

## PRINCIPLES OF OPERATION

Most ordinary sounds are complex combinations of individual frequency components or harmonics with a wide range of frequency and intensity. A spectrogram is simply a plot of the frequency components of such an audio signal as a function of time. In this Spectrogram program, digital audio recordings (PCM format) are analyzed to produce a plot of frequency versus time, with harmonic intensity represented by a variable color scale. These spectrograms reveal the fascinating hidden frequency structure of audio signals and can be used for identifying or classifying particular sounds.

Spectrogram uses a mathematical Fast Fourier Transform (FFT) to perform the frequency analysis. FFT's are usually specified by the number of input data points used in each calculation, which are always powers of two (512, 1024, 2048, etc). The frequency resolution of the spectrogram is always the digital sampling rate of the audio signal divided by the number of FFT data points. The greater the number of FFT data points, the finer the frequency resolution of the spectrogram. The maximum frequency computed by the FFT and the upper frequency limit of the spectrogram will be half the digital sampling rate.

The choice of sampling rate depends entirely on the highest frequencies in the audio signal. The rule of thumb is to use a sampling rate that is twice the highest frequency in the audio signal. That is, if you expect to have no frequency components above 11KHz, then a sampling rate of 22KHz is adequate. If you examine a spectrogram and see that all of the signal is concentrated in lower frequency components at the bottom of the display then it is a good bet that the recording was sampled at too high a rate, wasting a significant amount of memory. By varying the sampling rate and the number of FFT input data points, the frequency resolution and frequency span of the spectrogram can be chosen to best fit the audio signal of interest. Don't fall into the habit of automatically recording everything at a 44 KHz sampling rate. Lower rates often result in much better spectrogram displays

Spectrogram provides two basic modes of operation, "Analyze" and "Scan." The Analyze File and Analyze Input modes store the audio signal and spectrogram display bitmap in RAM and allow manipulation of the analysis parameters to achieve the best possible display of the data. To record a wave file using your sound card, use the Analyze Input mode for simultaneous spectrum analysis and recording. The Scan File and Scan Input modes do not store the audio signal or display bitmap and so do not allow manipulation of the analysis parameters. However, scanning of either a data file or audio input will provide a real-time high resolution display of audio data of unlimited length.

## THE SPECTROGRAM DISPLAY

The spectrogram display reveals the audio signal as a frequency versus time plot with signal amplitude at each frequency represented by intensity (or color). The display can be configured for either dual channel or single channel audio with a wide selection of frequency resolutions and either linear or logarithmic frequency scales. In dual channel operation, the spectrogram window is split into left and right halves with separate scrolling spectrograms for the left and right audio channels.

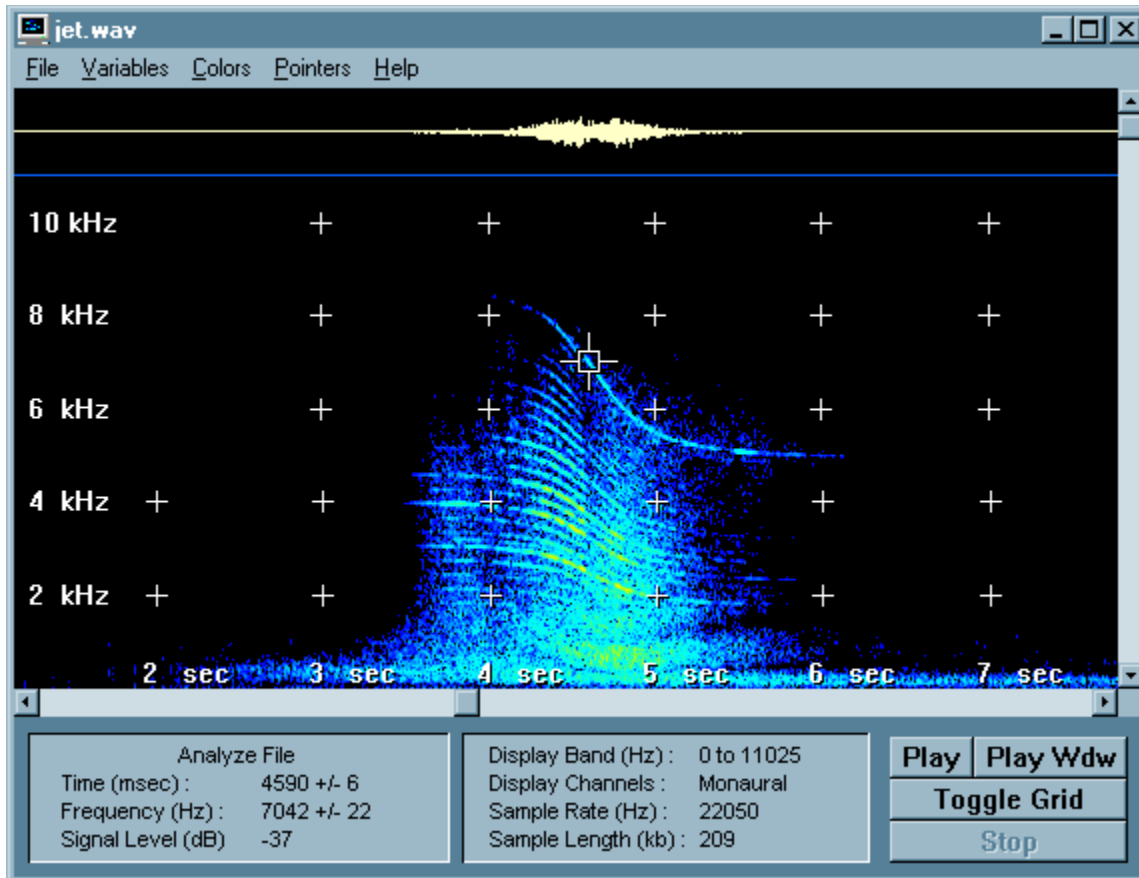
A continuous readout of time (milliseconds), frequency (Hz) and signal level (dB) at the position of the mouse pointer (cursor) is displayed at the bottom left of the display. A coordinate grid can also be added or removed by clicking the "Toggle Grid" button at the bottom right of the display.

In the Analyze modes, Spectrogram can play back the audio sample through your sound card when you click the "Play" or "PlayWdw" buttons at the bottom right of the display. PlayWdw replays only the segment of the spectrogram visible in the display window; whereas, Play replays the entire width of the spectrogram.

The width of the spectrogram display is limited only by the display screen. Maximizing the spectrogram window will expand the display horizontally to fill the screen. If the spectrogram width is greater than display width, you can use the horizontal scroll bar at the bottom of the display to position the spectrogram side-to-side.

The height of the spectrogram display is determined by the mode of operation and the size of the FFT chosen for analysis. The Analyze Input, Scan File, and Scan Input modes of operation each allow a maximum display height of 256 frequency points. These modes must update the scrolling spectrogram display in real time which limits the maximum allowable display height. The Analyze File mode of operation is used for very detailed frequency analysis and does not require real-time display update. The maximum height of the Analyze File mode is 1024 frequency points. If the spectrogram height is greater than the display height, you can use the vertical scroll bar at the right of the display to position the spectrogram top-to-bottom.

To maximize the Spectrogram window to fill the entire screen, choose the Analyze File mode, select a 1024 point FFT (or greater) and then maximize the Spectrogram window. Only by selecting an FFT of 1024 points or greater will the display fill the entire screen.



