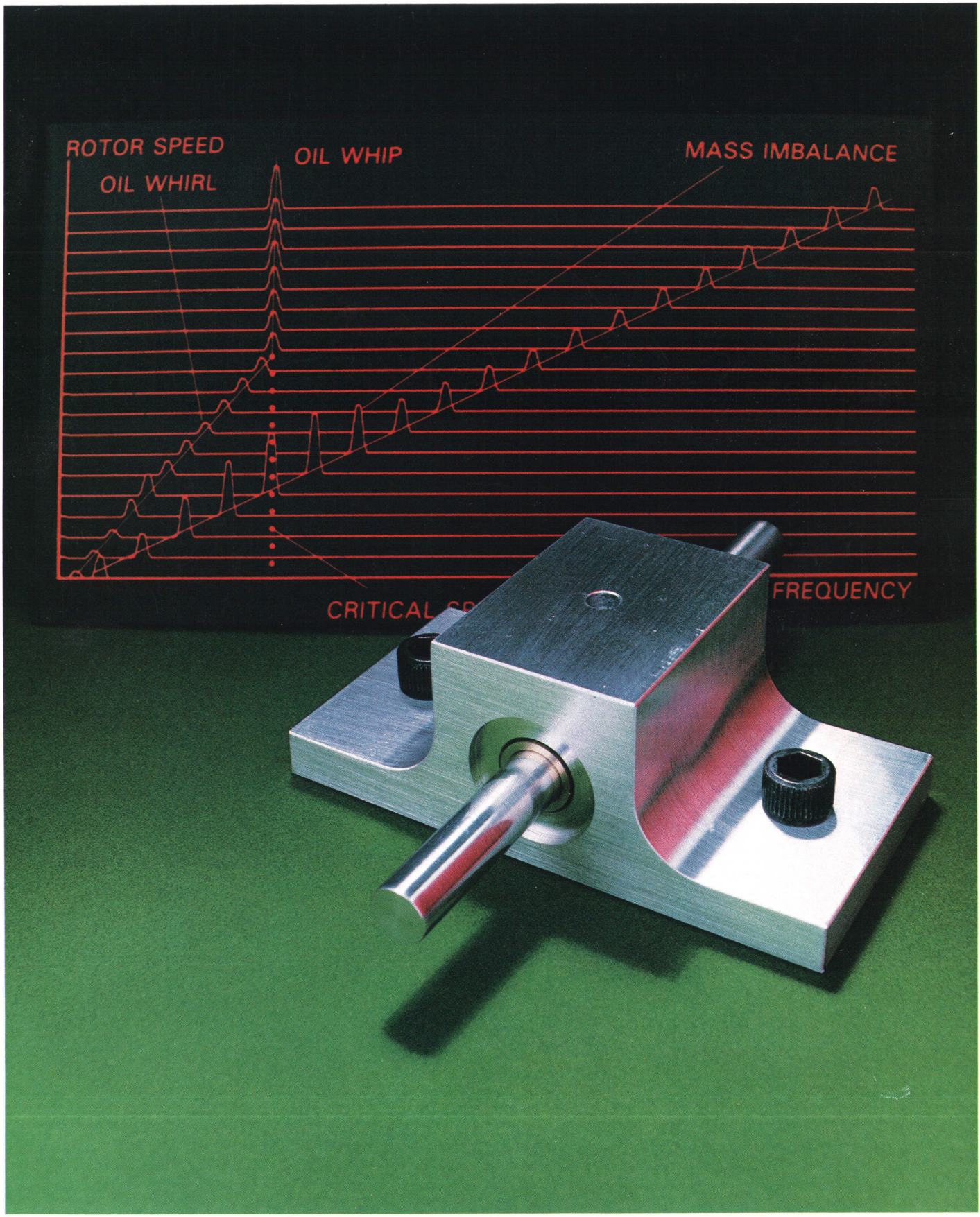
HEWLETT-PACKARD
JOURNAL

DECEMBER 1984

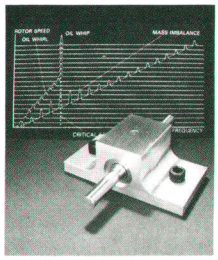


HEWLETT-PACKARD JOURNAL

December 1984 Volume 35 • Number 12

Contents

In this Issue



In 1965, V.W. Cooley and J.W. Tukey published their landmark "algorithm for the machine calculation of complex Fourier series," which quickly became known as the fast Fourier transform, or FFT. The FFT reduced from hours to seconds the time required for a digital computer to calculate the frequency spectrum of a signal. Hewlett-Packard's first FFT analyzer, the HP 5450A Fourier Analyzer, made its debut in 1970. This general-purpose instrument could sample input signals between dc and 25 kHz, store the sample values in its memory, and later compute frequency spectra or perform other mathematical manipulations of the data.

Over the years, new twists and refinements have made the FFT ever faster, and advancing technology has increased processing speeds by a combination of faster digital circuits and parallel processing techniques such as pipelining. In this issue you'll find the design story of a new FFT analyzer that uses the latest technology, plus a few of its designers' own ideas, to sample signals between dc and 100 kHz and compute frequency spectra almost in real time—as many as ten spectra per second, fast enough that the user can watch the displayed spectrum change as changes occur in the system being analyzed. The HP 3561A Dynamic Signal Analyzer is designed especially for electronic applications such as spectral analysis, network analysis, and waveform recording, for mechanical vibration applications such as predictive maintenance, dynamic balancing, and spectral analysis, and for acoustic applications such as noise reduction and speech signal processing. Our cover illustrates one use of the HP 3561A's near-real-time processing. The plot shown is variously known as a spectral map, a stack plot, or a waterfall diagram. It's a series of successive spectra displayed simultaneously to show the results of a change of some parameter such as the speed of a motor. In this case the system is a fluid-film bearing (like the bearings on your car's crankshaft) and the spectral map shows the phenomenon called oil whirl becoming oil-whip instability as the shaft speed increases. For predictive machinery maintenance, the HP 3561A is used just as a physician uses an electrocardiogram. A machine's vibration spectrum is compared with a normal one taken earlier to reveal changes that might signal an impending breakdown because of wear or internal damage.

The HP 3561A's capabilities are discussed in the article on page 4. On page 6 you'll find descriptions of some of its applications. Its hardware and software designs are described in the articles on pages 12 and 17, and the special aspects of its FFT implementation are outlined on page 20. The article on page 28 tells how the HP 3561A's digital filters process so much data so quickly.

Here's an invitation for HP end users and OEMs. Most of our articles are about the design of HP products and the contributions made by these products and designs. We'd like to publish an occasional article describing how HP hardware and/or software contributions have helped you make a contribution in your field. We're looking for original technical material that's interesting to more than just a few readers. If you have an idea for an article you'd like to write, send an outline to Editor, Hewlett-Packard Journal, 3000 Hanover Street, Palo Alto, California 94304 U.S.A.

-R.P. Dolan

What's Ahead

The design of a new line of general-purpose fiber optic test instruments will be featured in the January 1985 issue. The HP 8150A Optical Signal Source produces modulated, calibrated light signals for stimulating fiber optic devices. The HP 8151A Optical Pulse Power Meter measures various parameters of light signals, including peak levels, mesial power level, and extinction ratio. A third instrument, the HP 81519A Optical Receiver, is a stand-alone optical-to-electrical transducer that gets users into fiber optic testing at the lowest possible cost.

Versatile Instrument Simplifies Dynamic Signal Analysis at Low Frequencies

Analysis of low-frequency signals has many uses in electronic design, vibration studies, and acoustic measurements. This easy-to-use analyzer covers the range from 125 μHz to 100 kHz and displays the data in several useful formats.

by James S. Epstein

REAL-TIME ANALYSIS of low-frequency electrical, mechanical, or acoustic phenomena can be of great value to design and maintenance personnel. Until the advent of digital signal processing, such analysis was difficult if not impossible to do. As the complexity and performance of digital ICs increased, it became possible to study low-frequency phenomena in more detail.

Using fast Fourier transform (FFT, see box on page 20) digital analysis techniques, HP introduced several low-frequency signal analyzers in the 1970s—the HP 5450A Fourier Analyzer, the HP 5420A Digital Signal Analyzer, and the HP 3582A Spectrum Analyzer. Combining many of the features of these earlier instruments with new capabilities made possible by VLSI circuit technology, the HP 3561A Dynamic Signal Analyzer (Fig. 1) is the latest member of HP's family of FFT analyzers.

