WHAT TO DO IF BITTEN -- PRACTICAL TACTICS FOR FREQUENCY MEASUREMENT

Eric L. Blomberg
Cogent Design

ABSTRACT

A newcomer to the area of precise frequency measurement can shorten the process of learning to produce repeatable and credible results by developing a critical perspective. By discussing practical systems and their pitfalls, this paper hopes to establish such a perspective -- one with which the user can check his actual results and procedures against the large background of data and experience which the PTTI community has accrued.

INTRODUCTION

There is only one purpose in making precise frequency measurements. It is not just the characterization of the device or system in question. Rather, it is to make a characterization so that the results are repeatable and directly relatable to other users within the community. It is often the case that data is devalued by an inexact or poorly defined measurement process. We shall look at some of the particular processes used and point out what is needed to insure the usefulness of the results.

There is both good and bad news. The good news is that, at least in my own experience, the domain of frequency measurement is a 10% to 15% world -- in terms of repeatability, transportability and agreement with physics. The bad news is that this obligates us, as practitioners, to use procedures that close the loop at this level. When our results don't agree we can no longer claim that Black Magic didn't work today. We must actually review our procedures and data until we locate a cause of the discrepancy.

MEASURING PERFORMANCE VS DOCUMENTING ERRORS

Less than half the effort of a precise frequency measurement is spent on the actual characterization of device performance. A good deal of the effort must be spent on ascertaining the limitations and flaws of the measurement system. There are actually three kinds of data in each measurement (Fig. 1).

1) The actual device characterization data -- which tries to match predicted and measured data. Since this almost always involves a noise process, these results tend to be the statistical treatment of an ergodic or clearnly defined set of ergodic processes.

2) Measuring the "systematics" of the device -- all those specific cause-and-effect processes which mask and corrupt the underlying noise processes of interest. These effects cannot be treated statistically. They consist mainly of sensitivity coefficients to external and often unknown stimuli.

3) Measuring both the noise process and the corrupting systematics of the measuring system itself.

Since a precise frequency measurment is often functiong at the state-of-the-art, it is unlikely that the desired data will stand clearly above these

Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **DEC 1985**	2. REPORT TYPE	3. DATES COVERED **00-00-1985 to 00-00-1985**
4. TITLE AND SUBTITLE **What to do if Bitten - Practical Tactics for Frequency Measurement**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Cogent Design,London EC1N 8TS,**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

Proceedings of the Seventeenth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Washington, DC, 3-5 Dec 1985

14. ABSTRACT

see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**12**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

obscuring effects. What we can do and must do for reputable measurements is to measure and document these effects as well as the data itself.

BASELINES AND REFERENCES

All frequency measurements are relative to a baseline of some kind. An ideal laboratory might have a Hydrogen Maser or other near-ultimate reference standard whose performance would exceed the device under test by several orders of magnitude. A reference is only part of the story. In order to guarantee that the inherent performance of the reference is maintained, some sort of measurement system baseline must be established. This is often some sort of closed loop end-to-end test which includes everything except the device under test. Such a baseline can establish the credibility and performance level of the entire measurement system.

It is not so important that the baseline be of a certain ultimate level as it is that this baseline be well known. One traditional form of baseline is the common-mode type (Fig.2 &3). For measurement systems with two input channels (such as a phase comparator), inject the same reference into both inputs. The net output will be the internal phase variations of the measurement system itself. Clearly this baseline data must be taken over the same conditions (environmental, averaging time,etc.) as those for the device under test. This particular kind of baseline is differential -- that is, only the difference between channels is observed. Thus, it is less sensitive to the absolute behavior of the common mode source. It is then reasonable to use as a source, not a super reference, but the device under test. A well behaved differential baseline can improve upon its source by at least an order of magnitude. Some common sense needs to apply here. For the example of the Dual Mixer Time Difference system, the immunity to common mode effects is proportional to the smallness of the raw phase offset.