## THE SCOPE DISPLAY

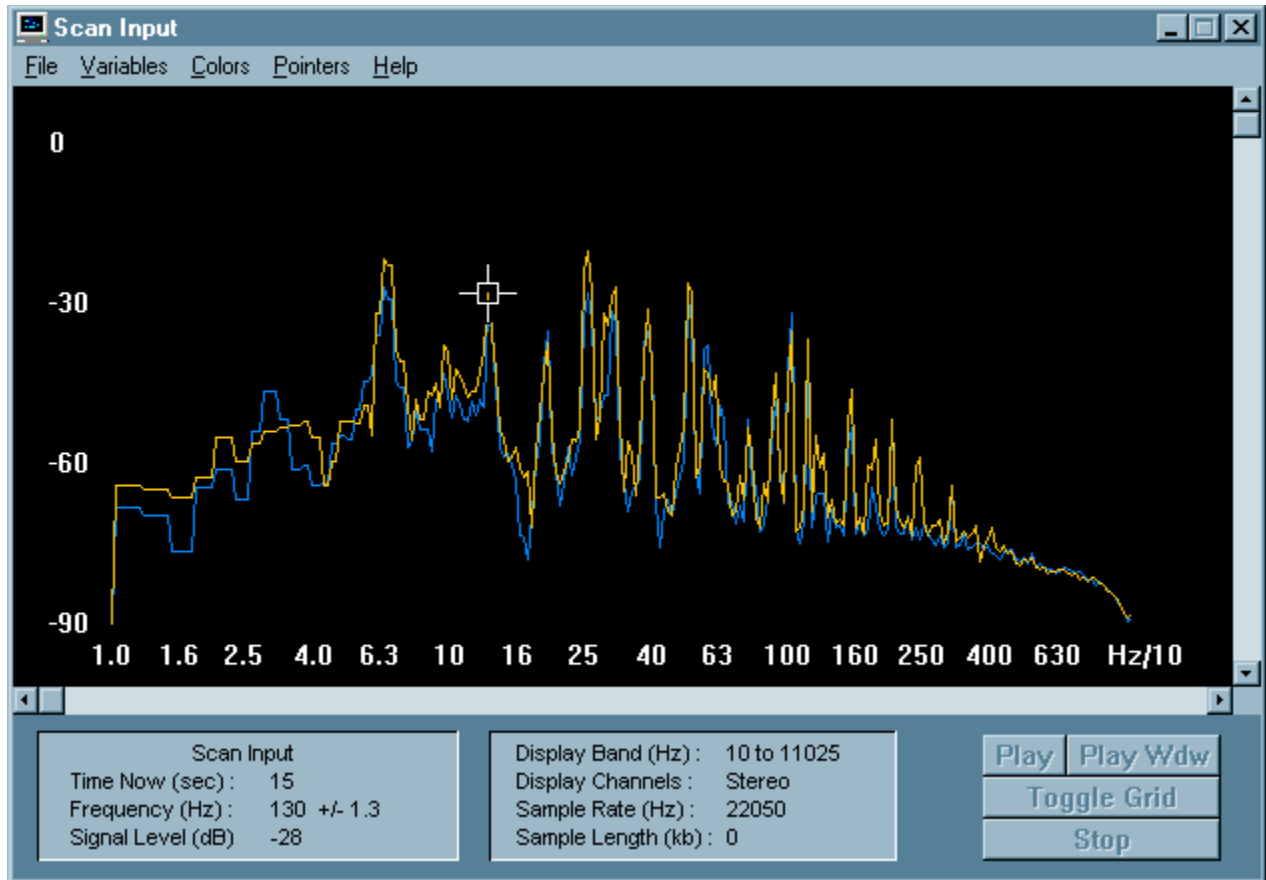
The Analyze Input, Scan File, and Scan Input modes also provide a spectrum analyzer scope display as an option for viewing the sound spectrum. The scope display can be configured for either dual channel or single channel audio with a wide selection of frequency resolutions and either linear or logarithmic frequency scales. In dual channel operation, the left channel data is plotted in blue and right channel data is plotted in yellow allowing an evaluation of the spectral differences between the two channels.

The scope display can be configured as either a histogram bar display where component amplitude is represented at each frequency by a vertical bar, or as a continuous line plot of component amplitude at each frequency.

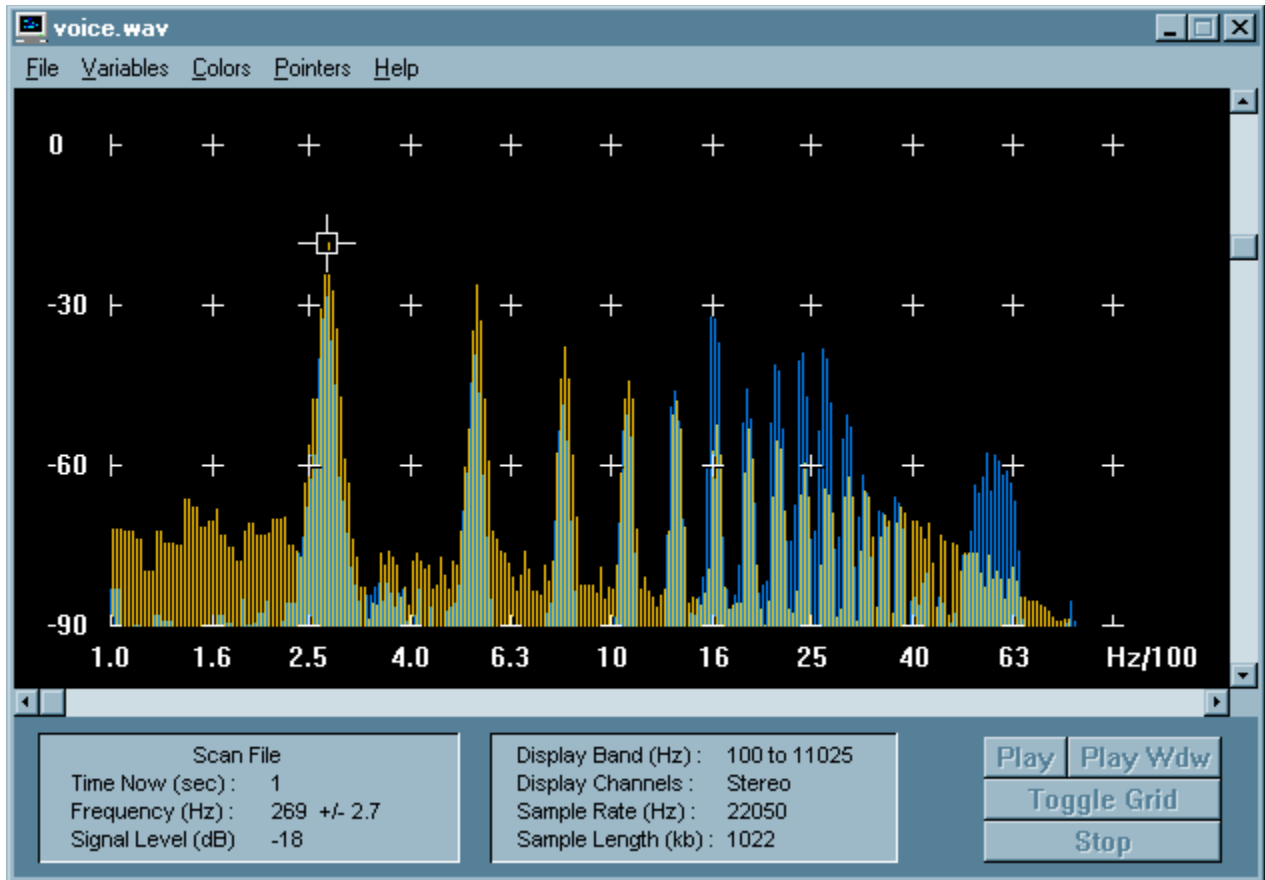
A continuous readout of time (milliseconds), frequency (Hz) and signal level (dB) at the position of the mouse pointer (cursor) is displayed at the bottom left of the display. A coordinate grid can also be added or removed by clicking the "Toggle Grid" button at the bottom right of the display.

You can calibrate the scope display using the vertical scroll bar at the right side of the display window. The scope display shows a fixed 90 dB signal range. However, many interacting factors will affect the absolute level of frequency components shown on the display. While the relative levels of frequency components to each other will always be correct, the gain of your sound card and the combination of frequency analysis parameters selected will affect absolute measurement. Use the vertical scroll bar to adjust the amplitude of a known frequency component to its exact value. Then absolute measurement of all other components will also be correct.