The HP 3561A is designed to provide high-performance dynamic measurement and analysis capability in a small portable package. With the increased computational capability that exists in today's microprocessors, many features can be now incorporated into an instrument that previously required a companion controller or computer. Analysis of dynamic signals has only recently been available because it awaited the development of suitable high-speed digital signal processing capability. This capability is incorporated in the HP 3561A in the form of multiple microprocessors and custom digital integrated circuits.

Conventional swept-frequency analyzers cannot analyze a dynamic signal very well because the signal may be changing during the sweep time. In dynamic signal analyzers, the signal is first digitized quickly and then analyzed after the data is captured. The analysis of the data may consist of examining the time waveform directly or examining the frequency spectrum at evenly spaced intervals or

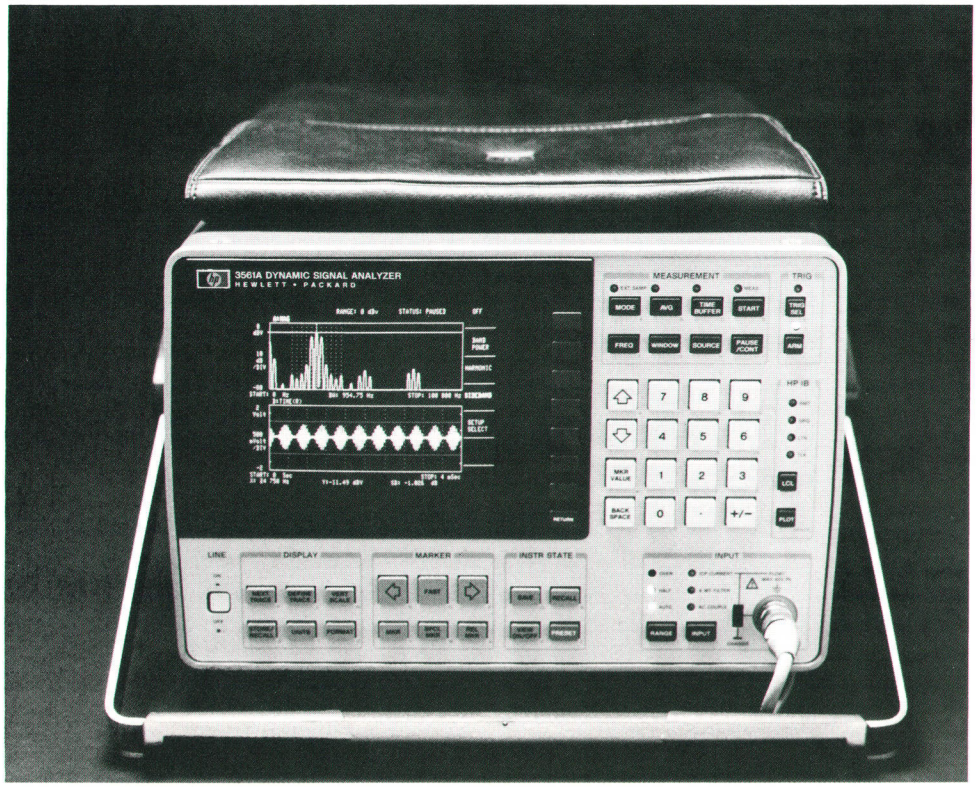


Fig. 1. *The HP 3561A Dynamic Signal Analyzer combines high-performance hardware and versatile menu-driven software in a portable package for measurement and analysis of low-frequency waveforms in electronic, acoustic, and vibration applications. Besides spectral analysis, the HP 3561A can perform network measurements and time domain waveform recording.*

in ⅓- or full-octave intervals. In addition, mathematical operations can be performed on the data to look at differences, ratios, and other computed results. After the data has been captured and the desired results computed, the next step is presenting the data for display. The HP 3561A has several display formats, some of which allow up to sixty spectra to be viewed simultaneously.

Key Features

The HP 3561A can measure signals over the frequency range from 125 μHz to 100 kHz. This covers the range of slow mechanical and geophysical vibrations to well above the 20-kHz audio band. One of the important specifications of a dynamic signal analyzer is its ability to resolve signals that are closely spaced in frequency. This allows a user to detect close-in sidebands on audio signals or vibrations caused by two different sources that are close in frequency. The HP 3561A can analyze signals with a resolution of 640 μHz over its entire 100-kHz range.

Another important consideration in dynamic signal analysis is measuring small signals in the presence of large signals, such as the distortion of a sine input. A measure of an analyzer's ability to do this is called dynamic range; it specifies the maximum difference in signal levels that can be measured and guarantees that any internally generated distortion and spurious products will be below the minimum measurable level. In the HP 3561A the dynamic range is 80 dB, which means that a primary signal and a secondary signal one ten-thousandth of the amplitude of the primary signal can be measured simultaneously. The accuracy of the measurement is also important. Using an internal calibration process, the HP 3561A has an amplitude accuracy of ±0.15 dB over the entire HP class B environmental range (0 to 55°C, <95% relative humidity at 40°C, and altitudes <4570 m).

Applications

The applications for a dynamic signal analyzer are very diverse. However, the HP 3561A Dynamic Signal Analyzer is specially designed to cover three areas of application.

The first area is general electronic measurements in the audio frequency range. One common measurement is the distortion of a pure tone. The HP 3561A can measure distortion products that are 80 dB below the fundamental and has a built-in marker function that computes the total distortion of up to 20 harmonics when the marker is placed on the fundamental. Another electronic measurement is frequency response. The HP 3561A has an internal source that, combined with the internal trace arithmetic, can be used to measure both amplitude and phase as functions of frequency. General spectral analysis is another measurement done in the audio band to determine the exact frequency and amplitude of an input signal or to determine the frequency and amplitude of a spurious input. Transient analysis is another measurement often needed in the electronic and mechanical areas. The 40,960-word buffer in the HP 3561A allows 0.8 second worth of data to be captured for signals in the 0-to-20-kHz frequency range.

Another major application area for dynamic signal analysis is measuring mechanical vibration using a transducer that converts either displacement, velocity, or acceleration into an electrical signal that can be measured by the analyzer. The analysis of vibration can be extremely useful in the maintenance of machinery. By examining the vibration in the frequency domain, small vibrations caused by worn-out parts can be detected even in the presence of large vibrations that are normal to the operation. In this manner, the machine can be shut down for repair only when necessary, saving costly down time. In addition to detecting when a problem has occurred, potential causes can be identified. (See box on page 6 for an example).

A third application area for which the HP 3561A is particularly well suited is acoustic measurements. Measuring acoustic signals requires special techniques. The human ear has its own frequency response, which must be considered when making measurements. An internal A-weighting filter allows the analyzer to duplicate the weighting applied by the human ear, thereby allowing the noise levels measured to be compared to their levels as perceived by a human observer. Another important consideration in acoustic analysis is frequency resolution. The human ear cannot detect two frequencies separated by a fixed difference at high frequencies as well as it can at low frequencies. Its response is logarithmic and is best characterized by using ⅓-octave analysis bands that become wider as the center frequency increases. These analysis bands have been used for a long time to measure acoustic signals and many noise specifications are written in terms of ⅓-octave intervals as defined by ANSI specifications. The HP 3561A can display its results in ⅓-octave, full-octave, or linear frequency increments to allow the user to determine which mode is most useful for a particular application.

