MEASUREMENT NOTES

Note 24

DETERMINISTIC ERROR ANALYSIS APPLIED TO EMP SIMULATOR DATA ACQUISITION

June 1977

Larry D. Scott
Mission Research Corporation
Albuquerque, New Mexico 87108

Abstract

Six error sources found to be dominant contributors to EMP simulator data error are identified, formulated and illustrated from a deterministic viewpoint. The reduction of the classical list of experimental error sources to these six is justified if a rigorous quality control program is part of the data acquisition and processing activities of a test program. The levels of error illustrated are representative of state-of-the-art EMP data acquisition and processing.

The derivation of frequency domain noise error from noisy uniformly sampled time domain data is shown in Appendix A and a computer algorithm for determining noise error for more general data processing situations is described in Appendix B.

CONTENTS

LIST OF ILLUSTRATIONS

1. INTRODUCTION

The purpose of this paper is to formulate and illustrate, from a deterministic viewpoint, the frequency domain dominant error terms or sources found in an EMP simulator data acquisition program.

By requiring a rigorous quality control (QC) program be applied to the instrumentation configuration, calibration and deployment as well as to the data processing and storage, we can limit the classical list of test data error sources as found in reference 1 to a reasonable number.

In addition to applying this rigorous QC to the instrumentation and processing, we will assume also that errors such as sensor placement, test item interaction, and numerical processing will be controlled to a reasonable level (i.e., below levels typically introduced by the dominant sources).

The six dominant error contributors to be discussed are:

(1) amplitude
(2) sweep speed
(3) base line shift
(4) base line rotation
(5) noise
(6) truncation.

Each of these error sources, even in a well QC'd program, can contribute significant errors singly and together over the frequency spectrum.

Let us now consider briefly some limitations that apply to these and closely-related error sources.

2. INHERENT LIMITATIONS OF INSTRUMENTATION

Consider the case of a "perfect" polaroid photograph, that is, assume that the instrumentation caused an exact replica of the driving function to be recorded on the photographic emulsion. Now, considering this ideally recorded signal to occupy about seventy-five percent of the oscilloscope screen (bipolar trace with base line midscreen), a resolution of one part in thirty has been found to be the best that can be obtained by careful quality control of the entire process. This corresponds to a 30 dB signal-to-noise ratio. For a monopulse signal, it would be possible to achieve a maximum of 36 dB signal-to-noise ratio, but at the cost of taking at least two shots to appropriately offset the base line and set the oscilloscope vertical deflection.

This 30 dB vertical or amplitude limitation tends to bound the contributions of base line shift and rotation in addition to setting the floor on noise arising from trace width and subsequent digitization noise. This 30 dB amplitude resolution bound also places a lower bound on truncation error in a well implimented program.

In considering horizontal resolution, typically, two horizontal divisions (major) are digitized with a resolution of one part in forty, or 2.5 percent. This error appears in the frequency domain as an expansion or contraction of the frequency scale; thus, the amplitude errors incurred by this source are signal-dependent and can become enormous for highly peaked functions.

Time-tie errors in combining multispeed records on a single shot, when bounded by tight QC, are also bounded by the three major amplitude error contributors above. That is, noise at 30 dB below the signal, base line shift and rotation.

Loss of portions of each trace at the time-tie points is minimized by resubmittal by the QC activity when errors greater than those stated above are found in the resulting amplitude or time values.

These error levels have been introduced to set the stage for the subsequent formulation of the specific kinds of error determined to be dominant.

Most of the calibration errors found in the various elements of instrumentation fall into the category of amplitude uncertainties (possibly frequency dependent) and sweep-speed or time-scale uncertainties.

3. DATA ACQUISITION AND PROCESSING FLOW

As shown in figure 1, the path that a particular piece of data may take will be dependent on the data content and purpose of subsequent analysis (frequency ranges, early time, late time, etc.).

Various kinds of error are injected at each block in the data flow diagram. The error analysis section to follow will address an error model to combine these errors and produce an overall error expression for the output data.

4. ERROR MODELS

The error models formulated below are universally applicable to screen box data recorded on polaroid film and then manually or automatically digitized for processing and storage or to data acquired by transient digitizers via a microwave data link and stored in a digital format on magnetic tape or disk for processing and storage.

Their validity is limited by the accuracy of the error parameter determination as set by the QC program.

We start with a recorded function, f(t), containing certain error or noise terms. Here we are concerned with how these errors affect the frequency domain spectrum of the recorded function.

Once the major error contributors have been identified, it is relatively easy to obtain deterministic error expressions.

We define the Fourier transform as

$$F(\omega) = \int_{-\infty}^{\infty} f(t)\, e^{-j\omega t} dt \ .$$

Let us now consider the error for the major error contributors found in a measurement program having a high level of QC.

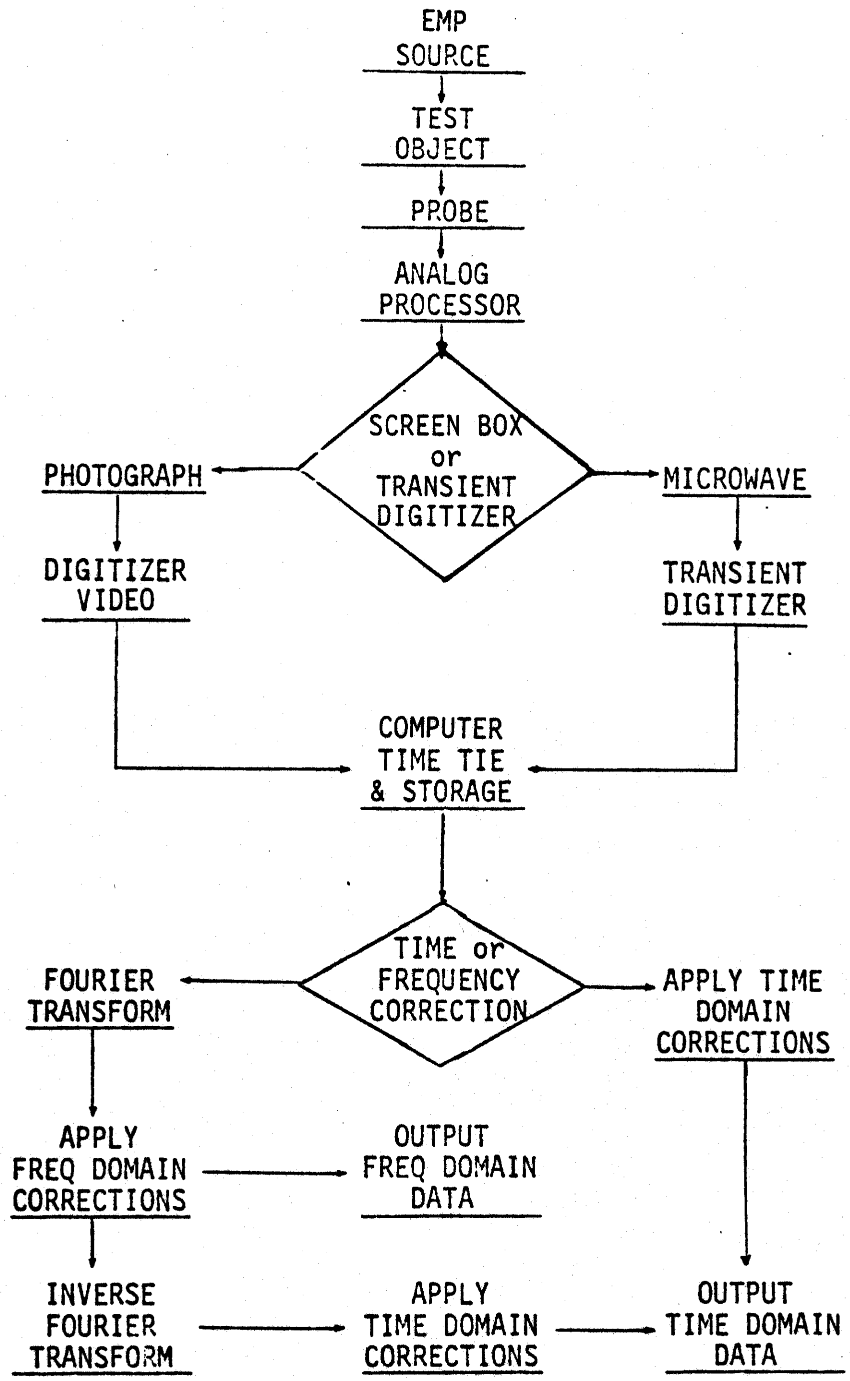


Figure 1. Typical Data Acquisition and Processing Flow Chart

(a) Gain uncertainty - in this instance the gain error, ε_{gain}, may be a function of frequency, so that

$$F_m(\omega) = F(\omega)\ (1 \pm \varepsilon_{gain}(\omega)).$$

$F_m(\omega)$ is the measured or apparent function of frequency.

