Introduction to dynamic and structural measurements for the data consumer[1]

William C. Shust, P.E.
Objective Engineers, Inc.
363 Venturi Drive
Pueblo West, CO 81007
(719) 547-2604

ABSTRACT

Engineering tests that do not impact business decisions waste time, budgets and morale. This paper will consider input and output information for the test process—specifically in vehicle dynamics and structural integrity labs. Uncontrolled test processes are unlikely to fulfill expectations. Conversely, the process may be closely defined in a widespread specification, and yet not answer the question at hand.

Potential downstream impacts of loosely defined technical details will be discussed. Topics include selection of transducers, recording parameters, filtering and analysis techniques. Physics of common environmental loading are presented, along with implications for both the test plan and later use of the test data.

INTRODUCTION

This paper compiles several years of explanations exchanged among and between test professionals and customers. It is not intended to replace the many excellent textbooks available [see References], but rather as a succinct review that one might offer to a busy engineer working outside the lab.

Dynamic and structural test labs have similar tests, analyses, and terminology across industries (defense, transport, utilities, etc.). However, the design engineer may have no formal training to help interpret such jargon: "How does 0.08 g*g/Hz tell me if my air conditioner bracket will break?" This paper will present several common forms of test results seen in many lab reports, as well as the fundamental information that a test customer should find therein.

BACKGROUND

Why The Dynamics Or Structural Lab Exists

Whether internal or external, the lab exists to fill gaps in knowledge about either the user environment or the product response. Either or both of these issues may lead to a test request as discussed below.

Sometimes the environment is well understood, but a design's performance in it is not. One subset of this case is the acceptance test—usually a predefined specification ("spec"). In this case, the lab exists to simulate an environment and verify acceptable operation of the design throughout. Unfortunately upon close examination, some industry-accepted specifications do not uniquely identify one and only one environment. Usually as long as the spec-authoring organization continues to perform all such tests, the methodology will remain constant and will produce comparable results over time. For design personnel, these tests-to-spec are usually the easiest to outsource.

Sometimes the design is well known (via finite element work or other rigorous analyses), but the environment is not. This may be due to new or unexpected use of an existing product (e.g. the screwdriver as an ice pick). In this case, the lab is asked to document the field stresses anticipated. Such a request should state whether median or worst-case users are to be examined. It is then the responsibility of the lab to measure the phenomena in question without altering it (either physically or via human behavior). In these situations, volunteer subjects can be given spurious reasons for a test while hidden sensors actually measure the phenomena of interest (e.g. typical door-slam forces).

Environment qualification tests are probably hardest for the design engineer to outsource. The oversight work will be at least twice, and perhaps 5 times more than a test-to-spec. In any case, the test requester should inform the lab regarding the underlying need for the data, and the form of data presentation required. It is not "overkill" to actually include dummy plots of the results desired, and the ramifications of results which exceed critical decision levels (pass, fail, redesign, redeploy). A brief editorial: test professionals do not remain in the lab for the vast fame and fortune, rather they derive satisfaction from answering questions that really matter. Testing for the sake of testing wastes time, money, and morale.

Review of Physics Between Frequency and the Transducer

The physical relationships between force, acceleration, velocity, displacement, strain, and noise often control the test frequencies of interest. Force is directly related to sum of the products of acceleration and mass, velocity and damping, and displacement and stiffness. For sinusoidal motion, each time domain derivative introduces a product of frequency ($2\pi f$):

$$\begin{aligned} x &= A\sin(2\pi ft) && \textit{displacement} \\ \dot{x} &= -(2\pi f)A\cos(2\pi ft) && \textit{velocity} \\ \ddot{x} &= (2\pi f)^2 A\sin(2\pi ft) && \textit{acceleration} \end{aligned}$$

Therefore, velocity transducers tend to emphasize high frequencies more than displacement measures, and accelerations even more yet.

For physical yielding, fatigue, or breakage, the best measure is often strain. Accelerometers are often the "default" motion sensor used because they are simple to install and rugged. However for problems related to fatigue, their high frequency emphasis may be a drawback, requiring additional lowpass filtering or integration. Since strain is a spatial derivative of displacement, not velocity or acceleration, it naturally emphasizes lower frequency data [2]. In this manner, the measurand itself can be a form of filter.

END-USER'S GUIDE TO 5 LAB MEASUREMENTS

This section is intended to guide the non-test professional through the most common graphical results obtained from a structural or dynamics lab. The sample data plots herein are sometimes treated as required but irrelevant content in a test report. That is, as long as the plots exist, the (sometimes preconceived) conclusions are believed—regardless of any particular plot's shape. With the limited review of fundamentals herein, it is hoped that the test consumer will do more than check for their existence.

All such plots reflect measurement of a physical quantity, generically known as an "Engineering Unit" or EU. Examples in structural or dynamics labs are g's, m/s, m, MPa, and με (microstrain). Axes types for the most common plots reported by test labs are listed in Table 1. These will be further discussed in three broad categories based on how they represent the data: chronologically (often called the time domain), by wavelengths found therein (the frequency domain), or by relative size of events (the amplitude domain).

Table 1. Common test laboratory data plots.

VERT. AXIS	HORIZ. AXIS	COMMON NAMES
EU Amplitude	Time	Time History, Peak/Valley History
EU Amplitude	Frequency	Spectrum, PSD, FFT, Waterfall Plot
Occurrences	EU Amplitude	Histogram, PDF, CDF
Occurrences	Rainflow Range	Rainflow Matrix, Damage Matrix
(Output/Input)	Frequency	FRF, Transfer Function
Correlation	Frequency	Coherence

I. TIME DOMAIN REPRESENTATIONS

Time History

This is typically the most basic plot of all, and may therefore be omitted in many test reports. It should not be. It's interpretation is the most intuitive, because humans act and react in this time domain. Furthermore, this simple format allows the most basic comparison with other data acquired at other times (and perhaps with other recording parameters). This may be the only plot that "speaks" to the customer. However simple, the test customers should pose several questions to themselves when examining time-domain data:

- Is this data within the range of the data collection system? (Including too small, or too large to be adequately measured)
- Can I visually see dominant frequencies and/or DC (steady state) levels?
- Is an inappropriate sampling rate or filtering process hiding or distorting important effects?
- Does the time history contradict observations from more complex plots?