Friendliness

With all of the capability that can be packed into an instrument such as the HP 3561A, it could be difficult to use these features if the instrument were not specifically designed for ease of use. A measurement should be easy to make without requiring a lot of setup to get accurate and quick results. This is made possible in the HP 3561A by

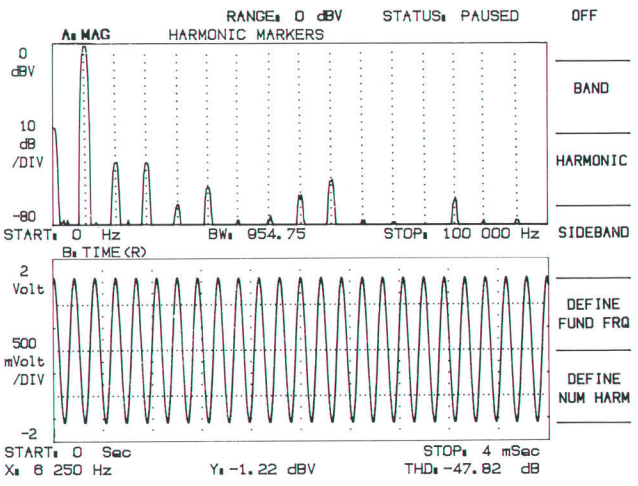


Fig. 2. *A special marker function displays the location of the harmonics of the selected center frequency using a dotted vertical grid. The lower trace shows the actual signal.*

Dynamic Signal Analysis for Machinery Maintenance

Using the HP 3561A Dynamic Signal Analyzer to analyze the vibrational signature of machinery can allow identification of problem spots before they cause machinery shutdown. In the time domain, vibrations from many sources in a machine can add to create complicated waveforms. When viewed in the frequency domain, however, individual problems such as imbalances, gear-mesh anomalies, or bearing defects can usually be identified (Fig. 1). By saving the spectra of a healthy machine in the HP 3561A's optional one-megabit bubble memory, a set of baseline measurements is formed that can be used for later comparison. Fig. 2a shows an example of a normal baseline measurement for a machine. The same machine at a later time has developed a damaged gear as shown by the excessive response at the gear-mesh frequencies (Fig. 2b).

In addition, the HP 3561A has a spectral map display format that is particularly useful in servicing rotating machinery. These spectral maps can be created by adding a third dimension to spectral measurements, such as time, r/min, or load. The HP 3561A's map mode can be used to identify changes or trends in the spectra of machines, circuits, or other devices. In the example shown in Fig. 5 on page 10, a collection of measurements made during a machine runup forms a spectral map, illustrating changes in machine vibration with an increase in rotational speed.

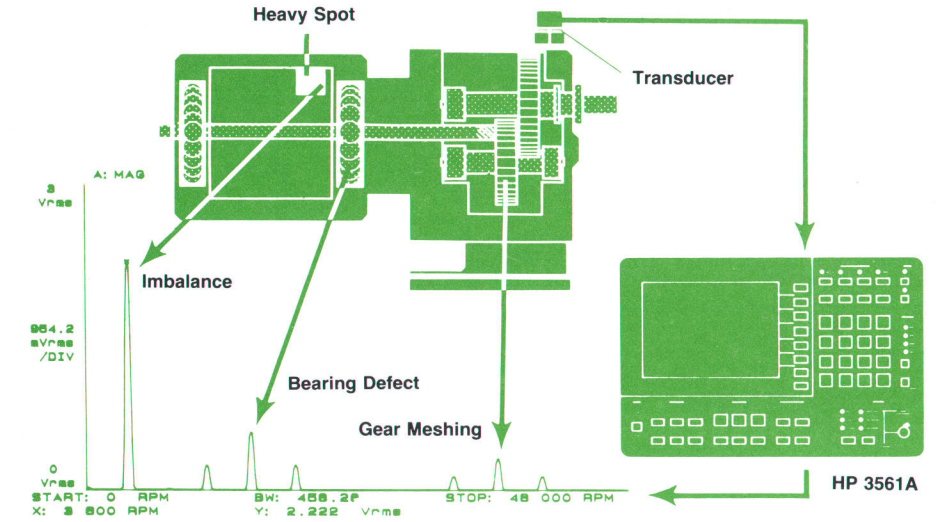


Fig. 1. *The individual components of vibration are shown in dynamic signal analysis displays of amplitude versus frequency (r/min).*

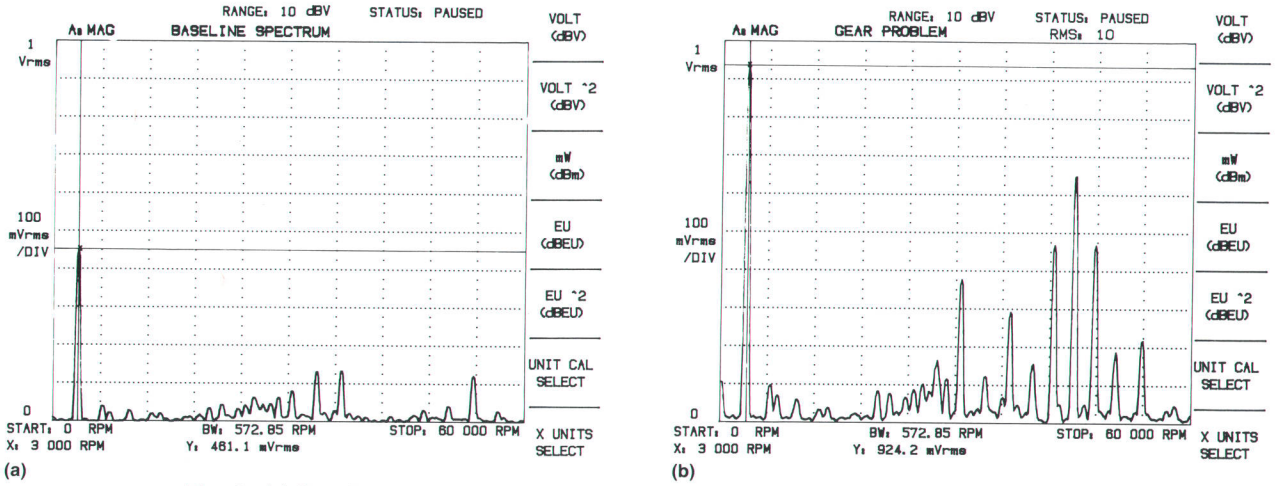


Fig. 2. *(a) Baseline spectrum for a machine. (b) Later spectrum for machine of (a) showing the excessive response at the gear-mesh frequencies caused by a damaged gear.*

several features coupled with softkey menus. An autorange algorithm selects the correct range to place the input signal within 4 dB of full scale. Autocalibration is periodically performed so that the measurement accuracy is maintained within the ±0.15-dB specification. The first measurement covers the entire 100-kHz frequency range so that all frequencies can be examined. Then the user can zoom in on a particular range of frequencies by using the HP 3561A's marker function to identify the center frequency and pressing the DEFINE CENTER softkey in the menu obtained by pressing the **FREQ**uency key on the front panel. After the center frequency is selected, the desired analysis band can be selected in the **FREQ** softkey menu by pressing the DEFINE SPAN softkey.

Several special marker functions are provided to make the analysis of the measurement easier. The relative marker function allows any signal to be selected as the reference for either amplitude or frequency or both. When this marker is used, all measurements are scaled relative to the selected reference. This function is particularly useful when measuring how far the amplitude of a spurious signal is below that of the desired signal.

Another special marker function will mark up to twenty harmonics on the display screen with a special grid after the fundamental frequency is entered by the user (Fig. 2). In addition to marking the harmonic components on the screen, the total harmonic distortion is computed for the harmonics marked. A similar marker function is provided for marking sidebands and computing the relative power in the sidebands with respect to the carrier.

A very common type of analysis is computing the power in a specified frequency band. This is done when analyzing noise power in different frequency regions. The band-power marker function allows the user to position right and left markers to denote the band of interest and have the power in the band computed.

Saving Instrument States

Some specialized measurements may require many of the HP 3561A's special features to be selected. When this is the case, the storage and recall of instrument states in nonvolatile memory can be a very handy feature. Once the instrument is set up as desired, its state can be stored in one of six locations. If the optional bubble memory is installed, up to 127 instrument states can be saved. At a later time, say at a remote measurement site, the desired state can be quickly recalled when it is time to make the measurement. This is not only convenient, but can help prevent errors in setting up the instrument.

Narrowband Measurements

The narrowband measurement is the basic analysis mode of the HP 3561A. This mode provides a frequency spectrum of 400 spectral lines equally spaced over the chosen analysis band. The analysis band can be centered anywhere in the instrument's frequency range from dc to 100 kHz with 0.25-Hz resolution and the frequency span for a zoomed measurement can be as small as 256 mHz. This gives a resolution of 640 μHz anywhere in the spectrum up to 100 kHz. In addition, when making a baseband measurement (not zooming around a center frequency), the analysis band, or frequency span, can be reduced to 10.24 mHz, which gives a maximum resolution of 25.6 μHz.