### Spectrum Analyzer Continuous Line Scope Display



Spectrum Analyzer Bar Scope Display



## **LINEAR AND LOG FREQUENCY SCALES**

Both the Spectrogram Display and the Scope Display have the option of displaying data with either a linear or a logarithmic frequency scale. The linear scale divides the frequency axis into equal intervals of frequency, where the logarithmic scale divides the frequency axis into equal intervals of the logarithm of frequency. The log frequency scale gives more prominence to low frequencies by expanding the display space for low frequencies at the expense of high frequencies.

The linear scale is less of a computational load since its frequency resolution is constant over the entire frequency band. However, the log scale requires very high frequency resolution at low frequencies. The result is that use of a log scale requires a much larger FFT in order to achieve the needed frequency resolution. Because of the computational load of a large FFT, scanning of wave files or audio input may run more slowly when using a log frequency scale.

When using a log frequency scale, choose the smallest FFT which gives acceptable frequency resolution on the spectrogram or scope at low frequencies. If the low frequency regions on a log frequency display appear noticeably blocky, increase the low frequency resolution by choosing a larger FFT size.

## **SIGNAL READOUTS**

A continuous readout of time (milliseconds), frequency (Hz) and signal level (dB) at the position of the mouse pointer (cursor) is displayed at the bottom left of the program window. The spectrogram display provides signal level measurements only after analysis or scanning is complete. The scope display can provide signal level measurements during all analysis or scanning.

A coordinate grid can also be added or removed by clicking the "Toggle Grid" button.

You can also use the mouse pointer to measure time and frequency differences on the spectrogram display. Click and hold down the right mouse button to establish a reference point in frequency and time. Then as you move the mouse, the differences in frequency and time at the mouse position will be displayed at the bottom left of the Spectrogram window. Releasing the right mouse button returns the mouse cursor to normal operation.



## **DISPLAY MARKERS**

In the real-time scanning modes (Analyze Input, Scan Input, and Scan File) you have the option of adding a fixed reference cross mark to either the scrolling spectrogram display or the scope display. This marker can be useful for identification of amplitude or frequency peaks in the audio signal. Click the left mouse button to set this cross mark at the cursor position on the display.

In addition, you have the option of adding one or two fixed frequency markers to the either the scrolling spectrogram display or the scope display. These markers can be useful when using Spectrogram to tune an audio source to a particular frequency. To activate the frequency markers, choose "Cursors - Freq Mark" from the main menu. Enter the frequency values (Hz) for each marker, and they will then appear on the screen as fixed lines at the chosen frequencies.

## **TO ADJUST OR FREEZE PARAMETERS**

Each time you exit the Spectrogram program, an initialization file (gram.ini) containing the final settings of analysis parameters is saved in the directory with Spectrogram. If you continuously use the same analysis parameters with Spectrogram, then you can choose to automatically use these saved parameters so that you will not have to enter them each time you analyze or scan a file or audio input. Choose "Variables - Freeze" from the Spectrogram menu to automatically use the analysis parameters stored in the gram.ini file. Choose "Variables - Adjust" to manually adjust the analysis parameters each time you perform a spectrum analysis. You can also use the F5 and F6 function keys to switch between adjustable and frozen analysis parameters.

Note that if you make changes in the analysis parameters while running the program, these revised values rather than those in the gram.ini file will be used as the frozen analysis parameters. This allows you to configure the analysis in any way you like, and to use the same parameters again and again without entering them each time.

## ANALYZE FILE DIALOG BOX

The Analyze File mode allows you to compute and display a spectrogram from a prerecorded digital audio sample or wave file. Choose "Analyze File" from the File menu (or Function Key F1) to load a digital audio sample file. Once a file has been selected, you will be presented with a dialog box entitled "Analyze File" for selecting the parameters of the frequency analysis. To select the default values, just press the OK button. You can also make changes in any of the parameters to customize the analysis. Click on any control in the Analyze File dialog box below for a detailed description.

The "Analyze File" dialog box is divided into two main sections: "Sample Characteristics" and "Frequency Analysis".

**Sample Characteristics:**

- Sample Rate (Hz): 44100 (with a "New" button)
- Begin - End (msec): 0 to 3193
- Resolution:  8 bit,  16 bit
- Type:  Mono,  Stereo

**Display Characteristics:**

- Channels:  Dual,  Lt,  Rt
- Attenuation (dB): 0
- Palette:  CB,  CG,  BW,  WB

**Frequency Analysis:**

- Freq Scale:  Linear,  Log
- FFT Size (Points): 512, 1024, 2048, 4096, 8192, 16384
- Freq Resolution (Hz): 10.8
- Band (Hz): 100 to 22050
- Time Scale (msec): 12
- Spectrum Average: 1

Buttons: OK, Cancel

## ANALYZE FILE, SAMPLE CHARACTERISTICS

These controls are available from the Analyze File Dialog Box. Choose “File-Analyze File” from the main menu to enter the Analyze File mode and bring up this dialog box.

- **Sample Rate Edit Box** - You can choose any value of sample rate from 5513 Hz to 44100 Hz. If you have selected a wave file (.wav), the sample rate displayed will be the rate used in the original recording. If you have selected a raw data file, a sample rate of 11025 Hz will be initially assumed, and you should enter the correct value if necessary. In order to enter a sample rate value, first click the "New" button in the Sample Characteristics group. Then enter the new value in the text box and click the "OK" button. Note that entering sample rate here does not result in any change to the data file, but simply changes the sample rate used in calculation, display, and playback of the spectrogram.
- **Begin/End Edit Box** - You can also choose the beginning and ending location in the selected file (in milliseconds) to be analyzed. Initially, the starting and ending location of the entire file will be displayed. If you make no change here, the entire file will be analyzed. For a very large file, it may be necessary to select a smaller time interval for analysis in order not to exceed available RAM memory.
- **Resolution and Type Radio Buttons** - You also have a choice of 8 bit or 16 bit data resolution and mono or stereo operation. Pick the value which you know corresponds to the data file you are analyzing. If you are loading a wave file, the correct value will already be shown. If this is a raw data file, 16 bit data and monaural operation will be assumed, but it is up to you to specify the correct value.

## ANALYZE FILE, FREQUENCY ANALYSIS

These controls are available from the Analyze File Dialog Box. Choose “File-Analyze File” from the main menu to enter the Analyze File mode and bring up this dialog box.

- Freq Scale Radio Buttons - You have a choice of using either a linear or logarithmic frequency scale for computing a spectrogram. A linear scale spaces frequency components equally across the entire spectrum, while a logarithmic scale expands the low frequency region of the spectrogram and compresses the high frequency region on the display. Experiment with these scales to choose the one best suited to your analysis.
- FFT Size Push Buttons - You can compute a spectrogram using 512, 1024, 2048, 4096, 8192, or 16384 point Fast Fourier Transforms (FFTs). The highest possible frequency resolution of your spectrogram will be the digital sampling rate divided by this FFT size. Use the larger FFTs only for high resolution analysis or with the logarithmic frequency scale. The higher resolution FFTs require more time to compute the spectrogram. For this reason it is sometimes preferable to decrease sampling rate when recording audio data if increased frequency resolution is needed, rather than to use a higher resolution FFT.
- Freq Resolution Slider - Adjustment of this slider, in conjunction with selection of the FFT size described above, determines the frequency resolution of the spectrogram. For usual narrow band frequency analysis this control should be set to the minimum (leftmost position) in order to achieve the highest possible frequency resolution. Where broadband analysis is required (such as for observation of speech formants) this control can be adjusted to set a wider analysis bandwidth.
- Freq Band Slider (Linear) - If you have chosen a linear frequency scale, this slider allows you to choose the frequency band to be displayed in the spectrogram. The highest resolution spectrograms may not fit entirely in the display window which has a maximum height of 1024 points. In this case, the Freq Band Slider allows you to choose which portion of the spectrum to display.
- Freq Band Slider (Log) - If you have chosen a logarithmic frequency scale, then the Freq Band Slider allows you to choose the lower frequency cutoff of the spectrogram. A logarithmic scale will always allow displaying the entire frequency band. However, you may find it useful to eliminate the very low frequencies to reduce clutter or improve the appearance of the spectrogram.
- Time Scale Slider - You can also select a time scale in milliseconds which corresponds to the time interval between the calculation of each FFT. To obtain greater time resolution in the spectrogram, choose a smaller value for the time scale. Each vertical line in the spectrogram display represents the output of one FFT calculation. The FFT data input window is stepped sequentially through the data, performing an FFT calculation at each step. The time scale selected determines the length of the step between each FFT and thus the total width of the Spectrogram display. The time scale can be assigned any value between 1 ms and 125 ms. Experiment with these values to pick the time scale which best displays the audio signal of interest.
- Spectrum Average - Use this slider to determine the number of sequential spectrum measurements which are averaged together before display on either the scrolling spectrogram display or the scope display. Averaging is particularly useful in recovering weak periodic signals from a noisy background, but is probably not of much interest in analysis of speech or other rapidly varying signals. You can choose averaging of one (no averaging) to 128 spectrum measurements.

