001 db would require ~ 1.9×10^{10} samples; ~ 600 years at $\Delta t = 1$ second).

At this point we have a method for simulating 1/f noise over many decades in the digital computer. It seems, therefore, to be an excellent opportunity to test the spectral estimator used in evaluating the experimental data by applying it to a data array with known statistics. It may also be constructuve to observe this function in the time domain as compared to the experimentally obtained data.

Producing a very long data record (so that the simulated noise could be analyzed over several decades in exactly the same manner as the experimental data) required a large amount of computing time (although not a particularly sophisticated program). A program was, therefore, written for an EMR model 6130 computer for which no accounting was required (the program required several hours of machine time). The program listed in Figs. A.1a and A.1b consists of a random number generator subroutine which provides numbers drawn from an exponential distribution (with variable time constant) and a main program which keeps track of the switching times and states of each of the burst components.

The random number generator[20] is composed of two parts. A sequence of random numbers, u_i, which are uniformly distributed over the range 1 to 2^{30} is generated by the formula $u_{i+1} = (au_i + c)_{\text{mod } 2^{30}}$ using

```
      INTEGER U0,U1,H02,DATA,AMP,I,J,K
      DOUBLE PRECISION PI,TT,CT,T,DT,A
      DIMENSION A(20),T(20),AMP(20),DATA(2232)
      PI=3.1415926
      DT=1.0
      H02=1
      U0=578
      U1=783
      TT=0.0D0
      A(1)=2.0*DT/PI
      DO 1 I=2,20
1        A(I)=A(I-1)*2.0

      DO 2 I=1,20
         AMP(I)=H02
2        CALL NEGEXP(T(I),A(I),U1,U0)
      DO 7 I=1,500
         I1=0
         DO 6 J=1,2000
              TT=TT+DT
              DO 4 K=1,20
3                 IF(T(K),GE,TT) GO TO 4
                  I1=I1+1
                  CALL NEGEXP(CT,A(K),U1,U0)
                  IF(CT.GT.1.0D6) CT=1.0D6
                  T(K)=T(K)+CT
                  AMP(K)=-AMP(K)
                  GO TO 3
4                 CONTINUE
              DATA(J) = AMP(1)
              DO 5 K=2,20
5                DATA(J)=DATA(J)+AMP(K)
6             CONTINUE
         DATA(2001)=I
         DATA(2002)=I1
         WRITE(6,END=7,ERR=100) DATA
7        CONTINUE
100   ENDFILE 6
      STOP
```

Figure A.1 a Computor program to produce simulated 1/f noise.

```
      SUBROUTINE NEGEXP(T,AA,X1,X0)
      DOUBLE PRECISION DP,T,AA
      INTEGER X0,X1
      INTEGER Y0,Y1,C0,C1,A0,A1
      A0=5
      A1=4096
      C0=12449
      C1=9587
/I
      LDA    A1
      MPY    X0
      STE    Y1
      LDA    A0
      MPY    X1
      MOV    E,A
      ADD    Y1
      MOV    A,E
      AND    =$77777
      STA    Y1
      LDA    A0
      MPY    X0
      ADD    Y1
      MOV    A,E
      AND    =$77777
      ADD    C1
      MOV    A,E
      AND    =$77777
      STA    Y1
      LDA    A0
      MPY    X0
      MOV    E,A
      ADD    C0
      STA    Y0
      JLZ    )100
      LDE    Y1
      JMP    )110
100   MOV    A,E
      AND    =$77777
      STA    Y0
      LDE    Y1
      MAO    E,E
110   AND    =$77777
      STA    Y1
/F
      X0=Y0
      X1=Y1
      IF(Y1.EQ.0) Y1=1
      DP=DFLOAT(Y1)
      DP=DP/3.2768D4
      DP=DLOG(DP)
      T=-DP*AA
      RETURN
```

Figure A.1 b Exponential distribution random number generator program.

the values a = 134217733 and c = 314159265. The modulus 2^{30} results from the 16 bit word size of the computer and the method by which double precision arithmetic was implemented by the program. The assembly language sequence in the subroutine (Fig. A.1b) is necessary to perform the modulus 2^{30} arithmetic since the FORTRAN compiler had no provisions for double precision integers. A random number, t_i, drawn from an exponential distribution is then found by evaluating $t_i = -\tau \, \ell n(u_i/2^{30})$. The main program then constructed an array of 10^6 points which is the sum 20 ideal burst components with characteristic frequencies $f_i = \frac{1}{\Delta t} 2^{-i}$ for i = 1,...,20.

This array was recorded on magnetic tape and is plotted in Fig. A.2. The vertical scale in this figure assumes that the peak to peak magnitude of each popcorn component, h, is 0.025 volts. In order to accurately simulate the conditions under which the experimental data were analyzed, a bias of 1.0 volt (at this arbitrary scale) was added to the data. A sampling period of Δt = 10 seconds was assumed for the initial data and spectral estimates were made at Δt = 10,100,10^3, 10^4, and 5 x 10^4 seconds (~10^4 points are used for all spectral estimates except ~10^3 points at $\Delta t = 10^4$ sec and ~200 pts at $\Delta t = 5 \times 10^4$ seconds) using the same resampling filters used with the experimental data. The expected spectral density is

$$10 \log S(f) \cong 10 \log \frac{h^2}{4\pi f} \sum_{j=-\infty}^{\infty} \frac{1}{2^j + 2^{-j}}$$