Speed. Speed was an important consideration in designing the hardware to make narrowband measurements. When making adjustments to electronic equipment, such as reducing distortion components in a device under test, the display must be updated fast enough to give useful feedback to the person doing the adjustments. When averaging many spectra, the total measurement time is greatly dependent

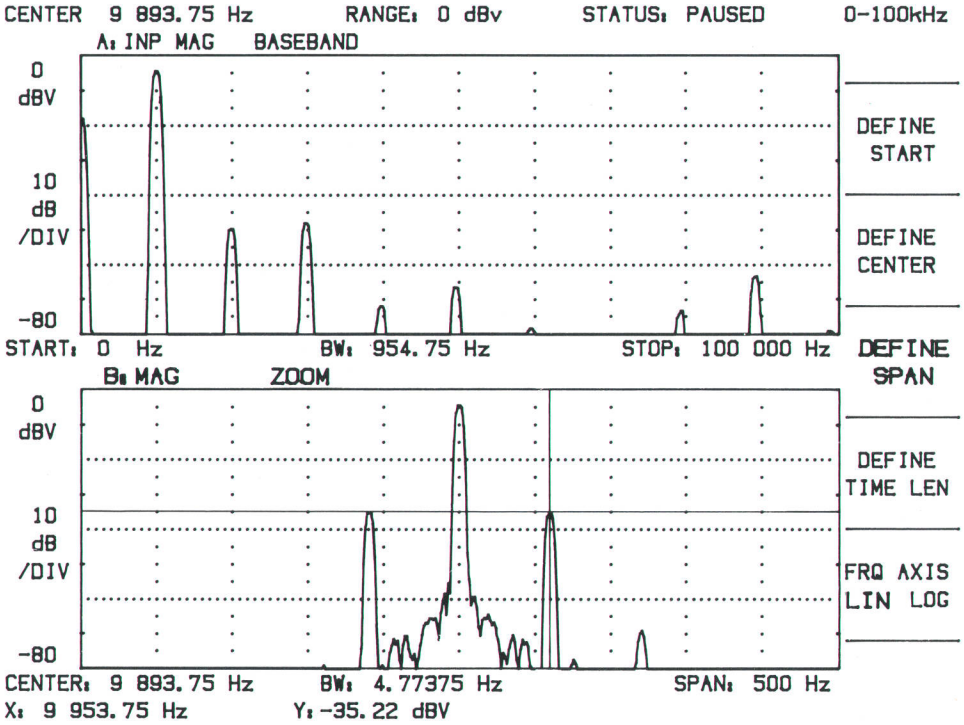


Fig. 3. *Simultaneous display of baseband (upper trace) and zoom (lower trace) spectra can be selected for the HP 3561A's CRT.*

upon the time required to generate each spectrum. The HP 3561A can update a spectral display 10 times per second and can average up to 20 spectra per second. See the article on page 12 for design details about how this performance was achieved.

Windows. The shape of the analysis band is determined by the window applied to the data. In dynamic signal analysis, the band shape is selected by windowing—multiplying the data in the time domain by a weighting function. This is equivalent to filtering the data in the frequency domain. Just as it is useful to have different passband shapes in a conventional swept-frequency analyzer, it is also useful to have different windows in a dynamic signal analyzer. The HP 3561A offers a choice of four window functions—flattop, uniform, Hanning, and exponential. The flattop window gives maximum amplitude accuracy at the expense of frequency resolution. Its bandwidth is approximately four times that of the uniform window, which weights all of the time samples equally. The uniform window is used for signals that are periodic or are fully contained within the analysis band (such as a transient or an output of one of the HP 3561A's internal noise sources), and it provides maximum frequency resolution.

The Hanning window provides a good compromise between frequency resolution and amplitude accuracy. The window flatness is within 1.5 dB, compared to 0.01 dB for the flattop window. Its bandwidth, however, is only 1.5 times the maximum available resolution.

The exponential window is useful when measuring transients. Coupled with the HP 3561A's internal impulse noise source, this window allows the HP 3561A to make impulse response measurements. The shape of the exponential window can be adjusted to fit the data that has been collected.

Measuring Phase. Another distinction between dynamic signal analyzers and conventional swept-frequency analyzers is the ability of the former to measure phase. There are also applications for measuring phase with respect to a

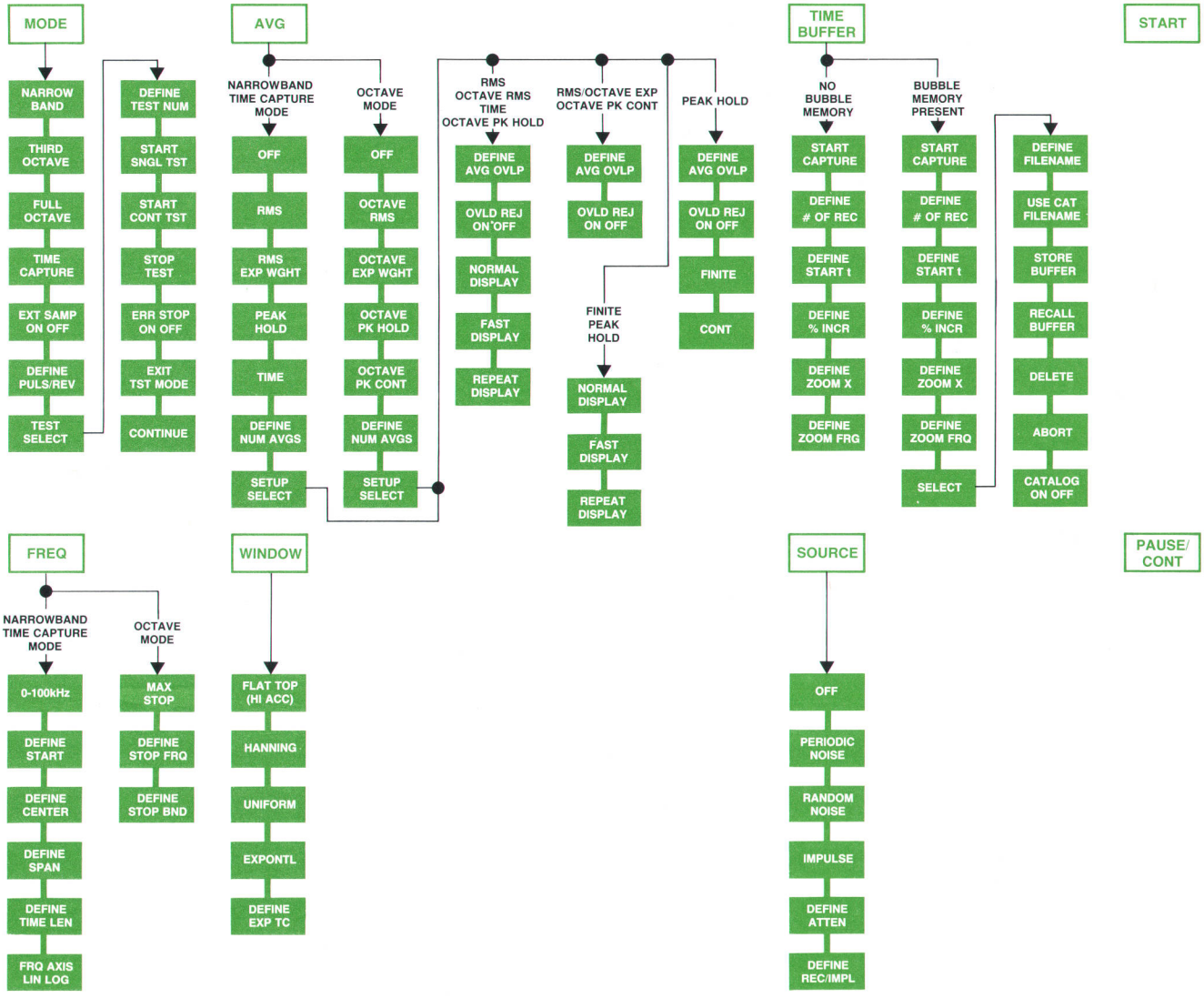


Fig. 4. *Softkey menu structure obtained when the user presses the keys in the Measurement group on the front panel.*

trigger, such as when balancing rotating machinery. Here, the measured phase is used to determine where in the rotation cycle an out-of-balance condition occurs. By using the HP 3561A's internal calibrator, the measured phase value is corrected for an accuracy of ±2 degrees up to 10 kHz and ±10 degrees from 10 kHz to 100 kHz.