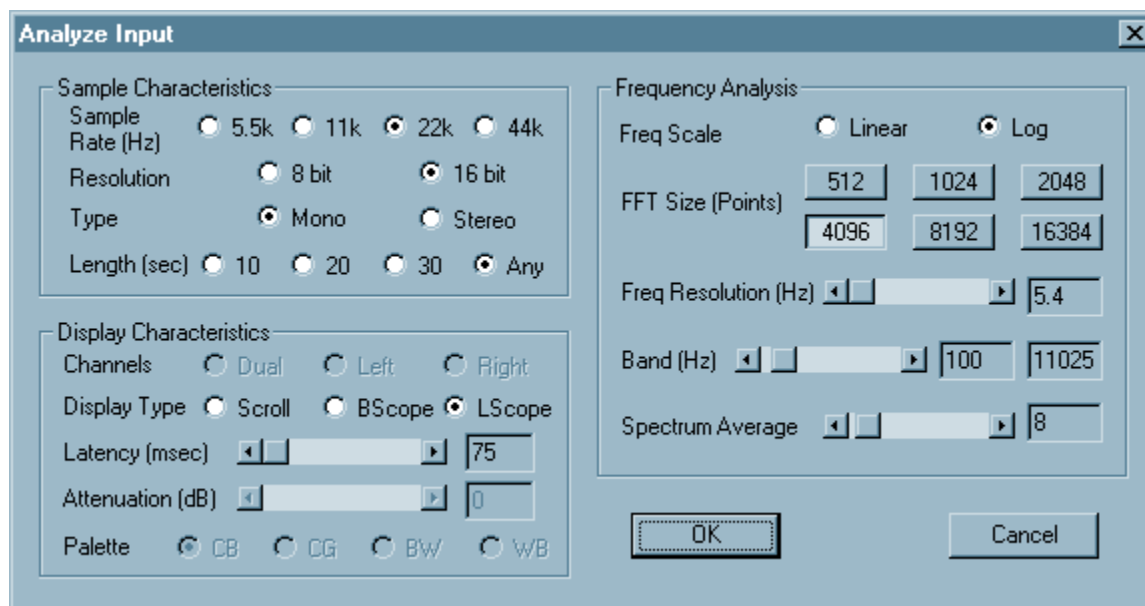
## ANALYZE FILE, DISPLAY CHARACTERISTICS

These controls are available from the Analyze File Dialog Box. Choose “File-Analyze File” from the main menu to enter the Analyze File mode and bring up this dialog box.

- Channels Radio Buttons - If you are analyzing a stereo wave file, you will be able to select left, right, or dual channel operation using these buttons. This selection will not be available for a monaural wave file.
- Attenuation Slider - You are given a choice of display attenuation in order to reduce clutter in noisy digital recordings. Attenuation can be selected to be any value between 0 and 18 dB. Use a threshold of 0 dB regularly, and select greater attenuation only if necessary to reduce clutter.
- Palette Radio Buttons - You also have a choice of four color palettes; color on a black background (CB), color on a gray background (CG), black on a white background (BW), and white on a black background (WB). For a CB or CG display, red represents the strongest frequency components and dark blue the lowest. For a BW display, darker black represents the strongest components, while for a WB display, brighter white represents the strongest components.

## ANALYZE INPUT DIALOG BOX

The Analyze Input mode allows you to record a digital sample as a wave file and to simultaneously produce a spectrogram analysis of the digital sample. Choose "Analyze Input" from the File menu (or Function Key F2) and you will first be presented with a file selection dialog box for choice of a wave file name, followed by a dialog box titled "Analyze Input" for selecting the parameters of the frequency analysis. To select the default analysis values, just press the OK button. You can also make changes in any of the parameters to customize the analysis. Click on any control in the Analyze Input dialog box below for a detailed description.



## ANALYZE INPUT, SAMPLE CHARACTERISTICS

These controls are available from the Analyze Input Dialog Box. Choose "File-Analyze Input" from the main menu to enter the Analyze Input mode and bring up this dialog box.

- **Sample Rate Radio Buttons** - These buttons give you a choice of sampling rates of 5.5K, 11K, 22k, and 44k samples per second for recording of audio samples. Use the lowest sampling rates possible, taking into account that the sampling rate should be at least twice the frequency of the highest frequency component in the audio sample. Lower sampling rates can be used for Spectrograms with linear frequency scales. Logarithmic frequency scales will require higher sampling rates.
- **Resolution and Type Radio Buttons** - You also have a choice of 8 bit or 16 bit data resolution and monaural or stereo recording. 16 bit data resolution should be used for all high resolution spectrograms. Monaural recording is also recommended where memory is limited.
- **Length Radio Buttons** - If you choose a sample length of 10, 20, or 30 seconds, your sample will be both stored on disk and stored in RAM for immediate spectrum analysis. You can also choose "Any" sample length for a sample which is limited only by the size of your hard drive. In this case, the sample will be stored on your hard drive but not stored in RAM for immediate spectrum analysis. To analyze this sample, you must process it separately using the Analyze Audio mode.



## ANALYZE INPUT, FREQUENCY ANALYSIS

These controls are available from the Analyze Input Dialog Box. Choose “File-Analyze Input” from the main menu to enter the Analyze Input mode and bring up this dialog box.

- Freq Scale Radio Buttons - You have a choice of using either a linear or logarithmic frequency scale for computing a spectrogram. A linear scale spaces frequency components equally across the entire spectrum, while a logarithmic scale expands the low frequency region of the spectrogram and compresses the high frequency region. Experiment with these scales to choose the one best suited to your analysis.
- FFT Size Push Buttons - You can compute a spectrogram using 512, 1024, 2048, 4096, 8192, or 16384 point Fast Fourier Transforms (FFTs). The highest possible frequency resolution of your spectrogram will be the digital sampling rate divided by this FFT size. Use the larger FFTs only for high resolution analysis or with the logarithmic frequency scale. The higher resolution FFTs require more time to compute the spectrogram. For this reason it is sometimes preferable to decrease sampling rate when recording audio data if increased frequency resolution is needed, rather than to use a higher resolution FFT.
- Freq Resolution Slider - Adjustment of this slider, in conjunction with selection of the FFT size described above, determines the frequency resolution of the spectrogram. For usual narrow band frequency analysis this control should be set to the minimum (leftmost position) in order to achieve the highest possible frequency resolution. Where broadband analysis is required (such as for observation of speech formants) this control can be adjusted to set a wider analysis bandwidth.
- Freq Band Slider (Linear) - If you have chosen a linear frequency scale, this slider allows you to choose the frequency band to be displayed in the spectrogram. The highest resolution spectrograms may not fit entirely in the display window which has a maximum height of 256 points. In this case, the Freq Band Slider allows you to choose which portion of the spectrum to display.
- Freq Band Slider (Log) - If you have chosen a logarithmic frequency scale, then the Freq Band Slider allows you to choose the lower frequency cutoff of the spectrogram. A logarithmic scale will always allow displaying the entire frequency band. However, you may find it useful to eliminate the very low frequencies to reduce clutter or improve the appearance of the spectrogram.
- Spectrum Average - Use this slider to determine the number of sequential spectrum measurements which are averaged together before display on either the scrolling spectrogram display or the scope display. Averaging is particularly useful in recovering weak periodic signals from a noisy background, but is probably not of much interest in analysis of speech or other rapidly varying signals. You can choose averaging of one (no averaging) to 128 spectrum measurements. In general, the scope display needs averaging of at least four spectrum measurements to produce a stable display.

