

Analyze Menu

Analyzing a data set implies extracting information not visible by simply viewing a graph of the raw data. Analysis can often mean creating a version of the data set in a completely different dimension: instead of values (e.g. heart rate) vs time, the analysis looks at transformed values (heart rate spectrum) vs frequency. FREQ presents two such analysis tools, the traditional Fourier Transform (plus its inverse transform) and the somewhat untraditional Frequency Search, or Fast Orthogonal Search.

[Frequency Search](#)

[Fast Fourier Transform](#)

[Inverse Fourier Transform](#)

[Approximate Entropy](#)

[Correlation](#)

[Convolution](#)

Analyze/**Frequency Search**

A Frequency Search is the most powerful analysis tool in FREQ, and the reason FREQ was written. There is no other comparable technique for researchers to determine what is in their data sets. Fourier series are difficult to interpret, have limitations, and often require that finite-length (that is, all) data sets be extensively modified prior to application. Read these sections first to familiarize yourself with the background and methodology:

[Fast Orthogonal Search](#)

[Preparing for a Frequency Search](#)

[Interpreting Frequency Search Results](#)

On selecting the Analyze/**Frequency Search** menu option, FREQ will present the user with the File/Open Period Table dialog for the period table and if necessary again for the data file: File/Open Data File. If period table text windows and/or data set graph windows are already open, FREQ assumes that the last active windows of each type are to be used in the Frequency Search.

A Search Progress dialog is displayed during the search, displaying current search results.

Search Progress

During the Frequency Search a dialog box appears on-screen displaying the progress of the search: the last identified period, its relative MSE (i.e. "power"), the computed amplitude and phase for the ID'd period, and the cumulative Mean Square Error Reduction.

The candidate period currently being tested is also shown. Hitting Cancel closes the dialog, but the search continues. Hit Abort Search to stop the search at any point.

Analyze/**Fast Fourier Transform**

See also [Interpreting the FFT](#)

With an active data set graph window, an FFT will produce a plot of the "spectral power", or magnitude of the transform, and a plot of the spectral phase. The FFT operates on a "complex" data set: the Y values are assumed to be complex numbers, of the form

$$Y = Y_{re} + i Y_{im}$$

where i is the notorious square root of minus 1. Y_{re} and Y_{im} are the "real" and "imaginary" parts of the complex number Y . The important fact here is that for the data sets we normally use, all the Y values are "real", so **FREQ** loads an "imaginary" array with 0 values for the FFT. If concepts such as "complex", "real", and "imaginary" are foreign to you, learn more about them from a math textbook if you want to progress in real-world research and analysis, no matter what your field.

The N -point data set ordinate values $Y = y(t)$ in the time domain become the real part of a complex array of values: $(Y + 0i)$ which when fed to the FFT are "rearranged" into a complex spectrum $S(f) = (S_{re} + i S_{im})$ in the frequency domain.

For $y(t)$, read "y of t" and for $S(f)$ read "S of f". Functions y and S are 'inverse' functions of one another, as the FFT and the inverse FFT map y into S and *vice-versa*.

The spectrum magnitude is

$$\text{sqrt}(S_{re}^2 + S_{im}^2)$$

and the spectrum phase is

$$\text{arctan}(S_{im} / S_{re}).$$

The magnitude and phase spectra are displayed by **FREQ**. The real and imaginary components are used by the inverse transform, but have not been re-ordered for display.

In addition to the difficulties interpreting the discrete Fourier transform, there are three problems with using an FFT to learn the characteristics of your data.

First, the FFT algorithm requires that the data set contains X values which are in strict arithmetic order, e.g. 0, 1, 2, 3, ... or 0.25, 0.50, 0.75, ... This is a

limitation of the technique, and irregularly spaced data sets must be somehow "filled in", or interpolated to yield (X,Y) points whose X values are precisely regular.

A far worse limitation for analysis purposes is that the FFT requires the number of data set points to be a power of 2: ... 256, or 512, or 1024, ..., 65536, etc. Data sets are truncated to the highest power of 2 less than or equal to the data set size, which means that the 'right end' of most data sets is left out of the analysis. Pre-processing tricks to allow using the entire data set involve techniques like 'padding' or 'shaping' or 'windowing' the data set. All these tricks suffer from modifying the data under analysis. Elaborate arguments exist for justifying these techniques, and in the hands of very skilled analysts their effects on the data can be minimized, but the Frequency Search used by FREQ obviates the necessity for them.

A third limitation is that the frequencies used for the FFT be in an arithmetic series, each a multiple of the fundamental frequency, $1 / T_{\max}$, where T_{\max} is the difference between the first and last abscissa points in the data set: $T_{\max} = T_N - T_0$.

The *quid pro quo*, of course, is that to use FREQ's Frequency Search, the researcher must pre-load the period table with periods known or suspected to be in the data set. The usual approach is to use a large table initially, which contains periods from large to small. As **Frequency Search** runs are made, the periods in the .TBL file are changed until they reflect reality, but it is always a good idea to leave a number of periods in over the entire possible range for your data set. As the Fast Orthogonal Search algorithm is exceptionally good at resolving two nearly-identical periods, consider clustering a number of periods about regions of interest. In this regard, the FFT may serve to determine the width of these regions.

Note that the FFT spectral magnitude is symmetric (and the spectral phase is anti-symmetric) about zero frequency for real data sets. An asymmetric spectral magnitude indicates a non-zero imaginary component to the data. You can see this for yourself, and maybe get some ideas how to "package" your data sets, if you try the Graph Window options Select Data set as Real Part and Select Data set as Imaginary Part available on the System Menus for two data sets that seem to have a lot in common, though they may not be from identical sources.

Interpreting the Fourier Transform

The FFT phase and magnitude plots are first "shifted" before graphing the results. That is, the FFT produces a transform which needs to have the right and left sides exchanged before we can look at them. The FFT real and imaginary part plots are "unshifted", so that they can be used directly in the inverse transform.

Two things emerge immediately when you first examine the FFT magnitude graph: first, there is usually a lot of action near the origin (low frequencies) and progressively less power in the higher frequencies. Second, the plot is nearly exactly symmetric about the Y axis.

Use the Synthesize/Generate Synthetic Data option to create a time series with a single sine wave. Give it a name, choose 1024 points, an interval of 1, and ask for 2 sine waves. Assign an amplitude of 80, a period of 4, and 0 phase offset to the first sine. Assign amplitude 50, period 32, phase 0 to the other sine. Apply the FFT using Analyze/**Fast Fourier Transform**. Look at the FFT magnitude plot, maximizing the graph window by clicking on the "up arrow" in the upper right corner. Because there are more points in the FFT than pixels across the screen, you may not see all four of the spikes. Re-size the graph (by dragging a corner of the graph window) until all four spikes are visible. Their x-values should be plus and minus 256 for the first sine, and plus and minus 8 for the second sine. Notice that the power in the second sine is less than the first's. Note that 256 is $1024 / 4$, and 8 is $1024 / 32$: this reciprocal relationship using the number of points in the series suggests an easy way to mentally move between the time and frequency domains. Be aware that for time series of length not a power of two, some points on the "right end" of the series will be omitted from the transform. That's why this experiment used 1024 points. If you apply the option Analyze/Inverse Fast Fourier Transform to a transformed time series you get back the original series, minus any part that exceeds a power of 2 in length.

Re-run this test, using a period of 7 for one of the sines. Note that the FFT magnitude plot now shows a pair of spikes at frequencies plus and minus 146, but the spikes have a broadened base. The period 7 is "incommensurate" with the frequencies used in the discrete Fourier transform, which are all of the form $N / (1/2) = 2 N$, where N can range from 1 to 512. The transform redistributes the power in the period 7 sine wave among a number of frequencies in the vicinity of 146: a period of 8 produces a pure spike in the FFT magnitude, but a period of 7 produces a spread spectrum.

FFT Formulas

The transform and its inverse transform are defined by the Fourier series (the discrete transform)

$$F(\omega) = (1/N) \sum_{n=0}^{N-1} f(t_n) e^{-i \omega t_n}$$

with $\omega = 2\pi / (m T)$, and $t = t_m - t_{m-1}$, and m ranges from 0 to N .
The inverse FFT is

$$f(t) = (1/N) \sum_{n=0}^{N-1} F(\omega_n) e^{i \omega_n t}$$

Note that the data set Y values must be defined over an abscissa X at regular intervals,

or $t_n = t_0 + n t$, where t_0 is the starting value and t is the interval between points.

This is a practical limitation of the discrete Fourier series, as is the requirement that the frequencies be in harmonic relation: $\omega_n = n \omega_0$.

Analyze/**Inverse Fourier Transform**

After Analyze/Fast Fourier Transform has been applied to a data set, the Inverse Fourier Transform will return a pair of new graph windows. The real part graph is essentially the original data set Y values, usually to better than 0.1% accuracy. The imaginary graph wiggles around, but examination shows the values are on the order of less than a millionth of the Y values, or essentially 0. Once you verify this, the imaginary graph can be closed.

What has this round-robin process accomplished? The information of value to most researchers is in the spectrum magnitude, where the spectrum shape often indicates how energy migrates through a physical system, and the location of peaks indicate the existence of periodic behavior in the time series. The inverse transform's congruence with the original data set verifies that the forward transform was valid.

The Graph Window System Menu options Select Data set as Real Part and Select Data set as Imaginary Part allowed you to associate one data set with the real part, and another data set with the imaginary part, of a complex data set. If the two data sets are not identical, the FFT spectral magnitude plot will be more or less asymmetric about zero frequency. The inverse transform applied to this complex FFT will yield back the original data sets as the IFFT real and imaginary parts. The IFFT magnitude and phase are also plotted in new graph windows.

Analyze/Approximate Entropy

NOTE: This analysis procedure is very SLOW for large data sets or run lengths: over a minute for 1000 points, proportional to the square of the data set length.

Entropy is the statistical measure of disorder, or lack of regular behavior, in a system. For a data set, regularity means reasonably long runs of points that don't jump out of a pre-determined range of values. Researcher Steven M. Pincus has derived a measure of regularity which quantifies this concept in a manner consistent with the mathematical and physical definition of entropy.

Basically, given a run of points which fall into a pre-determined band of values (usually stated as a percentage of the data set standard deviation), the user states the run length (number of data points, **M**) over which the "regularity" of the data set is to be estimated. We then calculate the number of cases for which this regularity is achieved. Every batch of **M** consecutive points is checked against all other **M**-sized groups.

The "regularity" criterion **R** is defined as a percentage of 1 standard deviation (s.d.). For **R** = 2%, we check to see if during a run of **M** points any of them exceed the preceding value by more than **R** = 0.02 s.d.

The result is called Approximate Entropy, or ApEn, and is near zero for very regular data sets, increasing for progressively more disordered data sets. Pincus shows electrocardiogram (EKG) data for an infant who is experiencing great cardiac stress (Sudden Infant Death Syndrome, or SIDS) while in the cradle, but who survives the episode, a so-called "aborted SIDS event". Comparison of the EKG with that for a normal infant shows The SIDS infant heart is curiously devoid of great beat-to-beat variability. Other diseases show the opposite relationship: normal behavior exhibits greater "disorder" than the pathological case.