Averaging. Averaging can be used to increase the usefulness of the narrowband measurement. Several types of averaging are available. When measuring the spectral characteristics of random data, the accuracy of the results is subject to the statistical properties of the data. One method of reducing the variance of the results is to perform rms or power averaging of the data. Another type of averaging that is useful for reducing the amount of noise on the signal is time averaging. This type of averaging requires a synchronization signal, which can be either an external trigger or a part of the signal being measured. Two other types of averaging provided by the HP 3561A are rms exponential and peak hold. The first is a type of rms averaging that weights earlier averages less than the current measurement. This allows a continuous update of the display while averaging. The second is used when the maximum value in each frequency bin is desired for a series of measurements.

Triggering. Triggering is important when measuring phase and in time averaging. Therefore, the HP 3561A has several triggering modes. The source of the trigger can be the input signal, an external TTL signal, the internal source, or the internal trigger (free run). In addition, the instrument can be armed for a single trigger or armed to respond each time a trigger is received. Pretriggering of up to 8 records in narrowband mode or 40 records in time capture mode (discussed later) is provided. Post-triggering of up to 1023 records can be selected. Pre- and post-triggering are very useful when a time difference exists between a known event and the data that is desired. When testing loudspeakers, for example, a time delay exists between the time the impulse excitation is generated by the instrument and when the impulse response is detected at the speaker.

Simultaneous Display. The HP 3561A has the ability to display a full 100-kHz baseband spectrum and a zoom spectrum simultaneously (Fig. 3). This capability is provided by the use of custom digital filters, which are discussed in the article on page 28. One way that this display format can be used is to have the full 100-kHz spectrum on the upper half of the display and the zoom spectrum on the bottom. Moving the marker to the desired signal on the upper display and pressing the CTR FREQ softkey in the marker menu will move the center frequency of the zoom spectrum to that of the desired signal. This can be performed repeatedly if desired to zoom in on several different signals for closer examination.

Time Capture Measurements

Time waveform capture can be used for transient analysis and the analysis of nonstationary signals such as speech. It also can be used to increase the real-time rate of the measurement to 100 kHz until the buffer is filled.

The major difference between narrowband and time capture measurements is that in time capture mode the entire time record is captured and stored in memory before any processing is done. If the bubble memory option is installed, the time record can be retained even while the instrument is turned off. The time record can be defined to be from 1 to 40 records of 1024 samples each. The full triggering capability of the HP 3561A can be used with time capture mode so that up to 40 records of pretrigger data can be taken, allowing the examination of data before the occurrence of a particular event. The ability to preserve the original time data before processing allows the center frequency and span for zoom analysis to be selected after collecting the data. This can not only save a lot of data collection time, but may be the only way to collect data when the input signal is of a transient nature.

In time capture mode, the data is first collected within the frequency range specified for collection. If the data is collected in baseband mode, it can be translated to any center frequency up to the maximum frequency collected and zoomed up to a factor equal to the number of records collected.

The postcollection translation can occur as many times as desired. In addition, the type of window applied to the data can also be changed to allow trade-offs between frequency resolution and accuracy. The postcollection processing is done by selecting the time capture data memory as the input to the digital filters. Because the data has already been collected, it can be run through the digital filters at maximum speed. Different start points in the data can be selected so that only desired data is analyzed and changes in spectral parameters can be examined at different points in the record. In addition, an average magnitude display can be selected to examine the average of all the data in the time capture memory at one time.

When the time capture mode is combined with other HP 3561A features, very useful measurements can be made. For example, the average magnitude display can be used in the upper half of the display to show the full frequency range of the captured data, and the lower half of the display can then be set up to scroll through the captured data with or without zoom. When a user desires to look at the entire time buffer at one time, the spectral map display mode (discussed later) can be used. Whether scrolling through the data or mapping it, an increment can be set up to specify the percent of the data record to be incremented for the next display. In this manner, a spectral map can be displayed so that each new trace in the map is offset by one to 1024 data points. This can be important when examining the decay of selected frequencies.

Third- and Full-Octave Analysis

For certain applications, most notably in the acoustic field, proportional frequency resolution is desirable. In this type of spectrum analysis, the signal is analyzed by a bank of filters with logarithmically spaced center frequencies and bandwidths proportional to the center frequency. The most widely used analysis of this type is at ⅓-octave and full-octave intervals. The HP 3561A computes the ⅓-octave and full-octave spectra digitally by combining narrowband frequency data taken on three separate analysis spans. To do this, the custom digital filters are set up to collect data at three sample rates: f_s, $f_s/10$, and $f_s/100$. The three data buffers are then transformed using the FFT, and a weighted sum of adjacent FFT filter bins is computed to form each

⅓-octave or full-octave filter bin.

The frequency response of the octave filter shapes is determined by the windowing function used, the number of FFT bins summed, and the weighting function applied to the FFT bins. In the HP 3561A, the Hanning window function is used for all of the ⅓-octave and full-octave filter shapes. The number of bins summed varies with the center frequency of the octave filter. The weights applied to the adjacent FFT bins equal unity except for the first and last bins.

The octave filter shapes are designed to meet ANSI and IEC standards. Because of the number of different octave spans allowed in the HP 3561A, 168 special filter shapes had to be designed and verified. Since this would have been a formidable task if it had been done manually, a computer program was written to find the filter weights. This program optimized the filter weights based on the noise equivalent bandwidth and the geometric mean center frequency specifications, which were the severest design constraints. A closed-form solution was used for the noise equivalent bandwidth and the geometric mean center frequency was optimized by using a binary search. The program completed the design task after a month of unattended operation on an HP 9000 Series 200 Computer!

Front Panel

The front panel on the HP 3561A is specifically designed for ease of use and ruggedness. The user interface with the instrument is through keys grouped by function and associated softkey menus that direct the user through the available options. The major functional areas are Measurement, Trigger, HP-IB, Display, Marker, Instrument State, and Input. These groups are clearly indicated on the front panel so that the user knows which keys to press for any desired function. In addition to physical grouping, the different types of keys are color coded so that related functions are closely associated. Within each functional group are specific keys that either perform a specific action or bring up softkey menus for further selections (Fig. 4). The softkey menus displayed along the right edge of the CRT screen show only the choices that are available for the mode selected.

The keyboard is made up of a new conductive rubber key set that is custom molded for the HP 3561A's front panel. This keyboard allows multiple colors and very sharp lettering to be molded into the keys. The keys have a tactile feedback that indicates that a key has been pressed. In addition to their attractive appearance, the conductive rubber keys provide a ruggedness that is commensurate with the portability of the instrument. These rubber keys are very difficult to damage, even if objects are dropped on them. Furthermore, they provide resistance to dust and moisture and are very easy to install, reducing the assembly time and the cost of the instrument.