THE RULE OF THREE

All precise frequency/phase measurements consist of linear frequency differences. This may take the form of simple frequency differences against a sound reference or short-term phase differences between two identical but lower quality oscillators. Whether the process we are viewing is a noise process or systematic response to stimuli, our viewpoint is still differential. Since we are also dealing with small proportional differences, a linear first-order view is entirely appropriate. This entitles us to extend our simple common-mode view to 3 sources. In a measurement system where the reference is not head and shoulders above the device under test, we are entitled to any performance inferences from the three pair-wise measurements of a group of three somewhat equal sources. This is particularly useful when trying to pin down drift and other systematic effects (Fig. 4).

A BASELINE EXAMPLE

In this example, an attempt to measure oscillator phase noise, L(f), we see how we can be bitten by a baseline (Fig. 5). The set-up is the familiar one, locking the oscillator under test to a super-oscillator with a loop bandwidth of less than one Hertz. The phase noise of the oscillator pair (we assume dominated by the oscillator under test) can be determined for Fourier frequencies greater than fifty Hertz simply by measuring the noise voltage at the mixer output. The system, by virtue of its spectrum analyzer is ideally set up to work in noise density (i.e. volts/Hz translatable to phase noise in dBC).

Since phase noise density, L(f), is simply the ratio of power in the carrier to power in a one Hertz bandwidth at a Fourier frequency, our baseline can be self-calibrating. If we unlock the loop and permit a beat between the oscillators, we can observe this sine wave at the mixer output. If the mixer is not saturated we can take this as our zero reference for the carrier (suppose 1 volt peak-to-peak or 0.35 volts RMS). If we lock the system and then measure the noise voltage density, then the simple ratio yields L(f), with three db to be subtracted to allow for the folding over of the other sideband.

Now let's look at the problem of establishing a baseline. Here the issue is determining the system noise floor. Looking at the block diagram, we see the inherent effective input noise density of each element. We also see the effective input noise appearing at the pre-amp input. Our goal is for the system noise floor to be at least ten dB below the expected device noise. If we calculate the noise voltage corresponding to an L(f) of -153 dBC for a good oscillator, we expect 10nv/Hz$^{1/2}$ at the mixer output (this has already the 3 dB DSD to SSB conversion factor). This is 10 dB below the pre-amp input noise. So far, so good. After the pre-amp gain of 1000, the expected noise is 10 uv/Hz$^{1/2}$. This is 20 dB below the analyzer's equivalent input noise. We would then conclude that we have proven a system baseline capable of measuring L(f) of -153 dBC.

wrong. A calculated baseline is not sufficient. Measuring the baseline (with both oscillators off) will not show a flat noise trace at an equivalent of -163 dBC. The noise floor will start to droop off starting at 10 KHz. This is due to the open loop gain limitation of the op-amp. A signal injection experiment will show pre-amp gain to fall off as well. While our calculated baseline will be born out for Fourier frequencies of less than 8 KHz, it collapses at the higher frequencies. At a Fourier frequency of 100 KHz this sytem will yield a 20 dB error. One important lesson here is that baselines must be measured and documented as fully as experimental data.

POWER AND GROUNDING

These two ares of vulnerability have scuttled many precise frequency measurements. Since some frequency measurements need 5 to 10 day uninterrupted runs, the window of vulnerability is large.

Most labs will not go more than one day without a major power glitch. The traditional, and relatively inexpensive, solution is to run both devices under test and key test equipment from a battery-backed-up source consisting of Sears Diehards and commercial grade DC to 60 Hz inverters. Considerable, but hard to trace, errors can come from free running inverters. They should be synchronized to the AC line (Fig. 6).

Another source of error can come from operating components of the test set-up from different lab AC circuits. I have observed over 100 mA of current forced through signal lines which bridged two supposed AC grounds.

The use of coax forces a single ended ground system upon us. Test equipment should have cases bonded by heavy straps. Critical high frequency lines should be broken for DC with wide-band shielded transformers with impedance matches properly maintained.