It is assumed throughout that competent test professionals have reported results to the test consumer. Even so, the data customer should interpret time histories with one eye on the collection parameters (often listed in a table elsewhere in the report).

Data consumers should remain wary whenever a physical event amplitude was large relative to the allowable full scale values. It is possible that the raw input over-ranged the measurement system (electronically or physically), but waveforms erroneously look reasonable after a low pass filter. In fact some accelerometers are advertised for their "built in electronic filter to prevent the potential for high frequency overload." Filtered or not, high frequency clipping can and will distort the desired lower frequency measurement [3].

Crude Frequency Content Observations

After mentally noting the overall amplitude suitability as defined above, the measurement customer should again look at the time history for rough patterns of frequency content. Zero frequency (DC) behavior should be consistent with one's preconceived reference frame. That is, should a non-zero mean value exist throughout the time history? In the case of a strain measurement, the mean value can indicate: zero-shift either due to a change in the physical boundary conditions (perhaps additional lading after instrumenting the vehicle), a loose gage, test component yield, or a wiring problem. A mean value from an accelerometer should indicate net overall acceleration or deceleration of the vehicle (including perhaps centripetal acceleration found during a long curve), or it may indicate that the datum value of earth's gravity being used is 1g (rather than 0g). Regardless, the mean value should make sense to the end user, not left to unspoken convention.

Dominant frequency components can often be seen in the time histories, and can be estimated by counting peaks or valleys. For example, engine idle vibrations on a highway tractor should show highly periodic data. Inflection points and/or zero crossings can be counted within a known period to relate the data to the physical system. In this case, the operating rpm of the 4-cycle internal combustion engine will reflect the 1st and 3rd orders of the engine firing frequency. For strongly periodic content, such cycle counting methods (employed across 5 to 40 wave periods) provides frequency identification that is quicker and at least as accurate as more complicated methods.

The end user should be suspect of components at 60 Hz (AC line frequency), harmonics and sub-harmonics (30, 120, 180, 240 Hz). Finally after the aforementioned rough correlations, the data customer is prepared to examine a "strange looking" event elsewhere in the data, or to move on to more elaborate analyses.

II. FREQUENCY DOMAIN REPRESENTATIONS

Human hearing functions as an excellent frequency analyzer. The ear processes high versus low frequencies in different regions of the cochlear membrane, and we can all discern multiple tones simultaneously. Several types of frequency domain plots, or spectra, are used to provide similar information regarding engineering measurements and will discussed in this section.

The Fast Fourier Transform (FFT) as Useful Shorthand

Mathematicians figured out hundreds of years ago that sequences of numbers could be approximated in shorthand by solving for a series of coefficients multiplied by certain mutually independent functions. Any group of such functions is called a basis set. For common high-school curvefits, one typical basis set is the polynomial expansion:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_n x^n$$

Therefore if one is willing to prescribe a level of approximation, a large set of numbers can be replaced by a smaller set of coefficients and the known basis set. For many physical problems, such a substitution also provides insight as to the underlying behavior.

In a similar fashion, any finite time history can be replaced by a basis set of sinusoidal waves starting with the lowest possible frequency (i.e. one cycle across the entire time record), up to as many cycles as the sampling permits (one-half the sampling rate, per Nyquist). The family of possible sinusoids are commonly known as spectral components, or spectral lines. These spectral lines are equally spaced by a frequency increment equal to the inverse of the time segment length to be converted (known as Δf, the frequency resolution).

Many technical explanations of the FFT conversion process exist [4], but a simple question will remind the novice of certain limitations caused by the inherent theory used:

- **"In order to approximate this particular segment of a time history, what combination of steady-amplitude sinusoidal waves could be simultaneously added together?"**

To illustrate the results, Figure 1 is the answer for this question (i.e. the FFT) of a 328 second time history. The first few seconds of the time signal are shown in Figure 2. The time history is noticeably periodic, but also contains smaller random features. The FFT peaks clearly show it to be dominated by four sinusoidal components at 2, 4, 8, and 12 Hz.

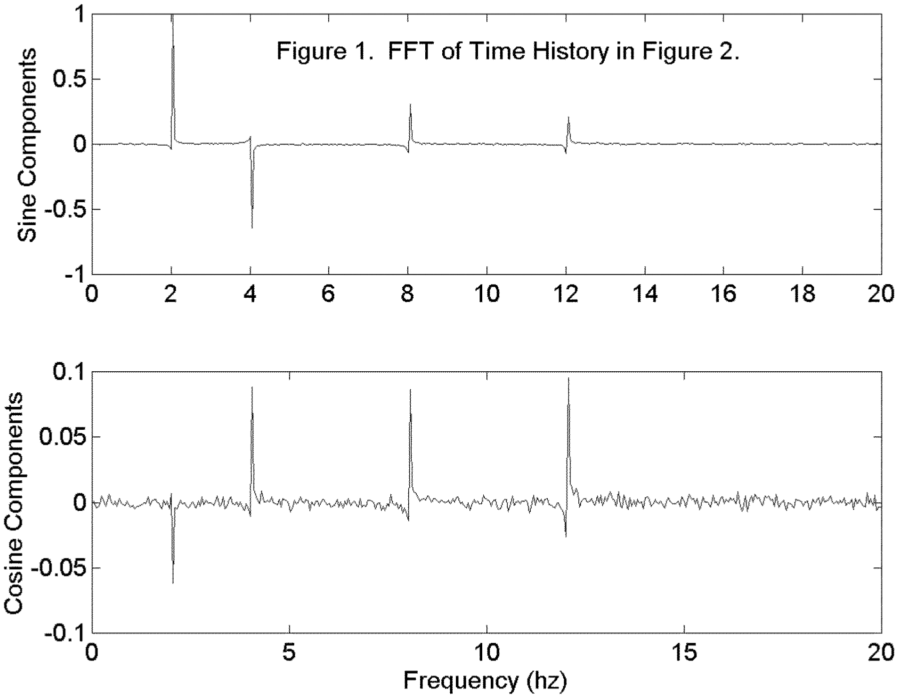


Figure 1. FFT of Time History in Figure 2.