Softkeys

The softkey menu interface allows the user to be guided through the selections required for a setup. When a key, such as the **MODE** key in the Measurement group, is pressed,

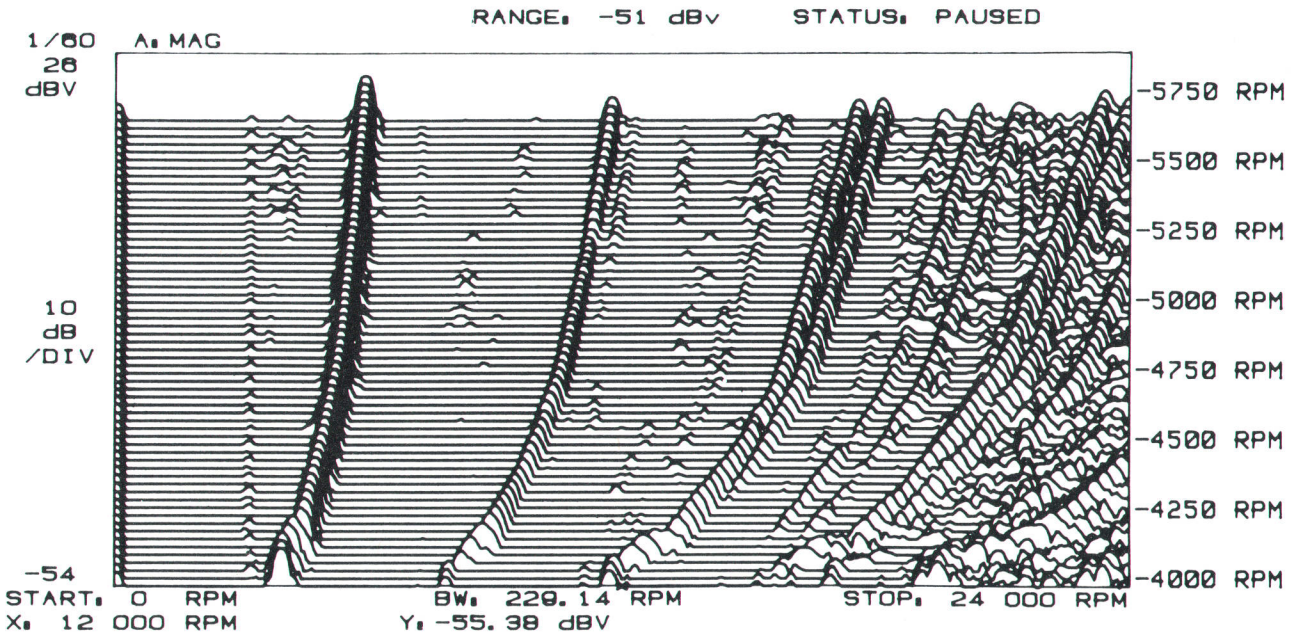


Fig. 5. *Spectral map display. Up to 60 magnitude traces can be displayed simultaneously for studies of changes in the spectrum as a function of time, r/min, etc. These maps can be saved and recalled later from the optional nonvolatile bubble memory.*

the associated choices then available to the user are displayed along the right edge of the display (see Fig. 1). For the **MODE** key, the resulting choices are NARROW BAND, THIRD OCTAVE, FULL OCTAVE, TIME CAPTURE, EXT SAMP ON/OFF, DEFINE PULS/REV, and TEST SELECT. This is a good representation of the different types of softkeys available. The first four softkeys make up a group of selections, only one of which is allowed at a time. When the selection is made, the associated softkey label is intensified so that it is obvious which softkey is presently in effect. The next type of softkey is the on/off softkey or two-state softkey. This softkey allows the user the choice of having the labeled function either on or off. For example, the EXT SAMP ON/OFF softkey allows the user to supply an external sample clock instead of using the internal one. When an on/off softkey is pressed, it toggles between on and off with its current state indicated by either the ON or OFF in its label being intensified. Another type of softkey is typified by the DEFINE PULS/REV softkey. This softkey requires a numeric entry, as do all softkeys that have the word DEFINE in their labels. When the label for this type of softkey is intensified, any numeric entries made will be for the parameter indicated on its label. If more than one DEFINE softkey label is available on a menu, only one will be intensified at a time. The units for the entry will be displayed on the softkey menu after any numeric key of the front panel's keypad or the intensified DEFINE softkey is pressed.

The final type of softkey available is the select softkey as represented by the TEST SELECT softkey. When a softkey with SELECT in its label is pressed, another menu is displayed with more selections. (See article on page 17 for more detail about the HP 3561A's software.)

Display Formats

Once the data is collected, the next step is to display it in a manner that best allows the user to analyze the results. Because of the many different types of measurements possible, the type of analysis performed will vary depending upon the application. For this reason, several different display modes are provided. We have already discussed the upper/lower display format in which two separate traces of any type can be displayed simultaneously, one above the other. There are many uses for this and it is the format that the instrument powers up in. Some examples are looking at magnitude and phase at the same time, observing the time waveform along with its magnitude or phase spectrum, or viewing the full 100-kHz baseband spectrum along with a zoom spectrum. In the time capture mode, viewing a compressed time record and a magnitude spectrum together allows the user to see which part of the large time record is currently being processed.

A single display format is also provided so that a display can be expanded to take up the entire screen when only one trace is necessary to display the measurement. Another display format that is very useful is the front/back display. This display also allows any two traces to be viewed simultaneously. However, they are both expanded to fill the entire screen with one positioned behind the other. The back trace is half the brightness of the front trace so that the two traces can be easily compared. The front and back traces can be switched by pressing the **NEXT TRACE** key.

The most versatile display format is the spectral map display (Fig. 5) obtained by pressing the MAG MAP/CLEAR softkey in the display format menu. Up to 60 magnitude traces can be simultaneously displayed, one behind the other, with a hidden-line algorithm applied so that the display is more viewable. The traces can be positioned either directly behind each other or slightly offset along the horizontal axis so that data that may be blocked by the hidden-line algorithm can be seen. Through the use of the **NEXT TRACE** key and the Marker group keys, the full accuracy of the instrument is available while viewing any one of the 60 traces. The **NEXT TRACE** key allows stepping through the map with the selected trace intensified. The marker function then reads out the x and y values for the location of the marker to full precision. If it is desirable to expand the selected trace to fill the screen, a toggle key allows the user to switch between the spectral map and the selected trace. The data in the map is preserved in the HP 3561A's memory until a new map is generated by pressing the MAG MAP/CLEAR softkey again. If the bubble memory option is installed, the entire map can be stored in nonvolatile memory for recall later.

Portability

Dynamic signal analysis applications are not always confined to the lab bench. These measurements are often made in remote locations such as on aircraft or at remote communications sites. Vibration analysis can be performed on machines that are located on shop floors, in power plants, in chemical factories, or in submarines. Hence, the HP 3561A is packaged much the same as a portable oscilloscope. It weighs 33 pounds. A convenient carrying pouch is attached to the top of the instrument to store cables, manuals, and other equipment.

Portability implies more than just a package; if a lot of additional equipment is required, portability is decreased. For example, when the instrument is carried to a measurement site, it may be desirable to bring the data back to a central computer for further analysis or storage. The optional bubble memory for the HP 3561A allows up to 127 traces to be stored and retained even when the power is turned off. The traces can be recalled at any time for further analysis or transfer to another computer.

When making vibration measurements, some kind of transducer is necessary. To reduce the number of transducers that have to be carried around with the instrument, inexpensive integrated circuit piezoelectric (ICP) accelerometers can be left attached to the machine under test. The current source required to power them is built into the HP 3561A as a standard feature.

Acknowledgments

Larry Whatley and Howard Hilton provided leadership and Jack Armo provided valuable marketing assistance for the development of the HP 3561A. Glenn Engel, Jerry Weibel, Ramona Johnson, John Ketchum, Dave Larson, Rob Mitchell, and Dennis O'Brien worked on the firmware design. Glen Purdy and Jerry Ringel worked on the firmware and hardware.

Thanks go to all the people who tested the final software, and to Mark Grosen, for their contributions.

Hardware Design for a Dynamic Signal Analyzer

by James S. Epstein, Glenn R. Engel, Donald R. Hiller, Glen L. Purdy, Jr., Bryan C. Hoog, and Eric J. Wicklund

DESIGNING THE HARDWARE for the HP 3561A Dynamic Signal Analyzer involved the constraints of portability, adequate shielding of the signal processing circuits against electrical noise, a compact power supply, high-speed digital circuitry, and an accurate clear display of data and calculated results. Supporting and simplifying some portions of this hardware design are an extensive softkey menu structure and a sophisticated operating system (see article on page 17).

Hardware Design

A block diagram of the HP 3561A is shown in Fig. 1. The front end performs the signal conditioning required to scale and filter the data so that it can be digitized by the analog-to-digital converter (ADC). After the signal is digitized, it is further processed by the digital filters to provide frequency translation and increased resolution. The main processor controls the entire instrument to perform the tasks required and communicates with the display to show the results of the measurement. The power supply provides the power to run all the internal circuitry and the nonvolatile memory stores traces, setups, and time records for further analysis. An optional bubble memory can be added to extend the available memory.