(b) Sweep speed uncertainty - for this case $f_m(t) = f(t[1 \pm \varepsilon_{swp}])$. By a simple change of variables we have

$$F_m(\omega) = \frac{1}{1 \pm \varepsilon_{swp}}\ F\left(\frac{\omega}{1 \pm \varepsilon_{swp}}\right).$$

(c) Base line shift - since all data recordings are finite in time, a base line shift error is simply an added step function, truncated at the end of the record. Letting the shift have an amplitude of $\pm h$ and the maximum record length of T, then base line shift error is

$$\varepsilon_{base} = \pm h \left[\frac{1 - e^{-j\omega t}}{j\omega}\right].$$

(d) Base line rotation - this error source is limited by the QC activity and generally appears as a rotation about the origin. The offset or error amplitude at the trace end (time T) is $\pm h$.

Evidently, this form of error and distortion of the waveform can be approximated by a simple ramp $\frac{h}{T}t$ for the small rotations found in a well-executed test program. The frequency-domain expression for this error is

$$\varepsilon_{rot} = \pm \frac{h}{\omega^2 t} \left\{1 - e^{-j\omega t}(1 + j\omega T)\right\}.$$

(e) Noise - this error is a combination of various noise sources, but overall effects may be lumped together. By using normally (Gaussian) distributed noise to represent this error source in the time domain, it is shown in Appendix A that the expressions below bound the one sigma expected relative errors in the magnitude of the frequency domain function

$$\varepsilon(\omega)|_{max} = \sqrt{1 + \frac{2\sigma\{|R(\omega)| + |I(\omega)| + \sigma\}}{R(\omega)^2 + I(\omega)^2}} - 1$$

$$\varepsilon(\omega)|_{min} = \sqrt{1 + \frac{2\sigma\{-|R(\omega)| - |I(\omega)| + \sigma\}}{R(\omega)^2 + I(\omega)^2}} - 1$$

$R(\omega) = \text{Re } F(\omega)$

$I(\omega) = \text{Im } F(\omega)$

σ = standard deviation of frequency domain noise derived from time domain noise.

Absolute error is simply $\varepsilon_{abs}(\omega) = |F(\omega)|\, \varepsilon_{rel}(\omega)$.

In this error model, σ is a function of the following four parameters:

1. the number of sample points in $f(t)$;
2. the length of each time increment Δt_i;
3. the total record length T_{max};
4. the particular Fourier transform algorithm used to process the data.

Appendix B describes an algorithm for determining σ from the S/N ratio of the original data.

For uniformly sampled time domain data that is numerically Fourier transformed using the rectangular rule, a relatively simple, but rigorous expression can be written for the frequency domain standard deviation $\sigma(\omega)$. It is a function of the time domain standard deviation and the sampling parameters Δt and n (see Appendix A). That is

$$\sigma(\omega) = \sigma(t)\ \Delta t\ \sqrt{\frac{n}{2}}$$

where

Δt = uniform time increment

n = number of samples.

Note that $n\Delta t = T_{max}$, the sample window.

(f) Truncation error - we consider first the error introduced by truncating the test function $f(t) = e^{-at}\sin bt$, such that the maximum amplitude of f(t) is $\leq 1/30$ of the peak of the function. Thus

$$f(t) = e^{-at}\sin bt \qquad \text{for } 0 \leq t \leq T$$

$$f(t) = 0 \qquad \text{for } T \leq t \leq \infty\ .$$

The Fourier transform is then

$$F(\omega) = \frac{b}{(a + j\omega)^2 + b^2} - e^{-j\omega t} \quad .$$

$$\left\{ e^{-at}\cos bT \left[\frac{b}{(a + j\omega)^2 + b^2}\right] + \right.$$

$$\left. e^{-aT}\sin bT \left[\frac{a + j\omega}{(a + j\omega)^2 + b^2}\right] \right\} \quad .$$

By truncating at a zero crossing (i.e., $bT = n\pi$ for n = integer), we can minimize the error and the transform simplifies to

$$F(\omega) = \frac{b}{(a + j\omega)^2 + b^2}\left(1 - e^{-j\omega T}\ e^{-aT}\right)$$

so that fractional error becomes

$$\varepsilon_{trun} = -e^{-j\omega T} e^{-aT} .$$

If we want the maximum envelope, then

$$|\varepsilon_{trun}| = e^{-aT} .$$

The QC discussed above would limit this to one part in thirty or 3.3 percent. This same argument and formula also holds for the decaying exponential $f(t) = e^{-at}$.

5. OVERALL ERROR MODEL

In combining these error contributors we have observed that for our limited but representative test cases of damped sine and doubly-delayed double exponential, using the QC limited error levels above, the set of significant error contributors can be further reduced to four, which are:

(1) gain uncertainty
(2) sweep speed uncertainty
(3) base line shift
(4) noise.

If less stringent QC limits are employed, the error contribution for truncation can become significant and perhaps should be incorporated, as in the examples to follow.

Evidently, waveform distortions introduced by the slight rotations permitted by good QC cause smaller errors than the ramp introduced by this rotation. Further, the ramp error is always bounded by a base line shift error of the same amplitude for these cases. Thus our model need only consider the base line shift.

Since we are interested in an error model envelope, some further simplifications are made. For the base line shift we have

$$\varepsilon_{base} = \pm h\left(\frac{1 - e^{-j\omega t}}{j\omega}\right) .$$

It is evident that the envelope of the magnitude of this function can be expressed by the following smooth function (i.e., no zeros):

$$\left|\varepsilon_{base}\right| = h\left|\frac{1 - e^{-j\omega t}}{j\omega}\right| \qquad \text{for } 0 \leq \omega \leq \frac{\pi}{T}$$

$$\left|\varepsilon_{base}\right| = \frac{h2}{\omega} \qquad \text{for } \frac{\pi}{T} \leq \omega \leq \infty \quad .$$

Since all of the above errors are uncorrelated, it is reasonable to add them in a root-sum-square manner. Thus the total absolute error model is:

$$\varepsilon^2_{abs} = \Big[|F(\omega)| \quad \varepsilon_{gain}(\omega)\Big]^2$$

(gain error)

$$+\left[|F(\omega)| - \frac{1}{1 \pm \varepsilon_{swp}}\left|F\left(\frac{\omega}{1 \pm \varepsilon_{swp}}\right)\right|\right]^2$$

(sweep speed error)

$$+\begin{cases} \left[h\left|\frac{1 - e^{-j\omega t}}{j\omega}\right|\right]^2 & \text{for } 0 \leq \omega \leq \frac{\pi}{T} \\ \left[\frac{2h}{\omega}\right]^2 & \text{for } \frac{\pi}{T} \leq \omega \leq \infty \end{cases}$$

(base line error)

$$+ \left[|F(\omega)| e^{-aT} \right]^2$$

(truncation error)

$$+ \left[|F(\omega)| \left\{ \sqrt{1 + \frac{2\sigma \left(|R(\omega)| + |I(\omega)| + \sigma \right)}{R(\omega)^2 + I(\omega)^2}} - 1 \right\} \right]^2$$

(noise error upper bound)

where

$F(\omega)$ = the frequency domain function

a = decay constant in damped signal being truncated

T = record length

$R(\omega)$ = real part of $F(\omega)$

$I(\omega)$ = imaginary part of $F(\omega)$

σ = standard deviation of noise in frequency domain based on S/N ratio in raw data and established by QC standards

ε_{gain} = instrumentation amplitude error defined by instrumentation calibration activity

ε_{swp} = sweep speed error is limit allowed based upon equipment calibration and audited by QC procedures

h = base line shift accepted based upon established QC standards.