$$\cong -10 \log f - 39.48 \quad \text{for } h = 0.025 \text{ volts.}$$

At 10 log f = -40, 10 log S(f) $\cong$ 0.52. This agrees quite well with the

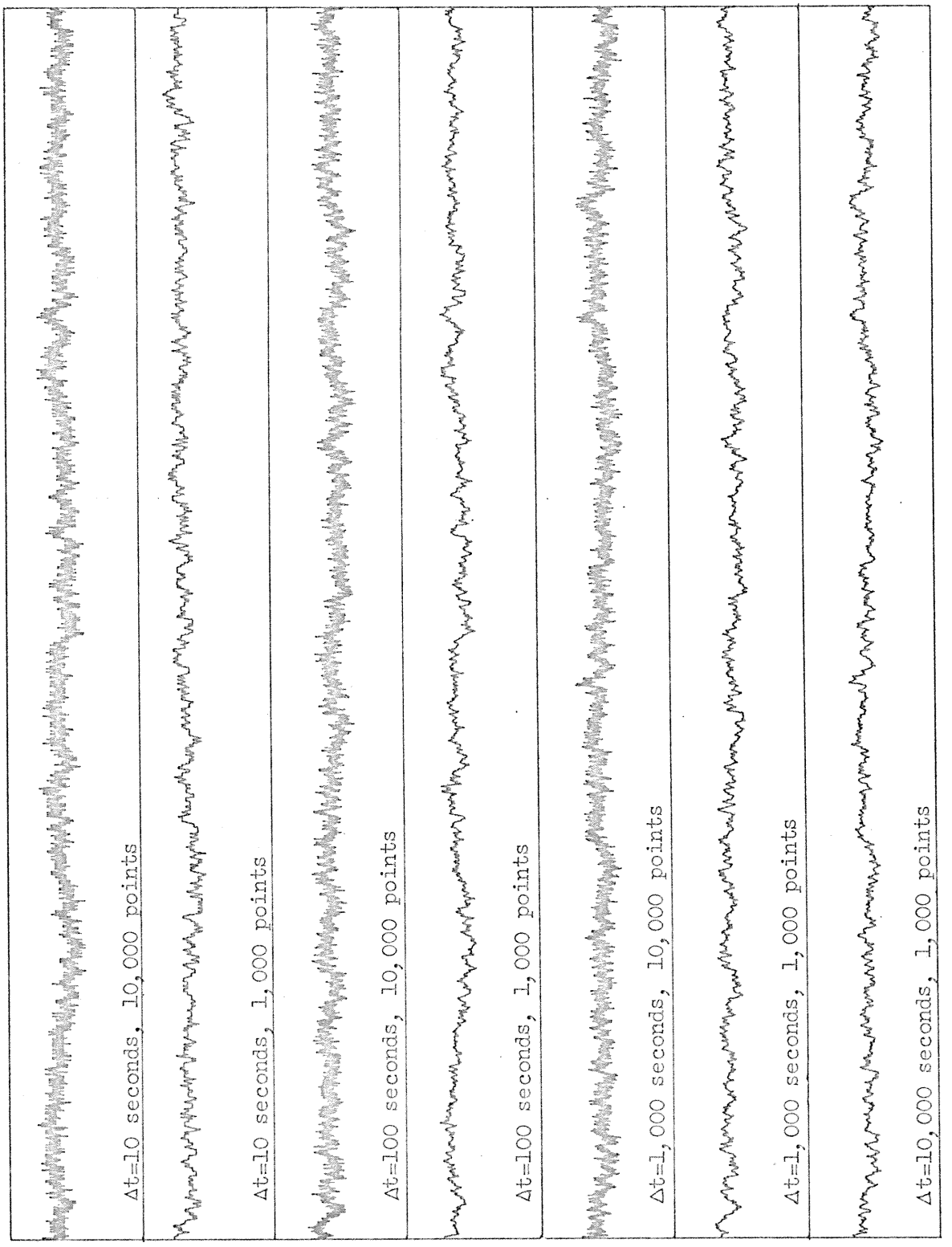


Figure A.2 Time domain plots of simulated 1/f noise.

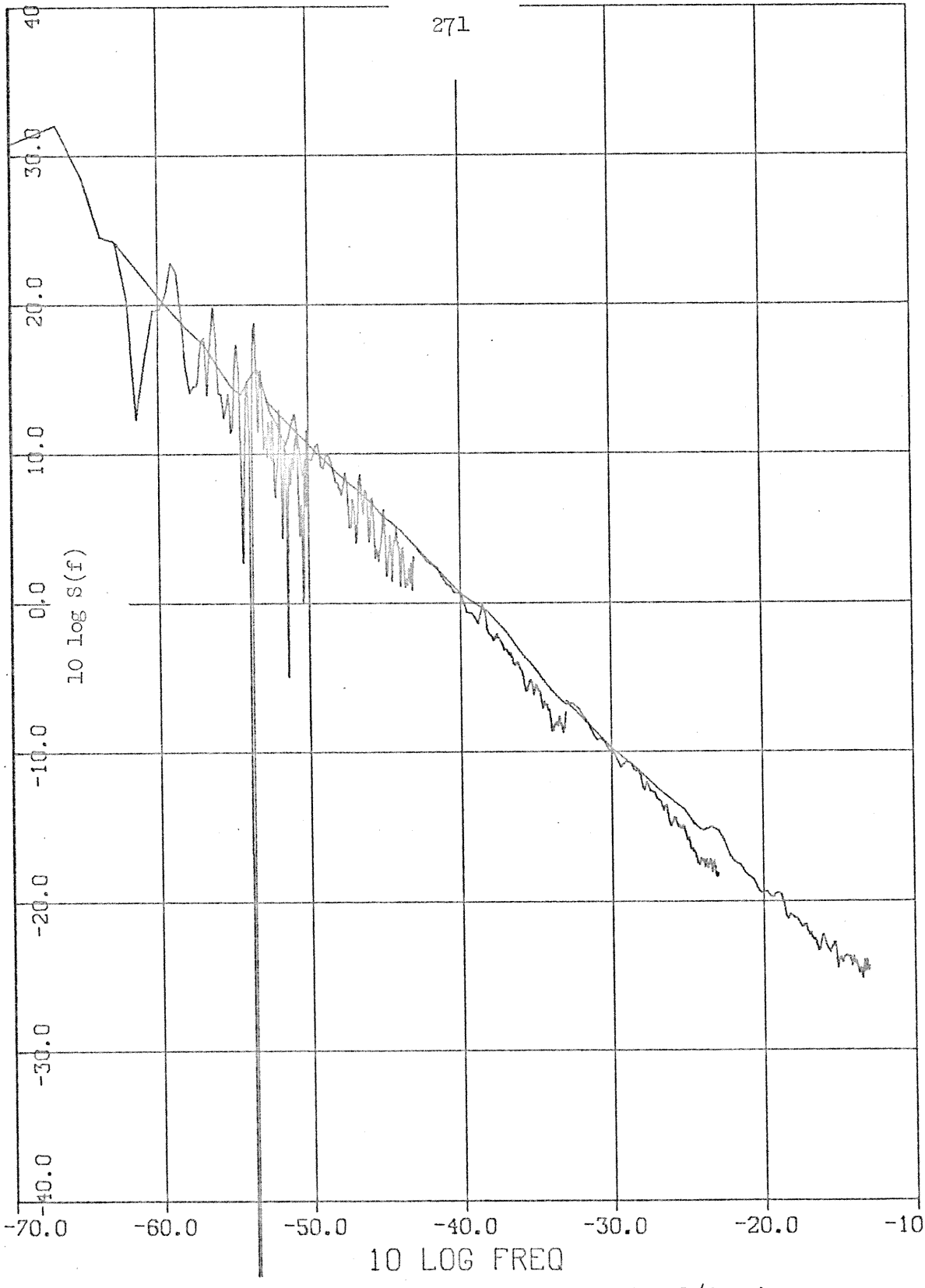


Figure A.3 Spectral density estimates for simulated 1/f noise ($-70 \leq 10 \log f \leq -13$).

estimates shown in Fig. A.3.

For power spectral density estimates over two decades of frequency, the expected standard deviation is $\sigma \sim \sqrt{\frac{100}{N'}}\ S(f)$ where $N' \sim N - 30$. Therefore

$$\left\{ \begin{array}{l} 10 \log \left(1 + \frac{\sigma}{S(f)}\right) \\ \\ 10 \log \left(1 - \frac{\sigma}{S(f)}\right) \end{array} \right.$$

are expected to be

$$\left\{ \begin{array}{l} 0.4 \text{ db} \\ -0.5 \text{ db} \end{array} \right. \quad \text{for } N = 10^4 \quad ,$$

$$\left\{ \begin{array}{l} 1.2 \text{ db} \\ -1.7 \text{ db} \end{array} \right. \quad \text{for } N = 10^3 \quad , \text{ and}$$

$$\left\{ \begin{array}{l} 2.5 \text{ db} \\ -6.3 \text{ db} \end{array} \right. \quad \text{for } N = 200 \quad .$$

Figure A.3 appears to agree quite well with these predicted variances and demonstrates no unexpected biases (the low estimates at the high frequency end of each estimation run are the result of the low pass resampling filter frequency response).