Front End

The input amplifier circuitry (Fig. 2) in the HP 3561A provides a floating, shielded input connection for rejection of common mode and ground loop induced signals. A grounding switch is included for chassis-referenced measurements. The signal path does not become ground referenced until it leaves the ADC and crosses an isolation transformer. Therefore, all circuitry through the ADC runs on separate floating power supplies.

The input signal passes through a 0-, 20-, or 40-dB attenuator (depending on range) into a JFET input amplifier stage. Four fixed-gain amplifiers, each preceded by a programmable attenuator, provide a total gain of 38 to −40 dB in 1-dB steps. This gives the HP 3561A a range of full-scale input values from 27 to −51 dBVrms in 1-dBVrms steps.

The full-scale rms output voltage level of the input amplifier is 228 mV. Gain is distributed through the amplifier chain to maintain a reasonable signal level for best noise figure and distortion performance. The resulting noise floor specification is -141 dBV/$\sqrt{\text{Hz}}$ for frequencies less than 1 kHz and -150 dBV/$\sqrt{\text{Hz}}$ above 1 kHz. Distortion is less than −80 dB relative to full scale. Internal offset voltages are corrected by an eight-bit DAC-controlled autozero circuit, which sums an opposing offset voltage into the last amplifier stage. A binary search algorithm is used by the firmware to set the D-to-A converter.

Following the input amplifier is a passive nine-pole antialiasing filter with a rolloff above 100 kHz. The signal is then buffered, sampled, and digitized by the ADC.

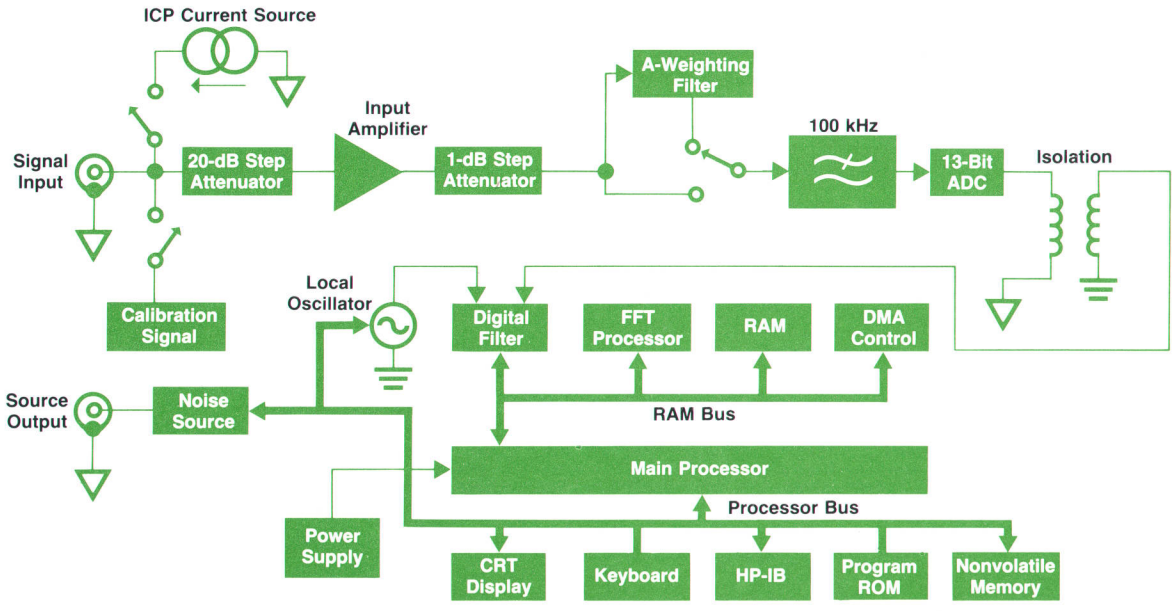


Fig. 1. *Hardware block diagram of the HP 3561A.*

Additional circuitry includes the 4-mA current source for use with external integrated circuit piezoelectric (ICP) accelerometers. An internal A-weighting filter can be selected to provide an ANSI standard simulated "human ear" transfer function in the amplifier path.

A-to-D Converter

Requirements of the ADC are set by the bandwidth and dynamic range specifications of the instrument. The anti-aliasing filter technology used allows bandwidths up to approximately 0.4 times the sample frequency, which is fairly close to the theoretical Nyquist limit of 0.5. This results in a sample rate of 256,000 samples per second for the instrument bandwidth of 100 kHz. Dynamic range considerations imply a resolution of 13 bits and maximum nonlinearity of ¼ LSB (least-significant bit).

The HP 3561A's ADC achieves its high accuracy by using a technique known as residue conversion. The final result is formed by performing successive passes or partial A-to-D conversions of increasing accuracy and decreasing range. Each new pass starts by converting the ADC output (the result of the previous passes) to an analog signal using a DAC and then subtracting it from the ADC input. The resulting residue represents the errors in the previous approximation. The residue is then converted with a resolution and accuracy much greater than for the previous pass. The result is added to the previous approximation of the ADC input to produce a much more accurate approximation.

A partial block diagram of the ADC is shown in Fig. 3. This two-pass converter uses the same local ADC (an ADC within an ADC) to perform the conversions for each pass. During the first pass the ADC input signal is routed to the local ADC through an 8:1 attenuator. The result of the local ADC conversion is the first-pass approximation of the input and is stored in latch L1. This first-pass approximation is the input to the upper eight bits of the 13-bit DAC, which converts with a linearity equivalent to ¼ LSB of 13 bits.

The second pass begins as the first-pass approximation is latched and the DAC begins to settle. The DAC output is subtracted from the ADC input signal and the result is the residue representing the errors in the first-pass approximation. The residue is multiplied by four by an amplifier stage before the local ADC performs the second-pass approximation. The gain ratio between the first and second pass is 4/(1/8) or 32, so the second pass resolves with 32 times greater accuracy, but has a range 32 times less. The second-pass approximation is then added to the first-pass approximation in adder A1, but is shifted down five bits to account for the gain ratio of 32 between the first pass and the second pass. This final result is stored in latch L2 before being transmitted to the digital filters.

The most important aspect of a residue converter is the cancellation of first-pass errors. If an error E is made in the first pass, the DAC is loaded with the correct approximation to the input plus E. This is subtracted from the input so that the residue is only −E. This is then converted and added to the first-pass approximation to produce the final result so that −E, the residue, cancels the E term contributed by the first-pass approximation.

An important parameter in a residue converter is the overlap between passes. The HP 3561A's ADC has three bits of overlap between the first and second passes, meaning that the upper three bits of the second-pass approximation have the same binary weight as the lower three bits of the first-pass approximation. The overlap is a measure of the magnitude of residue that can be converted by the second pass without overloading. If the first-pass approximation were perfect, then only quantization errors would be present in the first-pass approximation and no overlap would be required.

The significant sources of linearity errors in this design are the DAC errors, the ADC errors on the second pass, and the analog gain matching between the first and second pass (how exactly the gain ratio of 32 is held). The gain matching determines how much the first-pass errors contribute to the nonlinearity of the converter. Since, in the best case,