To find relative error ε_{rel} we use

$$\varepsilon_{rel} = \frac{\varepsilon_{abs}}{|F(\omega)|} \quad .$$

6. EXAMPLES OF OVERALL ERROR

In this section we have selected two sets of examples to illustrate the overall processing error to be found in, first, a rigorous quality controlled data processing system and, second, a more typical level of quality control.

Two different waveforms were selected for this example, the damped sinusoid which is characteristic of a class of data commonly seen in test object response testing, and a sum of delayed double exponentials which is characteristic of certain fields seen in the ATHAMAS I simulator field mapping.

a. Sample Waveforms

Figures 2 and 3 show the unperturbed time domain traces to be used and their Fourier transforms.

Let us consider now the following cases:

(1a) Improved Screen Box System with Video Digitization (no filtering, no common mode)

Error Type	Error Level
amplitude	± 0.238 dB
base line shift	≤ 1:30
truncation	≤ 1:30
sweep speed	≤ 1:25
noise/signal	≤ 1:30

Figure 4 shows how these errors combine for the damped sinusoid and double exponentials to produce an overall error envelope in the frequency domain.

(1b) Typical Screen Box System with Video Digitization (no filtering, no common mode)

Error Type	Error Level
amplitude	= .46 dB
base line shift	≤ 1:15
truncation	≤ 1:15
sweep speed	≤ 1:20
noise/signal	≤ 1:15

Figure 5 shows the combined frequency domain errors for this system.

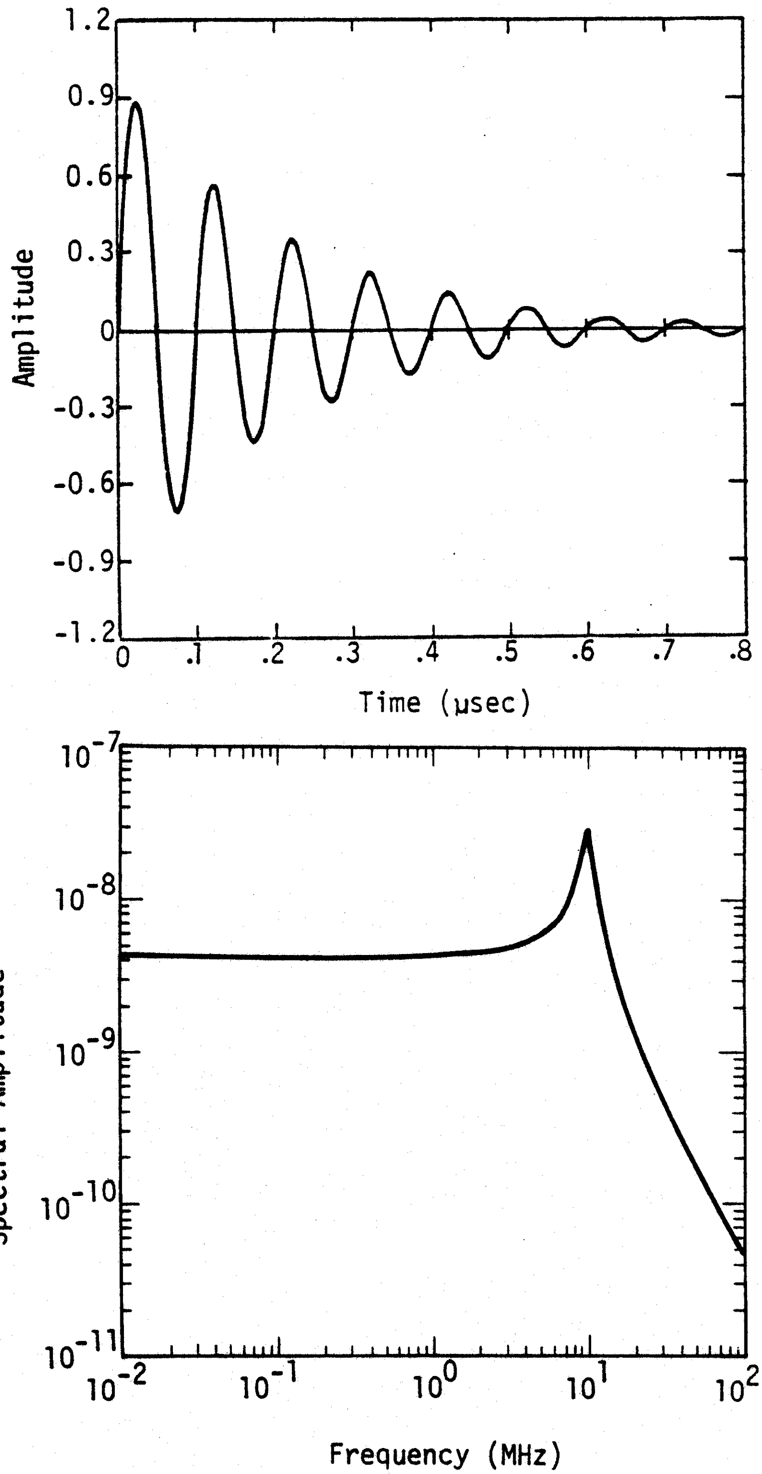
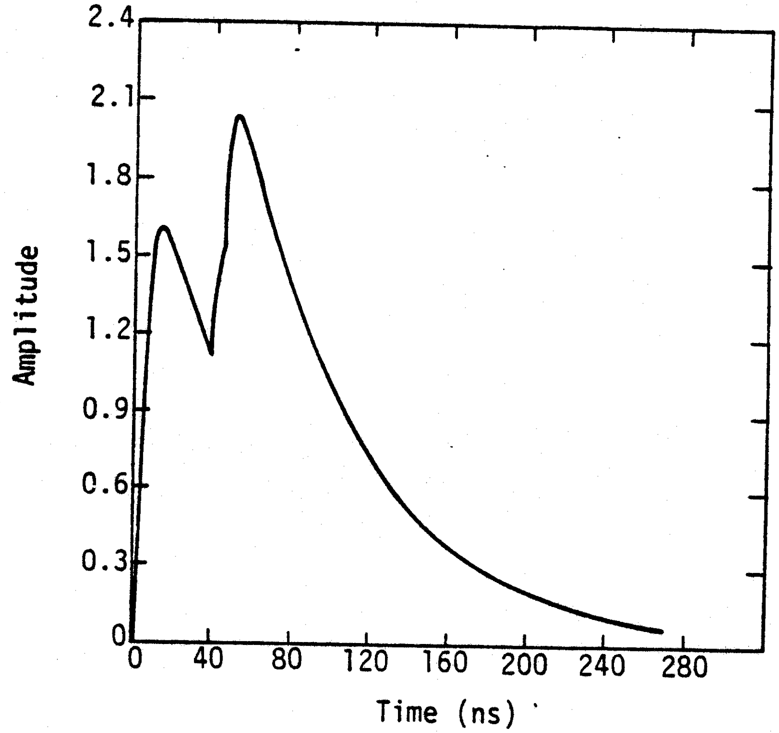


Figure 2. Damped sinusoid; time and frequency domain plots.



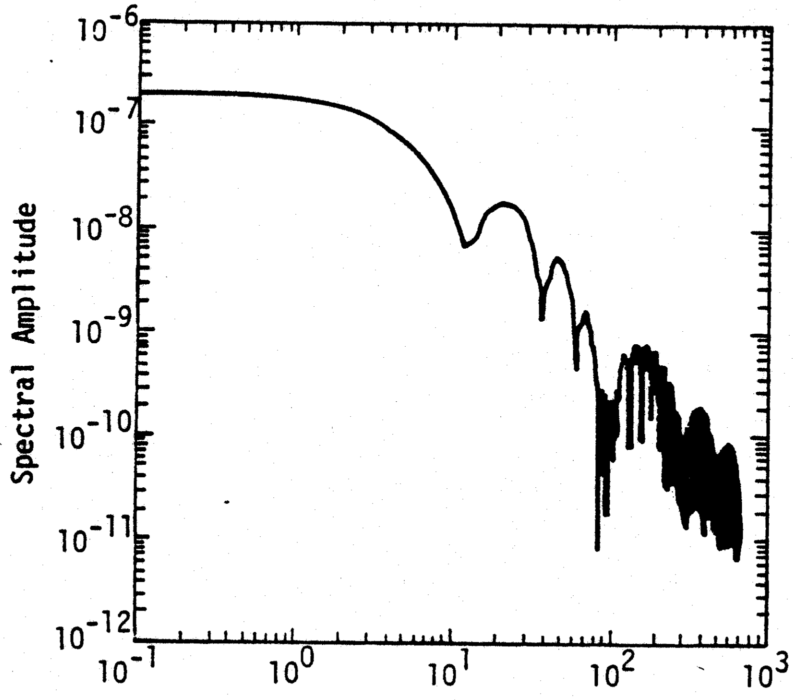


Figure 3. Sum of double exponentials; time and frequency domain plots.