A simple diagnostic is to measure between supposed ground points in the system with a low range DVM. More than a few millivolts of DC or AC is cause for alarm.

WHOLISTIC DATA TAKING

Data taking tactics need to deal not so much with the desired experimental data as with what is going to go wrong. The beat period data which is the heart of most precise frequency measurements should take up only 10% of the data volume. The pessimistic assumption here is that the frequency data will be corrupted by systematic effects. When both environmental and device parametric data are taken concurrently, there is some chance of removing these effects during data processing (Fig.7 and 8). It often occurs that the calibration of these sensitivity coefficients as at least as important as the frequency stability data itself.

Small computers and multi-channel D/A systems are inexpensive and easy to program. Total data volume can be limited by common sense. Temperatures, for example, probably don't need up-dating more than once a minute. Since this is a coarsely sampled sort of telemetry, it is futile to use this method to catch transients. Parameters where transients are expected are best viewed on analog strip charts. Correlation with frequency behavior is important, so the frequency data should be converted to an analog representation and placed on the same strip chart.

One problem with precise frequency measurements is that we never seem to be able to do just one of them. It is absolutely vital to preserve the integrity of the raw data and tie each to its respective measurement and observed conditions. Successful frequency measurements stem directly from our ability to learn from history -- the detailed history of successful experiments.

A vital tactical decision is the length of the data run. A good rule of thumb is 100 data points for each averaging interval. This can be somewhat long if the desired data interval is 100,000 seconds. Dave Allan and others in the PTTI community have done some work in getting more use from skimpy data sets, but this compromise goes in the direction of establishing reasonable bounds on estimated performance and does not improve the actual measurement certainty. There is no real substitute for enough data. A second consideration is the effect of systematics. I have found thata data run should be at least 5 times longer than the period of the slowest systematic effect. This has been born out for me while operating in labs with 20 minute air conditioner cycles.

PROCESSING THE DATA

The first and most important step is DON'T (Fig. 9). The raw data usually provides the most abundant clues to device performance. The best practise is to simply plot the raw data (suitably normalized and scaled) to look for systematic anomalies. The next useful step is visual correlation of frequency data and the various telemetry channels. The analog strip charts will have already plotted this. It turns out that the eye is one of the best detectors of correlation.

There is a distinct irony here. Processing normally is used to extract some distinct signal or signature by suppressing an overlying noise process. It seems that our task is to suppress the distinct signals in order to more clearly observe the noise process.

REPAIRING THE DATA

There is rarely an effective way to repair a data run corrupted by systematic effects. Even where there is clear evidence from telemetry of cause and effect, it is diffcult to cancel the effect. Temperature is a good example since its signature is usually very clear. Unfortunately, the driving function must often pass through various time constants and non-linearities before affecting the frequency. The most useful procedure is to make a second run with an exaggerated systematic to determine its detailed signature, and then attempt to mathematically extract it from the data. One fortunate feature of most processing of the Allan variance type is its tolerance of transient oddities given a large enough sample set.

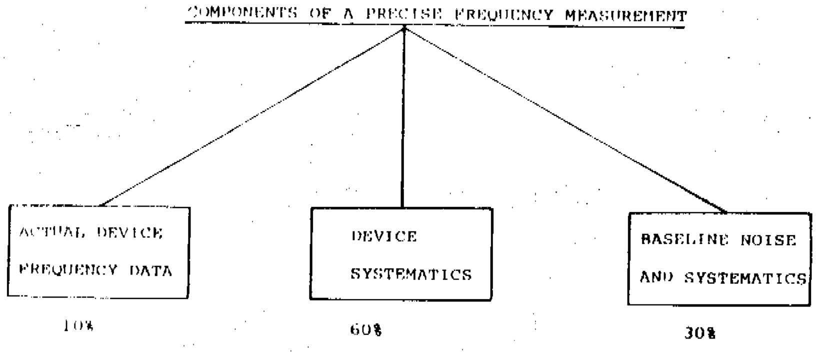
There are a few repairs possible when the effect is well determined. One HP counter which I have used would usually average 100 periods when set to do so. However, it would occasionally average 103 or 104. Since the noise process being examined was so small with respect to this counting error, it was possible to recompute what the original points must have actually been and sustitute these values into the data set. This does some theoretical violence to the continuity of the data, but since the occurence rate was low, the length of the data set absorbed the impact, leaving the final calculated Allan variance relatively unscathed.