ApEn is able to quantify this variability, or lack of it. ApEn is a function of two parameters, the run length **M** and the regularity criterion **R**. This is written ApEn (**M**, **R**) and the values of **M** and **R** are selected by the researcher to best illustrate the regularity difference between the normal and abnormal case. After a suitable selection of M and R, the researcher would, of course, seek to establish that the ApEn criterion applies to a number of normal vs abnormal test cases.

More prosaically, ApEn can ascribe a "signature" to certain classes of data sets, by which they may be quantitatively differentiated, without have to

examine the traces visually.

Analyze/**Correlation**

Analyze/**Convolution**

Fast Orthogonal Search

Also see

[Preparing for a Frequency Search](#)

[Interpreting Frequency Search Results](#)

Researchers and analysts often need to know if experimental data (a time series) contains significant amounts of signal at frequencies of interest to them. These frequencies may correspond to driving forces, environmental constraints (e.g. boundary conditions), or system responses, and may result from intrinsically non-linear processes. Many times, a mathematical or phenomenological model exists to explain some observed behavior, and experimental data is collected to verify whether or not the model is correct. A "frequency search", or decomposition, is a fundamental approach to performing these analyses.

The usual approach is to pre-process the data, apply a Fast Fourier Transform to the series, and plot the power spectrum. Peaks in the FFT spectrum may correspond to interesting frequencies. However, it is difficult for anyone but a signal processing expert to know how much power in the time series is actually accounted for by a given frequency. It is even harder to resolve nearby, overlapping broad peaks. More often than not, "noise" dominates the data and cannot be de-coupled from signals of interest. Pre-processing (filtering and windowing) of time series data sets is a demanding discipline, and many of the rules to assure the validity of pre-processing operations are difficult to apply.

There is an alternative, well-researched and nearly painless method available to perform such frequency analyses. You first prepare, as a text file, a table of candidate periods (see below, File Preparation; FREQ contains a rudimentary editor for just this purpose). On initial creation, the table usually contains a fair number of entries, as there may be no prior knowledge of what is really present in the data. Periods may be longer than the time series, or as short as twice the time interval between points. But the researcher often knows what to look for based on theory or existing work, and the table will contain several periods in the regions of interest.

An algorithm called the Orthogonal Search Method is applied to a time series, using the table of candidate periods, and the precise power, amplitude, and phase of sine waves corresponding to entries in the table is displayed. The objective is to determine if frequencies of interest to the researcher are present in significant measure. The measure employed is the

"Mean Square Error Reduction" or MSER. The MSER for all candidate frequencies is computed, and the most significant are reported in a text .OUT file after the run.

The Fast Orthogonal Search (FOS) algorithm is an adaptation of the original Orthogonal Search Method developed by Michael J. Korenberg and his group at Queens University, Kingston, Ontario. FOS searches a user-supplied table of candidate periods and selects the best candidate periods to fit an input time history. Results are stored to a .OUT file formatted to be read by the user. Results can also be saved as a data set by invoking the menu option Synthesize/Reconstruct Data Set From Search to make a reconstructed time history.

FOS analyzes a time series stepwise, reducing it to a combination of sines. The program examines a table of candidate or search frequencies created by the investigator, determines the ability of each period to explain a portion of the total variance (mean square error, or MSE) and orthogonally removes the sine explaining the largest percentage of the time series variance. The process is repeated on the residuals until there is no further error reduction or until a specified number of periods have been identified.

The results are written to the screen and to an output file. (DATAFILE).OUT presents the results of a single run, while FREQ.SUM accumulates the summaries of several runs.

Fast Orthogonal Search (FOS) allows a researcher to select from a time series some assumed periods of interest, which are tested by FOS for the Mean Square Error (roughly the 'power') accounted for by sinusoids of computed amplitude and phase. The researcher specifies a number of candidate periods in a text file (NAME).TBL, and a time-series (NAME).DAT formatted as (time, value) pairs or simply as a list of values. Note that the times need not be uniformly spaced, and there may be gaps. Absolute spacing of columns is not important. The 'time', of course, may be any monotonic physical variable characterizing the 'data', e.g. position.

FOS is reasonably noise-tolerant, as all data elements are used only in series-wide averages over the orthogonal basis functions.

NOTE that **Frequency Search** execution time is proportional not only to the length of the time-series, but to the number of candidate periods in PERIOD.TBL and the number of spectral features to search for.

Interpreting Frequency Search Results

Also see [Reading the OUT File](#)

The Frequency Search option selects periods from your .TBL table to satisfy its requirement to remove as much of the "Residual Mean Square Error" from the data set as possible. Each period is tested sequentially. The one which removes the largest error (fits the data best) is selected and the process is repeated on the remaining periods until completion criteria you chose are met.

To aid in interpreting results, you can look at the MSE, or Mean Square Error, as $1/2$ times the squared amplitude of the wave (e.g. the 'power' in the series if measuring current or voltage). This is strictly true only if either an integral number of cycles of the selected period sine waves fit in the data set, or if there are many such cycles. But it is a convenient way to visualize the ability of selected periods to "explain" your data.

Probably the most insight into your data set can be achieved by watching the reconstruction form on your monitor screen as the search progresses. During a Frequency Search, you will see the yellow superposed sine curves begin to hug the data closer and closer, eventually lying on top of the data curve. The first wave drawn corresponds to the period in your table responsible for the most 'power', or MSE reduction. Notice that it tends to fit the side-slopes, or "trends" in your data set, and pretty much ignores peaks and valleys unless they are very broad. The eye tends to see the peaks, but the power resides in the slopes. Successive, usually shorter period selections do a better job of snaking their way up into the peaks and valleys. If your table contained a good enough portrayal of the periods actually in the data set, you may account for more than 50% of the power. If not, think about possible reasons, enter new or different periods into the .TBL file using the `FREQ` text editor or your own, and re-run the search. Periods that didn't show up in the initial passes probably are not found in the data and can be removed from the table; this will markedly improve run speed.

You can use the period, amplitude and phase offset of each identified sine wave from a Frequency Search to make a new, 'synthetic' data set, virtually automatically. This may help assess the effect of noise in real data sets. Assure that the data set graph on which the Frequency Search was performed is active, by clicking the mouse anywhere over it or by hitting `Ctrl-Tab` repeatedly until the graph window is highlighted (See [Activating Graph Windows](#)). Then select the [Synthesize/Reconstruct Data Set from Search](#) menu option. A new graph window appears, holding the data set reconstruction comprised only of sine waves from the Frequency Search.

If you subject this reconstructed data set to a Frequency Search using the same Period Table as for the original search, the yellow reconstruction curve will finally lie very nearly on the data plot, in fact you may not see any difference. A nice bit of closure...

FREQ is not a curve-fitter, nor is it an FFT. It searches the data using the supplied candidate periods in the .TBL file, selecting those periods which account for the most "power". The terminology for power in use now is Mean Square Error, but MSE is not necessarily intuitive to a physiological researcher. Strictly, the MSE is wave power = $1/2$ (amplitude squared) only if many cycles or an integral number of cycles exist in the dataset. Only those Fourier coefficients for selected periods are saved.

FREQ uses an adaptation, called Fast Orthogonal Search (FOS), of the Orthogonal Search Method developed by Michael J. Korenberg and his group at Queens University, Kingston, Ontario, in the late 1980's. FREQ searches a user-supplied table of candidate frequencies and selects the best candidate frequencies to fit an input time series. The algorithm analyzes a time series stepwise, reducing it to a linear combination of sines whose periods are read from the table.

FREQ examines the table of candidate or search frequencies created by the investigator, determines the ability of each period to explain a significant portion of the total variance (mean square error, or MSE) and orthogonally removes the sinusoid explaining the largest percentage of the time series variance. This process is repeated on the residuals until there is no further error reduction or until a specified number of periods have been identified.

FOS doesn't do much with spikes and sharp peaks in data. The sharper the peak, the less power it contains (for finite amplitude) and FOS looks to explain where the power is. Also, recall that a sharp spike contains a large number of frequencies over a very wide band. FOS does best with rolling hills and valleys. It's country code.

Make a single sawtooth wave using the Synthesize/Generate Synthetic Data selection. Make a period table with periods extending from vary large (much larger than the length of the data set) to very small values (on the order of the interval between data points). Run Analyze/Frequency Search on the sawtooth. The code is very likely to choose a single very long period, so that the slope of the wave coincides with the slope of the sawtooth. Not what one would expect from an FFT, which would say the sawtooth contains a large number of harmonics of the fundamental sawtooth frequency. Is FOS wrong? No, it matched the data. Is it right? Not many researchers would say they matched their data with the initial rising slope of a sine wave 50 times longer than their data set. But they could...

CoDebris will eventually modify the criterion for selecting periods from a reliance on MSE alone, to one which (optionally) weights short period components preferentially after most of the power has been accounted for in the low-frequency components. This will allow replication of more high-frequency features. For now, you can set the MSE cutoff value, the second number in a .TBL file, to 0.01 to get more detail, with a commensurate increase in running time.

The derived amplitudes are roughly $\sqrt{2}$ times the power at the associated period}. Only periods in the .TBL having the most power are used. Researchers commonly 'shotgun' a wide range of candidate periods chosen from their experiments or theory, which works but is time-intensive to compute. Run time is proportional to the product of:

- the number of candidate periods (32 in REAL.TBL),
- the number of data points (1000 in REAL.DAT),
- the number of periods to select (10 in REAL.TBL), or the MSE cutoff (2.0 in REAL.TBL), in percent of the total power in the data, whichever occurs first.

A lower cutoff will cause FREQ to spend more time finding low-power data set components.

The Fast Orthogonal Search algorithm is capable of much greater time resolution than a Fourier transform, and is not limited to harmonics of a fundamental frequency. The algorithm is also quite insensitive to noise, as all data elements are used only in series-wide averages over the orthogonal basis functions, especially compared to techniques requiring a very narrow band filter to isolate the signal. In many nonlinear or biological systems, the signal frequencies move, or breathe, as the system evolves, so specifying a narrow band filter is not adequate. Finally, FOS even tolerates missing data points, or irregularly-spaced data sets. An amazing piece of work by Korenberg, I think you'll agree.

See the [References](#).

Reading the .OUT File

The .OUT file is named for its corresponding data set file, and is created in the same directory as the data file.

Here's a FREQ output (.OUT) file for an typical data set. Note the run time is for an 80386/33. An 80486/66 is 4 times faster (CPU clock is twice as fast, and the I/O bus bandwidth is twice as wide) and a Pentium90 is about 8 times faster.

FREQ -- SEARCH RESULTS

FREQ searched data file: C:\PROJ\FRQ\AAA2BOHR.INT
Using period table: C:\PROJ\FRQ\PERIOD.TBL
Run ended Wed Jan 25 00:57:44 1995 Elapsed time 316 seconds
Candidate periods: 20 Features identified: 5
MSE cutoff: 0.20 % Total MSE: 1.568e+004
Data points: 4810 Offset or Mean: 190.142
MSE explained: 6.742e+003 or 43.00%

Feature	Period	Amplitude	Phase (deg)	MSE	%TMSE
1	1200.0	115.54	184.3	2.905e+00 3	18.53
2	800.0	88.76	23.7	1.811e+00 3	11.55
3	400.0	53.44	143.66	1.230e+00 3	7.84
4	58.0	15.87	174.64	7.527e+00 2	4.80
5	10000.0	3347.26	334.64	4.337e+00 1	0.28

The values for Feature 5 are not errors. FOS simulates a linear trend if it has to. Note the low MSE value, and the trig approximation implicitly used here:

$$\sin(x) = x \quad \text{for } x \ll 1$$

For this very long wave, only a very small piece (the initial, linear rise) is actually in the data set. A linear trend or a constant offset can be easily removed with Synthesize/Pre-Processes menu options.