## ANALYZE INPUT, DISPLAY CHARACTERISTICS

These controls are available from the Analyze Input Dialog Box. Choose “File-Analyze Input” from the main menu to enter the Analyze Input mode and bring up this dialog box.

- Channels Radio Buttons - If you are recording a stereo sample, you will be able to select dual, left, or right channel operation using these buttons. This selection will not be available for a monaural sample.
- Display Type Radio Buttons - While recording, you can choose to display incoming data in real time using either a scrolling spectrogram display, or a spectrum analyzer scope display. The spectrogram display consists of a scrolling 256 point frequency vs. time spectrum for either single or dual channels. The scope display consists of real-time amplitude vs. frequency display in typical scope format for either single or dual channels with 256 frequency points. Choosing BScope provides a histogram bar scope display, and choosing LScope provides a continuous line scope display.
- Latency Slider - Latency here is the length of time (in milliseconds) that it takes for your sound card to process a data sample and record it in memory. If the latency value is set too low for your particular sound card, the spectrogram display during recording will be distorted. A spectrogram display which shows a series of closely spaced vertical spikes, or which fades in and out over time, probably has the latency value set too low. Choose the lowest value of latency that produces an undistorted spectrogram display. Most sound cards operate with a latency of less than 100 msec. However, there are exceptions, and you should experiment with this value to obtain the best spectrogram display. Also note that if the latency value is too large, there will be a noticeable delay between the occurrence of a sound event and its appearance on the spectrogram display.
- Attenuation Slider - You are given a choice of display attenuation in order to reduce clutter in noisy digital recordings. Attenuation can be selected to be any value between 0 and 18 dB. Use a threshold of 0 dB regularly, and select greater attenuation only if necessary to reduce clutter.
- Palette Radio Buttons - You also have a choice of four color palettes; color on a black background (CB), color on a gray background (CG), black on a white background (BW), and white on a black background (WB). For a CB or CG display, red represents the strongest frequency components and dark blue the lowest. For a BW display, darker black represents the strongest components, while for a WB display, brighter white represents the strongest components.

## SCAN FILE DIALOG BOX

The Scan File mode allows you to scan a prerecorded wave file of any length and display its spectrum using either a scrolling spectrogram display or a spectrum analyzer scope display. Choose "Scan File" from the File menu (or Function Key F3) to select a file for scanning. Once a file has been selected, you will be presented with the "Scan File" dialog box for selection of scanning parameters. In this mode, no audio data is saved in RAM for detailed analysis. Click on any control in the Scan File dialog box below for a detailed description.

The purpose of the Scan File mode is to allow the selection of smaller time intervals for analysis from within very large wave files. If a wave file is too large to fit in RAM for detailed analysis, the file can be scanned in real time until the "Stop" button is clicked. A detailed spectrogram analysis can then automatically be calculated at that point. The "Scan File" dialog box gives you the option to enable or disable this detailed analysis at the scan file stop point. Choose "Analysis On" or "Analysis Off" buttons from the "Analysis at Stop" group box.

The image shows a screenshot of the "Scan File" dialog box. The dialog is titled "Scan File" and has a close button (X) in the top right corner. It is divided into several sections:

- Sample Characteristics:** Includes a text field for "Sample Rate (Hz)" with the value "22050" and a "New" button. Below it are radio buttons for "Resolution" (8 bit, 16 bit) and "Type" (Mono, Stereo).
- Display Characteristics:** Includes radio buttons for "Channels" (Dual, Left, Right), "Display Type" (Scroll, BScope, LScope), a slider for "Attenuation (dB)" with the value "0", and radio buttons for "Palette" (CB, CG, BW, WB).
- Frequency Analysis:** Includes radio buttons for "Freq Scale" (Linear, Log). Below are buttons for "FFT Size (Points)" (512, 1024, 2048, 4096, 8192, 16384). Below that are sliders for "Freq Resolution (Hz)" with the value "5.4", "Band (Hz)" with the value "100", and "Spectrum Average" with the value "8".
- Analysis at Stop:** Includes radio buttons for "Analysis On" and "Analysis Off".

At the bottom of the dialog are "OK" and "Cancel" buttons.

## SCAN FILE, SAMPLE CHARACTERISTICS

These controls are available from the Scan File Dialog Box. Choose "File-Scan File" from the main menu to enter the Scan File mode and bring up this dialog box.

- **Sample Rate Edit Box** - You can choose any value of sample rate from 5513 Hz to 44100 Hz. If you have selected a wave file (.wav) for scanning, the sample rate displayed will be the rate used in the original recording. If you have selected a raw data file, a sample rate of 11025 Hz will be initially assumed, and you should enter the correct value if necessary. In order to enter a sample rate value, first click the "New" button in the Sample Characteristics group. Then enter the new value in the text box and click the "OK" button. Note that entering sample rate here does not result in any change to the data file, but simply changes the sample rate used in calculation, display, and playback of the spectrogram.
- **Resolution and Type Radio Buttons** - You also have a choice of 8 bit or 16 bit data resolution and monaural or stereo operation. Pick the values which you know corresponds to the data file you are analyzing. If you are loading a wave file, the correct values will already be shown. If this is a raw data file, 16 bit data and monaural operation will be assumed, but it is up to you to specify the correct value.

## SCAN FILE, FREQUENCY ANALYSIS

These controls are available from the Scan File Dialog Box. Choose “File-Scan File” from the main menu to enter the Scan File mode and bring up this dialog box.

- Freq Scale Radio Buttons - You have a choice of using either a linear or logarithmic frequency scale for computing a spectrogram. A linear scale spaces frequency components equally across the entire spectrum, while a logarithmic scale expands the low frequency region of the spectrogram and compresses the high frequency region on the display. Experiment with these scales to choose the one best suited to your analysis.
- FFT Size Push Buttons - You can compute a spectrogram using 512, 1024, 2048, 4096, 8192, or 16384 point Fast Fourier Transforms (FFTs). The highest possible frequency resolution of your spectrogram will be the digital sampling rate divided by this FFT size. Use the larger FFTs only for high resolution analysis or with the logarithmic frequency scale. The higher resolution FFTs require more time to compute the spectrogram. For this reason it is sometimes preferable to decrease sampling rate when recording audio data if increased frequency resolution is needed, rather than to use a higher resolution FFT.
- Freq Resolution Slider - Adjustment of this slider, in conjunction with selection of the FFT size described above, determines the frequency resolution of the spectrogram. For usual narrow band frequency analysis this control should be set to the minimum (leftmost position) in order to achieve the highest possible frequency resolution. Where broadband analysis is required (such as for observation of speech formants) this control can be adjusted to set a wider analysis bandwidth.
- Freq Band Slider (Linear) - If you have chosen a linear frequency scale, this slider allows you to choose the frequency band to be displayed in the spectrogram. The highest resolution spectrograms may not fit entirely in the display window which has a maximum height of 256 points. In this case, the Freq Band Slider allows you to choose which portion of the spectrum to display.
- Freq Band Slider (Log) - If you have chosen a logarithmic frequency scale, then the Freq Band Slider allows you to choose the lower frequency cutoff of the spectrogram. A logarithmic scale will always allow displaying the entire frequency band. However, you may find it useful to eliminate the very low frequencies to reduce clutter or improve the appearance of the spectrogram.
- Spectrum Average - Use this slider to determine the number of sequential spectrum measurements which are averaged together before display on either the scrolling spectrogram display or the scope display. Averaging is particularly useful in recovering weak periodic signals from a noisy background, but is probably not of much interest in analysis of speech or other rapidly varying signals. You can choose averaging of one (no averaging) to 128 spectrum measurements. In general, the scope display needs averaging of at least four spectrum measurements to produce a stable display.