SALVAGING THE DATA

There are three ways to proceed here. One is to just do a Sigma-Tau plot on the entire data run as it is. Systematics and residuals will appear as bumps and swellings in a plot which we would otherwise expect to to follow one of several straight-line power laws. We are performing an undesired spectrum analysis. The second is to use segments of the data run selected from areas where the systematics are known to be quiescent or at least stable. This is limited only by data length concerns discussed above. The last is a rather special case known as the Boston Pothole Method. Some highly periodic, but short duration systematics (we include measurement system disfunction) can put holes in the data. That is, the data may have up to 5% of its points destroyed, but with each incident (often a single point) flanked by good data. In this case, I feel comfortable filling each pothole with the average value of the adjacent points. I have not analyzed the effect of this practise, but experimenting with 100 point data sets has convinced me that the Allan variance is virtually unaffected.

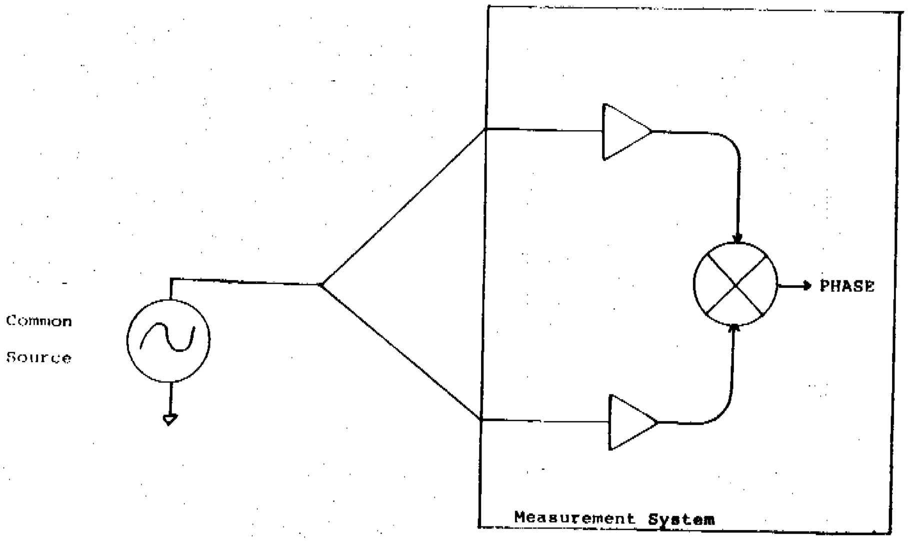
USING THE DATA

The best way to use data is to communicate it (Fig.10). While the contract deliverable may be a Sigma-Tau plot, sharing the data with systematics. and a good definition of the measurement set-up may be the best diagnostic of all. The PTTI community has an enormous historical data base which can prevent the expensive re-discovery of familiar effects.