Preparing for a Frequency Search

See also

[Data File Format](#)
[Period Table Format](#)
[Frequency Search Parameters](#)

The user is responsible for maintaining two types of external files: a table of candidate periods and a data file, or time series. The tables usually have a .TBL extension. The data set files ordinarily have a .DAT extension, but that extension (or any extension) is optional.

Period tables can be read into FREQ and edited, but they **MUST BE SAVED** (File/Save or File/Save As) before use in a Frequency Search, as the table is read from disk each time it is used. After selecting a Period Table for analysis, you are presented with a dialog to choose the number of candidate periods to select (the default is all the periods in the table) and for a "MSE cutoff". The MSE cutoff (default value 2%) is simply the lowest "power" sinusoid FREQ will use in building a 'reconstructed' data set (the yellow dotted curve which builds up during the Frequency Search). Typically, a couple of long-period sinusoids contain most of the power in a data set and define its trends, while shorter period waves above the MSE cutoff build detail and structure into the reconstruction.

If you are interested in determining whether a given set of periods is found in your data, you can include only those numbers in the Period Table. The amplitudes, %MSE, etc. will be correct, but the reconstruction won't look very convincing. Some users load a larger number of candidate periods into the table, running from several times the length of the time series to be searched all the way down to a few times the interval between data points, and covering the region between in a quasi-logarithmic way: say 5 points above 1000.0, 5 points in the hundreds, 5 more from 10.0 to 100.0, etc. After a preliminary run (perhaps on an abridged data set if it is very long, or one interpolated to have fewer points (see Synthesize/Pre-Process/Interpolate) examine the .OUT file for likely candidates to refine, and others to drop from further consideration.

Note that data set files are NOT directly editable as a FREQ menu option, but they (or copies of them, actually) may be manipulated in the pre-processor (menu selection Synthesize/Pre-Process). This allows outlier (artifact) removal, partitioning, smoothing, etc. to be

applied to named data sets. Data files CAN be opened as text **FOR VIEWING ONLY** in an editable window by choosing "All Files (*.*)" from the file type list box in the File/Open Period File or the Open Output File dialog. But **DO NOT** try to save the data set from the text window: you will end up with a truncated data set file if its size was larger than the 32kByte limit imposed for text windows. Editing data sets manually is something of a No-No anyway.

Frequency Search Parameters

FREQ uses two user-specified criteria to decide when to conclude a Frequency Search. In the **Frequency Search Parameters** dialog, the user enters the number of periods to search for, and the Mean Square Error ("power") cutoff value, in percent of the total (100%) "power" in the data set.

The dialog initially contains the number of candidate periods in the .TBL file selected by the user, and a Search Cutoff value of 2%.

The MSE cutoff (default value 2%) is simply the lowest "power" sinusoid FREQ will use in building a 'reconstructed' data set (the yellow dotted curve which builds up during the Frequency Search). Typically, a few long-period sinusoids contain most of the power in a data set and define its trends, while shorter period waves above the MSE cutoff build detail and structure into the reconstruction.

Selecting appropriate values depends greatly on the characteristics of your data set and the Period Table (see Period Table Format) you have created. If your data set is long, you may want to practice with a lower-resolution version of the data: use Synthesis/Pre-Process/Interpolate and enter, for the number of data points, no more than (say) 1000. You'll get a "rough copy" of the data set to refine your analysis skills; save it to disk with File/Save As with an easily-recalled name, including the DAT extension so it appears in the File/Open Data File when you repeat your tests.

Period Table Format

A Frequency Search requires a table of candidate periods in a text file FILENAME.TBL to be created by the user (FREQ contains a rudimentary text editor for creating, altering and saving files). A typical PERIOD.TBL has the form (for example):

```
30.0
20.0
15.5
2.25
8.85
.
.
.
```

where

```
30.0 is the first candidate period,
20.0 is the next candidate period,
.
.
.
```

The units for period are up to the user, but must be consistent with the data abscissa values in dimensioning (e.g. seconds, heartbeats, inches). The periods may be in any order, not necessarily monotonic. To help you associate a period table with a particular data file, create it with the same filename, followed by the .TBL extension (e.g. EXPR33.DAT, EXPR33.TBL).

If you are only interested in determining whether a given set of periods is found in your data, you can include only those numbers in the Period Table. The amplitudes, %MSE, etc. will be correct, but the reconstruction won't look very convincing. Some users load a larger number of candidate periods into the table, running from several times the length of the time series to be searched all the way down to a few times the interval between data points, and covering the region between in a quasi-logarithmic way: say 5 points above 1000.0, 5 points in the hundreds, 5 more from 10.0 to 100.0, etc. After a preliminary run (perhaps on an abridged data set if it is very long, or one interpolated to have fewer points (see Synthesize/Pre-Process/Interpolate) examine the .OUT file for likely candidates to refine, and others to drop from further consideration.

Selecting appropriate values depends greatly on the characteristics of your data set and the Period Table (see Period Table Format) you have created. If

your data set is long, you may want to practice with a lower-resolution version of the data: use Synthesis/Pre-Process/Interpolate and enter, for the number of data points, no more than (say) 1000. You'll get a "rough copy" of the data set to refine your analysis skills; save it to disk with File/Save As with an easily-recalled name, including the DAT extension so it appears in the File/Open Data File when you repeat your tests.

Data File (Time Series) Format

You must have a time-series formatted as ASCII (time, value) pairs or simply as a list of values. Any valid DOS filename will work, although **FREQ** first filters files with a .DAT extension. "Time", of course, can be any variable relevant to you work (e.g. angle, position, heartbeat), not necessarily just physical time; the usage "time series" is traditional.

Note that for a Frequency Search the times need not be uniformly spaced, and there may be gaps. Times, however, **MUST** be uniformly spaced with no gaps (or made so) for a Fast Fourier Transform to work. This limitation on the FFT is one of the key advantages of the Frequency Search technique.

Absolute spacing of columns is not important. All entries are normal, ASCII (man-readable as well as computer) text, not binary-coded. Integer, decimal, exponential notations are acceptable. If you are getting your data from a spreadsheet or data base, creating the "ordinate only" format below should not be a problem. An example of the "Abscissa-Ordinate" format:

0.0	-47.3
0.5	-22.4
1.0	14.6
.	.
.	.
.	.

Or, if your data set is of the form "Ordinate only",

-47.3
-22.4
14.6
.
.
.

you will be asked to supply the starting time and the time interval between points as the file is being loaded into memory. **FREQ** accepts negative start times.

Remember that although the terms 'time' and 'time series' are used, the X variable in a (X,Y) data set can be any quantity on which the Y values depend. However, be **VERY** careful over your choice of units (degrees, radians; or seconds, minutes, heartbeats)... it matters greatly in your

analysis. Strive for internal consistency during a FREQ session, and throughout your work. Try "walking through" a data processing session on paper, using printout of computer input and output files for raw material: check that numerical answers are appropriate to your assumptions. FREQ can be of great assistance here, as you can design your own dataset in the Synthesize/Generate Synthetic Data dialogs. Anything you don't understand about your data can be simulated and run through a Frequency Search or FFT. You will learn a lot about nuts and bolts stuff: units, intervals, bandwidths, sampling rates, etc.



FREQ Help System Contents

FREQ Version 2.1 Copyright (c) 1995 CoDebris
711 Barbara Avenue Solana Beach, CA 92075 (619) 755-4492

Overview

[Registration](#)
[Introduction](#)
[Who needs FREQ?](#)

Menu Selections

[Help](#) [File](#) [Edit](#) [Synthesize](#) [Analyze](#) [Window](#)

This is the shareware version of FREQ, with complete Frequency Search capabilities, including all the file I/O, printing, editing and graph options of the registration version.

To acquaint you with the more complete "waveform spreadsheet" capabilities in the [Registration](#) version, the extended Synthesis and Analysis options distributed only with the [Registration](#) version are documented in the Help system you are now reading.

All topics are also accessible through the Help/Search dialog.

Introduction

FREQ Version 2.1 Copyright (c) 1995

...Beyond Fourier Transforms...

Frequency Search and Analysis from CoDebris

711 Barbara Avenue Solana Beach, CA 92075 (619) 755-4492

FREQ is a data analysis tool which determines what sine waves make up a data set or time series: periods, amplitudes, phases, percent relative power. You specify which periods you believe are present (or specify a whole range of periods), and FREQ tests those periods and graphically assembles selected sine waves into a revealing portrait of your data: an educational, evolving visualization of your data you can't get from a Fast Fourier Transform without a whole lot of work and specialized training. Click for more information:

[Time Series Concepts](#)

[Frequency Search Introduction](#)

[Program Environment](#)

[Machine Environment](#)

[Program Navigation](#)

[Association and Visualization](#)

[On-line Help](#)

[Custom Upgrades](#)

[References](#)

Overview...

Researchers and analysts often want to know if experimental data (a time series) contains significant amounts of signal at frequencies of interest to them. These frequencies may correspond to driving forces, environmental constraints (boundary conditions), or system responses, and may result from intrinsically non-linear processes. Often, mathematical or phenomenological models exist to explain some observed behavior, and experimental data is collected to verify whether or not the model is correct. "System Identification" is a name given to attempts to determine the underlying dynamics which generate a data set.

FREQ is a new, fundamentally unique tool for performing these analyses. FREQ not only identifies the period, phase and amplitude of sine waves which represent your data, but then plots each new wave component over a graph of your data. A plot of the original data and reconstruction can be printed, and the reconstruction can even be saved as a data set, on the same scale as the original data.

If you don't have a good math background, FREQ is an educational experience. Use the Synthesize/Generate Data Set option to learn about sine waves, square waves, noise, trends, offsets, etc. in a graphical environment. FREQ does all the work, you just do the thinking to set up new "experiments" and analyze "why". In the Registration version, there are more options under Synthesize/PreProcess to alter, combine and "pre-process" the "synthetic" data sets you make. What happens when two sine wave graphs are added or multiplied together? Some of the graphs you can produce are pretty wild, verging on computer "art".

On first running FREQ, choose Analyze/Frequency Search from the menu and select a (supplied) table of candidate periods (use REAL.TBL the first time) and a supplied data file (REAL.DAT). The .TBL files can be edited and saved using FREQ's built-in editor.

REAL.DAT is a physiological data set. The synthesized curve is not a bad approximation to the data, as you can see while the reconstruction builds before your eyes, even though only 55% of the "power" is accounted for by the seven identified periods, amplitudes and phases, analyzed in under 1 minute. Refining the REAL.TBL file might find even better (closer) periods to use, but the periods supplied are pretty good, or they wouldn't be used in a demo.