The simulated data in the time domain (Fig. A.2) resemble the experimental data to an extraordinary degree (particularly after filtering and resampling) even though it is composed of only 21 discrete

levels at the base sampling rate (note, in particular, the resemblance to noise source T5, the only source not demonstrating obvious popcorning, in Figs. 5.1,...,5.7). This, of course, leads to no conclusions except that the sum of popcorn component hypothesis has not been shown to be inconsistent with the observed data.

Appendix B

STANDARD DEVIATION FOR HISTOGRAMS

When attempting to estimate the probability distribution for a sample function we often produce a histogram showing the number, n, of samples falling in a given range of values. If we have N independent samples of the function, then we assume that the actual probability of the function taking on a value in this range for any given sample is $P \sim \frac{n}{N}$. This procedure is used several times in Chapter V for estimating the waiting time distribution of burst noise and the probability distribution of the difference function,

$$D_i = \frac{1}{L} \sum_{j=1}^{L} \left\{ N[(i - j + L)\Delta t] - N[(i - j)\Delta t] \right\} .$$

If the samples are independent, then

$$P\left\{ \begin{matrix} \text{k samples falling in given} \\ \text{range out of N samples} \end{matrix} \right\} = \binom{N}{k} p^k (1 - P)^{N-k}$$

(the well known binomial distribution).

$$E\{k\} = \sum_{k=0}^{N} k \binom{N}{k} p^k (1 - P)^{N-k}$$

$$= \sum_{k=0}^{N} \frac{N!}{(N-k)!(k-1)!} p^k (1 - P)^{N-k} \quad \text{let } m = k - 1$$

$$= \sum_{m=0}^{N-1} \frac{N(N - 1)!}{(N - 1 - m)!\, m!} p^{m+1} (1 - P)^{N-1-m}$$

$$= NP \sum_{k=0}^{N-1} \binom{N-1}{m} P^m (1 - P)^{N-1-m}$$

$$= NP \qquad \text{as we expected.}$$

$$\sigma^2_k = E\{k^2\} - E\{k\}^2$$

$$= E\{k^2 - k\} + E\{k\} - E\{k\}^2$$

$$= NP - N^2P^2 + \sum_{k=2}^{N} \frac{N!}{(N-k)!\,(k-2)!} P^k (1 - P)^{N-k} \qquad \text{let } m = k - 2$$

$$= NP - N^2P^2 + \sum_{m=0}^{N-2} \binom{N-2}{m} N(N-1)\, P^{m+2} (1 - P)^{N-2-m}$$

$$= NP - N^2P^2 + N(N-1)P^2 = NP - NP^2$$

$$\sigma_k = [NP(1 - P)]^{\frac{1}{2}}$$

In some cases we may wish to estimate σ_k after fitting a curve to the $\frac{k}{N}$ data for the various ranges considered. In other cases we will assume that our best estimate of P for each interval is $\frac{k}{N}$ in which case we will estimate σ_k to be

$$\hat{\sigma}_k = [k(1 - \tfrac{k}{N})]^{\frac{1}{2}} \quad .$$

Appendix C

DIGITAL DEALIASING FILTER

In Chapter II we require a digital low pass filter with a cut off frequency (3 db point) of approximately 0.1 x $\frac{1}{2\Delta t}$ (Δt is the present sampling period) so that the data may be resampled with a new sampling period of 10 Δt. If we consider x(t) and y(t) to be the input and output respectively of the filter, then the general form of the filter which we wish to consider is

$$y(t) = \sum_{k=0}^{N_1} a_k \, x(t-k\Delta t) - \sum_{k=1}^{N_2} b_k \, y(t-k\Delta t) \qquad \text{[Eq. C.1]}$$

Since the output at time t involves input and output values only at times $t-k\Delta t$ for $k = 0, 1, \ldots,$ it is sufficient to consider this continuous filter and separately consider the effects of sampling y(t) (which has already been done in Section 2.5), instead of considering the input as a sampled function as is actually the case when the filtering is performed in the digital computer.

If we assume that x(t) and y(t) are stationary processes then

$$\sum_{k=0}^{N_2} b_k \, y(t-k\Delta t) = \sum_{k=0}^{N_1} a_k \, x(t-k\Delta t) \quad \text{where} \quad b_0 = 1 \qquad \text{[Eq. C.2]}$$

$$\sum_{k=0}^{N_2}\sum_{j=0}^{N_2} b_k b_j\, y(t-k\Delta t)\, y(t+\tau-j\Delta t) = \sum_{k=0}^{N_1}\sum_{j=0}^{N_1} a_k a_j\, x(t-k\Delta t)x(t+\tau-k\Delta t)$$

$$\sum_{k=0}^{N_2}\sum_{j=0}^{N_2} b_k b_j\, C_y(\tau+(k-j)\Delta t) = \sum_{k=0}^{N_1}\sum_{j=0}^{N_1} a_k a_j\, C_x(\tau+(k-j)\Delta t)$$

$$S_y(f)\sum_{k=0}^{N_2}\sum_{j=0}^{N_2} b_k b_j\, e^{i2\pi f(k-j)\Delta t} = S_x(f)\sum_{k=0}^{N_1}\sum_{j=0}^{N_1} a_k a_j\, e^{i2\pi f(k-j)\Delta t}$$

$$|H(f)|^2 = \frac{S_y(f)}{S_x(f)} = \frac{\sum_{k=0}^{N_1}\sum_{j=0}^{N_1} a_k a_j\, e^{i2\pi f(k-j)\Delta t}}{\sum_{k=0}^{N_2}\sum_{j=0}^{N_2} b_k b_j\, e^{i2\pi f(k-j)\Delta t}} \qquad \text{[Eq. C.3]}$$

$$= \frac{\sum_{k=0}^{N_1} A_k \cos(2\pi f k\Delta t)}{\sum_{k=0}^{N_2} B_k \cos(2\pi f k\Delta t)}$$

$$= \frac{\sum_{k=0}^{N_1} \hat{A}_k \cos^{2k}(\pi f\Delta t)}{\sum_{k=0}^{N_2} \hat{B}_k \cos^{2k}(\pi f\Delta t)} \qquad \text{[Eq. C.4]}$$

In order to assure stability of the filter we must insist that all of the s-plane poles of the transfer function lie in the left half plane. Taking the Laplace transform of Eq. C.2 we have

$$H(s) = \frac{\sum_{k=0}^{N_1} a_k e^{-ksT}}{\sum_{k=0}^{N_2} b_k e^{-ksT}} \qquad \text{[Eq. C.5]}$$

If we let $z = e^{sT}$, then the poles of H(s) occur at the zeroes of $\sum_{k=0}^{N_2} b_k z^{N_2-k}$ and at $z = 0$. If z_k $(k = 1,\ldots,N_2)$ are the zeroes in the z-plane, then the zeroes in the s-plane occur when

$$e^{sT} = z_k, \; s = \frac{1}{T}\left[\ln|z_k| + i(\arg z_k \pm n\,2\pi)\right] \quad .$$

There are, therefore, an infinite number of poles in the s-plane for each pole in the z-plane and they all lie along a vertical line in the left half s-plane if and only if $|z_k| < 1$. The filter is stable if and only if all of the zeroes of $\sum_{k=0}^{N_2} b_k z^{N_2-k}$ lie inside the unit circle.