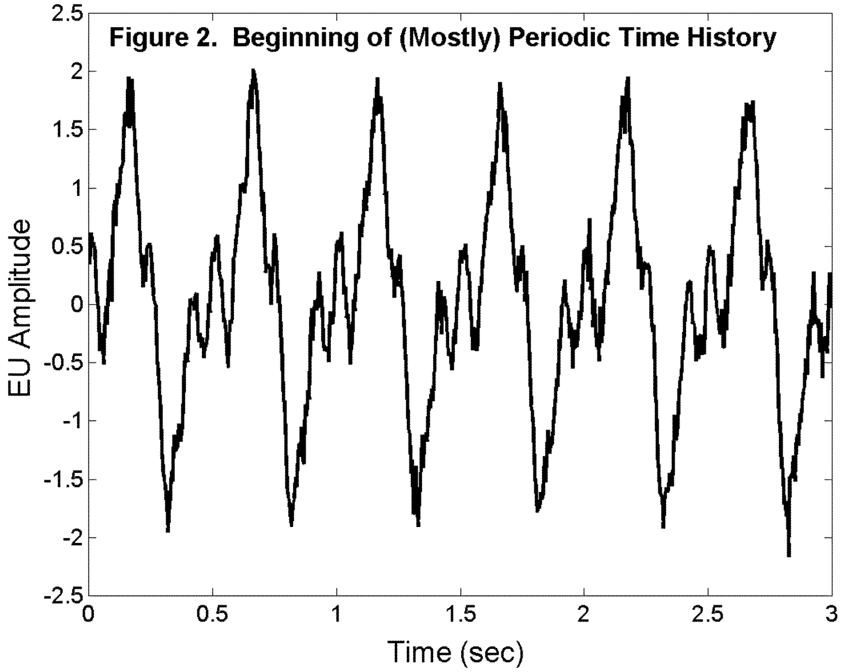


Figure 2. Beginning of (Mostly) Periodic Time History

The FFT format used in Figure 1 plots two sets of coefficients, one set for sine waves and one set for cosines. However, since summing various sine and cosine waves of the same frequency merely serves to shift the wave left or right (forward or backward in time), this data can also be shown more intuitively as Figure 3.

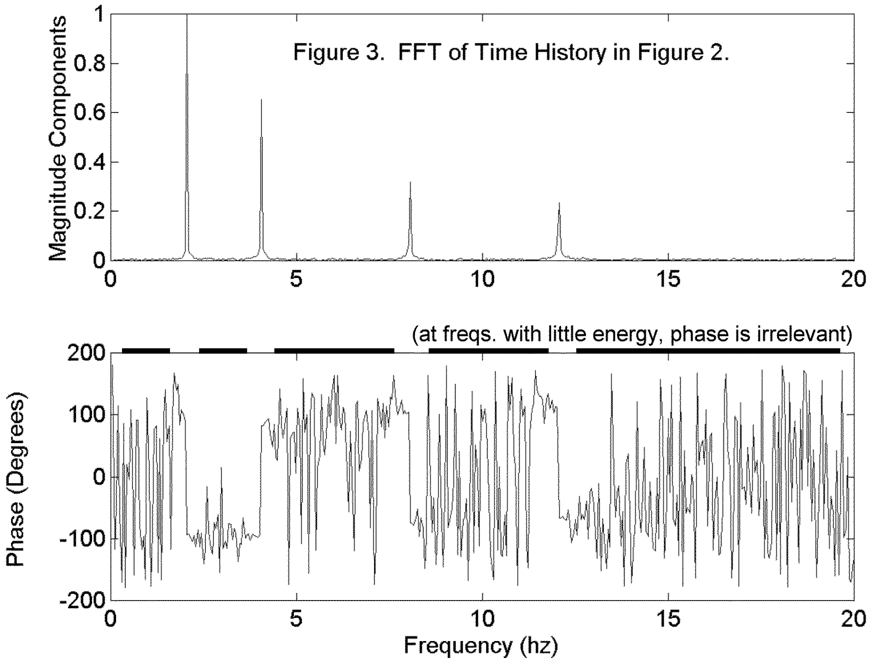


Figure 3. FFT of Time History in Figure 2.

(at freqs. with little energy, phase is irrelevant)

This is the same FFT shown in Figure 1, only presented differently. The format in Figure 3 shows the resultant magnitude of each sinusoid pair (square root of the sum of squares for both sine and cosine component), and a phase term (to represent sliding the resultant wave left or right relative to components at other frequencies).

An FFT transforms the original data from the time domain into the frequency domain. When examining any one record of data, this transformation can be done and undone again, returning to the original time history. However, transforming a lengthy segment of time data in one FFT operation does not exploit its potentially random behavior [5]. At this point it is useful to ask another question:

- **"During several independent inspections of this signal, do particular frequencies repeatedly account for much of its energy?"**

The answer is obtained by breaking the input data into many independent (successive) time records, asking the sinusoidal approximation question for each segment (perform a FFT), and averaging the resulting magnitude coefficients. After averaging, the results are commonly known as an autospectrum or power spectrum. If the results are further divided by the frequency resolution (merely a constant), a normalized Power Spectral Density (PSD) is obtained.

Figure 4 shows eight overlaid sets of magnitude coefficients for the time data in Figure 2. Initially the full time history was segmented into eight records of 41 seconds each. Then the question was asked "What are the amplitudes of sinusoidal waves that can be summed to approximate the time data between 0 and 41 seconds? Between 42 and 82 seconds? And so on for each of the eight records.