The important thing is watching the highest-power, long period sines define major trends in the data and seeing the short-period sines fill in the peaks

and valleys. On seeing this for the first time, you can achieve a very intuitive sense for what the data says, and of Fourier series in general.

As you feed FREQ other data sets of interest to you, or generated from the rich variety of Synthesis/Generate options in FREQ, your understanding of what your data is telling you grows immediately, in a very visual way.

Arrange your data for input to FREQ by creating a file consisting of two columns (X, Y) and as many rows as there are points in the data set. Alternatively, you can just feed FREQ a single column of Y values only, and tell FREQ the initial X value and the interval between successive values.

FREQ will line-plot the actual data in green, and the synthesized data (the "reconstruction") from the selected periods as a yellow dotted line. Plots are auto-rescaled on window resizing, actually recomputed every time the window is painted to optimize the displayed data information content.

Each time a new period is selected as containing the next most significant amount of power, the yellow curve is redrawn to include the new information. You can get a good sense for periods to include in the .TBL file, and which to exclude, by watching the analysis and then examining the .OUT file that pops up on the screen after each analysis is complete.

In fact, watching the earlier, higher-power sinusoids define the major features of your data is an insightful experience; one researcher termed it a "spectacular visualization" of his data. The side slopes of the sine curve will lie along major trends in the data, and the peaks won't necessarily correspond to peaks in the data. That is left to shorter period components, which when properly phased with the long period components will ride up and down peaks in the data with increasing fidelity.

You will definitely learn to think nearly simultaneously in so-called "dual spaces": time and frequency. You will learn to think carefully about what are the proper "units", or "dimensions" of your data. The supplied PERIOD.TBL is designed to encompass a wide range of possible values, but it takes longer to analyze a data set. After running it on your data, examine the associated .OUT file and create a "tailored" table of periods to run with that particular sort of data. Use it to test assumptions about what is really in your data.

Background...

The usual frequency analysis approach is to pre-process your data, apply a Fast Fourier Transform to the series, and plot the power spectrum. Peaks in

the FFT spectrum may correspond to interesting frequencies. However, it is difficult for anyone but a signal processing expert to know how much power in the time series is actually accounted for by a given frequency. It is even harder to resolve nearby, overlapping broad peaks. More often than not, "noise" dominates the data and cannot easily be de-coupled from signals of interest.

Pre-processing (filtering and windowing) of data sets is a very demanding discipline, and many of the rules to assure the validity of pre-processing operations are difficult to apply, as such operations actually modify the characteristics of the manipulated data.

Further, most researchers with a need for waveform analyses do not have formal training in the subject, and many feel uncomfortable with having to use a host of implicit assumptions.

There is an alternative, nearly painless method available to perform such frequency analyses. The researcher first prepares (as a text file) a table of candidate periods. On initial creation, the table usually contains a fairly large number of entries, as there may be no prior knowledge of what is really present in the data. Periods may be longer than the time series, or as short as twice the time interval between points. But the researcher often knows what to look for based on theory or existing work, and the table will contain several periods in the regions of interest.

FREQ is not just a curve-fitter, nor is it simply an FFT. To use it, you should know something about your data, but you need nearly no data analysis background. FREQ searches your data using the supplied candidate periods in a .TBL file which you prepare, selecting those periods which account for the most "power".

FREQ uses an adaptation, called Fast Orthogonal Search (FOS), of the Orthogonal Search Method developed by Michael J. Korenberg and his group at Queens University in the late 1980's. The algorithm is applied to your data set, using an associated table of candidate periods. The precise power, amplitude, and phase of sine waves corresponding to entries in the table is displayed. The objective is to determine if frequencies of interest to the researcher are present in significant measure, and report the results.

The algorithm analyzes a time series stepwise, determining the ability of each period to explain a significant portion of the total variance (mean square error, or MSE: roughly, the data set's "power"). It then orthogonally removes the sinusoid explaining the largest percentage of the time series variance. This process is repeated on the residuals until there is no further significant error reduction or until a specified number of periods have been identified.

The algorithm is capable of much greater time resolution than a Fourier transform, and is not limited to harmonics of a fundamental frequency. It is also quite insensitive to noise, as all data elements are used only in series-wide averages over the orthogonal basis functions. Finally, it tolerates missing data points, irregularly-spaced data sets, and short data segments. In many nonlinear or biological systems, the signal frequencies move, or breathe, as the system evolves, so short segments are necessary for system identification.

FREQ is a Multiple Document Interface application, so it will display the results from multiple Frequency Searches on-screen at the same time, including the reconstructions (dotted yellow plots) and text output files. A rudimentary FFT is included, mostly for contrast.

Time Series Concepts

A time series is a succession of data values at stated intervals. Typically, a time series consists of pairs of numbers: one member of each pair is the value, the other member is the time at which the value was measured. But the values can be from just about any source:

an instrument recording sound, heartbeats, light intensity -- any phenomena at measured intervals of time, of position, or any event at all which can be recorded. The expectation is that the recording will contain information to assist understanding or controlling the phenomenon. Time series analysis is the collection of techniques to extract that information.

The intervals do not have to be regular: a data value may be recorded only on the occurrence of a heartbeat, which may be irregular, or even skipped. Many (older) analysis procedures log, or record, data in this way. In **FREQ**, the menu option **Synthesize/Pre-Process/Interpolate** may be used to regularize a data set.

A data set is made of pairs of X, Y values -- so-called ordered pairs, as the X value is always stated first. The relationship is usually expressed $Y = Y(X)$ and the parentheses here imply Y is to be considered a function of X, to depend on X.

X is usually considered the independent variable, Y the dependent variable: each value in a succession of recorded events depends on the time it was recorded. The set of X,Y pairs which comprise a data set is, in graph parlance, referred to the abscissa and the ordinate of a graph: the X values run out along the horizontal axis, and each corresponding Y value is plotted above its X, at a height proportional to the Y value.

In **FREQ**, the graph is the fundamental element, not the individual X,Y pairs, or even the X or Y values themselves: a data set exists as a collection, and the graph window is the simplest visualization of the data.

The main menu option **Synthesize/Generate Synthetic Data** allows a researcher to emulate real data using superposed pure sines, square waves, sawtooth waves, squirt waves, a linear trend (constant sloping line), a constant (often called DC, for Direct Current) offset, and additive noise. The researcher first specifies a name for the contrived test data set. The number of data points, temporal resolution, number of each type of wave, and the periods, amplitudes and phases of the waves are then entered. A

new graph window appears on screen.

Abscissa

The set of X values of a graph window. The horizontal "time" axis of a time series. Every X value has a corresponding Y value in a data set.

Ordinate

The set of Y values of a graph window. The vertical axis of a time series. Every Y value corresponds to a definite X value in a data set.

Association and Visualization

You don't need to scan columns of numbers to see very subtle differences between two data sets. By juxtaposing two data sets in some way (not necessarily just on the screen or in print), relationships between data sets from very different sources may be seen which would not likely be discovered through conventional analyses.

Sometimes, simply tiling the data set's graphs appropriately is sufficient: use Window/Tile Vertically to look at differences in peak and trough values between two graphs, or Window/Tile Horizontally to look at differences between two graphs in the locations of X-axis crossings (zero crossings) and the X-coordinate (abscissa) values at peaks and valleys.

FREQ plots a data set in green, and the synthesized data (the "reconstruction" from a Frequency Search) for the selected periods as a yellow dotted line. Plots are auto-rescaled on window resizing, actually recomputed every time the window is painted to optimize the displayed data information content. Each time a new period is selected as containing the next most significant amount of power, the yellow curve is redrawn. You can get a good sense for periods to include in the .TBL file, and which to exclude, by watching the analysis and then examining the .OUT file which pops up on the screen after each analysis is complete.

In fact, watching the earlier, higher-power sinusoids define the major features of your data is an insightful experience. The side slopes of the sine curve will lie along major trends in the data, and the peaks won't necessarily correspond to peaks in the data. That is left to shorter period components, which when properly phased with the long period components will ride up and down peaks in the data with increasing fidelity as more sinusoids are identified.

Select Data Set as Abscissa

Select Data Set as Ordinate

These features enable the user to associate two data sets, using one as the abscissa, or independent variable, and the other as the ordinate, or dependent variable. Click the System Menu in the upper left corner of each data set graph window, click again on "Select Data Set As Abscissa" or "Select Data Set As Ordinate" (selections are checkmarked on the menus). As soon as a pair has been selected, a new window will be opened showing the relationship. In general, a positive (upward) sloping elliptical blob of green line segments implies a positive correlation between the two data

sets. A downward-sloping trend indicates anti-correlated sets. A more or less uniform disk indicates a lack of correlation.

Note when running Analyze/Fast Fourier Transform on a data set that the spectral magnitude is symmetric (and the spectral phase is anti-symmetric) about zero frequency for real data sets. An asymmetric spectral magnitude indicates a non-zero imaginary component to the data. You can see this for yourself, and maybe get some ideas how to "package" your data sets, if you try the Graph Window options Select Data Set as Real Part and Select Data Set as Imaginary Part available on the System Menu for two data sets that seem to have a lot in common, though they may not be from identical sources.

See also: Lissajous Figures

Program Environment

Due to the modular, low-level nature of each processing function, FREQ is an in-depth tool-kit for educational purposes as well as serious work . It is not a trivial program to learn to use productively, but the feedback is immediate and visual.

FREQ was written using the Windows Multiple Document Interface to allow any number of graph and text windows to appear on-screen at once, either overlapped or tiled, up to memory limits. Functions may be applied to Synthesize or Analyze a data set or time series, and the results will be left on-screen as one or more new Graph Windows for assessment. Original data is never modified unless you expressly save a Graph Window by overwriting the original data; see Edit/Rename Data Set.

The intermediate data set graphs can themselves be operated on and displayed simultaneously on-screen, allowing a degree of controllability not available in programs which allow only one graph on-screen at a time.

Any graph may be saved to disk with a couple of mouse clicks or keystrokes, as an internal naming system keeps track of the operation which generated each intermediate graph. Graphs can be iconized, tiled, or stretched to fill the entire screen.

Machine Environment

NOTICE: We do not modify ANY of your system files, and de-installing FREQ is accomplished by simply deleting the FREQ Program Group and the files, which are usually found in C:\FREQ.

CAUTION: Disable any Screen Savers (Program Manager /Control Panel /Desktop) before running a Frequency Search on a long data set. Screen Savers suck up large amounts of CPU time.

To use FREQ productively, your hardware environment should be at least a 386 running at 33 MHz, with 8 MB of 70 - 80 ns DRAM, a hard drive with better than 20 - 25 ms seek time, and a Windows-compatible mouse. A math co-processor (Intel 80x87 or equivalent) is required: although FREQ will run without one, it would be too slow to be useful on a large data set. Processing speeds are noted (subjectively) where relevant.

Windows 3.1 is required: it's much better than 3.0: more stable, faster, less quirky. Any DOS version 3.21 or later should work. Time series data processing is intrinsically memory-intensive, so having at least 8 megabytes of fast DRAM installed is not a luxury. FREQ video display utilizes the standard Windows video drivers as well as any specialized drivers which are supplied with video/graphics adapters, as long as they were registered with Windows Setup. Even EGA/VGA 16 color video modes are supported.