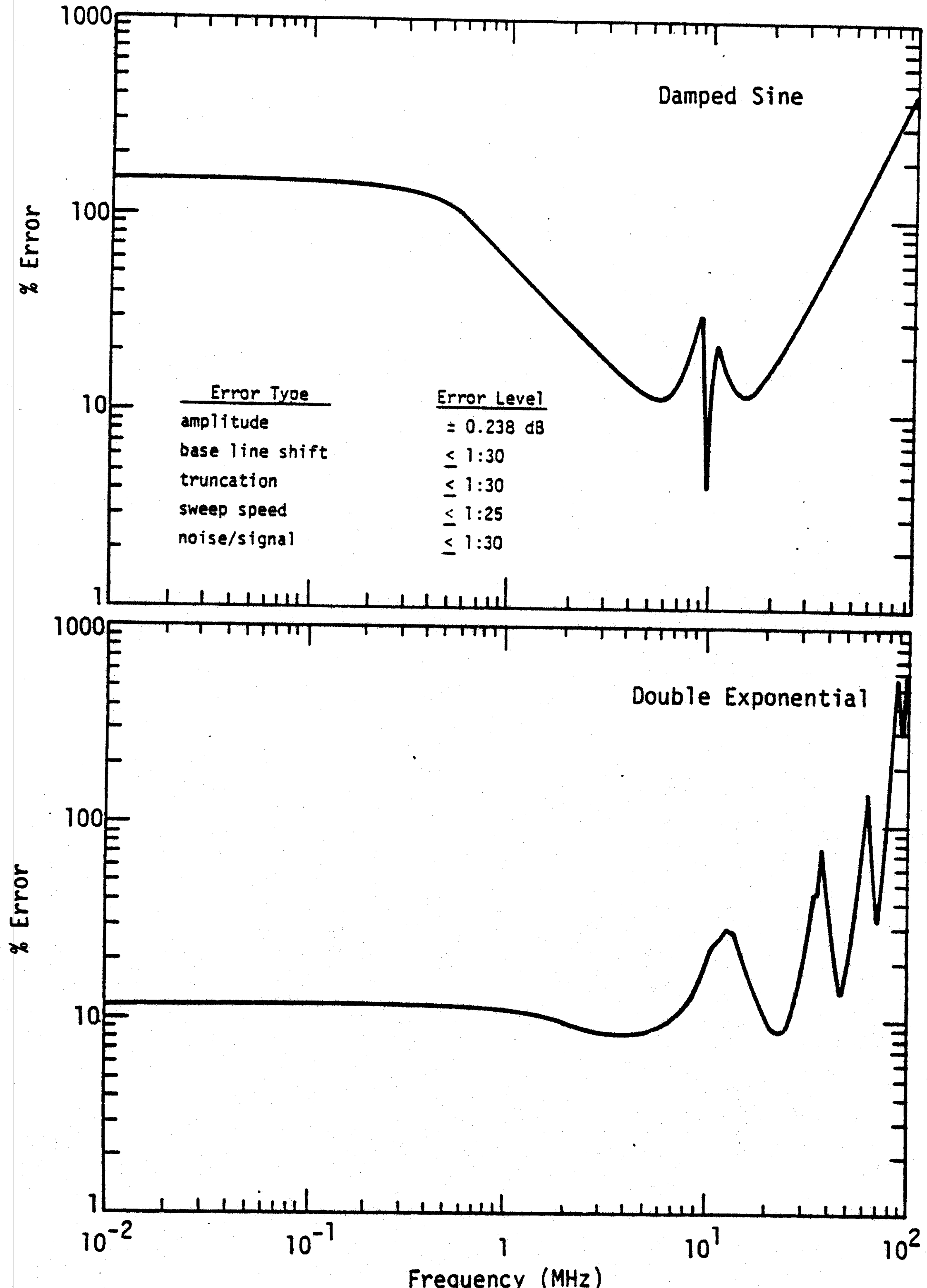


Figure 4. Case 1a: Overall Error Envelope

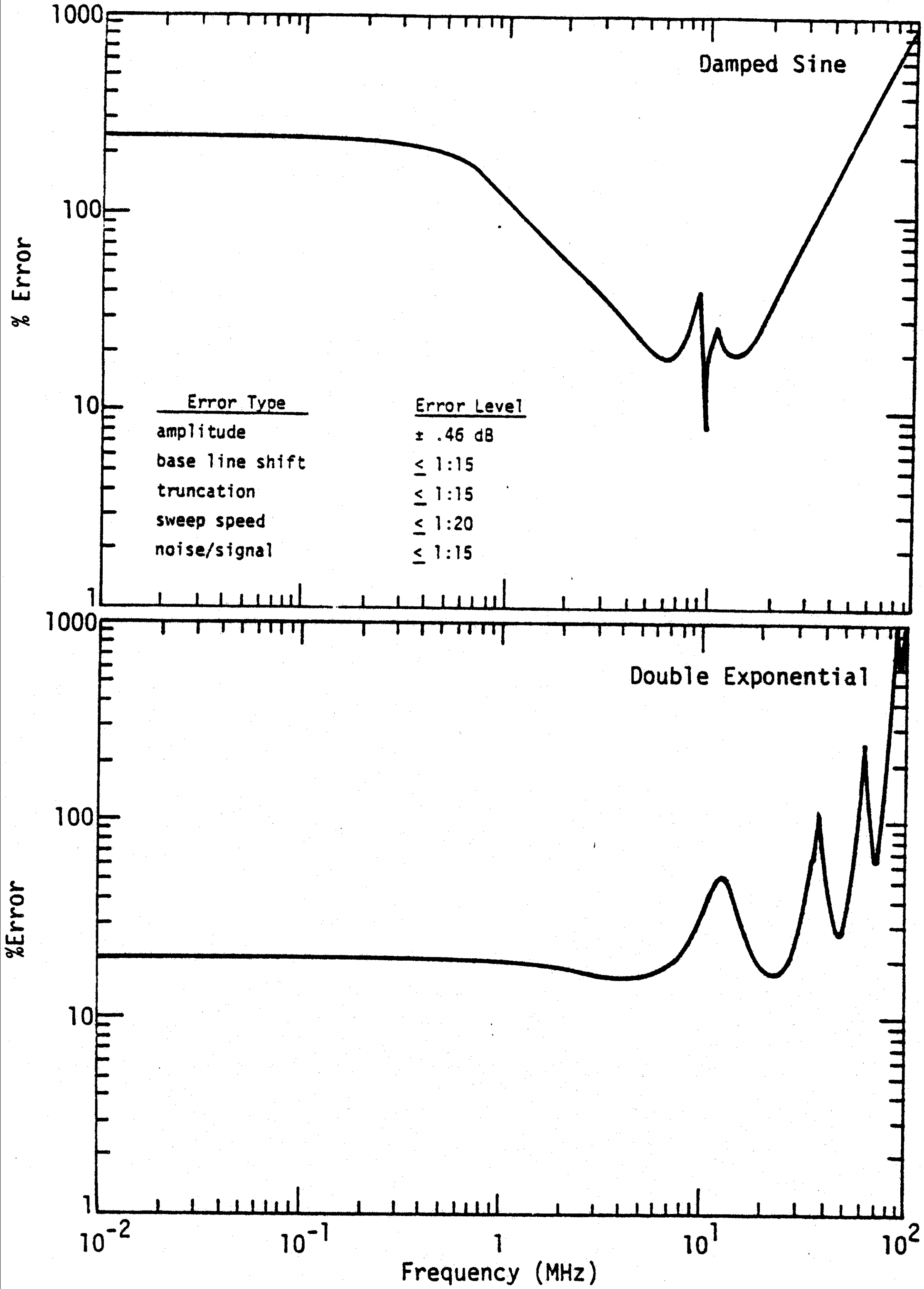


Figure 5. Case 1b: Overall Error Envelope

(2a) Transient Digitizer/Microwave System (no filters, using AutoCal*, upper limit 10^8 Hz)

Error Type	Error Level
amplitude	± .5 dB
base line shift	$\leq$ 1:30
truncation	$\leq$ 1:30
sweep speed	$\leq$ 1:100
noise/signal	$\leq$ 1:30

Figure 6 shows the combined frequency domain errors for this system.