## SCAN FILE, DISPLAY CHARACTERISTICS

These controls are available from the Scan File Dialog Box. Choose “File-Scan File” from the main menu to enter the Scan File mode and bring up this dialog box.

- Channels Radio Buttons - If you are scanning a stereo file, you will be able to select dual, left, or right channel operation using these buttons. This selection will not be available for a monaural sample.
- Display Type Radio Buttons - While scanning, you can choose to display incoming data in real time using either a scrolling spectrogram display, or a spectrum analyzer scope display. The spectrogram display consists of a scrolling 256 point frequency vs. time spectrum for either single or dual channels. The scope display consists of real-time amplitude vs. frequency display in typical scope format for either single or dual channels with 256 frequency points. Choosing BScope provides a histogram bar scope display, and choosing LScope provides a continuous line scope display.
- Attenuation Slider - You are given a choice of display attenuation in order to reduce clutter in noisy digital recordings. Attenuation can be selected to be any value between 0 and 18 dB. Use a threshold of 0 dB regularly, and select greater attenuation only if necessary to reduce clutter.
- Palette Radio Buttons - You also have a choice of four color palettes; color on a black background (CB), color on a gray background (CG), black on a white background (BW), and white on a black background (WB). For a CB or CG display, red represents the strongest frequency components and dark blue the lowest. For a BW display, darker black represents the strongest components, while for a WB display, brighter white represents the strongest components.

## SCAN FILE, ANALYSIS AT STOP

These controls are available from the Scan File Dialog Box. Choose "File-Scan File" from the main menu to enter the Scan File mode and bring up this dialog box.

- Analysis On and Off Radio Buttons - These buttons give you the option to enable or disable detailed analysis (using the Analyze File Mode) at the point at which scanning of a large wave file is stopped. This feature allows you to scan a large wave file for features of interest, stop scanning at that point and perform a detailed spectrum analysis. Choose "Analysis On" or "Analysis Off" buttons from the "Analysis at Stop" group box to enable or disable this automatic feature.

## SCAN INPUT DIALOG BOX

The Scan Input mode allows you to scan the input audio signal from your sound card and display its spectrum in real time using either a scrolling spectrogram display or a spectrum analyzer scope display. Choose "Scan Input" from the File menu (or Function Key F4) to bring up the "Scan Input" dialog box for selection of scanning parameters. Click on any control in the Scan Input dialog box below for a detailed description.

The image shows a software dialog box titled "Scan Input" with a close button (X) in the top right corner. The dialog is divided into several sections:

- Sample Characteristics:** Includes radio buttons for Sample Rate (5.5k, 11k, 22k, 44k), Resolution (8 bit, 16 bit), and Type (Mono, Stereo).
- Display Characteristics:** Includes radio buttons for Channels (Dual, Left, Right), Display Type (Scroll, BScope, LScope), a slider for Latency (msec) set to 75, a slider for Attenuation (dB) set to 0, and radio buttons for Palette (CB, CG, BW, WB).
- Frequency Analysis:** Includes radio buttons for Freq Scale (Linear, Log), a grid of buttons for FFT Size (Points) (512, 1024, 2048, 4096, 8192, 16384), a slider for Freq Resolution (Hz) set to 5.4, a slider for Band (Hz) set to 100, and a slider for Spectrum Average set to 8.
- Quick Save:** Includes radio buttons for On and Off.

At the bottom of the dialog are "OK" and "Cancel" buttons.



## SCAN INPUT, SAMPLE CHARACTERISTICS

These controls are available from the Scan Input Dialog Box. Choose “File-Scan Input” from the main menu to enter the Scan Input mode and bring up this dialog box.

- Sample Rate Radio Buttons - These buttons give you a choice of sampling rates of 5.5K, 11K, 22k, and 44k samples per second for scanning of audio input. Use the lowest sampling rates possible, taking into account that the sampling rate should be at least twice the frequency of the highest frequency component in the audio sample. Lower sampling rates can be used for Spectrograms with linear frequency scales. Logarithmic frequency scales will require higher sampling rates.
- Resolution and Type Radio Buttons - You also have a choice of 8 bit or 16 bit data resolution and monaural or stereo scanning. 16 bit data resolution should be used for all high resolution spectrograms. Monaural scanning is also recommended where memory is limited.

## SCAN INPUT, FREQUENCY ANALYSIS

These controls are available from the Scan Input Dialog Box. Choose “File-Scan Input” from the main menu to enter the Scan Input mode and bring up this dialog box.

- Freq Scale Radio Buttons - You have a choice of using either a linear or logarithmic frequency scale for computing a spectrogram. A linear scale spaces frequency components equally across the entire spectrum, while a logarithmic scale expands the low frequency region of the spectrogram and compresses the high frequency region. Experiment with these scales to choose the one best suited to your analysis.
- FFT Size Push Buttons - You can compute a spectrogram using 512, 1024, 2048, 4096, 8192, or 16384 point Fast Fourier Transforms (FFTs). The highest possible frequency resolution of your spectrogram will be the digital sampling rate divided by this FFT size. Use the larger FFTs only for high resolution analysis or with the logarithmic frequency scale. The higher resolution FFTs require more time to compute the spectrogram. For this reason it is sometimes preferable to decrease sampling rate when recording audio data if increased frequency resolution is needed, rather than to use a higher resolution FFT.
- Freq Resolution Slider - Adjustment of this slider, in conjunction with selection of the FFT size described above, determines the frequency resolution of the spectrogram. For usual narrow band frequency analysis this control should be set to the minimum (leftmost position) in order to achieve the highest possible frequency resolution. Where broadband analysis is required (such as for observation of speech formants) this control can be adjusted to set a wider analysis bandwidth.
- Freq Band Slider (Linear) - If you have chosen a linear frequency scale, this slider allows you to choose the frequency band to be displayed in the spectrogram. The highest resolution spectrograms may not fit entirely in the display window which has a maximum height of 256 points. In this case, the Freq Band Slider allows you to choose which portion of the spectrum to display.
- Freq Band Slider (Log) - If you have chosen a logarithmic frequency scale, then the Freq Band Slider allows you to choose the lower frequency cutoff of the spectrogram. A logarithmic scale will always allow displaying the entire frequency band. However, you may find it useful to eliminate the very low frequencies to reduce clutter or improve the appearance of the spectrogram.
- Spectrum Average - Use this slider to determine the number of sequential spectrum measurements which are averaged together before display on either the scrolling spectrogram display or the scope display. Averaging is particularly useful in recovering weak periodic signals from a noisy background, but is probably not of much interest in analysis of speech or other rapidly varying signals. You can choose averaging of one (no averaging) to 128 spectrum measurements. In general, the scope display needs averaging of at least four spectrum measurements to produce a stable display.

## SCAN INPUT, DISPLAY CHARACTERISTICS

These controls are available from the Scan Input Dialog Box. Choose “File-Scan Input” from the main menu to enter the Scan Input mode and bring up this dialog box.