Equation C.4 demonstrates that for any real set of weighting coefficients a_k and b_k (such that the b_k satisfy the stability condition) the squared magnitude of the transfer function may be written as a

ratio of polynomials in $\cos^2(\pi f\Delta t)$ where the numerator is of order N_1 and the denominator is of order N_2. Clearly, for all ratios of polynomials in $\cos^2(\pi f\Delta t)$ there cannot exist a set of weighting coefficients yielding a filter with this magnitude squared transfer function. Consider for example, $\frac{1}{2}-\cos^2(\pi f\Delta t)$ which is negative for some real values of f. It is possible, however, to show that any ratio of polynomials in $\cos^2(\pi f\Delta t)$ with real coefficients and which is non-negative and has no poles for all real valued f can be obtained as the $|H(f)|^2$ of a filter of the form of Eq. C.1.

It is sufficient to consider the problem of finding a set of real valued a_k such that

$$\left(\sum_{k=0}^{N} a_k\, e^{-i2\pi f\Delta tk}\right)\left(\sum_{j=0}^{N} a_k\, e^{i2\pi f\Delta tj}\right) = F(f) \equiv \sum_{k=0}^{N} A_k\, \cos^{2k}(\pi f\Delta t)$$

where the A_k are real and $F(f) \geq 0$ for f real.

Let $z = e^{i2\pi f\Delta t}$; then

$$F(f) = \sum_{k=0}^{N} A_k \left(\frac{z^{\frac{1}{2}} + z^{-\frac{1}{2}}}{2}\right)^{2k} \qquad \text{[Eq. C.6]}$$

$$= \sum_{k=0}^{N} 2^{-k}\, A_k \left(\frac{z + z^{-1}}{2} + 1\right)^{k}$$

$$= \sum_{k=0}^{N} 2^{-k}\, A_k (x+1)^k \quad \text{where } x = \frac{z+z^{-1}}{2} = \cos(2\pi f\Delta t) \qquad \text{[Eq. C.7]}$$

$$= C \prod_{k=1}^{N} (x-x_k) \quad \text{where} \quad C \text{ is a real constant.} \qquad \text{[Eq. C.8]}$$

From Eq. C.7 we know that any non-real valued x_k must occur in conjugate pairs. Since $C \prod_{k=1}^{N} (x-x_k) \geq 0$ for all real x such that $-1 \leq x \leq 1$, any real x_k such that $-1 < x_k < 1$ must occur in even multiples. Therefore, all x_k must occur in conjugate pairs except real valued x_k such that $|x_k| \geq 1$. We will, therefore, order the x_k such that

$$F(f) = C \prod_{k=1}^{N_r} (x-x_{r,k}) \prod_{j=1}^{N_c/2} (x-x_{c,j})(x-\bar{x}_{c,j}) \qquad \text{[Eq. C.9]}$$

where N_c is even, $N_r + N_c = N$, $x_{r,k}$ are real and $|x_{r,k}| \geq 1$.

$$F(f) = C \prod_{k=1}^{N_r} \left(\frac{z+z^{-1}}{2} - x_{r,k} \right) \prod_{j=1}^{N_c/2} \left[\left(\frac{z+z^{-1}}{2} \right)^2 - (x_{c,j} + \bar{x}_{c,j}) \frac{z+z^{-1}}{2} + |x_{c,j}|^2 \right]$$

$$= 2^{-N} C \prod_{k=1}^{N_r} \left(-x_{r,k} \pm \sqrt{x_{r,k}^2 - 1} \right) \left[z - \left(x_{r,k} \pm \sqrt{x_{r,k}^2 - 1} \right) \right] \left[z^{-1} - \left(x_{r,k} \pm \sqrt{x_{r,k}^2 - 1} \right) \right] \prod_{j=1}^{N_c/2} \left| -x_{c,j} \pm \sqrt{x_{c,j}^2 - 1} \right|^2 \left[z^2 - 2\mathrm{Re}\left\{ x_{c,j} \pm \sqrt{x_{c,j}^2 - 1} \right\} z + \left| x_{c,j} \pm \sqrt{x_{c,j}^2 - 1} \right|^2 \right]$$

$$\left[z^{-2} - 2\,\mathrm{Re}\left\{ x_{c,j} \pm \sqrt{x_{c,j}^2 - 1} \right\} z^{-1} + \left| x_{c,j} \pm \sqrt{x_{c,j}^2 - 1} \right|^2 \right] \qquad \text{[Eq. C.10]}$$

Clearly, we have succeeded in factoring F(f) into

$$F(f) = \left(\sum_{k=0}^{N} a_k\, z^{-k} \right) \left(\sum_{j=0}^{N} a_k\, z^{k} \right)$$

with real a_k if we can show that

$$\left(2^{-N}\right) C \prod_{k=1}^{N_r} \left(-x_{r,k} \pm \sqrt{x_{r,k}^2 - 1} \right) \prod_{j=1}^{N_c/2} \left| - x_{c,j} \pm \sqrt{x_{c,j}^2 - 1} \right|^2$$

is real valued and non-negative. But $C = 2^{-N} A_N$ where A_N is real and $x_{r,k}$ are real with $\left| x_{r,k} \right| \geq 1$ so it is only necessary to show that

$$\left(2^{-N}\right) C \prod_{k=1}^{N_r} \left(-x_{r,k} \pm \sqrt{x_{r,k}^2 - 1} \right) \geq 0$$

We know that $F(f) \geq 0$ for all real f. Therefore, by Eq. C.9 at x = 0,

$$C \prod_{k=1}^{N_r} (-x_{r,k}) \prod_{j=1}^{N_c/2} \left| x_{c,j} \right|^2 \geq 0$$

$$C \prod_{k=1}^{N_r} (-x_{r,k}) \geq 0$$

Since $x_{r,k}$ is real and $\left|x_{r,k}\right| \geq 1$, $(-x_{r,k})$ and $\left(-x_{r,k} \pm \sqrt{x_{r,k}^2-1}\right)$ have the same sign so that

$$2^{-N}\, C \prod_{k=1}^{N_r} \left(-x_{r,k} \pm \sqrt{x^2{}_{r,k}-1}\right) \geq 0$$

Since the zeroes of $\sum_{k=0}^{N} a_k\, z^{N-k}$ have magnitudes $\left|x_k \pm \sqrt{x_k^2-1}\right|$ (see Eq. C.10) and $\left|x + \sqrt{x^2-1}\right| = \dfrac{1}{\left|x - \sqrt{x^2-1}\right|}$ we can always choose the appropriate sign to achieve a stable filter (z-plane poles inside the unit circle) except when $\left|x_k \pm \sqrt{x^2-1}\right| = 1$. But this would require a real valued $-1 \leq x_k \leq 1$ which would imply a real pole in $|H(f)|^2$ (if F(f) represented the denominator of $|H(f)|^2$) contrary to our original assumption.