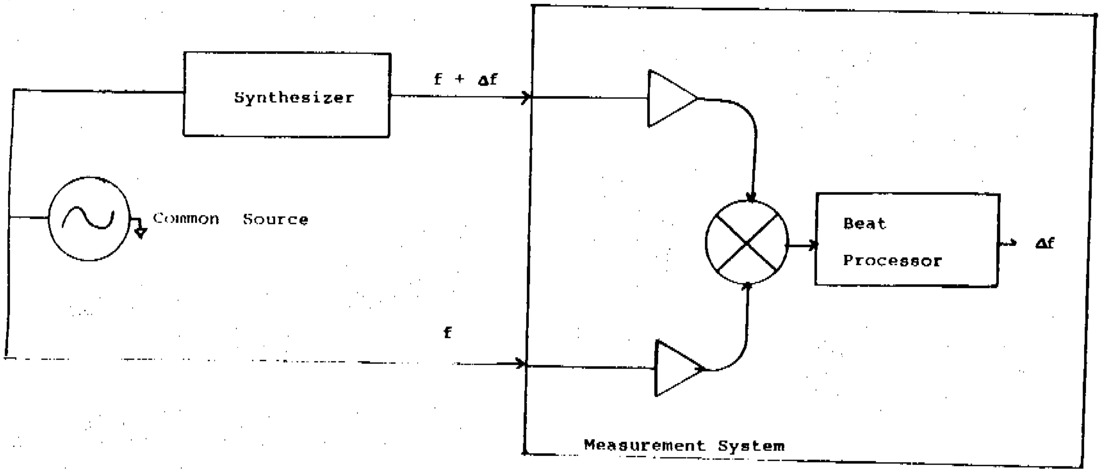


Cogent Design 1

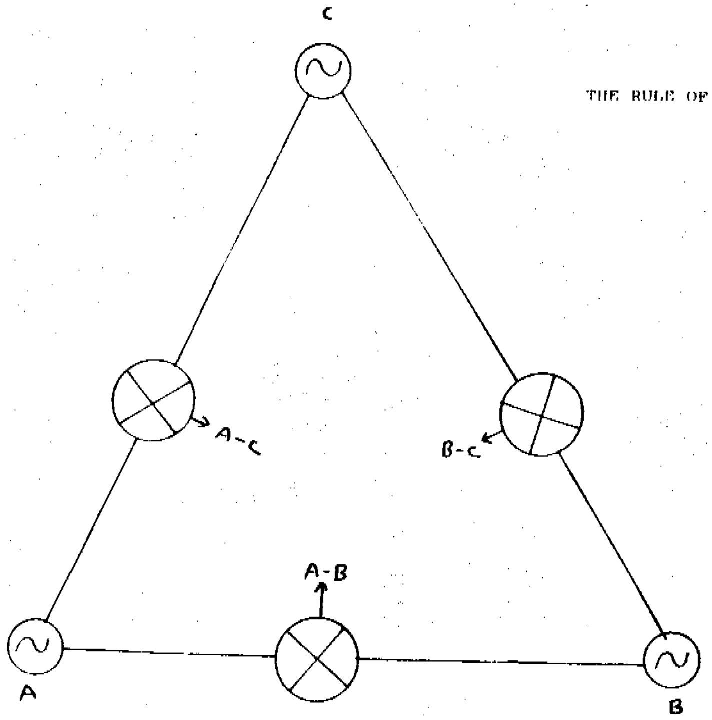
BASELINES == CLOSING THE LOOP



Cogent Design 2



Cogent Design 3



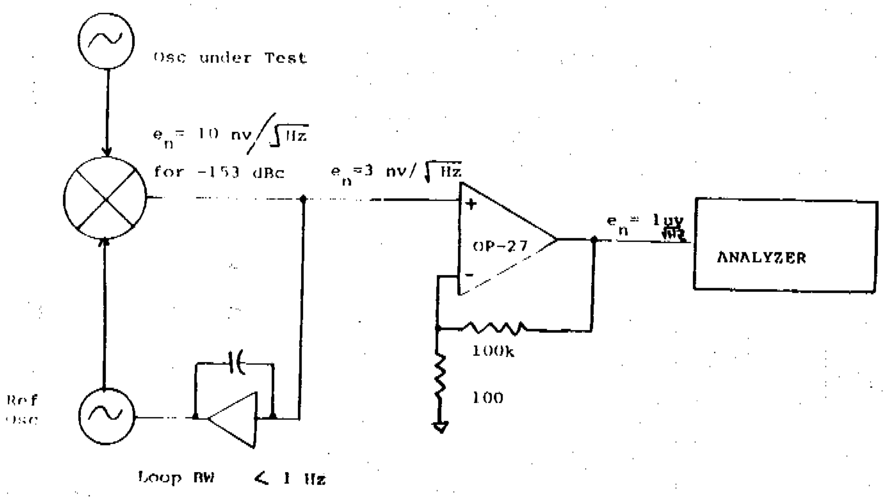
THE RULE OF THREE

Δ A will appear in A-C