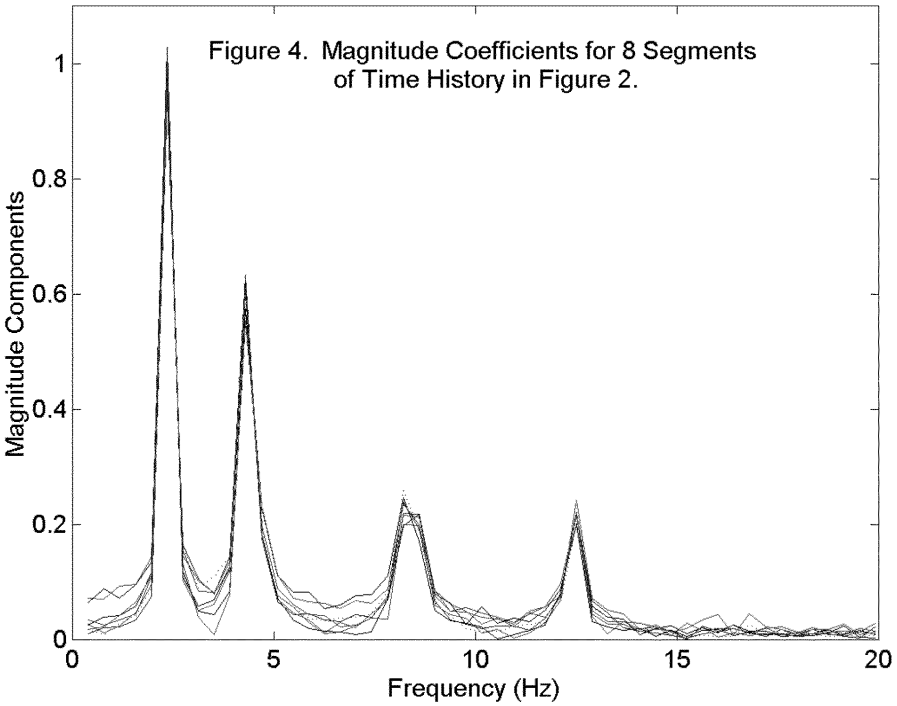


Figure 4. Magnitude Coefficients for 8 Segments of Time History in Figure 2.

Note that if the system is changing throughout the time history (e.g. vehicle speed increasing from 10 to 50 mph), then one would not expect the amplitudes of the sine waves to remain constant from record to record. Also, if the vehicle moved from a paved surface to a gravel surface partway through the data, one would not expect the sinusoidal amplitudes to remain constant from record to record. These situations violate underlying assumptions of Fourier theory requiring consistent input data (referred to as being stationary and ergodic [6]). Therefore, such segmentation and averaging should only be done with a valid understanding of the nature of the environment recorded, and limitations of Fourier theory.

The PSD is the most common single-channel analysis performed on random data. However because a PSD doesn't retain the desired phase information between the various frequency components, it's not best for periodic data. (e.g. Consider a PSD created from a square wave; it will show only the fundamental frequency plus odd harmonics. This information is not enough to later rebuild the proper time waveform because the critical alignment—or phasing—between harmonics is not kept by the PSD.) Therefore, the "one-shot" FFT analysis is more useful for periodic data.

In modern practice, the summation-of-sinusoids question is answered mathematically with the FFT algorithm. However the question could also be answered by looking at the original time record through many narrow-band filters. If each filter was chosen to pass only a specific spectral frequency, then each average squared output would equal a particular magnitude coefficient.

Whether achieved with the FFT, or with RMS measurements of narrow filter outputs, the frequency domain conversion is an exercise in estimating mean amplitudes over time. Therefore, the quality of the estimate will increase as lengthier data are considered. In other words, the magnitude of each sinusoidal component will vary from time segment to time segment, and will approach the true average value (i.e. the average found over the entire service life) only as larger and larger amounts of data are recorded. For a true random stationary ergodic process, the distribution of PSD variance is known [7]. The resulting random error is inversely proportional to the square root of the number of records considered. Therefore to reduce the random error associated with each spectral component by 1/2, four times more data must be collected.

What does all this mean to the end user? First it means that any plot of EU versus frequency is not a property of the system, but an approximation:

1) The FFT shows amplitudes of, and phase shifts between, steady-state sine and cosine waves that could be added together to build one data segment similar to the input record.

2) A Power (Auto) Spectrum or Power Spectral Density averages multiple FFTs together and shows an estimate of the average sinusoidal magnitude needed at this frequency to approximate energy in the input records.

When inspecting either plot type, if some frequency components are much larger than others, they may indicate a structural resonance (e.g. fundamental bending mode of a chassis) to random inputs, or they may reflect periodic input energy (e.g. engine firing frequency, or regularly spaced pavement joints), but without a resonant response. In the least desired situation, a resonance aligns with a periodic input.

Several additional comments regarding FFT and PSD (Power Spectral Density) plots have been included here for the interested reader.

Common PSD Amplitude Units

Energy spectra may have amplitude units of EU, EU RMS, EU^2, and EU^2/Hz. The EU and EU RMS units are equivalent. Conversion to units-squared is intuitive. However, to the both test and non-test professionals, the EU^2/Hz fraction can be troublesome. The popularity of this fraction can be illustrated by considering sending an input signal through a bank of narrow-band filters. For example, if a given analysis was performed using narrow filters at 5 Hz spacing (0-5, 5-10, 10-15 Hz, etc.) and another project used filters at 10 Hz spacing (0-10, 10-20, 20-30, etc.), the amplitudes of the sinusoidal components would not be equal even if the random input data were exactly the same. The wider filters would result in larger output energy. Therefore dividing the reported amplitude by the filter width used normalizes the data and allows amplitude comparisons from report to report (or lab to lab). Note: since a periodic signal could be passed equally by two such filter widths however, this normalization process is not suitable for sinusoidal energy.

Windowing Effects

Most segmented time data will not start and finish at exactly the same amplitude. Consequently the FFT algorithm will "see" a discontinuous step in the input data, erroneously adding to the "noise floor" of the PSD (smearing energy across the spectrum). This can be minimized by artificially tapering the data at the beginning and end of the segment, known as windowing. Various shaped windows are commonly used (Hamming, Hanning, Exponential), which remove different amounts of energy from the input data. PSD algorithms may or may not rescale the resulting spectrum in an attempt to reverse this energy removal. When comparing spectra obtained at different times and places, the data consumer must verify the similarity of windowing algorithms.