Each (X,Y) data point requires 16 bytes from the global heap (2 double-precision floating point numbers, in C) . Another 16 bytes per point, plus change, for a rough total of 30,000 points per MByte, is used for working arrays ONLY during the Analyze/Frequency Search. All working arrays are returned to the heap after each analysis. Stack usage is minimized by block declarations and dynamic allocations on the global heap, for most objects. Each graph window stores its own own auxiliary data, such as the data set mean, the extrema, the RMS value, etc. to save recomputing each time a data set statistic is wanted.

FREQ began as a DOS program; the main reason for porting to Windows was to get an 84,000 point datafile analyzed for a physiological researcher. FREQ will analyze a 200,000 point data set on an 8 Mbyte system, over half a million points in 16 Mbytes, without swapping to disk for virtual memory.

FREQ will display a number of files governed by the FILES= and BUFFERS= limits in CONFIG.SYS (typically 40 and 30), but be aware that Windows uses many files internally.

Help For FREQ Users

Extensive help is available on all features from the menu selection Help/Contents. For specific menu selections, a context-sensitive Help Bar appears at the bottom of the screen, and the ordinary context-sensitive help (Shift+F1 evokes a Help Cursor) accesses Help topics directly. Some multi-function tool windows also support a Help button. From within Help, the Search option allows the user to scroll through all FREQ help entries.

The Help Bar, shown by default on start-up, contains help on high-lighted menu items. Using the mouse (or the F10 and arrow keys) to highlight a selection, context-sensitive help for each topic is displayed at the bottom of the screen. The Status Bar can be toggled on/off via Ctrl-T or from the Help/Help Bar main menu selection if you want more screen real estate.

The usual context-sensitive help, normally available via the Shift+F1 key, loads the full WinHelp engine and the FRQ.HLP helpfile, opened to the section selected by the special Help Cursor. This is the mechanism preferred by Microsoft to display context-sensitive help, but it's a bit intrusive, as the entire screen is taken over. If the Help Cursor is no longer needed, hit the Esc key to restore normal cursor functioning.

Finally, some dialog boxes are supported by a Help button which displays a text window of help information.

Registered users of FREQ can call CoDebris (619-755-4492, 9am - 7pm Pacific Time) or leave E-mail at Compuserve (72074,772) with any questions about FREQ. We can probably steer you to the right selection or sequence of selections, or make arrangements for new features. Discussions about time series processing are usually illuminating to both parties. CoDebris will also recommend one of the available commercial packages if it becomes clear you need to spend that kind of money (typically several hundred dollars). We gain insight into what people really need this way, so new features and interfaces can be added to future versions. CoDebris is particularly interested in supporting more data acquisition devices: A/D converters, digital I/O boards, sound boards, etc. Let us know what you're using and if there is sufficient interest we will contact the device manufacturer for interface specifications.

Custom Upgrades

Call for information on making user specifications for a custom (within reason) version upgrade: special purpose filters, proprietary data formats, new features. Source code for FREQ is available, and many parts of FREQ can be exported as a Dynamic Link Library (DLL).

Registration

Registered users receive an upgraded version of FREQ, including (see below) a number of data set **Pre-Processing** and **Synthesis** functions you may find very useful in a work or research environment. More **Analysis** functions are provided, including a Fast Fourier Transform, and Approximate Entropy to estimate data set regularity. Phone assistance is also provided to registered users. Registration is \$40, payable in check, cash or M.O. to CoDebris, 711 Barbara Avenue, Solana Beach, CA 92075.

For a negotiable fee, any desired feature or customization tailored to your requirements can be incorporated. If the feature is of general utility, the fee may be small or even waived completely. We are very interested in enlarging the variety of data set formats FREQ can recognize: let us know if you want FREQ to read your spreadsheet or data base files or accept data from acquisition hardware: sound boards, DSP boards, A/D or digital I/O boards.

On registration we will either mail you a 3.5" floppy disk or send you e-mail with an attached FREQ.ZIP file containing the upgrade, and more tutorial data files with associated period tables.

If you run FREQ and like it, let us know (register it). If you run it and don't like something about it, let us know that too. FREQ is becoming a fairly comprehensive waveform analysis and synthesis package, and most features are added in response to user suggestions: in the works are digital filters, multi-variate data sets and non-linear dynamics tools.

In addition to everything in the shareware version, the registered version of FREQ (ver. 2.2) contains these additional functions:

Synthesize:

- Pre-Process Data Set...
- Interpolate
- Segment
- Smooth
- Remove Trend
- Remove Artifacts
- Subtract Mean
- Scale X or Y Data
- Offset X or Y Data
- RMS Partition
- Copy Current Data Set
- Two-Graph Operations...

Concatenate (Join) Data Sets
Add Data Sets Pointwise
Subtract Data Sets
Multiply Data Sets
Divide (Ratio) Data Sets
Reconstruct Data Set From Search Sines

Analyze:

Fast Fourier Transform
Inverse FFT
Approximate Entropy

These features are discussed further under the menu options Synthesize and Analyze.

The next upgrade to FREQ (ver. 2.4) will add these features:

Multi-variate data sets (e.g. simultaneous heart rate, EEG, cardiovascular output vs. time)

Log/linear, linear/log, log/log graphs

Windowing: Hamming, Hann

Filters: Bessel, Butterworth, Chebychev, Elliptical, Ideal
Low/High/Band Pass, Notch, Comb and Band Stop

Convolution, Correlation analyses

Non-linear analyses: Poincare plot, Second Return, Min Info, Lyapunov dimension

Full-resolution (not scaled bitmap) printed graphs, color optional, with settable titles, legends, and captioning.

Registrations received BEFORE July 1, 1995 can receive this version 2.4 upgrade deep-discounted during August. In July, the registration price will become \$60.

References ...and further reading

Ho, T.; Kwok, J.; Law J.; Leung (or Ljung), L. "Nonlinear system identification" ELEC-490 Project Rept., Dept. Elec. Eng., Queens Univ., Kingston, Ontario, Canada, 4/87.

Berry, Stephen J. , Daniel Franco, Craig Schwartz: "Characteristics of the FOS Method as Applied to Physiological Time History Analysis", Version 1.0 of FOS, Crump Inst. for Med. Engineering, UCLA, 1950 Sawtelle Blvd., Suite 330, Los Angeles, CA 90025 (213) 825-4111, June 23, 1988.

Korenberg, M. J. "Fast orthogonal identification of nonlinear difference equation and functional expansion models" Proc. 30th Midwest Symp. Cir. Sys. 1:270-276 8/87.

Korenberg, M. J. "A robust orthogonal algorithm for system identification and time-series analysis" Biol. Cybernetics, 60:267-276 1989.

Korenberg, M. J. "Identifying nonlinear difference equation and functional expansion representations: The fast orthogonal algorithm" Ann. Biomed. Eng. 16:123-142 1988.

Korenberg, M. J.; Bruder, S. B.; McIlroy, P. H. "Exact orthogonal kernel estimation from finite data records: extending Wiener's identification of nonlinear systems" Ann. Biomed. Eng., 1988.

Korenberg, M. J. "Orthogonal identification of nonlinear difference equation models" Proc. 28th Midwest Symp. Cir. Sys. 1:90-95 8/85.

Ljung, L. "System Identification: Theory for the User", Prentice-Hall, Englewood Cliffs, NJ 1987.

McIlroy, P. J. H. "Applications of nonlinear systems identification", M. S. Thesis, Queens Univ., Kingston, Ontario, Canada, 1986.

Pincus, Steven M. "Approximating Markov Chains", Proc. Natl. Acad. Sci. USA, 89:1-5, 5/92.

Lissajous Figures

Choose a sine wave graph as abscissa, and another, different sine wave graph as ordinate: from the two graph windows, click the upper left corner of one window using Select Data Set as Abscissa and the other using Select Data Set as Ordinate.

A new graph is created, where instead of the regular, monotonically increasing X values, the abscissa now is the Y value from the second graph. Striking swirling patterns called Lissajous Figures appear in the new window. They represent the interaction between two quantities which are not usually combined in this way.

Even more striking patterns are created by making a Fourier Transform from the new graph. Weird, but striking.

Frequency Search Introduction

See also:

[Fast Orthogonal Search](#)

[Preparing for a Frequency Search](#)

[Interpreting Frequency Search Results](#)

This is a new and greatly improved version of the FREQ Frequency Search and Data Synthesis program, version 2.1. Install FREQ in a convenient Program Group under Windows and click the icon to run, or select File/Run under Program Manager and fill in the blanks.

For a demonstration run, choose Analyze/[Frequency Search](#) and select a (supplied) table of candidate periods (.TBL) and a supplied data file (.DAT). The .TBL file can be edited and saved in FREQ.

REAL.DAT is a physiological dataset. The synthesized curve is not a bad approximation to the data, as you can see, even though only 55% of the "power" is accounted for by the seven identified periods, amplitudes and phases, analyzed in under 4 minutes.

The REAL5.DAT and REAL5.TBL data set is a "partitioned" version of REAL. It is formed from the RMS value of each 5 points from REAL, and was made using the Synthesize/[Pre-process Data Set](#)/RMS Partition menu option after choosing File/Open for REAL.DAT. In FREQ, the pre-processed data set is a copy of the original, the original is never unintentionally modified.

The time-stamped output file .OUT contains the selected periods, amplitudes and phases, and They can be input directly into FREQ for viewing or editing. Although the .OUT file is overwritten for each run on a given data file, a cumulative run summary FREQ.SUM file is also produced.

The summary file contains information relevant to a researcher for all runs over all data sets, including time & date. Edit or discard the .SUM file if it grows unwieldy. Full text editing for files smaller than 32kB is built into FREQ. Note that the actual .DAT data files are NOT directly editable as a FREQ menu option, but they may be altered in the pre-processor (menu selection Synthesize/Synthesize/[Pre-Process Data Set](#)). This allows outlier (artifact) removal, partitioning, segmentation, decimation, filtering, etc. to be applied to named data sets. Data files CAN be opened as text **FOR VIEWING ONLY** in an editable window by choosing "All Files (*.*)" from the

file type list box in the File Open dialog. But DO NOT try to save the data set from the text window: you will end up with a truncated data set file if its size was larger than the 32kByte limit imposed for text windows.

Edit Text Menu

Undo

Cut

Copy

Paste

Clear

Select All

Rename Data Set

X,Y Axis Names and Units

Edit/**Undo**

The effects of the last Edit operation are reversed. Last chance to recover from incorrect entries or deletes.

Edit/**Cut**

Highlighted text is removed from the window and placed on the Clipboard.

Edit/**Copy**

Highlighted text is copied to the Clipboard.

Edit/**Paste**

Text on the Clipboard is pasted to the text window at the location of the caret.

Edit/**Select All**

All text in the window is highlighted, including text outside the scroll region.

Edit/**Clear**

Clear the Clipboard of any text.

Edit/**Rename Data Set**

Give the active Graph Window data set a new title. It need not be a valid DOS file name, or have a .DAT extension; it can be mixed case and have spaces. The title will be printed with the graph, but you will be prompted for a valid file name in order to write the data set to disk (File/Save As).