and A - B but not in
B-C

Cogent Design 4

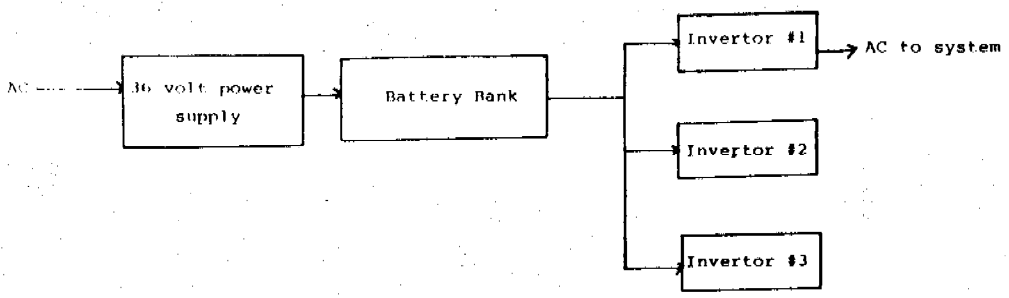
BASELINE EXAMPLE

$$L(f) = 20 \text{ Log} \left[\frac{\text{noise density at } f \text{ in volts}/\sqrt{\text{Hz}}}{\text{carrier beat in volts rms}} \right] - 3\text{dB}$$