Amplitude of Strong Periodic Components

If a pure (or very strong) sinusoidal component is present in the signal, it will likely not align with the center frequencies of available spectral lines. Therefore the energy of such a component will be split between the two nearest frequency lines, but neither will reflect the correct amplitude seen in the time domain. (e.g. a one-volt pure tone at 6 Hz, analyzed with 5 Hz frequency resolution will be reflected in both the 5 Hz and 10 Hz spectral lines, but neither will show 1 g). With some care, use of a flat-top window will "smear" the correct amplitude across more than one reported frequency. However, this should be performed with a Power Spectrum, not a Power Spectral Density.

Mean Value (DC) Effects

A strong DC component in a signal is likely to distort the lowest spectral lines in a PSD. Since strain gage data can have valid, but large, mean values (perhaps due to payload changes in vehicles after transducer zeroing operations), this mean value should be subtracted from the signal prior to an FFT analysis.

Overall Energy in the Signal

By summing amplitudes across the entire PSD and multiplying by the frequency resolution, one obtains overall EU-squared. This should match the mean square value of the original signal. If these do not agree, unaccounted effects such as energy removal via windowing are likely to be present in the signal processing algorithm.

Non-Stationary Conditions

Although a violation of underlying assumptions, non-stationary conditions are regularly examined with FFT analyses. Recording a third variable for each record transformed can help these. The third variable should reflect the changing operating condition, such as engine speed. Other possible variables might be time of collection, vehicle speed, gear ratio, or temperature. Slightly offset overlays of each successive spectrum are often called waterfall plots. The most common use is in rotating machinery studies, because input energy related to rpm will move across the frequency axis with subsequent spectra. However, invariant structural responses (natural frequencies) will remain steady at particular frequencies.

OUTPUT TO INPUT ENERGY RATIOS—THE FREQUENCY RESPONSE FUNCTION (FRF)

Although less common than single-channel frequency analyses, more insight can often be gleaned from comparisons of multichannel data. This paper will concentrate on simple two-channel relationships, such as a component's vibration resulting from a measured input force. However, modern signal processing packages can consider many response and reference channels simultaneously.

Figure 5 shows the linear system typically drawn when describing a Frequency Response Function (FRF, less accurately called a transfer function). Herein the word "linear" roughly means two conditions are met:

- A single-frequency input causes an output at (only) the same frequency.
- For any given excitation frequency, doubling the input amplitude will double the output amplitude.

Simple monolithic metal structures approach these conditions very closely. Real systems (with fasteners, intermittent gap closures, or elastomeric parts) will deviate more or less from linearity. This does not mean that multichannel FRF analyses cannot be performed, however the results lose validity as the nonlinearity increases.

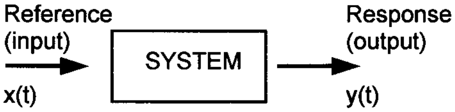


Figure 5. Block Diagram of Linear Time-Invariant System

Also, although commonly interpreted as relating the frequency content of a system output to a system input, the math can be computed for any two simultaneous streams of data. Therefore causality is not verified by FRF analysis, rather merely linear correlation between two channels (or lack thereof).

The single channel FFT techniques are employed as building blocks for dual channel analyses. However in the two-channel case, a robust analysis will include three quantities estimated at all frequencies of interest:

1) Magnitude ratios (How much response for a given reference data channel?)

2) Channel-to-channel phase (When does a sinusoidal component in the response channel peak, relative to its corresponding sinusoid in the reference channel)

3) Coherence (How well can a sinusoidal component in the response be predicted, given the reference channel?)

Items 1 and 2, the magnitude and phase characteristics may be displayed alternatively in complex notation as real and imaginary components, however this is less helpful for the typical test requester.

As stated above, FRFs can be computed for two data streams other than an energy source and its resulting effect. The two channels used may both be responses to a third factor. Again a question can highlight the Frequency Response Function (FRF) without implying causality:

- **"Can I estimate the size (magnitude) and position in time (phase) of the frequency components within one data channel, given another data channel?"**

The coherence calculation then shows the expected quality of such a prediction. With coherence near zero, the two channels cannot be related using a linear model. Conversely, coherence equal to 1 indicates perfect prediction of the response channel, given the reference channel. Note that both of these numerical results can and do occur simultaneously, but at different frequencies.

For example, skin temperature while sunbathing is highly coherent with the sun's brightness as clouds come and go throughout an afternoon (higher frequency). The magnitude units here would be (degrees Celsius/lumen) or its reciprocal, depending on which channel was chosen as reference. But the same temperature/brightness relationship does not hold from day to day or season to season (i.e. low coherence at much lower frequencies). This example also illustrates that neither channel (temperature or photon level) causes the other, and that one or both channels can be contaminated by "noise" such as an intermittent breeze in this case. Here, noise is any undesired and often uncontrolled influence on the data.

Similarly accelerometers fixed to a truck's axle and body are likely to be highly correlated at lower frequencies (1-15 Hz), but are likely not coherent at high frequencies (above 30 Hz), due to the suspension isolation between them.

Magnitude Characteristic of the FRF

At each FRF frequency line, the magnitude portion estimates what amplitude response sinusoid will be found as a ratio of the reference channel amplitude (for a given frequency). Sometimes the magnitude characteristic is called the gain.

As with single channel frequency analyses discussed earlier, each element in this magnitude array will have statistical errors associated with it. The random error standard deviation will again decrease as the number of time segments considered increases.

Regarding plot appearance for well-behaved (reasonably linear) mechanical systems, the FRF magnitude should not jump greatly at adjacent frequency lines. If the FRF magnitude appears very "hashy" it likely indicates that too few time segments were used, or that no significant linear relationship exists between the two channels.

When a causal system response is already known to exist, the FRF magnitude can be considered to show "how much out per unit in." However during nonlinear behavior (e.g. a rattling connection), a single frequency sinusoidal force will create many acceleration frequencies as output. At these new frequencies, the denominator of (out/in) will be zero, and the FRF will be invalid.