- Channels Radio Buttons - If you are scanning stereo input, you will be able to select dual, left, or right channel operation using these buttons. This selection will not be available for a monaural sample.
- Display Type Radio Buttons - While scanning, you can choose to display incoming data in real time using either a scrolling spectrogram display, or a spectrum analyzer scope display. The spectrogram display consists of a scrolling 256 point frequency vs. time spectrum for either single or dual channels. The scope display consists of real-time amplitude vs. frequency display in typical scope format for either single or dual channels with 256 frequency points. Choosing BScope provides a histogram bar scope display, and choosing LScope provides a continuous line scope display.
- Latency Slider - Latency here is the length of time (in milliseconds) that it takes for your sound card to process a data sample and record it in memory. If the latency value is set too low for your particular sound card, the spectrogram display during recording will be distorted. A spectrogram display which shows a series of closely spaced vertical spikes, or which fades in and out over time, probably has the latency value set too low. Choose the lowest value of latency that produces an undistorted spectrogram display. Most sound cards operate with a latency of less than 100 msec. However, there are exceptions, and you should experiment with this value to obtain the best spectrogram display. Also note that if the latency value is too large, there will be a noticeable delay between the occurrence of a sound event and its appearance on the spectrogram display.
- Attenuation Slider - You are given a choice of display attenuation in order to reduce clutter in noisy digital recordings. Attenuation can be selected to be any value between 0 and 18 dB. Use a threshold of 0 dB regularly, and select greater attenuation only if necessary to reduce clutter.
- Palette Radio Buttons - You also have a choice of four color palettes; color on a black background (CB), color on a gray background (CG), black on a white background (BW), and white on a black background (WB). For a CB or CG display, red represents the strongest frequency components and dark blue the lowest. For a BW display, darker black represents the strongest components, while for a WB display, brighter white represents the strongest components.

## SCAN INPUT, QUICK SAVE

These controls are available from the Scan Input Dialog Box. Choose “File-Scan Input” from the main menu to enter the Scan Input mode and bring up this dialog box.

- The Quick Save feature allows you to turn recording on and off while you are scanning audio input. If you select Quick Save On, you will be prompted for a file name. Then, while scanning you can click the “Save” and “Stop Save” buttons to turn recording on and off. Note that it may take a few seconds before recording can be started again after you have turned recording off. However each time you restart recording, the new data segment is added to the end of the wave file you have specified. This feature allows you to record interesting events without the need to record continuously for long periods.
- Please note that use of the Quick Save feature requires that your hard drive be continuously powered. Some computers can operate in power saving mode in which the hard drive power is turned off if the drive is not accessed for a fixed amount of time (usually 10 to 20 minutes). If you attempt to Quick Save after the hard drive power has been turned off, the time delay required to turn the drive on and bring it up to operating speed will disrupt the timing of the Spectrogram program and introduce this delay into data display and recording. If you intend to run Quick Save after an extended period of scanning, then you must disable your computer’s power saving mode so that the hard drive will remain continuously powered.

## MODIFYING SPECTROGRAMS

Once you have computed a spectrogram, you may want to make changes to its length, vertical or horizontal scale, threshold or color to improve the frequency analysis. Choose "Parameters - Change" from the File Menu to bring up a dialog box to "Modify Analysis Parameters." The parameters which can be changed here are identical to those used to define the original spectrogram.

Frequently you will want to select a portion of the spectrogram for modification rather than the entire length. You can drag select this section from the spectrogram display. Position the mouse pointer at the desired starting point, press the left mouse button and drag the mouse to the desired ending point and then release the mouse button. The dialog box will then appear with the starting and ending locations filled according to your selection.

You can return to your starting spectrogram prior to any modification by choosing "Parameters - Restore" from the File menu. The starting spectrogram display is established the first time you choose "Parameters - Change" after analyzing a new data file or recording a new sample. You can return to the starting spectrogram again and again until you analyze a new data file or record a new sample. Note that this function is not available if you have only 8MB of RAM memory.

You can change the palette of a completed spectrogram by choosing "Change" from the "Colors" menu. A dialog box will give you the same color choices available in Display Characteristics. Click the "Try It" button to preview your color selection. Click the "OK" button to implement your selection and return to the Spectrogram. Note that these menu controls are not available if you are operating with more than 256 colors.

## **SPECTROGRAM PLAYBACK**

If you have a windows compatible sound card installed, you will also be able to play back the spectrogram by clicking the 'Play' or 'Play Wdw' buttons. The Play button plays back the entire length of the .wav file, while the Play Wdw button plays back only that portion of the spectrogram which is visible in the Spectrogram Window.

## **SIMULTANEOUS OPERATION**

You can run the Spectrogram program multiple times simultaneously in order to compare different sound samples with each other. However, only one program at a time can access the sound card. This means that only one program at a time can record, playback, or scan audio data.

## **SAVING AUDIO AND BITMAP FILES**

You can save a .wav file of the digital audio of your spectrogram by choosing "Save Wave" from the File Menu.

You can also save a bitmap of the Spectrogram Window by choosing "Save Bitmap" from the File Menu. You can choose to save either the visible section of the display bitmap by choosing "Window Bitmap," or the entire display bitmap including area outside the display window by choosing "Entire Bitmap." The bitmap save feature is available only for single channel spectrograms.



## **PRINTING**

### **PRINTING FROM SPECTROGRAM DISPLAY**

With a graphics capable printer attached to your computer, you can print the spectrogram display from the Analyze File and Analyze Input modes by choosing "File - Print Window" from the main menu. You will be presented with the "Print" dialog box for selection of a printer and printer properties. Click "Properties" to change paper orientation or other print characteristics.

When printing a spectrogram to a black and white printer, it may be best to choose the black on white color palette (BW) for the spectrogram. Otherwise the printer will use a very large amount of black ink filling in the black background on the printed image.

### **PRINTING FROM THE SCOPE DISPLAY**

You can also print the scope display from the Scan File, Scan Input, and Analyze Input modes with a graphics capable printer. Stop the display at the desired point and then choose "File - Print Window" from the main menu. The scope display will always be printed with a white background rather than a black background in order to conserve printer ink.

## DATA LOGGING

### FULL SPECTRUM

This program provides an automatic data logging capability for researchers who wish to record the time, frequency, and harmonic level of events in an audio file. To save the amplitude and phase of every frequency point in a single channel spectrogram, choose "Log Data - Full Spectrum" from the File menu after computing a spectrogram. Data is saved in a text file which records each FFT output point, frequency, amplitude (16 bit), and phase. The amplitude levels in the full spectrum log are calculated without any averaging of sequential spectrum measurements. The log file for the entire spectrogram can be very large, so it is best to drag select a smaller segment of interest on the spectrogram before saving a full spectrum log.

### SINGLE SPECTRUM

To save the amplitude and phase of every frequency component at a single point in time, choose "Log Data - Spectrum" from the File menu after computing a single channel spectrogram. After you select a file name, a dialog box for recording of events will be presented. Data is saved in a text file which records each FFT output point, frequency, amplitude (16 bit), and phase. The amplitude levels in the single spectrum log are calculated without any averaging of sequential spectrum measurements. Click the left mouse button on the event of interest on the spectrogram display, and the spectrum will be displayed in the dialog box. Choose "Save/Exit" to close and save your data log file.

### POINTS

To save the dB signal level and time of individual points, choose "Log Data - Points" from the File menu after computing a single channel spectrogram. After you select a file name, a dialog box for recording of events will be presented. Data is saved in a text file which records an event identifier (usually a letter of the alphabet), the frequency, the time, and the harmonic level in dB (20 log amplitude) at the selected event on the spectrogram. Events to be recorded are defined by the researcher, and could be such things as signal start, signal stop, highest pitch, lowest pitch, highest level, or lowest level, etc. Click the left mouse button on the event of interest on the spectrogram display, and the corresponding data log values will be computed and entered. Enter an event identifier by clicking one of the buttons marked "a" through "h" or by typing an entry in the text box provided. Click "Enter Data" to record the event to your data log file, and then move on to the next event and repeat the process. Enter as many events as needed, and then choose "Save/Exit" to close and save your data log file. Opening an existing data log file will allow you to append data without starting a new file.

## **SYSTEM REQUIREMENTS & INSTALLATION**

Spectrogram requires Windows 95, Windows 98, or Windows NT 4.0. Spectrogram does not run properly under Win 3.1 or Win 32s.

Spectrogram runs best with 16MB or more of RAM memory. You can run Spectrogram with only 8MB of RAM. However, not all functions for modifying the Spectrogram display will be available.

Spectrogram is designed for use with 256 colors. Use of more than 256 colors is not recommended because of the speed penalty incurred in real-time display update.

Spectrogram processes PCM format digital audio data such as .wav sound files but cannot convert compressed audio data.