We have, therefore, shown that any

$$|H(f)|^2 = \frac{\sum_{k=0}^{N_1} \hat{A}_k \cos^{2k}(\pi f \Delta t)}{\sum_{k=0}^{N_2} \hat{B}_k \cos^{2k}(\pi f \Delta t)} \geq 0$$

for all real f, with real valued $\hat{A}_k$ and $\hat{B}_k$, and with no poles for f real, can be realized by a stable filter of the form

$$y(t) = \sum_{k=0}^{N_1} a_k\, x(t-k\Delta t) - \sum_{k=1}^{N_2} b_k\, y(t-k\Delta t) \quad .$$

Two such filters (which can be found in the literature[21], [22]) are

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\tan \pi f \Delta t}{\tan \pi f_c \Delta t}\right)^{2N}}$$

and

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{\sin(\pi f \Delta t)}{\sin \pi f_c \Delta t}\right)^{2N}}$$

Both of these filters (which will be referred to as the "Tangent" and "Sine" "Butterworth" filters) satisfy the criterion for being realizable filters. The 3 db point is f_c for both filters and they are easily factored to obtain a set of weights. For f, $f_c << \frac{1}{2\Delta t}$ the filters behave nearly the same for a given N. For the tangent filter, however, $N_1 = N_2 = N$ while for the sine filter $N_2 = N$ and $N_1 = 0$. For a given number of arithmetic operations we can perform the filtering process with a "Sine Butterworth" filter with N twice as large as for the "Tangent Butterworth" filter (the "Tangent Butterworth", however, holds the advantage for $f_c \sim \frac{1}{2\Delta t}$). A filter with a 3 db cutoff frequency at $f_c = \frac{1}{20\Delta t}$ and ~30 db attenuation for $f = \frac{2}{20\Delta t}$ can be realized with N = 5 (for the "sine" filter we have $10 \log |H(\frac{1}{20\Delta t})|^2 = 29.6$ db and $10 \log |H(\frac{1}{2\Delta t})|^2 = -80.6$ db). This filter requires only five multiplications and additions per input sample but has an important disadvantage.

The impulse response is infinite. This means that the transient response never goes to zero (although it dies off exponentially) so that great care must be exercised in deciding how many data points must be discarded at the beginning of the record after filtering so that the

transient response of the filter will not significantly bias the spectral estimates. A non-recursive filter (Eq. C.1 with $N_2 = 0$) has a finite impulse response. This assures that we will incur no bias due to transients if we discard the first N_1 data points out of the filter. It seems obvious that we will require $N_1 \geq 10$ for a low pass filter with cutoff frequency of $f_c = \frac{0.1}{2\Delta t}$. This suggests, on the surface, that a greater number of arithmetic operations will be required to filter the data. For the non-recursive filter, however, it is necessary to calculate the output samples only at the output sample rate. For a sample rate reduction of one tenth, N_2 could be as large as 50 and still require only 5 additions and multiplications per input sample.

The problem now is to find a suitable transfer function. For simplicity, the class of filters considered was limited to symmetric filters ($a_k = a_{N_1-k}$ in Eq. C.1). This provides filters with the often desirable property of linear phase (although this property is not important in this application).

Referring to Eq. C.5, the transfer function (after the delay is compensated for by redefining $t = 0$) may be written

$$H(f) = \begin{cases} a_0 + 2 \sum_{k=1}^{(N-1)/2} a_k \cos(2k\pi f\Delta t) & , \ N \text{ odd} \\ 2 \sum_{k=1}^{N/2} a_k \cos[(2k-1)\pi f\Delta t] & , \ N \text{ even} \end{cases} \qquad \text{[Eq. C.11]}$$

where the a_k are still the filter weights but have been renumbered from the center out. The original motivation for the following discussion comes from noticing that Eq. C.11 is the same equation as for the beam amplitude (as a function of angle) for a linear, equi-spaced, symmetric array of point source transmitting elements driven with equal phase but different amplitudes ($f\Delta t$ here corresponds to $\frac{d}{\lambda}\sin\phi$ in the antenna example where d is the element spacing, λ is wavelength and ϕ is angle from the perpendicular to the array). The following discussion results from a 1946 paper by C. L. Dolph[23] concerning sonar arrays.

Equation C.11 may be rewritten as an N-1 order polynomial in $\cos(\pi f\Delta t)$ with only even or odd ordered terms.

$$H(f) = \begin{cases} \displaystyle\sum_{k=1}^{N/2} A_k \cos^{(2k-1)}(\pi f\Delta t) & , \quad N \text{ even} \\ \displaystyle\sum_{k=0}^{(N-1)/2} A_k \cos^{2k}(\pi f\Delta t) & , \quad N \text{ odd} \end{cases}$$

$$= G_{N-1}\left(\cos(\pi f\Delta t)\right)$$

Without loss of generality we will normalize $G_m(x)$ such that $G_m(1) = 1$ ($H(0) = 1$). Note that $G_m(x) = (-1)^m G(x)$.

Consider the Chebyshev polynomial $T_n(x_0 x) = \cos(n \arccos(x_0 x))$ for $-1 \leq x_0 x \leq 1$ with $x = \cos \pi f\Delta t$. This is also an nth order polynomial in $\cos(\pi f\Delta t)$ with only even or odd ordered terms. Let

$M_n(x) = \frac{1}{a} T_n(x_0 x)$ where x_0 is defined by $T_n(x_0) = a$ ($x_0 > 0$ and $a > 1$). Then

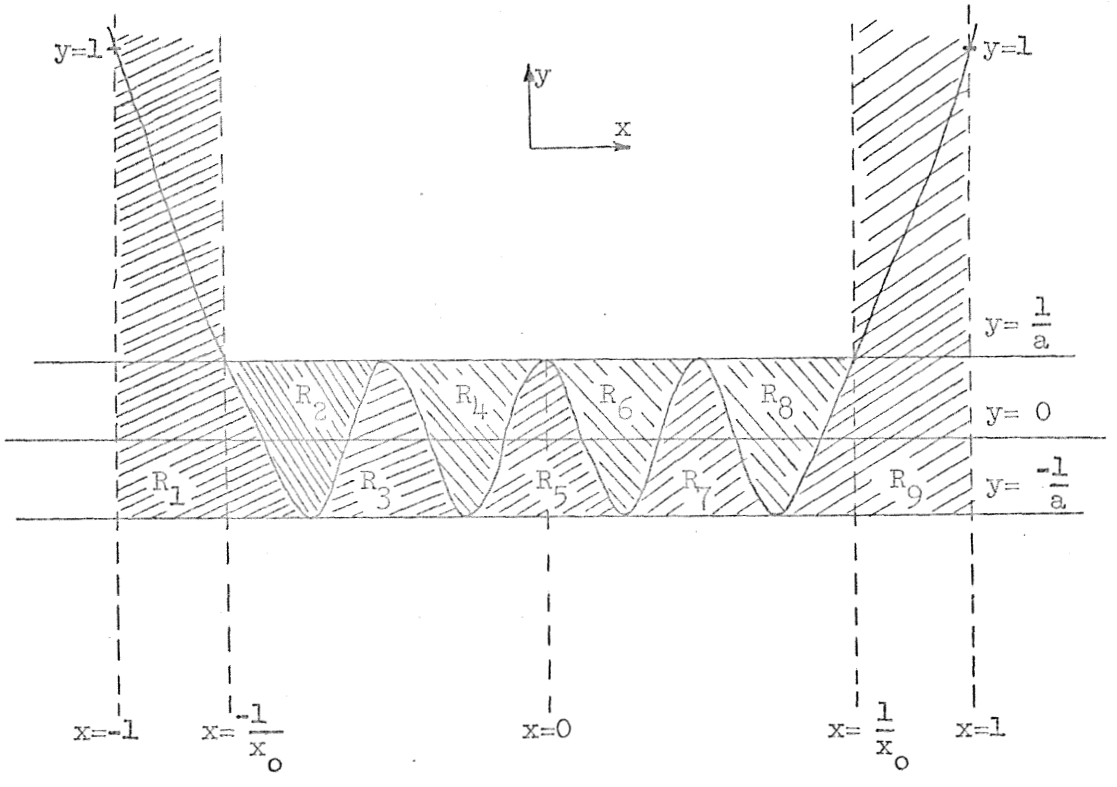
$$M_n(1) = 1$$

$$M_n(x) \leq \frac{1}{a} \quad \text{for} \quad -\frac{1}{x_0} \leq x \leq \frac{1}{x_0}$$