Phase Characteristic of the FRF

As with the FRF magnitude, the phase characteristic of a FRF relates two sinusoidal components found in the data at each frequency considered. The phase shows when a response channel sinusoid peaks relative to the reference channel. If both reference and response sinusoids peak at the same time, the phase between them would be zero. If one peaked exactly when the other was minimum (valley) the phase would be 180 degrees (or pi radians). Sometimes the phase is referred to as "lag." However, since the underlying FFT assumption is that time-varying wave segment can be built from a sum of many constant amplitude sinusoids, again the FRF does not prove causality or time delay. Rather it can merely estimate when a frequency component in one channel will peak, given knowledge of the other. As such, the FRF phase can be reported in various equivalent ways: commonly from 0 to 360 degrees, and from –180 to 180 degrees.

Regarding plot appearance, again many mechanical systems will prove to have phase relationships that do not vary greatly over nearby frequency lines. If the phase varies wildly from spectral line to spectral line, this indicates too few time samples or no correlation between the channels at hand (e.g. out/in denominator near zero). When a causal relationship is known to exist, the phase characteristic can be considered to show lag of output relative to input.

Coherence Function

Some professionals in the signal processing community feel that coherence should always be included as a separate third trace on FRF plots [8]. Coherence relates the existence and quality of gain and phase relationships between two channels. It is computed by comparing whether the gain and phase vary from time segment to time segment. If only one time segment is analyzed, the coherence will be reported as exactly 1.0 everywhere, but this is meaningless (similar to the average of a population of 1). If consistent relationships are found in all segments, the coherence will be 1. If no correlation is found because the relationships vary unpredictably, the coherence will be zero. Between these extremes the coherence indicates:

1) outside physical influences on one or both sensors

2) noise on one or both instrumentation channels

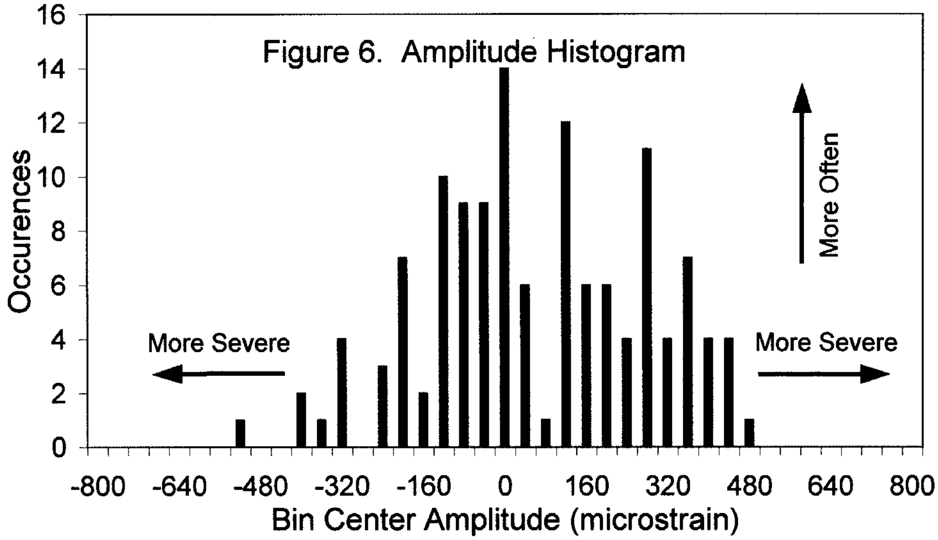
3) changing or nonlinear system behavior

Depending on which of these effects occur (perhaps documented via further test work), the data consumer may find use in the FRF data even if it shows low coherence. Examination of coherence between multiple channel pairs may help locate causal relationships.

III. AMPLITUDE DOMAIN REPRESENTATIONS

Some engineering questions cannot be answered by finding dominant frequency components. Perhaps the test lab is asked to measure the severity of events, and how many are likely in a typical service life. In such cases, a test consumer is likely to encounter an amplitude histogram, as in Figure 6. Possible amplitudes of the original signal are plotted as the horizontal axis and the vertical axis reflects a tally of events found within these certain amplitude categories.

To create a histogram one must first define the number of total amplitude bins, or the width of each. Wide bins usually result in many counts per bin but few possible amplitude categories, while narrow bins can result in only a few counts but many categories. This is a tradeoff between accuracy in the quantity of counts versus the precision of amplitudes discriminated. The test requester should indicate the degree of acceptable bin coarseness to be reported for the problem at hand. Note that when evaluating histogram data, engineers usually assume all events tallied within a bin occurred at the bin's center value. For wide bins, this can greatly underestimate or overestimate the amplitude of any particular event in the bin.



Again as with the frequency component estimates, more data allows greater confidence in extrapolating results from a limited recording to the larger (but impractical to measure) entire service life. The ratio of counts in each bin to the total counts in all bins is an estimate of the percentage of service life experiencing this severity. Some organizations multiply all reported amplitudes by a Load Reserve Factor in the range of 1.1 to 1.5 to account for the uncertainty of such extrapolations.

Again, what does this mean to the end user? First, the amplitude categories must be narrow enough to discern a significant difference between histograms measured on different machines. (What constitutes a "significant difference" is often a matter of contention, but in the authors' experiences with road data, reported amplitude differences must exceed 10% as a significance threshold.) Second, any two histograms must have enough underlying data such that the random error for each estimate is acceptably less than the population difference being sought.

The stair-step nature of histograms will diminish as the number of total data points is increased, and as bin sizes are narrowed further (eventually limited by analog-to-digital converter resolution). In the limit, the histogram becomes a Probability Density Function (PDF), or if one-sided exceedences are tallied, a Cumulative Density Function (CDF). These functions are easily compared from test to test, because as density functions they represent population percentiles.

SPECIAL HISTOGRAMS FOR FATIGUE ANALYSES ONLY (RAINFLOW)

Unlike the previous simple histogram of time history amplitudes, fatigue engineers are interested in variation in amplitude (often between strain peaks and strain valleys that are not adjacent to each other). Therefore, rainflow count results are not intuitive to most engineers. Although the details of the procedure are outside the scope of this document [9], typical tabular and graphical rainflow results will be explained.