(2b) Transient Digitizer/Microwave System (no filters, no Autocal, upper limit 10^8 Hz)

Error Type	Error Level
amplitude	± .5 dB
base line shift	$\leq$ 1:15
truncation	$\leq$ 1:15
sweep speed	$\leq$ 1:30
noise/signal	$\leq$ 1:15

Figure 7 shows the combined Transient Digitizer error in the frequency domain for this typical case.

7. CONCLUSIONS AND OBSERVATIONS

From the examples displayed in figures 2 through 7, several things become apparent:

(1) Sweep speed errors cause tremendous amplitude errors when the signal contains deep nulls or high peaks. For some purposes these errors can be neglected since they arise from a shift in the time base which in turn shifts the

*AutoCal - computer based calibration procedure

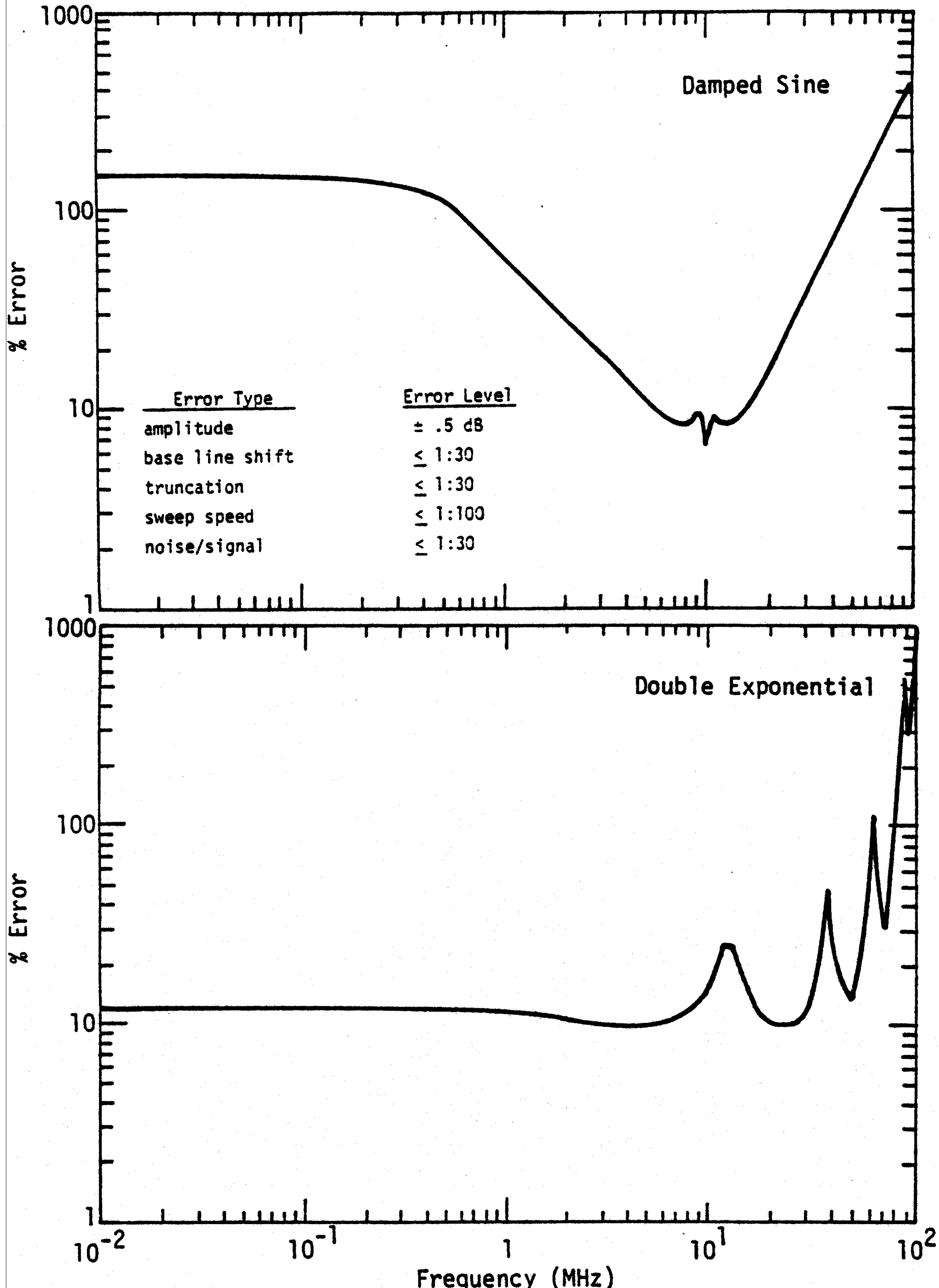


Figure 6. Case 2a: Overall Error Envelope

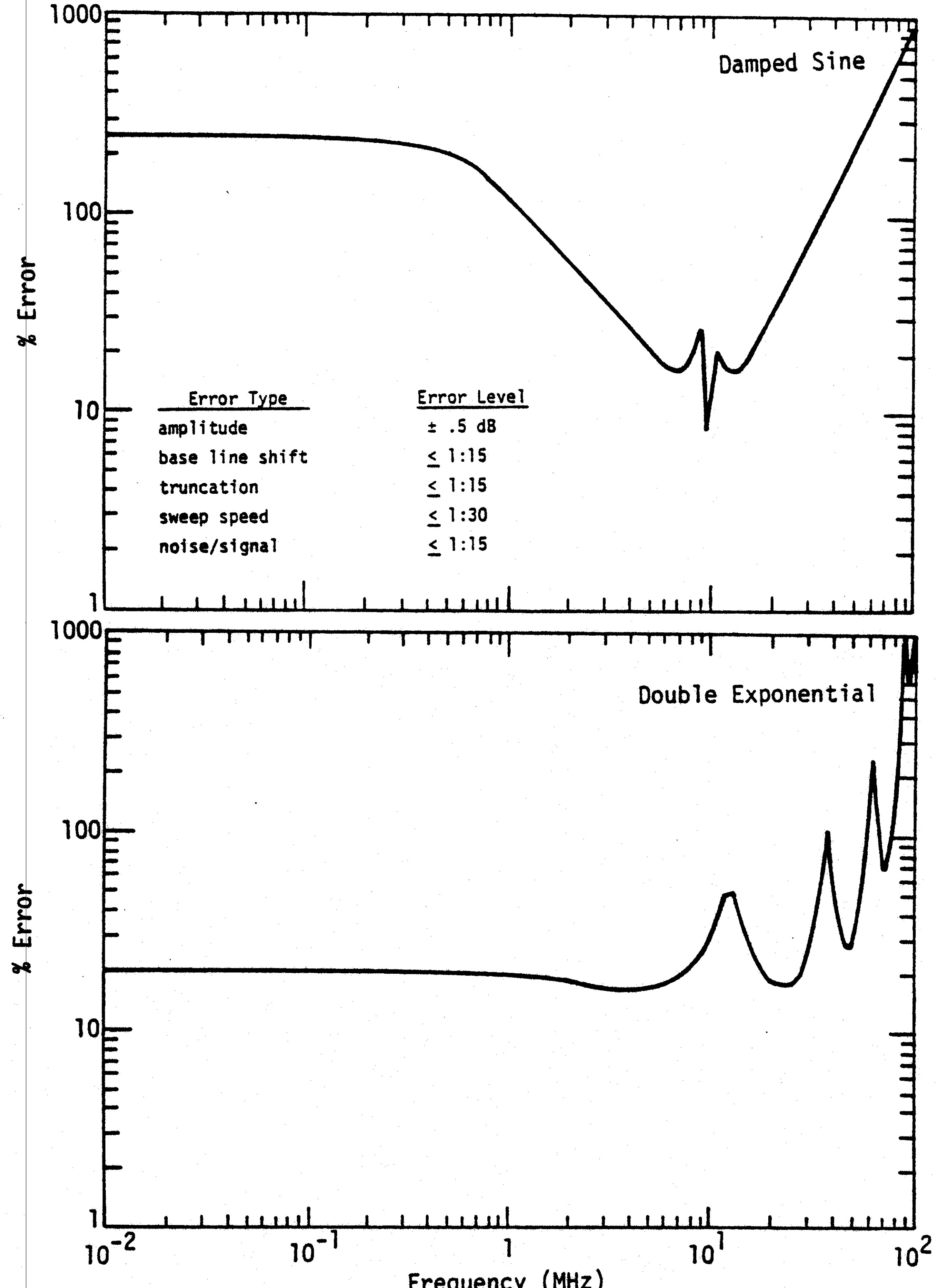


Figure 7. Case 2b: Overall Error Envelope

frequency base. The additional amplitude shift is relatively small. Where the exact frequency of a particular null or peak is not required, this form of error is not of concern. However, if for example, the data is to be ratioed with another similar function, these errors are real and must be considered since the coincidence or lack thereof may be the essential information from the ratio.