In order to record and play back sound samples, you will need a Windows compatible sound card installed. However, a sound card is not necessary in order to analyze and display audio spectrograms.

To install Spectrogram just copy gram.exe, gram.hlp, and gram.cnt to any convenient directory on your hard drive. You may then also create a shortcut to gram.exe for display on your desktop. Each time the Spectrogram program is closed, it will save an initialization file (gram.ini) containing analysis and display settings to be used the next time the program is run.

## LIMITATIONS

If you are using 16MB of RAM and you encounter a severe slowing of the program or lengthy disk activity, the program requires more RAM than you have available and Windows is attempting to use the hard drive as virtual memory. The only solution to this problem is to analyze only small sound samples of short duration, or to install more RAM.

Spectrogram runs best on systems using 256 colors. Because the program must update the screen many times a second, the large bitmaps of true color screens will slow Spectrogram significantly. Using 256 colors will reduce annoying flickering of your mouse pointer or cursor during analysis or scanning. Also, not all of the palette controls are available with more than 256 colors. So use 256 colors for the best performance in running Spectrogram.

If you are using a colorful screen background bitmap, Spectrogram can change the background to odd looking colors while running. This isn't a bug. Spectrogram must take control of all bitmap color assignments in order to produce its own display. Your background will be restored when Spectrogram is closed.

Any other programs running simultaneously can interfere with Spectrogram, particularly programs which access the hard drive while Spectrogram is scanning a file or recording to disk. For best performance, shut down these other programs.

Don't expect to be able to scan wave files direct from a CD ROM drive. Instead, copy any wave files to be scanned to your hard drive. You can analyze any file from a CD ROM drive. However, real-time scanning puts a heavy load on your computer and doesn't leave much time for data transfer from disk. CD ROM drives are just not fast enough yet. Of course, CD music disks can be scanned by connecting the audio output of your CD player to the input of your sound card and using the "Scan Input" function.

Likewise, don't expect to be able to scan wave files direct from a floppy disk, or to record direct to a floppy disk. You can analyze any file from a floppy disk, but real-time scanning or recording is not possible using floppy disks.

Please note that use of Quick Save during Scan Input requires that your hard drive be continuously powered. Some computers can operate in power saving mode in which the hard drive power is turned off if the drive is not accessed for a fixed amount of time (usually 10 to 20 minutes). If you attempt to Quick Save after the hard drive power has been turned off, the time delay required to turn the drive on and bring it up to operating speed will disrupt the timing of the Spectrogram program and introduce this delay into data display and recording. If you intend to run Quick Save after an extended period of scanning, then you must disable your computer's power saving mode so that the hard drive will remain continuously powered.

Please don't automatically record everything at 44kHz. Lower sampling rates are often a better choice to produce high resolution spectrograms.

## MEMORY USAGE

Spectrogram can require an enormous amount of memory, particularly for high resolution, two channel spectrograms using 24 bit display color. For example, a 30 second, two channel spectrogram, computed using a 2048 point FFT and a 5 msec time scale, will require 36 mbytes just for a 24 bit color bitmap. If you ever see the warning "Not Enough Display Memory," your computer cannot allocate the memory needed for the entire display bitmap.

Solutions to such memory shortages are:

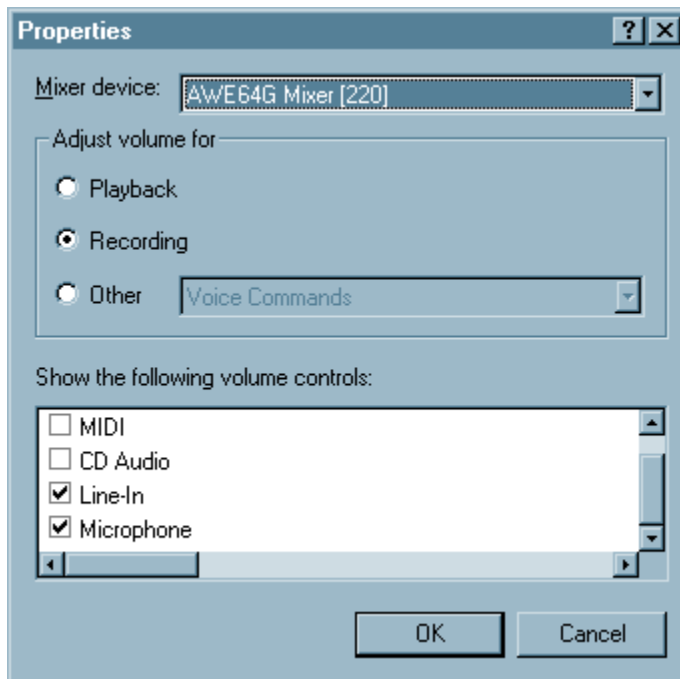
1. Use 8 bit (256 color) display color instead of 24 bit display color when running Spectrogram to reduce display bitmap memory by two thirds.
2. Run Spectrogram in monaural rather than stereo mode. A single channel spectrogram requires one half the display bitmap memory of a two channel spectrogram.
3. Increase the spectrogram frequency and time scales in order to reduce the size of the needed display bitmap.
4. Reduce the spectrogram time span.

Use of 256 colors also has the advantage of improving the performance of real-time scanning and recording of wave files or audio input from your sound card.

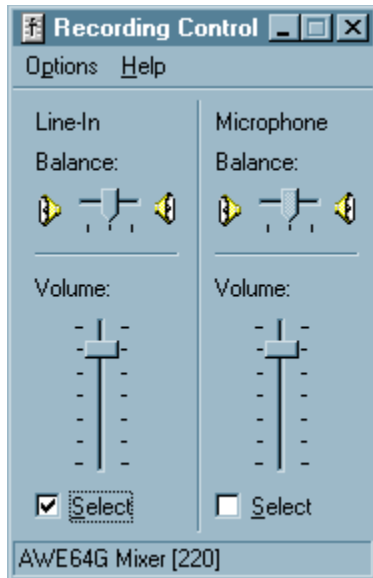
## SOUND CARD INITIALIZATION

First, connect an audio data source to the signal input jack of your sound card. Either Line-In or Microphone inputs may be used depending on the signal level expected. It is always best to turn off your computer before making these connections.

Then you must instruct your Windows sound card to process audio signals. Windows 95 provides a separate Volume Control program for this purpose. Choose "Project - Volume Control" from the Spectrogram menu. This will start the Volume Control program. Then choose "Options - Properties" from the recording control menu. Then mark the "Recording" button, and under "Show the following volume controls" choose "Line-In," and "Microphone."



Now when you click "OK" you will see a Recording Control allowing you to adjust the volume for Line-In and Microphone recording. Select either Line-In or Microphone inputs here depending on the audio source that you intend to use. It is best not to choose both inputs simultaneously, but to choose one or the other depending on the audio source.



Occasionally this Recording Control will also have an active menu item “Options – Advanced Controls.” If this menu item is available it should also be selected. This selection will make available an “Advanced” button on the Recording Control for turning Automatic Gain Control (AGC) on or off. AGC is often useful to keep the recording level automatically within the correct range.

## **SCOPE DISPLAY CALIBRATION**

You can calibrate the scope display using the vertical scroll bar at the right side of the display window. The scope display shows a fixed 90 dB signal range. However, many interacting factors will affect the absolute level of frequency components shown on the display. While the relative levels of frequency components to each other will always be correct, the gain of your sound card and the combination of frequency analysis parameters selected will affect absolute measurement. Use the vertical scroll bar to adjust the amplitude of a known frequency component to its exact value. Then absolute measurement of all other components will also be correct.



## **PROBLEM REPORTING & PROGRAM UPDATES**

Programs can only be improved if users provide feedback to the author. I can be reached via the Internet for you to report any bugs or to provide comments or feedback. I encourage anyone with a question or suggestion to contact Richard Horne at [rshorne@mnsinc.com](mailto:rshorne@mnsinc.com).

The Spectrogram Program is updated periodically to add features and fix bugs reported by users. You can always download the latest Spectrogram update from <http://www.monumental.com/rshorne/gram.html>.