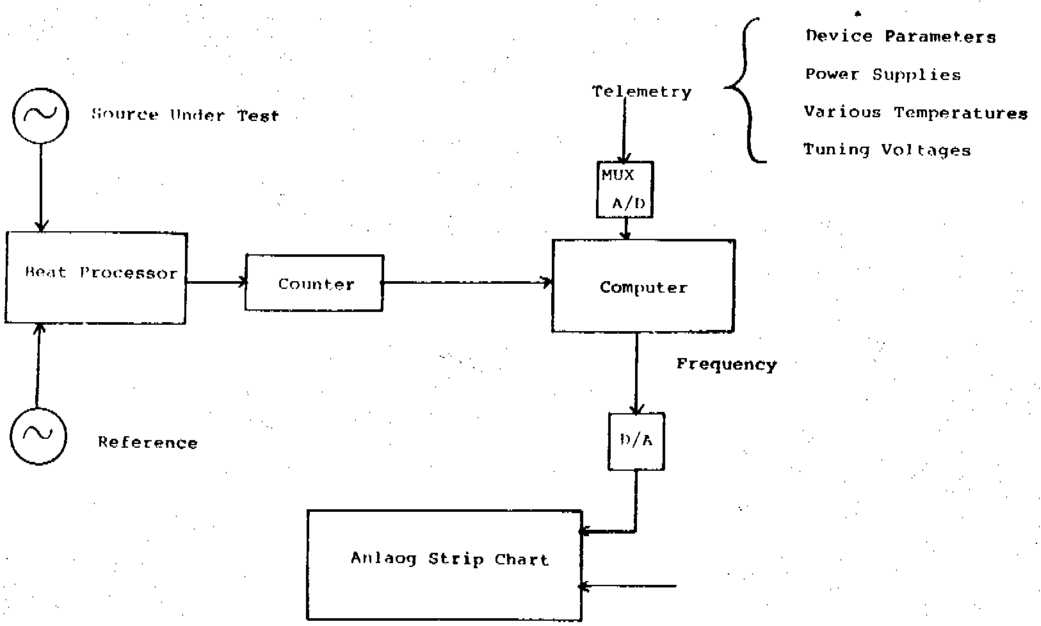
Cogent Design 5

BATTERY BACKUP



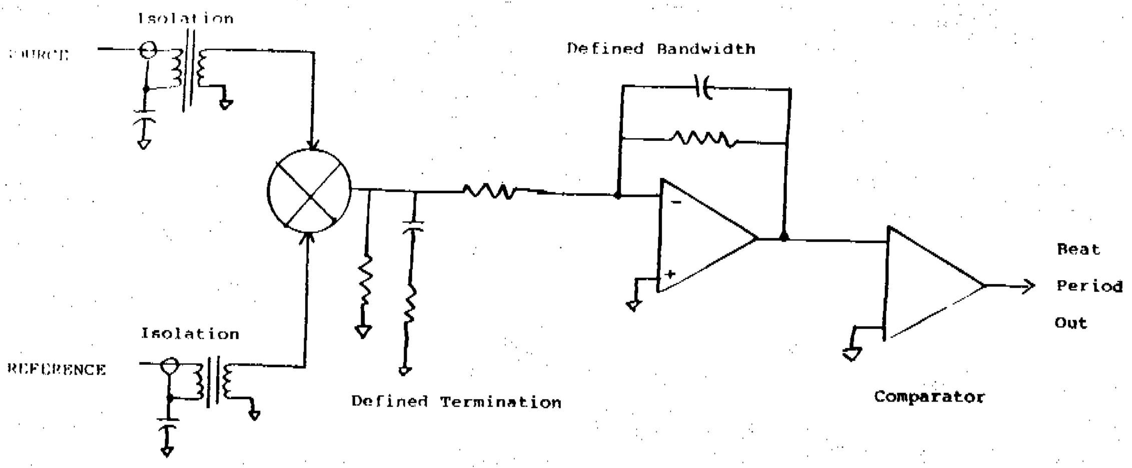
Cogent Design 6

TYPICAL MEASUREMENT SET-UP



Cogent Design 7

BEAT PROCESSOR



Cogent DEsign 8

DATA PROCESSING STEPS

1) DON T
2) Visual Inspection of Raw Data
3) Visual Correlation with Telemetry
4) Extraction of Systematics
5) Choosing Segments
6) Filling Pot Holes
7) Sigma Tau

Cogent Design 9

USING THE DATA

1) Communicate It

- -- Raw Data
- -- Systematics
- -- Baseline
- -- Measurement System Description

Take Advantage of the agglomerated intelligence of
The entire PTTI community

2) Deliver it

Cogent Design 10

QUESTIONS AND ANSWERS

DAVID ALLAN, NATIONAL BUREAU OF STANDARDS:
One other suggestion that is useful when you have systematics, and I fully agree with your concern about the importance of systematics, is that, if you have a periodic event, for example if that chart were night day air conditioning or whatever, then if you sample at the period of the event,. you can alias away that periodic function and look at long term stability and not be biased by that event, if you wish to look at system performance minus that systematic affect. We do this on GPS a lot by using the sidereal one day sample point with the same geometry and we alias away a lot of the other effects such as propagation problems that might be there. You can then look at the clock on board the space vehicle with much better accuracy of information than otherwise. It is also interesting along with Dr. Bloch's comment that with quartz oscillators that you worry about having human hands around. The same seems to be true of atomic clocks. The atomic clocks aboard the space vehicle seem to work much better than those down here where we can grab them.

MR. BLOMBERG:
I would just add that, in my view, the ability to correlate away, successfully, a systematic is useful in direct proportion to your exact knowledge of the systematic. If you are able to successfully remove that from the data, you can't lose because that implies that you must have done a good job of analysis in order to identify exactly what the systematic was, or its signature.