(2) As a rule of thumb, base line shift or rotation introduces error as the inverse of frequency so low-frequency data is most affected by this form of error. It should be mentioned that the above figures are illustrating the upper bounds of these error contributions. In this instance the error will introduce additional structure to the apparent signal due to the windowing of the ramp or step introduced by the error itself.

(3) Truncation error is like base line error in that it will introduce added structure to the signal. The envelope bounding this error as defined here is, however, constant with frequency.

(4) Noise error is best viewed in the real and imaginary parts of the frequency domain where it is a constant with frequency, providing the noise in the time domain is normal, has zero mean, and random.

REFERENCES

1. Ashley, C. and W. R. Graham (RDA), "Estimation of Errors in Calculating Transfer Functions from Pulse Data," Measurement Note 19, Air Force Weapons Laboratory, 5 September 1973.

APPENDIX A

DETERMINISTICALLY-DERIVED FREQUENCY-DOMAIN NOISE FROM INSTRUMENTATION/DIGITIZATION TIME DOMAIN NOISE

1. NOISE PROPAGATION OF DATA PROCESSED BY DIGITIZATION AND DISCRETE FOURIER TRANSFORMATION

Consider a signal f(t) contaminated by Gaussian noise n(t). The resulting function is $f_m(t) = f(t) + n(t)$. Analytically the Fourier transform is

$$F_m(\omega) = \int_{-\infty}^{\infty} [f(t) + n(t)]\, e^{-j\omega t}\, dt$$

or

$$F(\omega) + N(\omega) = \int_{-\infty}^{\infty} f(t)\, e^{-j\omega t}\, dt + \int_{-\infty}^{\infty} n(t)\, e^{-j\omega t}\, dt.$$

Thus, we need only consider the last integral in looking at the noise propagation.

In the real world, all data begins at some time $> -\infty$, typically at $t = 0$ and terminated at some finite time $t = T_{max}$.

So in processing data, the integral is required over a finite interval.

$$N(\omega) = \int_{0}^{T_{max}} n(t)\, e^{-j\omega t}\, dt\ .$$

Further, the signal being processed is digitized (i.e., sampled at discrete intervals) so that the integral becomes a sum

$$N(\omega) \doteq \sum_{i}^{N} n(t_i)\, e^{-j\omega t_i}\, \Delta t_i\ .$$

The details of the integration algorithm will differ depending on the Fourier transform code being used but the essential parameters are the same.

We see that the transfer function on the noise as we transform from time to frequency will in general be dependent on the following parameters:

a. N, the number of sample points

b. Δt_i, the duration of each time increment

c. $T_{max} = \sum_{i}^{N} \Delta t_i$, the signal time window

d. the particular discrete Fourier transform integration algorithm.

We also note in reference 1 that since n(t) is normal, has zero mean, is weakly stationary and a random function of time, then the real and imaginary parts of $N(\omega)$ are uncorrelated, have zero mean, are weakly stationary and are random functions of freqxency. These properties hold for both discrete and continuous Fourier representations.

With these properties in mind, reference 2 shows that the following is true for uniform sampling, where x(t) is normal with a variance of unity:

$$\text{variance of} \left\{ \sum_{i=0}^{n-1} x(t_i) \cos \omega t_i \right\} = \frac{n}{2} ; \qquad \omega \neq 0 \text{ or } \pi$$

or

$$\text{variance of} \left\{ \sum_{i=0}^{n-1} x(t_i) \sin \omega t_i \right\} = \frac{n}{2} ; \qquad \omega \neq 0 \text{ or } \pi .$$

The mean is zero and the covariance is zero for these two series.

For uniform sampling (Δt), this is directly applicable to numerical Fourier transforms. For example, using rectangular rule integration, if we let

$$x(t_i) = \Delta t n(t_i) ,$$

then

$$\sigma(\omega) = \sigma(t) \; \Delta t \sqrt{\frac{n}{2}} .$$

$\sigma(\omega)$ is the frequency domain standard deviation and $\sigma(t)$ is the time domain standard deviation.

2. COMBINING FREQUENCY DOMAIN NOISE WITH THE PROCESSED SIGNAL

We have seen above that when n(t) is Gaussian with zero mean, we expected that the processed noise $N(\omega)$ would be Gaussian also, and the real and imaginary parts would be uncorrelated and have zero means with the same sigma or standard deviations. Evidently this expectation is reasonable for well-implemented data processing.

Let us express the transformed noisy function of frequency as

$$R_e\{F_m(\omega)\} = R(\omega) + N_R(\omega)$$

$$I_m\{F_m(\omega)\} = I(\omega) + N_I(\omega)$$

where

$$N_R(\omega) = R_e\{N(\omega)\}$$

$$N_I(\omega) = I_m\{N(\omega)\} \ .$$

We note that when the noise is expressed in this manner, the noise in the transform of the measured function is simply $N_R(\omega)$ and $N_I(\omega)$ added to the real and imaginary parts of the unperturbed function.

Assuming a suitable method is available to determine $N(\omega)$ from n(t) (see Appendix B for an algorithm for general data processing), the problem of error determination in the frequency domain is straightforward. If one looks at only real and imaginary parts, the relative error equation can be written immediately in terms of one sigma error

$$\varepsilon_R(\omega) = \frac{\pm\sigma_R(\omega)}{R(\omega)}$$

and

$$\varepsilon_I(\omega) = \frac{\pm\sigma_I(\omega)}{I(\omega)} \ .$$

If this error is propagated into the function magnitude, the expression becomes a bit more complicated.

$$|F_m(\omega)| = \sqrt{R_e\{F_m(\omega)\}^2 + I_m\{F_m(\omega)\}^2}$$

expanding

$$|F_m(\omega)|^2 = \{R(\omega) + N_R(\omega)\}^2 + \{I(\omega) + N_I(\omega)\}^2$$

$$= R(\omega)^2 + I(\omega)^2 + 2[R(\omega)N_R(\omega) + I(\omega)N_I(\omega)] + N_R(\omega)^2 + N_I(\omega)^2$$

$$|F_m(\omega)|^2 = |F(\omega)|^2 \left\{1 + \frac{2\ R(\omega)N_R(\omega) + I(\omega)N_I(\omega) \quad + N_R(\omega)^2 + N_I(\omega)^2}{R(\omega)^2 + I(\omega)^2}\right\}$$

or

$$|F_m(\omega)| = |F(\omega)|\sqrt{1 + \frac{2\ R(\omega)N_R(\omega) + I(\omega)N_I(\omega) \quad + N_R(\omega)^2 + N_I(\omega)^2}{R(\omega)^2 + I(\omega)^2}} .$$

Thus relative error is

$$\varepsilon_R(\omega) = \sqrt{1 + \frac{2\ R(\omega)N_R(\omega) + I(\omega)N_I(\omega) \quad + N_R(\omega)^2 + N_I(\omega)^2}{R(\omega)^2 + I(\omega)^2}} - 1 .$$

Recall $N_R(\omega)$ and $N_I(\omega)$ have the same distribution as n(t), if n(t) is normal or Guassian. Hence they are constants in frequency and have the same sigma (σ).