$$M_n(x) \text{ has n roots in } -\frac{1}{x_0} < x < \frac{1}{x_0} \quad .$$

Consider the set of all polynomials of order n such that $P_n(x) = (-1)^n P_n(-x)$ and $P_n(1) = 1$. Assume that there exists a polynomial in this set such that $P_n(x) \leq \frac{1}{a}$ for all $|x| \leq \frac{1}{x_0}$. The difference polynomial $D(x) = M_n(x) - P_n(x)$ is at most degree n. $D(1) = 0$ and $D(-1) = 0$. Figure C.1 shows an example of $M_n(x)$ with $n = 8$ and n+1 regions $R_1, \ldots,$ R_{n+1} bounded by $\pm\frac{1}{a}$ and $M_n(x)$ for $|x| \leq \frac{1}{x_0}$. $P_n(x)$ must lie in the union of these sets for $|x| \leq 1$. Clearly, if $P_n(x) < \frac{1}{a}$ for $|x| \leq \frac{1}{x_0}$, $P_n(x)$ must pass through each of the regions giving n+2 zeroes for D(x) (including the two at $x = \pm 1$). If, however, we have $P_n(x) \leq \frac{1}{a}$ for $|x| \leq \frac{1}{x_0}$, $P_n(x)$ could pass from region R_i to R_{i+2} without entering region R_{i+1}. This apparently results in only one zero of D(x) in this interval instead of two. During this interval, however, D(x) does not change sign even though it has a zero. It must, therefore, have a zero of even order at this point again implying at least n+2 zeroes for D(x). However, D(x) is at most order n, which implies $D(x) = 0$ and $P_n(x) = M_n(x)$.

It has been shown, therefore, that for a symmetric non-recursive low pass filter with N coefficients (normalized so that $H(f = 0) = 1$) the peak value of $|H(f)|$ for $f_c \leq f \leq \frac{1}{2\Delta t}$ may be minimized by choosing



$y = M_8(x)$ (not to scale)

$P_n(x)$ must lie in the union of $R_1, \ldots, R_9$ for $-1 < x < 1$

Figure C.1 8th order Chebyshev polynomial.

$H(f) = \frac{1}{a} T_{N-1}(x_0 \cos\pi f\Delta t)$ where $\cos(\pi f_c \Delta t) = \frac{1}{x_0}$ and $T_{N-1}(x_0) = a$. For a given value of f_c, N may be increased until a sufficiently small value of $\frac{1}{a}$ is obtained to insure that the minimum value of N necessary to meet the peak ripple criterion has been obtained.

We need a filter which has a gain of less than -30 db for $\frac{1}{10\Delta t} < f \leq \frac{1}{2\Delta t}$. Therefore, $x_0 = \frac{1}{\cos\frac{\pi}{10}}$ and $a \geq 10^{1.5} = 31.6228$ where $T_{N-1}(x_0) = a$.

Using $T_n(x) = \frac{1}{2}\left[\left(x + \sqrt{x^2 - 1}\right)^n + \left(x - \sqrt{x^2 - 1}\right)^n\right]$ we find that $T_{13}(x_0) = 31.8188$. Therefore, a filter with N = 14 weights will give a low pass filter with a gain of less than -30.05 db for $f \geq \frac{1}{10\Delta t}$.

Since the transfer function is $H(f) = \frac{1}{a} \cos(n \arccos(x_0 \cos\pi f\Delta t))$ $(n = N-1)$ the n zeroes occur at

$$n \arccos(x_0 \cos \pi f\Delta t) = \frac{2k-1}{2}\pi \qquad k = 1,\ldots,n$$

$$\cos(\pi f\Delta t) = \frac{1}{x_0} \cos\left(\frac{2k-1}{2n}\pi\right)$$

$$z^{\frac{1}{2}} + z^{-\frac{1}{2}} = \frac{2}{x_0} \cos\left(\frac{2k-1}{2n}\pi\right)$$

$$z^2 + \left(2 - \frac{4}{x_0^2} \cos^2\left(\frac{2k-1}{2n}\pi\right)\right) z + 1 = 0$$

Therefore

$$H(z) = \begin{cases} C \prod_{k=1}^{n/2} \left[z^2 + \left(2 - \frac{4}{x_0^2}\cos^2\left(\frac{2k-1}{2n}\pi\right)\right)z + 1\right] & \text{if n even} \\ C(z+1) \prod_{k=1}^{(n-1)/2} \left[z^2 + \left(2 - \frac{4}{x_0^2}\cos^2\left(\frac{2k-1}{2n}\pi\right)\right)z + 1\right] & \text{if n odd} \end{cases}$$

where the coefficients of the polynomial in z are the weights, a_k, in Eq. C.1 (C is a normalizing constant).

If the polynomial is multiplied out for n = 13, $x_0 = \frac{1}{\cos \frac{\pi}{10}}$ and normalized so that H(z = 1) = 1, the following coefficients are obtained:

$$a_0 = a_{13} = 0.030172$$

$$a_1 = a_{12} = 0.037455$$

$$a_2 = a_{11} = 0.055338$$

$$a_3 = a_{10} = 0.073744$$

$$a_4 = a_9 = 0.090395$$

$$a_5 = a_8 = 0.10304$$

$$a_6 = a_7 = 0.10986 \quad .$$

One may also specify the exact value of a which is desired and find the value of n yielding a cutoff frequency as close as possible to the desired value. There was an existing program available in the Univac 1108 at the Naval Undersea Center at the time the data analysis was beginning which yielded the filter coefficients under these conditions. It we specify a peak ripple of -30 db then $a = 10^{1.5}$,

$$x_0 = \tfrac{1}{2}\left[\left(a + \sqrt{a^2-1}\right)^{1/n} + \left(a + \sqrt{a^2-1}\right)^{-1/n}\right] = 1.0513 \text{ for } n = 13.$$

$f_c = \frac{1}{10.014\ \Delta t}$. Since the cutoff frequency is so close to the desired value, these coefficients were actually used in the data analysis:

$$a_0 = a_{13} = 0.030301$$

$$a_1 = a_{12} = 0.037511$$

$$a_2 = a_{11} = 0.055371$$

$$a_3 = a_{10} = 0.073741$$

$$a_4 = a_9 = 0.090352$$

$$a_5 = a_8 = 0.102960$$

$$a_6 = a_7 = 0.109764$$

The transfer function is plotted in Fig. C.2 with $\Delta t = \frac{1}{2}$. This plot was made using Eq. C.3 and the actual coefficients used in the data analysis.