Edit/**XY Axis Names and Units**

Choose names for the X and Y axes, and units to associate with their respective data.

This is a valuable technique to use to remind you of what you're looking at in a Graph Window. Be very careful with units, misconceptions and mistakes can and have ruined months of labor for unwary researchers.

Ain't that the truth, David?

File Menu

Data set and text file acquisition:

- New
- Open Data File...
- Open Period Table...
- Open Output File...
- Close

Data set and text file archival:

- Save
- Save As...
- Save Graph Bitmap
- Print Text
- Print Graph
- Printer Setup

Program Control

- Exit
- About FREQ...

File/**New**

Opens a blank text window, which can receive text from the Clipboard, be edited, and saved.

File/Open Data File

Open a data file into a new Graph Window, where it appears as a bright green plot over black background. The data set in any active (highlighted title bar) Graph Window is directly available for any processing option in FREQ.

After selecting the data file name, you are presented with a Data Type dialog, where you are asked to specify the data format. The usual FREQ format has data arranged in rows of X Y pairs. Viewing the data file as text can help determine its format (number of columns, identity of each column).

Data files are not intended to be directly editable as text, but a minor subterfuge will allow you to bring one in for viewing as a Text Window. Just open the data file using the File Filter list box option "All Files (*.*)" in the File/Open... dialog and click on the file name. Or, enter the data file name directly in the file name edit box.

FREQ will begin accepting proprietary binary data file formats in the next version. Users with immediate requirements can contact CoDebris for assistance.

Data Type dialog

On opening a time series data file, you are asked to specify whether the data is arranged in two columns as pairs of X Y numbers, or as a single column of Y data values.

In the first format, FREQ assumes the X value (the abscissa) precedes the Y value (the ordinate) on each row.

The other format consists of a single number, the Y value, in each row. Many research and engineering analysis instruments produce data files in this format; you are expected to know that the X values are uniformly spaced, and what values to use for the initial X value and the interval between points. If you aren't sure which values to use, use the defaults (initial X = 0.0, interval = 1.0) and proceed. In most cases, the X values determine what numbers to choose as candidate periods in the Period Table (see menu option Analyze/Frequency Search) when performing a Frequency Search. The important consideration is to be internally consistent in your choice. Additional information:

[Frequency Search Introduction](#)

[Preparing for a Frequency Search](#)

File/**Open Period Table**

Opens a Table of Candidate Periods (.TBL) into an editable Text Window. If edited, the table must be saved (File/Save) before it can be used in a Frequency Search, as it will be read in from disk before the search.

File/Open Output File

Text output files are produced by Frequency Search (.OUT, .SUM), by Pre-Process Data Set (.PRP), and by Generate Synthetic Data or Reconstruct Data Set from Search (.GEN). These are informational files only, and are not used directly in any time series processing. As they are time-stamped and reasonably detailed, they can be used as a record of processing steps.

File/**C**lose

Closes the active Text or Graph Window. If data has been altered in a Text Window, you are prompted to save it before closing. A Graph Window is never "dirty", as processing always produces a new, titled Graph Window.

File/**Save**

Writes the contents of the active Text or Graph Window to disk. If the window title is not a valid DOS file name, you are presented with the File/Save As dialog to enter or select a filename.

File/**Save As**

Presents a dialog to select or enter a file name under which to save the active Text or Graph Window.

File/**Save Graph Bitmap**

The currently active (highlighted title) Graph Window is saved to disk as an uncompressed bitmap, in the .BMP format suitable for pasting into word processors and other applications, like PaintBrush (many such applications do not know how to deal with the various compression modes). The saved bitmap is colored for printing: white background, black plot, gray axes and labels. The size of the saved bitmap will be identical to the size of the screen Graph Window, allowing you to tailor its size and aspect for insertion without distortion.

After saving, the bitmap appears over its parent Graph Window. To recover the original screen coloring, re-tile the screen, or maximize the Graph Window.

File/**Print Text**

Print the contents of the active Text Window. Graph Windows will be printable as scalable, color-optional bitmaps in the next version of FREQ.

File/**Print Graph**

The currently active (highlighted title) Graph Window is printed with a white background, black plot, gray axes and labels. The size of the printed bitmap can be identical to the size of the screen Graph Window, allowing you to tailor its size and aspect for printing without distortion.

A dialog to select printing options includes Best Fit, Stretch To Page, and independent X, Y Scaling. Both Best Fit and Stretch To Page will probably distort the graph's aspect. To preserve the aspect, or to choose a size on the printed page, use the Scaling option. As long as the X Scale and Y Scale are the same, no distortion will occur. For printing a maximized Graph Window to nearly fill the printed page, use Landscape Mode (see Printer Setup) if supported by your printer (virtually all Windows printers do). The upper left corner of the graph will always lie near the upper left corner of the page.

During and after printing, the printed bitmap appears over its parent Graph Window. To recover the original screen coloring, re-tile the screen, or maximize the Graph Window.

File/Printer Setup

The standard Windows Printer Setup dialog is presented. Depending on your printer, you can select Portrait or Landscape orientation, Black Density, Print Quality, and other aspects of your printer's behavior.

The next version of FREQ will support user-selectable options for color printing.

File/**Exit**

Normal program termination, restores Windows resources. The program System Menu/Close in the upper left corner has the same behavior.

Use Window/Close All Windows (Alt W A) to clear the workspace without exiting from the program.

File/**About** **FREQ...**

FREQ logo and copyright. Shows current Windows mode (Enhanced Mode required for FREQ). Displays remaining memory (including swap file) in kilobytes, current disk drive capacity in bytes, plus User and Graphics Display Interface system resources as a percentage of maximum. The presence or absence of a math coprocessor is noted.

When remaining memory approaches the size of the virtual memory swap file, disk activity (swapping) dominates and performance deteriorates.

When system resources go below about 60%, performance suffers, and ultimately some activities will cease occurring (e.g., window repaints.)

Graph Window

A Graph Window contains a plot of the X,Y values from a data set as a bright green line plot on a black background. If a Frequency Search was performed on the Graph Window, the superposed, reconstructed sine waves identified during the search are plotted as a yellow dotted curve on the same graph.

Text Window

A Text Window contains editable text. The text can be altered and saved to disk. If the text is a Period Table (.TBL file), then any edits made to the entries **MUST** be saved to disk before a Frequency Search is performed.

Windows Keys

To review the keys used in Windows:

Cursor Movement Keys

Dialog Box Keys

Text Selection Keys

Text Editing Keys

Help Keys

Menu Keys

System Keys

Window Keys (including FREQ windows.)

Cursor Movement Keys

<u>Key(s)</u>	<u>Function</u>
DIRECTION key	Moves the cursor left, right, up, or down in a field. Moves the graph cursor across a graph window.
End or Ctrl+Right Arrow	Moves to the end of a field.
Home or CTRL+Left Arrow	Moves to the beginning of a field.
PAGE UP or PAGE DOWN	Moves up or down in a field, one screen at a time.

Dialog Box Keys

<u>Key(s)</u>	<u>Function</u>
TAB	Moves from field to field (left to right and top to bottom).
SHIFT+TAB	Moves from field to field in reverse order.
ALT+letter	Moves to the option or group whose underlined letter matches the one you type.
DIRECTION key	Moves from option to option within a group of options.
ENTER	Executes a command button. Or, chooses the selected item in a list box and executes the command.
ESC	Closes a dialog box without completing the command. (Same as Cancel)
ALT+DOWN ARROW	Opens a drop-down list box.
ALT+UP or DOWN ARROW	Selects item in a drop-down list box.
SPACEBAR	Cancel a selection in a list box; selects or clears a check box.
CTRL+SLASH	Selects all the items in a list box.
CTRL+BACKSLASH	Cancel all selections except the current selection.
SHIFT+ARROW key	Extends selection in a text box.
SHIFT+ HOME	Extends selection to first character in a text box.
SHIFT+ END	Extends selection to last character in a text box.

Editing Keys

Key(s)

Function

Backspace

Deletes the character to the left of the cursor.
Or, deletes selected text.

Delete

Deletes the character to the right of the cursor.
Or, deletes selected text.

Help Keys

Key(s)


Function

F1

Gets Help and displays the Help Index for the application. If the Help window is already open, pressing F1 displays the "Using Windows Help" topics.

In some Windows applications, pressing F1 displays a Help topic on the selected command, dialog box option, or system message.

SHIFT+F1

Changes the pointer to  so you can get Help on a specific command, screen region, or key.

You can then choose a command, click the screen region, or press a key or key combination you want to know more about. FREQ changes the pointer only when a menu selection is highlighted.

Menu Keys

<u>Key(s)</u>	<u>Function</u>
Alt	Selects the first menu on the menu bar.
Letter key	Chooses the menu, or menu item, whose underlined letter matches the one you type.
Alt+letter key	Pulls down the menu whose underlined letter matches the one you type.
LEFT or RIGHT ARROW	Moves among menus.
UP or DOWN ARROW	Moves among menu items.
Enter	Chooses the selected menu item.

System Keys

The following keys can be used from any window, regardless of the application you are using:

<u>Key(s)</u>	<u>Function</u>
ALT Hyphen	Opens the System (Control) Menu for a document window.
Alt+Esc	Switches to the next application window or minimized icon, including full-screen programs.
Alt+TAB	Switches to the next application window, restoring applications that are running as icons.
Alt+PrtSc	Copies the entire screen to Clipboard.
Ctrl+Esc	Switches to the Task List.
Ctrl+F4	Closes the active window.
F1	Gets Help and displays the Help Contents for the application. (See <u>Help Keys</u>)

Text Selection Keys

<u>Key(s)</u>	<u>Function</u>
SHIFT+LEFT or RIGHT ARROW	Selects text one character at a time to the left or right.
SHIFT+DOWN or UP	Selects one line of text up or down.
SHIFT+END	Selects text to the end of the line.
SHIFT+HOME	Selects text to the beginning of the line.
SHIFT+PAGE DOWN	Selects text down one window. Or, cancels the selection if the next window is already selected.
SHIFT+PAGE UP	Selects text up one window. Or, cancels the selection if the previous window is already selected.
CTRL+SHIFT+LEFT or RIGHT ARROW	Selects text to the next or previous word.
CTRL+SHIFT+UP or DOWN ARROW	Selects text to the beginning (UP ARROW) or end (DOWN ARROW) of the paragraph.
CTRL+SHIFT+END	Selects text to the end of the document.
CTRL+SHIFT+HOME	Selects text to the beginning of the document.

Window Keys

<u>Key(s)</u>	<u>Function</u>
ALT SPACEBAR	Opens the Control menu for an application window.
ALT Hyphen	Opens the Control menu for a document window.
Alt+F4	Closes the FREQ main (frame) window.
Ctrl+F4	Closes the active MDI graph or text window.
Alt+Esc	Switches to the next application window or minimized icon, including full-screen programs.
Alt+TAB	Switches to the next application window, restoring applications that are running as icons.
Ctrl-F6, Ctrl+TAB	Switches to the next FREQ graph or text window, "sinks" the formerly active window to the bottom of the stack of open windows.
Alt+ENTER	Switches a non-Windows application between running in a window and running full screen.
DIRECTION key	Moves a window when you have chosen Move from the Control menu. Or, changes the size of a window when you have chosen Size from the Control menu.