To develop the one sigma limits on $\varepsilon_R(\omega)$, we substitute $\pm\sigma$ for $N_R(\omega)$ and $N_I(\omega)$ and adjust the signs to achieve max and min of $\varepsilon_R(\omega)$ at any one frequency ω

$$\varepsilon_R(\omega)\Big|_{max} = \sqrt{1 + \frac{2\sigma\{|R(\omega)| + |I(\omega)| + \sigma\}}{R(\omega)^2 + I(\omega)^2}} - 1$$

$$\varepsilon_R(\omega)\Big|_{min} = \sqrt{1 + \frac{2\sigma\{-|R(\omega)| - |I(\omega)| + \sigma\}}{R(\omega)^2 + I(\omega)^2}} - 1 \quad .$$

We note that for $\sigma < |R(\omega)| + |I(\omega)|$ that $\varepsilon_R(\omega)\big|_{min} < 0$ and $\varepsilon_R(\omega)\big|_{max} > 0$ but they are not symmetric about zero. For $\sigma > |R(\omega)| + |I(\omega)|$, the one sigma error is biased positively so that the unperturbed function value is no longer contained between the upper and lower one sigma bounds.*

3. APPLICATION OF NOISE ERROR MODEL TO MEASURED DATA

Since measured and transformed data is already contaminated with noise it is a bit less straightforward to determine the expected error.

As discussed above, if we limit our attention to the real and imaginary parts of the function we have

$$R_e\{F_m(\omega)\} = R(\omega) + N_R(\omega) = R_m(\omega)$$

$$I_m\{F_m(\omega)\} = I(\omega) + N_I(\omega) = I_m(\omega)$$

where

$$N_R(\omega) = R_e\{N(\omega)\}$$

$$N_I(\omega) = I_m\{N(\omega)\} \quad .$$

From the noise transfer function model we can determine the noise term (a constant). Then to find one sigma bounds, we let $N(\omega) \rightarrow \pm\sigma$ and the relative error for the real part becomes

$$\varepsilon_R(\omega) = \frac{\pm\sigma}{R_m(\omega)\pm\sigma}$$

and for the imaginary part

$$\varepsilon_I(\omega) = \frac{\pm\sigma}{I_m(\omega)\mp\sigma} \quad .$$

*The distribution function is now Rayleigh, see reference 3.

To carry this through to the magnitude is a bit more involved, and leads to an expression for the error that contains singularities.

For small levels of noise ($\sigma < |I_m| + |R_m|$), the equations found in the previous section for relative error, on the signal magnitude, are useful. At greater noise levels, a rigorous expression for the magnitude error blows up.

APPENDIX A REFERENCES

1. Goodman, N. R., "Some Comments on Spectral Analysis of Time Series", Technometrics, Vol. 3, No. 2, May 1961.

2. "Fourier Analysis of Time Series: An Introduction", Bloomfield, Wiley, 1976, p. 109.

3. Freeman, Harold, Introduction to Statistical Inference, Addision-Wesley. Reading, Massachusetts, 1963, p. 155.

APPENDIX B

PROCEDURE FOR DETERMINING THE APPROXIMATE NOISE TRANSFER FUNCTION FOR FOURIER TRANSFORM DATA PROCESSING

We expect that the random properties found for the idealized numerical transforms (Appendix A) will tend to hold also for typical measured data being processed. With this expectation, a procedure may be devised for determining a noise transfer function that is approximately constant with respect to frequency.

1. Start with Gaussian noise that is truncated at 2σ (representative of instrumentation/digitization noise, see Appendix C for noise generation algorithm).

2. Pass this noise through the processing system using same Δt's, T_{max}(time window) and FFT, DFT or FIT algorithm that is to be used on the data to be evaluated. Use the same frequency band and frequency calculational points as well. Limit frequencies to less than one-fifth Δt (as you would normally for valid data processing).

3. Repeat this procedure until the accumulated standard deviation (SD) of both real and imaginary components of the resulting frequency domain noise are essentially equal. The SD is calculated for all resultant values over the frequency band stated above. Subsequent passes are accumulated to get the overall SD.

4. The resulting value for the SD can then be used to predict noise on a transformed noisy signal, f(t) + n(t) having a known signal-to-noise (S/N) ratio or absolute noise level, by appropriate scaling of the above SD.

RESULTS

This procedure has been implemented using several computers and using algorithms having the following properties:

a. N = 1001 sample points

b. $\Delta t = 8\times10^{-9}$ sec

c. $T_{max} = 8\times10^{-6}$ sec

d. rectangular rule integration algorithm and Simpson's rule integration (i.e., same as used for Upgrade Testing CDC 7600 Fourier transform).

e. frequency sampled at 28, 55, and 90 (9, 18 and 90 per decade) points from 10^4 to 10^7 Hz.

The resultant SD was about 1.43×10^{-7} for the truncated 2σ (max input deviation $\pm 2\sigma$) Gaussian noise input.

It is comforting to observe that the analytically predicted result for this same case, but without truncating the input noise, is 1.78×10^{-7} for the standard deviation. A number of additional computer-derived results for different Δt's are in similar agreement.

APPENDIX C

PROPOSED STANDARDIZATION OF FUNCTIONS USED TO SIMULATE NOISE IN MEASURED DATA †

1. INTRODUCTION

The presence of noise in data analyzed with pole extraction algorithms can significantly affect the accuracy of the derived poles. A measure of the accuracy of a particular algorithm is best obtained by using as inputs, analytical functions whose true pulse are known. A number of different methods have been used to simulate the presence of noise in such analytical functions. As a result there is no common measure of the relative accuracy of two given pole extraction algorithms. In this paper, a standard model simulating noise in analytical data is proposed which will facilitate comparisons of algorithm accuracy.

2. NOISE MODEL

To best approximate a physically real situation we assume the noise to be additive, and write

$$f_M(t) = f(t) + n(t)$$

where f(t) is an ideal (noise-free) function of time, n(t) is the noise (a zero mean, weakly stationary, random function of time)*, and $f_M(t)$ is the result of measuring f(t) in the presence of noise.

Denoting the Fourier transforms of $f_M(t)$, f(t), and n(t) as $\tilde{F}_M(\omega)$, $\tilde{F}(\omega)$, and $\tilde{N}(\omega)$, respectively we can write

$$\tilde{F}_M(\omega) = \tilde{F}(\omega) + \tilde{N}(\omega) \ .$$

These Fourier transforms are complex functions, and the real and imaginary parts of $\tilde{N}(\omega)$ are uncorrelated, zero mean, weakly stationary, random functions of frequency (ref. 1).

*that is, the covariance matrix of $n(t_1)$, $n(t_2)$, ..., $n(t_n)$ is invariant under time translation.

†Prewitt, J. F. and L. D. Scott, "Proposed Standardization of Functions Used to Simulate Noise in Measured Data," AMRC-N-58, Mission Research Corporation, May 1977.

Again since many physical processes are reasonably modeled by normal distributions we now assume that n(t) is Gaussian distributed (normal). Its probability density function is

$$p(n) = \frac{1}{\sigma_n\sqrt{2\pi}} \exp\,[-n^2/2\sigma_n{}^2].$$

Then the real and imaginary parts of $\tilde{N}(\omega)$ are also normally distributed with probability density

$$P(\tilde{N}) = \frac{1}{\sigma_N\sqrt{2\pi}} \exp\,[-\tilde{N}^2/2\sigma_N{}^2].$$

However, writing

$$\tilde{N}(\omega) = Ne^{i\phi_N},$$

neither N nor ϕ_N are normally distributed. In fact N is Rayleigh distributed (ref. 2), with probability density function

$$P(N) = \frac{N}{\sigma_N{}^2} \exp\,[-N^2/2\sigma_N{}^2] \qquad N \geq 0$$

$$= 0 \qquad N < 0$$

and ϕ_N is uniformly distributed on the interval $0 \leq \phi_N \leq 2\pi$ (ref. 3). The probability density functions p(n) and P(N) are illustrated in figure 1. Note that since the amplitudes $|\tilde{F}(\omega)|$ and $|\tilde{N}(\omega)|$ are not additive, the knowledge that N is Rayleigh distributed is not of primary interest.

Any process used to measure a given time (or frequency) domain signal has a number of inherent thresholds. Such thresholds have the effect of clipping the "tail" of the signal's noise probability density function. We propose to simulate this clipping by discarding all values of $n(t)[\tilde{N}(\omega)]$ for which $|n(t)|$ $[N(\omega)]$ is larger than $2\sigma_n[2\sigma_N]$.