A filter is also required to reduce the sample rate from $\Delta t = 10^4$ seconds to $\Delta t = 5 \times 10^4$ seconds when performing the spectral estimates in the $10^{-7} \leq f \leq 10^{-5}$ Hz range. The parameters chosen were peak side lobes of less than -30 db and $f_c = \frac{1}{5\Delta t}$ (the frequency at which the peak side lobe level is first reached).

$$x_0 = \frac{1}{\cos\left(\frac{\pi}{5}\right)}$$

$$20 \log T_7(x_0) = 35 \quad .$$

Therefore, choosing n = 7 yields an 8 weight filter with a peak side lobe level of -35 db. The normalized coefficients actually used are:

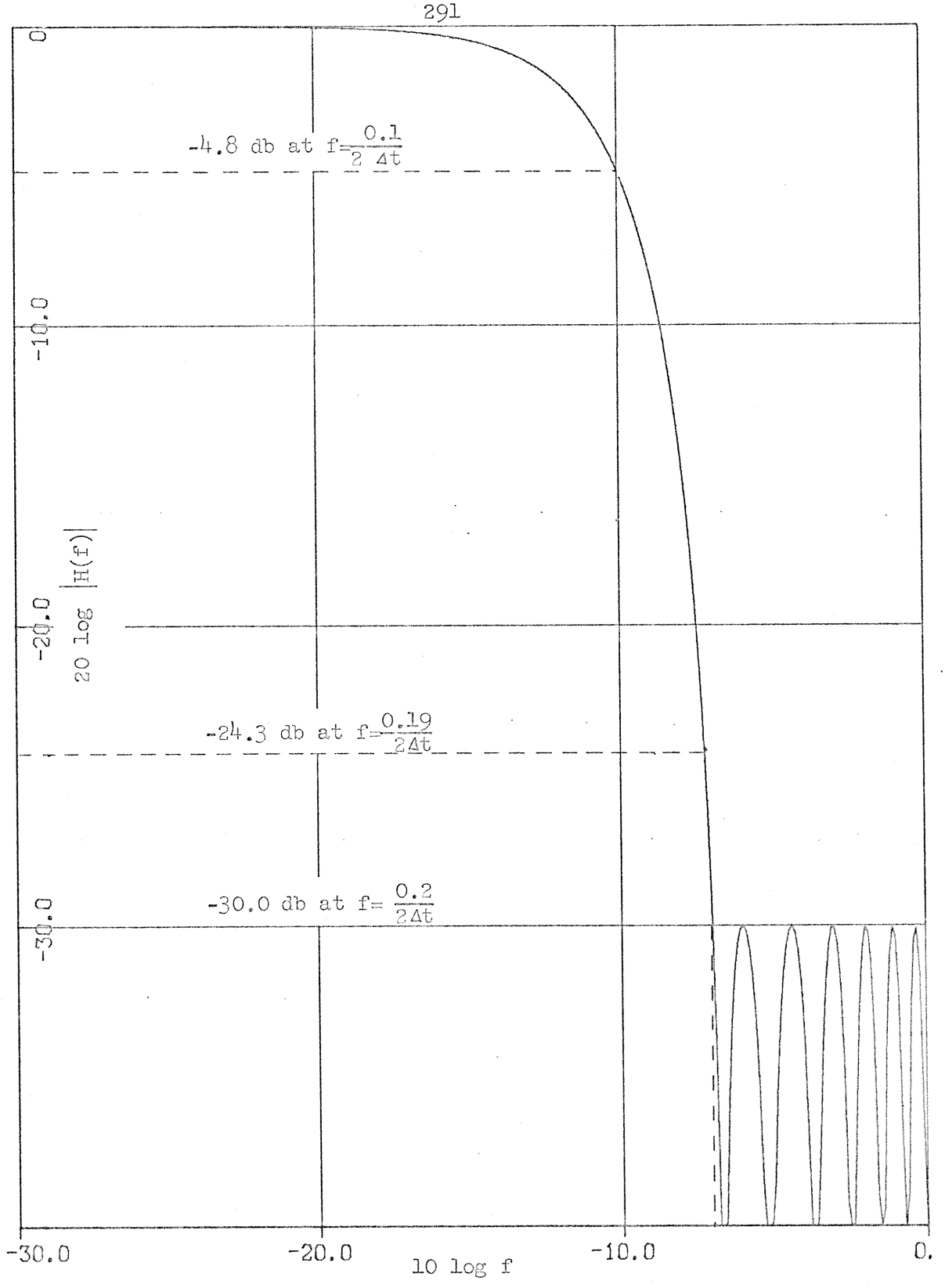


Figure C.2 One tenth sample rate reduction digital dealiasing filter ($\Delta t=0.5$).

$$a_0 = a_7 = .03930277$$

$$a_1 = a_6 = .09505141$$

$$a_2 = a_5 = .16073031$$

$$a_3 = a_4 = .20491551$$

The transfer function of this filter is plotted in Fig. C.3 for $\frac{1}{2\Delta t} = 1$ Hz.