Program Navigation

Finding a comfortable set of methods to use with complex software is usually a trial-and-error process, as documentation is usually specific to a given task and doesn't offer a section which looks at the software from the user's viewpoint. Here is FREQ from the keyboard, mouse button and cursor point of view:

[Keyboard Mouse Emulation](#)

[Graph Cursor](#)

[Graph Coordinates Display](#)

[Activating Graph Windows](#)

[Graph Window Sizing](#)

[Graph System \(Control\) Menu](#)

[MDI Keys](#)

[Running FREQ](#)

[Diagnostic Sounds](#)

Graph Cursor



A special cursor appears automatically over a Graph Window as the mouse is moved around. It is a cross-hair type, but with an open region in the center to allow precise alignment over any pixel. As long as the mouse left button or a keypad cursor key is depressed, the cursor cannot leave the plot surface, allowing you to get into corners or right up to plot edges without triggering the system move/size cursors.

Running FREQ

As with most new programs, use the defaults first time through and play with the options as experience accumulates. Using FREQ might be an educational experience, giving you great control over many standard time series processing elements, and introducing some new ones. See [Frequency Search Introduction](#) and [Approximate Entropy](#) for examples.

After loading a data set ([File/Open Data File](#), or creating a synthetic data set with [Synthesize/Generate Synthetic Data](#), exercise all options which seem even remotely applicable to your line of inquiry, and discard unsatisfactory intermediate graphs. Time series processing is often accompanied by surprises which yield real insight into your data, and is a whole lot more fun than word processing.

Save interesting results to file often for comparison with later results. You may want to create an empty subdirectory to ease cleanup and discarding of intermediate results which have been superseded by improved versions. If you don't want to use the built-in file naming convention, choose **unique** filenames which will allow easy cleanup: TT1.DAT, TT2.DAT, TT2.GEN, etc. are deleted with a single command DEL TT*.*.

Use the [Help/About FREQ...](#) menu option often while in the menu system.

Graph Window System (Control) Menu

In addition to the Main Menu, each graph window contains a so-called System, or Control, Menu accessible by clicking the box in the window upper left corner, next to the window title bar. The normal entries Restore, Move, Size, Maximize, Close, Next Window are joined by some special options incorporated in FREQ to allow the user to associate two data sets with one another:

Select Data Set as Abscissa
Select Data Set as Ordinate
Use Data set as Real Part for FFT
Use Data set as Imaginary Part for FFT.
Select Data Set as Operand 1
Select Data Set as Operand 2

The first special options are

Select Data Set as Abscissa
Select Data Set as Ordinate

These features enable the user to combine two data sets, using one as the abscissa, or independent variable, and the other as the ordinate, or dependent variable. Click the System Menu in the upper left corner of each data set graph window, click again on "Select Data Set As Abscissa" or "Select Data Set As Ordinate". As soon as a pair has been selected, a new window will be opened showing the relationship. In general, a positive (upward) sloping elliptical blob of green line segments implies a positive correlation between the two data sets. A downward-sloping trend indicates anti-correlated sets. A more or less uniform disk indicates a lack of correlation. See Association and Visualization for more information.

In a similar manner,

Use Data set as Real Part for FFT
Use Data set as Imaginary Part for FFT

are used to make a complex data set (one having both real and imaginary parts) from two real data sets for use in running the Analyze/Fast Fourier Transform. The FFT is run automatically as soon as a real, imaginary pair of

data sets has been selected.

The remaining special options

Select Data Set as Operand 1

Select Data Set as Operand 2

select data sets on which to perform the Two Graph Operations:
Concatenation, Addition, Subtraction, Multiplication, or Division.

Keyboard Mouse Emulation

With no mouse, most processing is difficult, but possible. The keypad arrow keys will move the cursor over a graph. For mouse emulation, make sure the Num Lock key is not active (unlit).

The keypad will also permit precise one-pixel-at-a-time position adjustments over a graph window when mouse movement is awkward or mouse response is "sticky". The rate of movement increases as an arrow key is held down. The arrow keys can always be used in conjunction with the mouse for greater positional control.

By the way, all menu responses can be invoked from the keyboard as an "Alt key sequence": Window/Tile Horizontally is selected with Alt W H. The mnemonic letters are not always initials on pop-up menus with several options, as naming conflicts often result. The relevant key code is underlined on each menu entry.

Several Menu selections also have dedicated "virtual keystrokes". Ctrl X, for example, is usually used to Edit/Cut highlighted text from a text window.

Graph Coordinates Display

Select Window/Display Graph Coordinates. Move the mouse over any graph window, and the x,y coordinates with corresponding pixel position can be noted. The keyboard cursor keypad can be used to move in precise increments across the graph. The X,Y values displayed are, of course, the data set values.

If Edit Text/X,Y Axis Names and Units was selected, the X,Y axes names and the physical units associated with each axis are also printed in the Coordinates box.

Right Mouse Button

In all situations, the right mouse button either has no effect, or produces the same response as the left button. Several features were originally controlled with the right button, but experience using FREQ showed this was not desirable or necessary.

MDI Keys

As a Multiple Document Interface (MDI) application, FREQ uses the Ctrl-key for some functions normally executed with the Alt-key. Ctrl-Tab will activate each open graph window successively; this is useful if a graph is covered by one or more others, as the "on top" windows don't have to be moved to see the underlying graph. Alt-Hyphen (the minus sign key) accesses the active graph window's Control Menu, usually accessed by clicking on the small box in the upper left window corner (Alt-Spacebar is used for all other windows. See [Help/Keyboard](#) for more information on using the keyboard in Windows applications.

Activating Graph Windows

When more than one graph window is open on-screen, clicking anywhere over the window surface is sufficient to activate it. Clicking over the Title Bar (also called the Caption Bar) and dragging the mouse is used to move a window.

Try **Ctrl-Tab** to activate graph windows in succession to "thumb through" a screenful of datasets. It can bring up graph windows from the bottom of the "z-order" window stack, and "sinks" the formerly active window to the bottom. See Window Keys.

Graph Window Sizing

You can, of course, drag the border edges or corners to resize a graph window. FREQ will completely recalculate the plot to optimize the information displayed. Keep in mind that a large data set may have many more (or quite a few less) points than there are horizontal pixels on the screen: you can't expect to see every datum unless you slice up some data sets. The Synthesize/ Pre-process/Interpolate option can help here.

The keyboard can also be used to move or re-size graph windows. See Help/Keyboard for more information on using the keyboard in Windows applications.

Diagnostic Sounds

Sounds from your system speaker during a FREQ session are usually intended to tell one of two things: either the system could not satisfy a request of some sort from you or the program, or something needs to be done by you and you're not doing it.

The occurrence of a "beep" from the speaker often indicates an action needs to be accomplished before you can move on. Maybe a dialog box or message window needs its "OK" or "Cancel" button pressed. Usually, hitting the keyboard "Enter" key will suffice, as the default option is the one you'd normally take, and it is highlighted with a small, dotted rectangle around the text printed on the key.

A sound like a "chirp", with some variation in the tone and duration, is used by FREQ to indicate a serious degradation in performance. Either too many system resources are in use, or requests for memory are failing. The solution is commonly to shut down any other applications running in background. If things have sunk low enough (windows are not being painted, are broken up or do not appear when expected), save your work, exit FREQ, close down everything else and exit Windows itself.

Monitor the status of your system memory, hard drive, and resources occasionally from the Help/About FREQ... menu selection, especially if operations are getting to be slow and uneven.

Synthesize Menu

Synthesis is the process of creating a data set from components, including components derived from analysis (see the Anlayze menu) or from another data set. Products from any of the synthesis operations below are a Graph Window plotting the data set, an optional data file of X,Y pairs in ASCII format (use File/Save As to create the data file), and (be advised) a reference text file FREQ.SYN is continually appended to, containing summary information on every synthesized data file: it's useful to review what went where, but edit it or delete the file if it gets too large.

Pre-Process Data Set
Copy Selected Data Set
Generate Synthetic Data
Two-Graph Operations
Reconstruct Identified Sines

Reasons for synthesizing a data set might include creating a noise-free test bed for modeling, "loading" a data set with noise to test extraction and filtering responses, or testing assumptions about the behavior of real data sets before they actually become available.

Synthesize/**Generate Synthetic Data**

FREQ can generate a data set, or "time series" from superposed sine waves, square waves, sawtooth waves, and 'squirt' waves.

You can add a constant (so-called 'DC', for Direct Current) offset, a linear trend (a sloping, linear rise or fall along the entire data set) and uniformly distributed noise.

These are the **Generate Synthetic Data** dialog entries:

Data set title	Free format, up to 32 characters.
# of Data Points (2-200000)	Use 1025 if you want 1024 intervals.
Initial X value (abscissa, + or -)	Usually 0, but any X offset is accepted.
Interval Between Points (> 0)	The difference between adjacent abscissa X values.
# of Sines to Generate (0-99)	Specify amplitude, period, phase for each.
# of Square Waves (0-99)	Specify Amplitude, period, phase for each.
# of Sawtooth Waves (0-99)	Specify amplitude, period, phase and positioning of peak.
# of Squirt Waves (0-99)	Specify amplitude, period, percentage of period for "pump".
Noise Amplitude	Uniformly distributed.
Specify Constant Offset (+ or -)	Also called DC offset.
Specify linear trend (+ or -)	Choose value to use as right endpoint, the amount by which to "tilt" data set.
Subtract Mean	After generating data set, remove the mean.
Shuffle Data Set	Randomly mix the X,Y pairs: 13th point might become 211th. Used for statistical tests.

All amplitudes (for waves, noise, Y offsets) are assumed in the same units, your choice. Ditto for periods. Remember that although the terms 'time' and 'time series' are used, the X variable in a (X,Y) data set can be any quantity on which the Y values, in some sense, depend.

However, be VERY careful over your choice of units (degrees, radians; or seconds, minutes, heartbeats)... it matters greatly in your analysis. Strive for internal consistency during a FREQ session, and throughout your work.

Try "walking through" a data processing session on paper, using printout of computer input and output files for raw material: check that numerical answers are appropriate to your assumptions. FREQ can be of great assistance here, as you can design your own dataset in the **Synthesize/Generate Synthetic Data** dialogs.

Many aspects of your data can be simulated and run through a Frequency Search or Fast Fourier Transform. You can learn a lot this way about nuts and bolts stuff: units, intervals, amplitudes, bandwidths, sampling rates, etc.

Synthesize/**Pre-Process Data Set**

Although not as important for a Frequency Search, pre-processing a time series can greatly affect the validity and interpretation of a Fast Fourier Transform.

On selection, the following pre-processing capabilities are listed in a sub-menu:

Interpolate

Smooth

Remove Trend

Remove Artifacts

Subtract Mean

Introduce X or Y Offset

RMS Partition

RMS Partition

RMS means root mean square: in a partition, sum the squared numbers, divide by one less than the partition size, take the square root. A partition is a piece of the data set: a 1000-point data set may be partitioned into 200 segments, each containing 5 sequential points.

The user is prompted for the number of points in each partition segment, 5 in this example.

The data set being partitioned is elevated so that its lowest value is on the X axis: no negative values, as the process of squaring each data point would cause data tending to go negative to suddenly sweep upward.