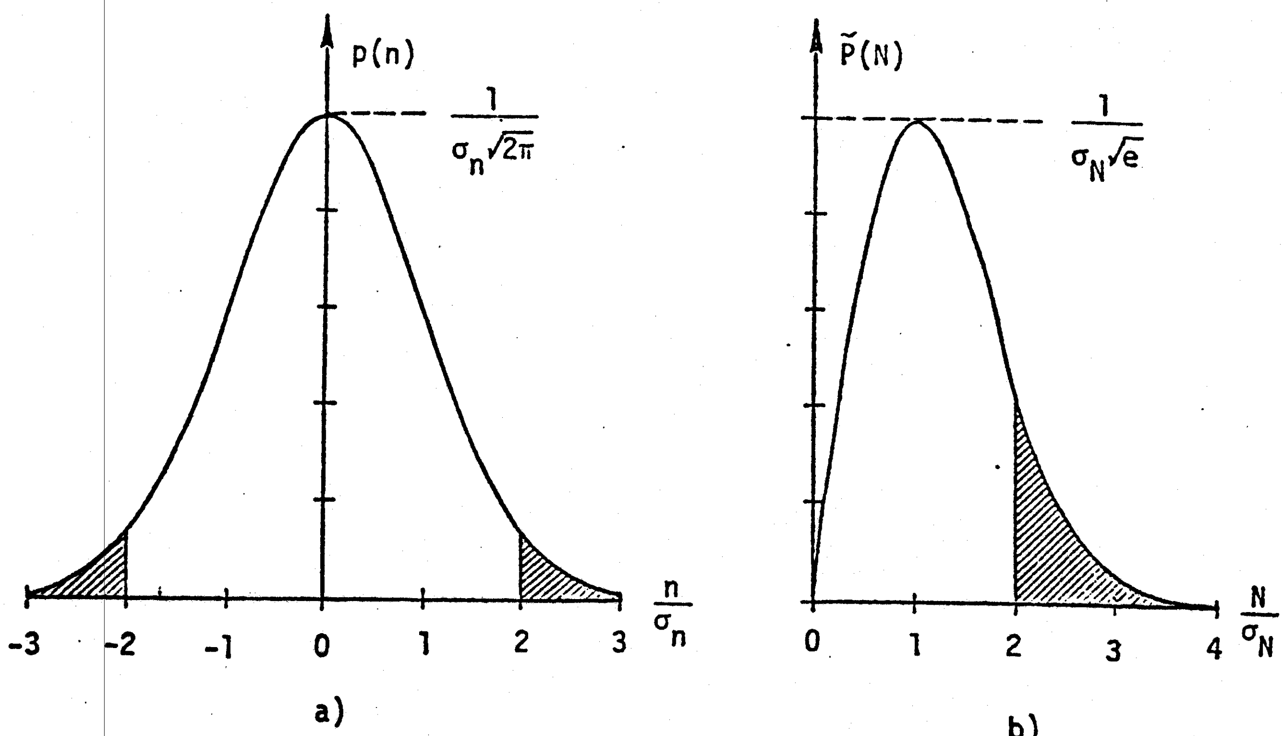


Figure 1. Probability Density Functions for:
a) a Gaussian Distribution
b) a Rayleigh Distribution.

The remaining parameter to be specified is a measure of the noise amplitude. It is porposed that noise amplitude be defined by specifying $2\sigma_n [2\sigma_N]$ as a fraction of the largest value of $|f(t)|$ $[|\tilde{F}(\omega)|]$. Thus the effective signal-to-noise ratio in the time or frequency domain is

$$S/N = \frac{|f_{max}(t)|}{2\sigma_n}$$

or

$$S/N = \frac{|F_{max}(\omega)|}{2\sigma_N} .$$

3. GENERATION OF GAUSSIAN DISTRIBUTED RANDOM NUMBERS

If a sequence of independent random numbers $\{x\}$ has a cumulative distribution function $f(x)$, then the solutions x_i of the equation

$$u_i = F(x_i) \qquad i = 1,2,....$$

where {u} is a sequence of random numbers distributed uniformly in the interval (0,1), are members of the sequence {x} (ref. 4). This statement is useful only if

$$x = F^{-1}(u)$$

can be obtained or approximated analytically. The cumulative distribution function is related to the probability density function p(x) by the integral

$$F(x) = \int_{-\infty}^{x} p(y)dy.$$

For the Gaussian distribution,

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-y^2/2\sigma^2} dy,$$

cannot be integrated analytically, and $u = F^{-1}(x)$ cannot be formed directly. However, note that for the complex random function

$$y = y_R + jy_I = xe^{i\phi}$$

$$= x(\cos\phi + j\sin\phi)$$

if y_R and y_I are zero mean Gaussian distributed functions, then x is Rayleigh distributed and ϕ is uniformly distributed on $(0,2\pi)$. Then

$$y_R = x\cos\phi$$

$$y_I = x\sin\phi ,$$

so if we can generate Rayleigh distributed numbers, we can obtain Gaussian distributed numbers from them.

For the Rayleigh distribution, the cumulative distribution function is

$$F(x) = \int_{0}^{x} \frac{y}{\sigma^2} e^{-y^2/2\sigma^2} dy .$$

Since

$$\frac{d}{dy}\left[e^{-y^2/2\sigma^2}\right] = -\frac{y}{\sigma^2}\ e^{-y^2/2\sigma^2} ,$$

$$F(x) = -\left[e^{-y^2/2\sigma^2}\right]_0^x$$

and

$$u = 1 - e^{-x^2/2\sigma^2}$$

has solutions

$$x_i = \sigma\sqrt{2\ell n(1 - \mu_i)}$$

which are Rayleigh distributed random numbers on the interval $(0,\infty)$. Note that is u_i is a unifromly distributed random number on (0,1), so is $(1 - u_i)$. Thus we can write more simply

$$x_i = \sigma\sqrt{-2\ell n\ u_i} .$$

Since ϕ is uniformly distributed on $(0,2\pi)$, we can write

$$\phi_m = 2\pi u_m,$$

where u_m is uniformly distributed on (0,1). Thus

$$(y_R)_k = \sigma\sqrt{-2\ell n\ u_i}\ \cos(2\pi u_m)$$

and

$$(y_I)_k = \sigma\sqrt{-2\ell n\ u_i}\ \sin(2\pi u_m)$$

both belong to a zero mean, Gaussian distribution with standard deviation s. These same two formulas, obviously, can be used to generate a pair of Gaussian random numbers in the time or frequency domains. These formulas are presented (without explanation) in reference 4.

Thus the proposed standard noise distributions can be generated as follows:

1. Calculate the Gaussian standard deviation from the desired signal-to-noise ratio

$$\sigma_n = \frac{|f_{max}|}{2(S/N)} .$$

2. Generate a uniform random number pair u_1 and u_2, $0 < u_i \leq 1$ (note exclusion of zero).

3. Calculate a Gaussian random number pair as

$$n_1 = \sigma\sqrt{-2\ell n\ u_1}\ \cos(2\pi u_2)$$

$$n_2 = \sigma\sqrt{-2\ell n\ u_1}\ \sin(2\pi u_2).$$

4. Discard n_i for $|n_i| > 2\sigma$.

5. Form

$$f_M(t_1) = f(t_1) + n_1$$

$$f_M(t_2) = f(t_2) + n_2$$

or

$$\tilde{F}_M(\omega_1) = Re[\tilde{F}(\omega_1)] + n_1 + j\{Im[\tilde{F}(\omega_1)] + n_2\}.$$

6. Return to Step 2.

7. The functions $f_M(t)$ and $\tilde{F}_M(\omega)$ contain Gaussian random noise truncated at $\pm 2\sigma$ (sepcified by the desired signal-to-noise ratio).

APPENDIX C REFERENCES

1. Goodman, N. R., "Some Comments on Spectral Analysis of Time Series", Technometrics, Vol. 3, No. 2, May 1961.

2. Freeman, Harold, Introduction to Statistical Inference. Addison-Wesley, Reading, Mass., 1963, p. 155.

3. Feller, William, An Introduction to Probability Theory and Its Applications, Volume II. John Wiley and Sons, New York, 1971, p. 101.

4. U. S. Department of Commerce, National Bureau of Standards, Handbook of Mathematical Functions, No. 55, Applied Mathematics Series, June 1964.