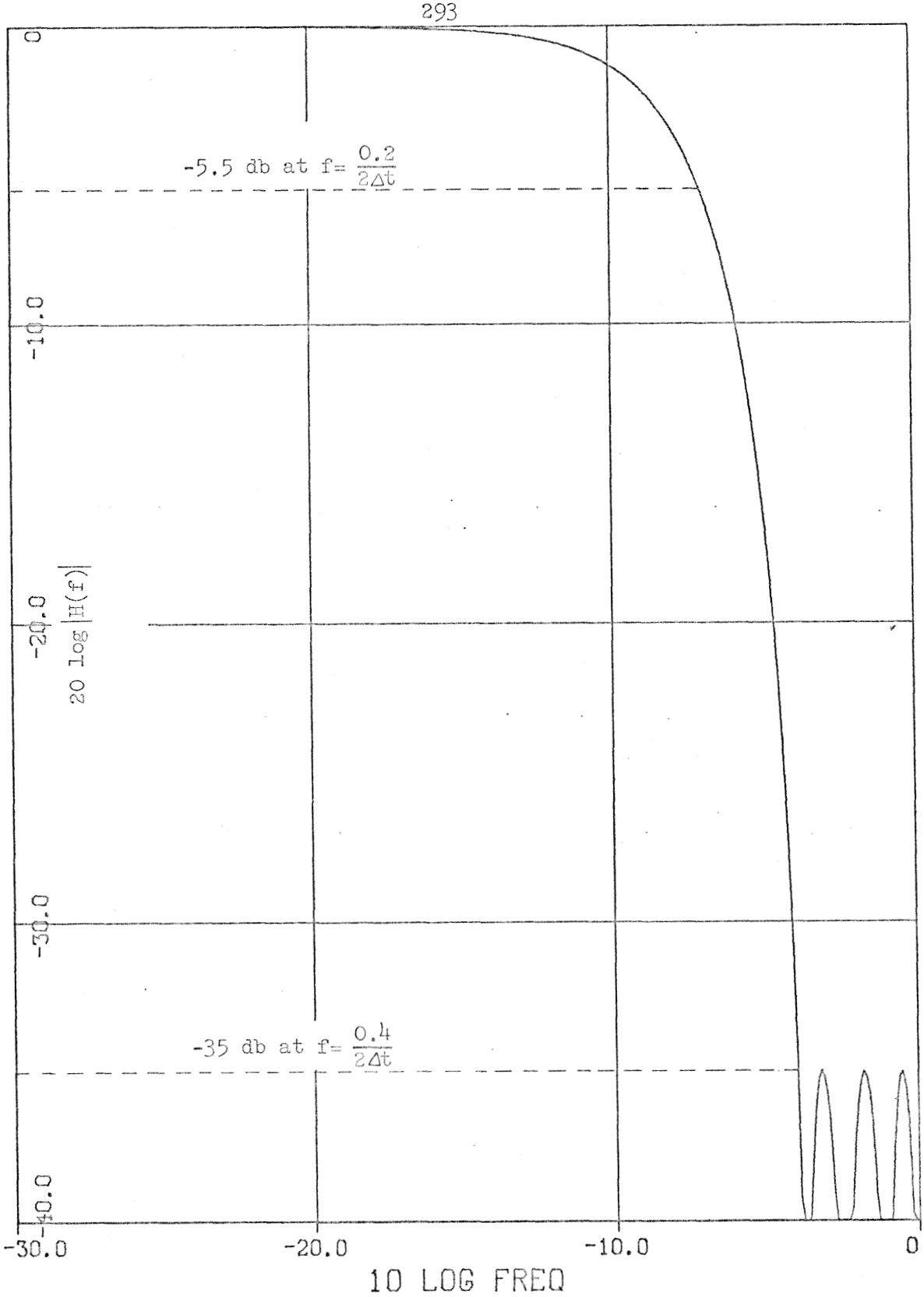


Figure C.3 One fifth sample rate reduction filter (Δt=0.5).

LIST OF FIGURES

REFERENCES

1. M. A. Caloyannides, "A Mathematical and Experimental Investigation of Microcycle Spectral Estimates of Semiconductor Flicker Noise," Ph.D. Thesis, Caltech, 1972.

2. A. L. McWhorter, Semiconductor Surface Physics, (edited by R. H. Kingston), University of Pennsylvania Press, 1956, pp. 207-229.

3. J. N. Puckett, Jr., "An Electrical and Statistical Study of Burst Noise," Ph.D. Thesis, Caltech, 1971.

4. J. B. Johnson, "The Schottky Effect in Low Frequency Circuits," Phys. Rev., Vol. 26, July 1925, pp. 71-85.

5. W. R. Attkinson, L. Fey, J. Newman, "Spectrum Analysis of Extremely Low Frequency Variations of Quartz Oscillations," Proc. IEEE, Vol. 51, Feb. 1963, p. 379.

6. D. Brouwer, "A Study of the Changes in the Rate of Rotation of the Earth," Astron. J., Vol. 57, Sept. 1952, pp. 125-146.

7. J. Bernamont, "Potential Fluctuations at the Ends of a Metallic Conductor of Small Volume Traversed by a Current", Annales de Physique, Vol. 7, Jan. 1937, pp. 71-140.

8. B. V. Rollin and I. M. Templeton, "Noise in Semiconductors at Very Low Frequencies," Proc. Phys. Soc. of London, Vol. 66B, 1953, pp. 259.

9. B. V. Rollin and I. M. Templeton, "Noise in Germanium Filaments at Very Low Frequencies," Proc. Phys. Soc. of London, Vol. 67B, 1954, pp. 271.

10. T. E. Firle and H. Winston, "Noise Measurements in Semiconductors at Very Low Frequencies," Journal of Applied Physics, Vol. 26, No. 6, 1955, pp. 716-717.

11. D. J. Blakemore, "Microcycle Spectral Estimation," Ph.D. Thesis, Caltech, 1966.

12. E. Baldinger and E. Nüesch, "Rauschen von Transistoren bei sehr tiefen Frequenzen," Helvetica Physica Acta, Vol. 41, 1968, pp. 313-322.

13. R. B. Blackman and J. W. Tukey, The Measurement of Power Spectra, Dover Publications, Inc., 1958.

14. C. A. Greenhall, "A Stochastic Process that Behaves Like Flicker Noise," unpublished, JPL Contract #NAS-7-100.

15. W. B. Davenport, Jr., and W. L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Co., 1958.

16. A. J. Perrone, "Deep-Ocean Ambient-Noise Spectra in the Northwest Atlantic," The Journal of the Acoustical Society of America, Vol. 46, 1969, pp. 762-770.

17. I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series, and Products, Academic Press, 1965.

18. The Linear Integrated Circuits Data Catalog, Fairchild Semiconductor (Fairchild Camera and Instrument Corporation), Nov. 1971

19. C. A. Mead, "Electronic Current Flow Through Ideal Dielectric Films," The Anodic Behavior of Metals and Semiconductor, Vol. I, Chapter 3, 1971.

20. D. E. Knuth, The Art of Computer Programming, Vol. 2 "Seminumerical Algorithms", Addison-Wesley Publishing Co., 1969.

21. C. M. Rader, and B. Gold, "Digital Filter Design Techniques in the Frequency Domain," Proc. IEEE, Vol. 55, Feb. 1967, pp. 149-171.

22. R. K. Otnes, "An Elementary Design Procedure for Digital Filters," IEEE Transactions on Audio and Electroacoustics, Vol. AU-16, No. 3, Sept. 1968, pp. 330-335.

23. C. L. Dolph, "A Current Distribution for Broadside Arrays which Optimizes the Relationship Between Beam Width and Side-Lobe Level," Proc. of I.R.E., June 1946, pp. 335-348.

24. J. Bernamont, "Fluctuations in the Resistance of Thin Films," Proc. Phys. Soc. of London, Vol. 49 (extra part), 1937, pp. 138-139.

25. A. van der Ziel, "On the Noise Spectra of Semiconductor Noise and of Flicker Effect," Physica, Vol. 16, 1950, pp. 359-372.

26. F. K. du Pre, "A Suggestion Regarding the Spectral Density of Flicker Noise," Physical Rev., Vol. 78, June 1950, pg. 615.

REFERENCES NOT SPECIFICALLY CITED

T. B. Watkins, "1/f Noise in Germanium Devices," Proc. of Phys. Soc. of London, Vol. 73, 1959, pp. 59-68.

A. Van der Ziel, "The Surface Recombination Model of p-n Diode Flicker Noise," Physica, Vol. 48, 1970, pp. 242-246.

R. A. Dell, Jr., "Experimental Study of 1/f Noise Stationarity by Digital Techniques," Ph.D. thesis, Northwestern University, 1971.

O. Mueller, "Thermal Feedback: A Source of 1/f-Flicker Noise in Transistors," General Electric (Telecommunications Products Department, Engineering Library, Lynchburg, Virginia) No. R72-TPD-5, Nov. 1972 .

S. T. Hsu and R. J. Whittier, "Characterization of Burst Noise in Silicon Devices", Solid State Electronics, Vol. 12, 1969, pp. 867-878.

S. T. Hsu, R. J. Whittier, and C. A. Mead, "Physical Model for Burst Noise in Semiconductor Devices," Solid State Electronics, Vol. 13, 1970, pp. 1055-1071.

O. Mueller, "Excess Noise in P-N Junction Diodes," General Electric (Telecommunications Products Department, Engineering Library, Lynchburg, Virginia) No. R72-TPD-6, Nov. 1972 .