Fatigue damage accumulates due to energy dissipated in completed stress/strain (hysteresis) loops. The pairing of proper peaks and valleys is commonly known as depopulation of a time sequence, and yields a histogram of rainflow cyclic strain amplitudes. We will examine tabular and graphical representation of rainflow results. Although outside the scope of this paper, a rainflow histogram is usually multiplied by material fatigue properties, resulting in a related but different "damage" histogram.

Figure 7 shows a short time history of four cycles, arranged to simplify cycle counting for the purposes here. This history has one loop each with ranges of 110, 220, 330, and 440 microstrain, and mean (equal to the peak minus the valley) values of 55, 110, 165, and 220 microstrain, respectively.

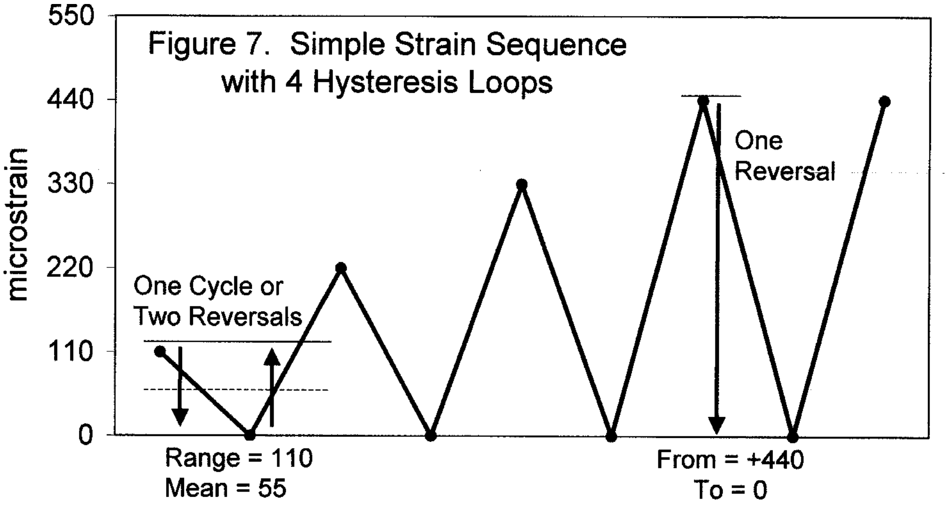


Figure 7. Simple Strain Sequence with 4 Hysteresis Loops

This will be used to show the three common types of rainflow data representation: range only, range-mean, and to-from. The most simplistic tally is range only. This analysis counts hysteresis loops in terms of their overall range only, and is shown in Table 2. The range-mean format carries additional data because each loop's mean value is saved as well, as shown in Table 3. This would be needed if using a fatigue model that considers the effects of mean stress. A range-mean matrix can be used to reconstruct a range-only matrix.

Table 2. Range-Only Tally of Four Simple Hysteresis Loops (Number of Loops per Bin)

STRAIN RANGE	COUNTS
500-599 με	0
400-499	1
300-399	1
200-299	1
100-199	1
0-99	0

Table 3. Range-Mean Rainflow Tally of Four Simple Hysteresis Loops (Number of Loops per Bin)

	MEAN 0-99 με	MEAN 100-199	MEAN 200-299
RANGE 400-499 με	0	0	1
RANGE 300-399	0	1	0
RANGE 200-299	0	1	0
RANGE 100-199	1	0	0
RANGE 0-99	0	0	0

The third format is to-from, which counts half-loops, called reversals. Each reversal is examined for its starting amplitude (the "from" value), and its ending amplitude (the "to" value), as shown in Table 4. Thus, a to-from matrix carries slightly more information than range-mean, and can always be used to reconstruct a range-mean matrix.

Note that the first two formats could also be plotted as reversals instead of cycles by doubling the loop counts. The data customer is cautioned that axes labels of "counts" and "amplitude" are ambiguous unless the report defines these terms further. It should be clear whether counts indicate complete loops or reversals, whether amplitude is peak-to-peak (range), or zero-to-peak, and whether reported labels indicate bin center values or upper or lower bin limits.

Table 4. To-From Rainflow Tally of Four Simple Hysteresis Loops, Number of Reversals per Bin

	TO 0-99 με	TO 100-199	TO 200-299	TO 300-399	TO 400-499
FROM 400-499 με	1	0	0	0	0
FROM 300-399	1	0	0	0	0
FROM 200-299	1	0	0	0	0
FROM 100-199	1	0	0	0	0
FROM 0-99	0	0	1	1	2

Each of the counting formats can also be displayed graphically. Figures 8 through 10 show the three plot types after rainflow counting several minutes of a typical strain history.

With regard to common engineering materials, rainflow counted strain signals accurately reflect the environment for most fatigue purposes. Unlike PSD theory, rainflow use does not assume stationary or ergodic data. Less accurately, some test specifications request particular PSD shapes to simulate service during shaker tests. Classic PSD control algorithms randomize the phase between frequency components, and will create a stationary and ergodic environment (even if the field conditions were not). Thus a particular PSD shape does not uniquely identify an environment's damage potential. If such lab conditions must be used, the resulting damage can be counted (for a given input and controller) using strain gages and a damage algorithm. However the resulting fatigue environment will be particular to the simulation at hand.