The RMS value of each (for example) 5 point segment is found, and the RMS values are used to make a new data set. The resulting data set is a somewhat "chunkier" distillation of the original. It finds increasing use among researchers attempting to capture large amounts of data over a long recording time.

Interpolate

Creates a new data set composed of all or part of an existing one. You can have more, fewer, or the same number of points as in the original. You choose the lower limit and upper limit X values to include in the interpolation from a dialog where the selected data set X limits and existing number of points is already in the edit control.

Interpolation is a nice way to **align** two data sets so they have the same minimum and maximum X values, and the same number of points. Aligning two data sets provides a valid method for choosing a pair of data sets to be the abscissa and the ordinate of a new graph window for visualizing the inter-relationship between them (see Select Data Set as Abscissa and Select Data Set as Ordinate). If the sets are not aligned, part of the longer data set is not included in the visualization, and the X values may not even lie in the same ranges. Your responsibility to use caution here is emphasized. A future update to FREQ will automate the alignment process to a couple of keystrokes or clicks, but for now you have to do some thinking in advance.

Note that selecting the entire data set for interpolation is not exactly the same as making a copy unless you use exactly the same number of points: it is a linear interpolation of those values, and will be very close, but to get an exact copy use Synthesize/Copy Current Data Set.

The X lower and upper limit values you choose must of course be part of the selected data set, or you're trying to do extrapolation (prediction) and not interpolation. Prediction is not a FREQ option at this time. There are several approaches for predictive models available; if you have a preference, let us know..

Smooth

Applies a "boxcar" moving averager to the selected data set. You choose the number of points over which to average. More subtle averagers are available, but most analysts use this technique for a first try at smoothing their data to remove high-frequency fluctuations.

For a 5-point boxcar, a given point is averaged with the two preceding and the two following points.

Remove Trend

The "tilt", or average slope, of a data set is calculated and used to subtract the "least-squares" best-fit straight line from the data set Y values. This will not necessarily result in a zero mean. The existence of linear trends is often an indicator of non-stationarity, which means that statistical parameters such as the mean and standard deviation are not constant. Non-stationarity can imply a meaningful "transition" between two states, or it can mean an uncompensated "drift" in electronics or data collection equipment. You should attempt to understand the source of a trend in your data: an uncompensated drift may eventually bump all your data into a dynamic range limit, or a "transition" may have been caused by the system being measured non-linearly crossing into a new behavioral regime. In FREQ, the slope and Y-intercept of the line is calculated from a "least-squares" fit to the data:

$$\text{Slope} = (\{X Y\} - \{X\} \{Y\}) / (\{X X\} - \{X\} \{X\})$$

$$\text{Intercept} = \{Y\} - \text{Slope} \{X\}$$

where $\{ \}$ is an average over the length of the data set.

Subtract Mean

The mean (averaged over the data set) value of Y is subtracted from each data set Y value.

Introduce X or Y Offset

The selected data set is shifted horizontally or vertically (or both at once) to a new origin. You choose the amounts (+ or -) by which to translate the data set.

Remove Artifacts

Artifacts, or "outliers", are sudden positive- or negative-going spikes in the data. You choose a relative level, in percent, above (or below) which artifacts will be "clipped" from the data set: artifacts extending more than, say, 200% above or below data values on either side are replaced by the average of the side values.

Introduce Jitter

A controlled amount of "jitter", or noise, in the X coordinates of your data set can be specified, as a percentage of the X interval (the distance between adjacent X values).

X values will still be monotonic, but not in a uniform sequence.

Copy Selected Data Set

A one-to-one exact copy of the selected data set is made and displayed as a graph window. It is assigned the data set name PREPxx, and is NOT saved to disk.

Synthesize/**Two-Graph Operations**

Concatenate, Add, Subtract, Multiply, Divide two data sets.

Select two data sets for the Two Graph Operation by opening the Graph System Menu from the square button at the upper left corner of each Graph Window. Then choose **Synthesize/Two-Graph Operations** from the main menu and select the specific action from a pop-up dialog.

Concatenate simply glues two data sets together at the end of the first one, with a "bridge" interval between them. No attempt is made to "smooth" the joint. Concatenation permits data sets to be "re-assembled" after having been separated for independent processing. The data sets may have different X-interval values; the "bridge" value between the final element of the first data set and the initial element of the second data set is the X-interval from the second data set. Each data set may be independently scaled by constant multipliers (see below) during the concatenation.

The four arithmetic operations are of the form

$$C1 \text{ [Data Set 1] } \text{ op } C2 \text{ [Data Set 2]}$$

where C1 and C2 are user-selectable constant multipliers, and "op" is one of +, -, x, or /. The data sets are checked for a region of overlap in their X values, and the operation is performed only on the overlap region. If the data sets share no common X values, no operation is performed. The number of points in the resultant data set is the maximum of the number of points from one of the data set's overlap region with the other. Data Set 1 is also called "Operand 1", Data Set 2 is "Operand 2".

For **division**, if the denominator has zero values in the region, the quotient is assigned the Y value of zero. Also, for **division** and **subtraction** note that Operand 2 is divided into (or subtracted from) Operand 1.

Synthesize/**Reconstruct Data Set from Search**

You can use the period, amplitude and phase offset of each identified sine wave from a Frequency Search to make a new, 'synthetic' data set, virtually automatically. This may help assess the effect of noise in real data sets, by creating a data set with the same overall structure, but noise-free. Make sure the Graph Window on which the Frequency Search was performed is active, by clicking the mouse anywhere over it or by hitting Ctrl-Tab repeatedly until the graph window is highlighted. Then select the **Synthesize/Reconstruct Data Set from Search** menu option. A new graph window appears, holding the data set comprised only of sine waves whose periods were selected from the Frequency Search.

Select Data Set as Abscissa

See also [Select Data Set as Ordinate](#)

These features enable the user to associate two data sets, using one as the abscissa, or independent variable, and the other as the ordinate, or dependent variable. Click the System Menu in the upper left corner of each data set [Graph Window](#), click again on "Select Data Set As Abscissa" or "Select Data Set As Ordinate". As soon as a pair has been selected, a new window will be opened showing the relationship. In general, a positive (upward) sloping elliptical blob of green line segments implies a positive correlation between the two data sets. A downward-sloping trend indicates anti-correlated sets. A more or less uniform disk indicates a lack of correlation.

Select Data Set as Ordinate

See also [Select Data set as Abscissa](#)

These features enable the user to associate two data sets, using one as the abscissa, or independent variable, and the other as the ordinate, or dependent variable. Click the System Menu in the upper left corner of each data set [Graph Window](#), click again on "Select Data Set As Abscissa" or "Select Data Set As Ordinate". As soon as a pair has been selected, a new window will be opened showing the relationship. In general, a positive (upward) sloping elliptical blob of green line segments implies a positive correlation between the two data sets. A downward-sloping trend indicates anti-correlated sets. A more or less uniform disk indicates a lack of correlation.

See an interesting application in [Lissajous Figures](#), with which anyone who has ever played with an oscilloscope will be familiar.

Select Data Set as Real Part

See also [Select Data set as Imaginary Part](#)

These features enable the user to associate two data sets, using one as the real part, and the other as the imaginary part, of a "complex" data set. Click the System Menu in the upper left corner of each data set [Graph Window](#), click again on "Select Data Set As Real Part" or "Select Data Set As Imaginary Part". As soon as a pair has been selected, a [Fast Fourier Transform](#) is automatically performed on the complex data set.

Select Data Set as Imaginary Part

See also [Select Data set as Real Part](#)

These features enable the user to associate two data sets, using one as the real part, and the other as the imaginary part, of a "complex" data set. Click the System Menu in the upper left corner of each data set [Graph Window](#), click again on "Select Data Set As Real Part" or "Select Data Set As Imaginary Part". As soon as a pair has been selected, a [Fast Fourier Transform](#) is automatically performed on the complex data set.

Select Data Sets as Operands

Select Data Set as Operand 1 and **Select Data Set as Operand 2** are special options from the Graph Window System Menu (button at upper left corner of a graph window), and are used to perform Two-Graph Operations on data sets: concatenation, point-wise addition, subtraction, multiplication, and division.

Window Menu

Window Control options are standard for Windows. Open windows are listed on the menu, and the active window has a check mark. An inactive windows can be activated by selecting its title from the menu.

Window Control

Tile Horizontally

Tile Vertically

Cascade

Arrange Icons

Close All Windows

Window information

Display Graph Info

Display Graph Coordinates

Window/**Tile Horizontally**

Arranges windows one above the next, each extending across the screen. This is the default mode for displaying graph and text windows in **FREQ**; the "newest" window is painted at the top.

Window/**Tile Vertically**

Arranges windows one to the left of the next, each extending down the screen. Several windows will tile in an overlapping rectangular array.

Window/**Cascade**

Arrange open windows along a left-to-right descending diagonal, leaving the title bar of covered windows showing.

Window/**Arrange Icons**

Arrange graph and text window icons along the screen bottom row. A window is 'iconized' by clicking on the small "down arrow" square near the window upper right corner, and when iconized can be restored to its former size by double-clicking on the icon.

Window/**Close All Windows**

Close all open windows, cleaning up all resources. Returns FREQ to its start-up configuration.

This selection DOES NOT query to save each open graph and text window before closing it, but asks if you want ALL the windows closed.

Window/**Display Graph Info**

A selection of data set information is displayed for the currently active Graph Window: data set title, full-path filename, number of data set points, X,Y extrema, Y mean and RMS (root mean square) values, first and last X,Y values, interval between the 1st and 2nd X points, and other items.

While Display Graph Info is active, click over other Graph Windows to view their information.

Window/**Display Graph Coordinates**

When selected, this option displays the data set X,Y coordinates corresponding to the cursor position whenever the cursor is over a graph window.

The Axis names and units, if chosen, are also shown, as well as the actual pixel coordinates under the cursor relative to the graph upper left corner.

Note that the keypad cursor keys can also move the graph cursor one pixel at a time over the graph. Holding down the cursor key accelerates the motion.

Help Menu

[Contents](#)

[Keyboard](#)

[Using Help](#)

[About FREQ](#)

While in the Help system, also try Help Search, with over 200 entries.

Help Contents

Presents a listing of hypertext topics for every menu selection, and a number of topics of interest to FREQ users. See [Contents](#).

Help/**Keyboard**

Standard Windows key assignments, including Multiple Document Interface (MDI) key assignments used in FREQ.

To see help sections for the keyboard now, click on this topic: [Windows Keys](#).

Help/**Using Help**

Help on navigating the Windows Help system. Well worth a pass, even if you feel you know your way around.

Help Bar

Help/**About** **FREQ**...

FREQ logo and copyright. Shows current Windows mode (Enhanced Mode required for FREQ). Displays remaining memory in kilobytes, current disk drive capacity in bytes, plus User and Graphics Display Interface system resources as a percentage of maximum. The presence or absence of a math coprocessor is noted.

When remaining memory approaches the size of the virtual memory swap file, disk activity (swapping) dominates and performance deteriorates.

When system resources go below about 60%, performance suffers, and ultimately some activities will cease occurring (e.g., window repaints.)