Figure 8. Range-Only Rainflow Histogram

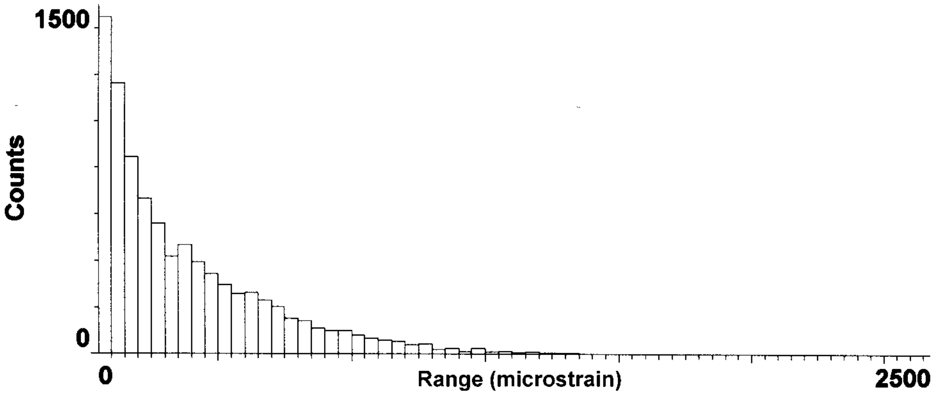
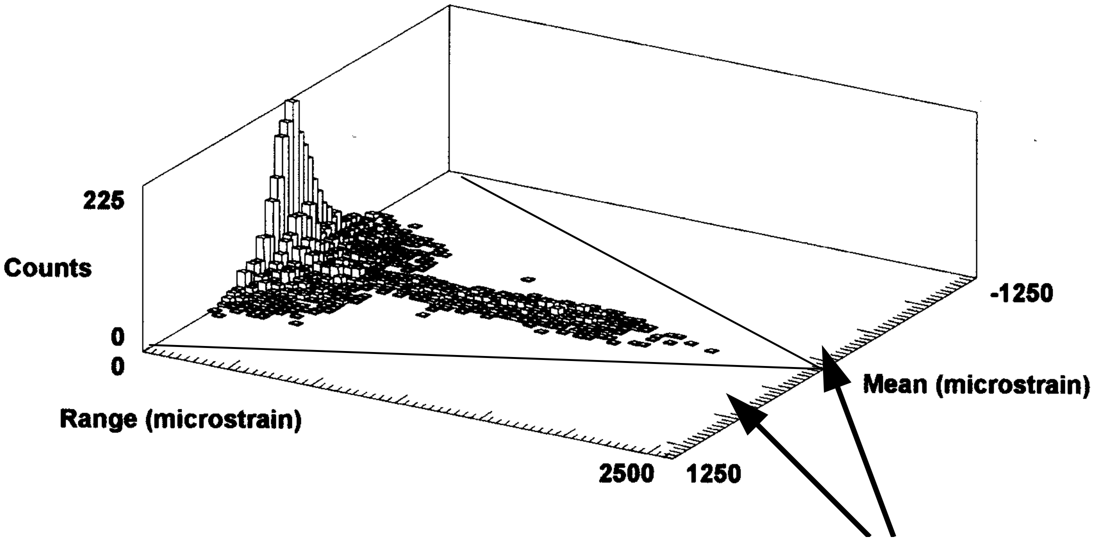
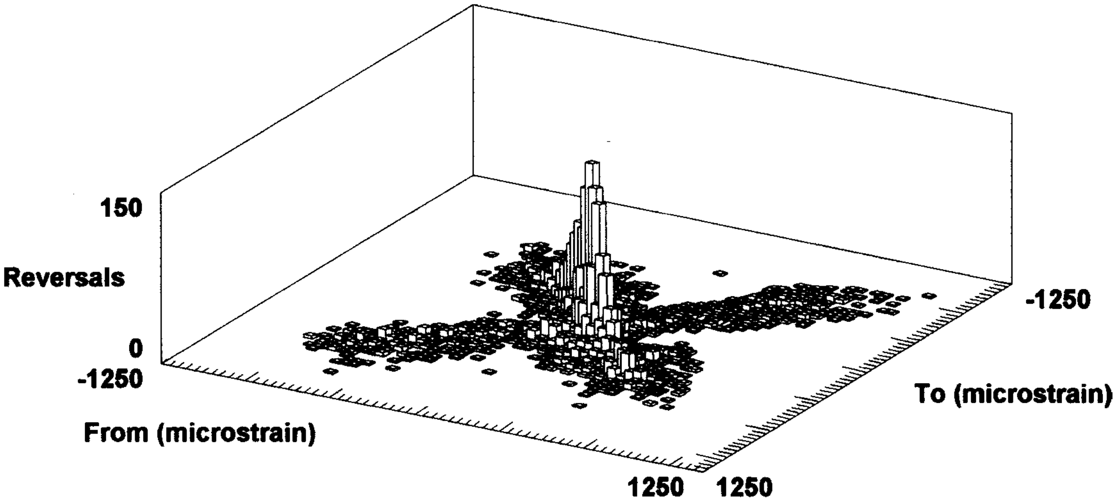


Figure 9. Range-Mean Rainflow Histogram



Triangular Offscale Areas
(will never contain data)

Figure 10. To-From Rainflow Histogram



CONCLUSION

This paper has discussed common structural and dynamic test results for the internal and external lab customer. Case studies have shown that a good test does not merely provide numbers, but converts engineering data into business decisions.

More technically-rigorous explanations are available in many excellent texts regarding signal processing and mechanical testing. However, it is hoped that this review is more accessible to the non-test professional. Furthermore, I hope that it provides added value due to a combination of multiple engineering viewpoints (signal processing, electrical and mechanical measurements, and material fatigue issues) in one end-user reference.

REFERENCES

1. Stein, P., “Experiment No. 6, The Response of Transducers to their Environment. The Problem of Signal and Noise,” *Lf/MSE Publ. No. 38*, Fall 1971, Stein Measurement Services, Phoenix, AZ, , p.14.
2. Vold, H., Morrow, D., “Compression of Time Histories Used for Component Fatigue Evaluation,” SAE Paper 930403, Society of Automotive Engineers, Warrendale, PA.
3. Wright, C., 1995, *Applied Measurement Engineering*, Prentice-Hall, New York, pp. 98-102.
4. Bendat, J., Piersol, A., 1971, *Random Data: Analysis and Measurement Procedures*, Wiley-Interscience, New York, pp. 301-309.
5. Wright, pp. 249-253.
6. Bendat & Piersol, pp. 10-12.
7. Bendat & Piersol, pp. 208-212.
8. *Experimental Dynamics and Measurements*, MacNeil/Schwendler Corp. Short Course, 1986, Dr. Julius Bendat.
9. *Fatigue Design Handbook*, 2nd Ed., 1988, SAE, Warrendale, PA, pp. 124-136.