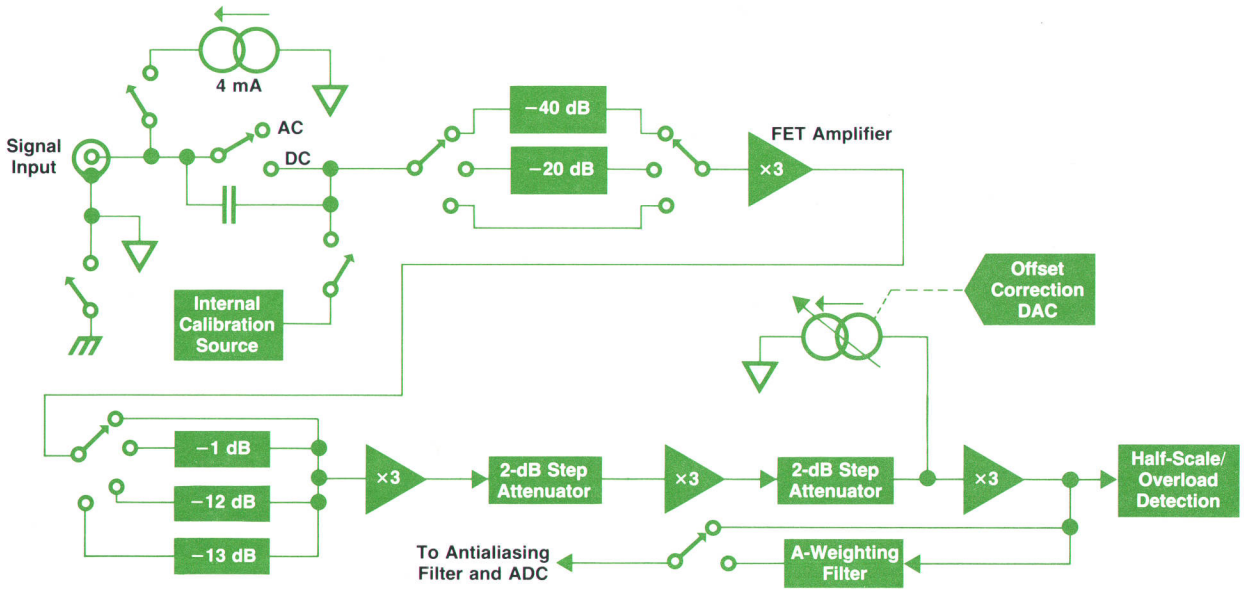


Fig. 2. *Input circuitry of the HP 3561A.*

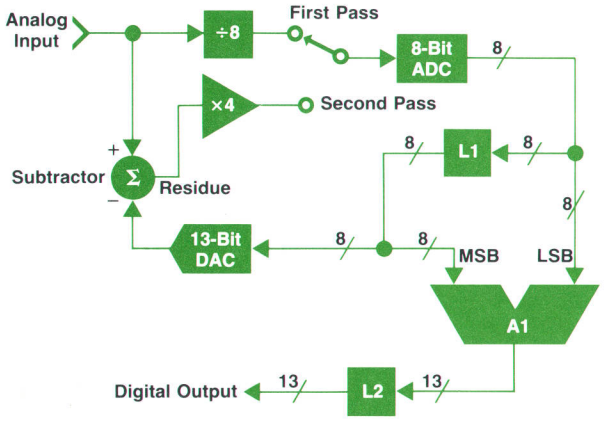


Fig. 3. *Partial block diagram of the ADC used in the HP 3561A's front end.*

the first-pass approximation still has quantization errors (the conversion is only eight bits, not 13), the gain matching is always important.

Most ADCs (and DACs) have particularly nasty forms of nonlinearity. Common forms are step and impulse voltage transfer function errors. A step error is a constant error that is only added to the output when the input is in a certain range. This can occur when the analog weight associated with a particular bit is slightly in error. A common occurrence is when the MSB (most-significant bit) of a DAC is in error. In this case a constant error is present whenever the input is greater than half scale. An impulse error is an error that is very localized. These nonlinearities result in distortion products that do not decrease (and can even rise) as the input level is lowered. This type of behavior is not desirable in a spectrum analyzer. One technique particularly effective for reducing this behavior is dithering (also discussed in the article on page 28).

Dithering is the addition of an extra signal to the input to "smear out" nonlinearities. This causes any particular input level to correspond to different points on the voltage transfer function at different times. In the long term the net transfer function at a particular input level becomes the weighted average over the neighborhood of that level in the original transfer function. The weighting function is the probability distribution function for the dither signal. Another way of looking at the effect of dithering is that harmonic products tend to be converted into intermodulation products between the input and dither signals when the nonlinearities are localized as typical ADC and DAC nonlinearities are. This intermodulation effect implies that the dither signal should be broadband enough to distribute these products into many frequency bins so that the absolute magnitude at any frequency is much lower than if the products were distributed over only a small group of frequencies. This is equivalent to randomizing the errors so that they are not correlated with the input signal.

A natural choice for a broadband dither signal is random noise. However, dithering only smooths out nonlinearities over a range the size of the dither signal, so distortion improvement is negligible for input levels greater than the dither level. This implies that the dither signal must be of a greater amplitude than the input signal to receive any distortion improvement. One way to increase the level of dither while not increasing the noise floor is to subtract the dither signal after the conversion. However, this rules out many sources of dither signals since the dither signal must be easily available in digital form. Hence, the use of pseudorandom noise is a natural choice.

The HP 3561A's ADC uses this form of dithering in a special way. For the dithering signal to be available both as an analog signal and as a digital signal implies that some sort of additional data conversion must be present, usually a DAC. It is possible to exploit the residue converter topology and achieve the same result without requiring another DAC. Fig. 4 shows the ADC with the dithering circuitry added.

Since the first-to-second-pass gain match is not perfect, it is necessary to dither the first-pass conversion to randomize first-pass nonlinearity errors. Here the dithering signal Nf looks just like any other first-pass error and so is subtracted out in the combination of the first-pass and second-pass conversions. Since this dither signal appears in the DAC input and hence the residue, the DAC and second-pass conversion are somewhat dithered. However, since the first pass quantizes Nf, the dither signal has only several values and is not able to smear out nonlinearities adequately. A more continuous range of dither values is provided by adding pseudorandom noise to the first-pass conversion using adder A2. This dither signal also acts like a first-pass error since it is added before latch L1. The result is that all of the sources of nonlinearity are dithered without the use of an extra DAC or ADC. The only extra circuitry required is a simple analog noise source Nf (actually one bit of pseudorandom noise passed through an RC filter to produce a more uniform probability distribution), the pseudorandom noise source, and adder A2, which performs only carries since the noise source has a zero mean. The circuitry within the shaded area in Fig. 4 is realized with a gate array in the HP 3561A, keeping the added cost of this dithering scheme very low.

Processor

The main processor of the HP 3561A is an 8-MHz 68000 microprocessor with 393,216 bytes of program ROM. This processor controls the hardware and the user interface and performs the display processing and much of the data processing (see Fig. 1). It interfaces to the other hardware subsystems via two different 16-bit buses: the processor bus and the RAM bus. The processor bus interfaces the main processor to the standard CMOS and optional bubble nonvolatile memories, the display hardware, the local oscillator (LO), the HP-IB (IEEE 488) interface, the source, and the keyboard.

The HP 3561A has 131,072 bytes of dynamic RAM, which is accessed through the RAM bus. This RAM is organized as 64K 16-bit words, but its data can be accessed as bytes by the main processor. This RAM is used for the processor's stack and program variables and for all the data buffers needed during the data processing. The RAM bus is controlled by a bus arbiter and the dynamic RAM controller. It has DMA ports to the dedicated FFT processor and the outputs of the digital filter control chip. The digital filter control chip performs DMA accesses using two address

counters which control where in RAM the data is written. The RAM bus also has some control and status ports used by the main processor for interfacing to the FFT processor, the digital filter control chip, the digital filter DMA address counters, timing logic, and the interface to the HP 3561A's front-end circuitry and ADC.

Nonvolatile Memory

The standard nonvolatile memory in the HP 3561A consists of 8K bytes of CMOS RAM. This RAM is used to store data for two traces and six instrument setups, with one of the setups being the last state before power-down. The CMOS RAM is backed up by a three-volt lithium battery, which allows it to retain data for up to five years.

There is also a bubble memory option that can add a megabit of nonvolatile memory. This allows users to store extended time records, as well as more setups and traces. The bubble memory can be accessed by a simple processor interface and can transfer data at a rate of 12.5K bytes/s. These speeds allow the user to retrieve an extended time trace of forty records in 0.7 second, worst case.

Noise Source

The noise source in the HP 3561A can generate three types of signals: periodic, impulse, and random. All three of the signals are band-limited and band-translated. Like the tracking oscillator in a swept analyzer, the noise source always tracks the selected frequency span.

The impulse source has the added capability of being triggered from the HP 3561A's front panel or via the HP-IB, and it can skip time records. These features are useful when transient inputs are required to settle before a new measurement is made. The impulse source can skip up to 32,766 records before generating another impulse.

Digital Filter

The digital filter section is based on custom ICs developed especially by HP for dynamic signal analysis. Controlled by the main processor, this section filters the digitized input waveform to form the appropriate data records for the FFT processor. See the article on page 28 for a detailed discussion of the design constraints and realization of this section.

Display

The HP 3561A uses a raster-scanned magnetic-deflection CRT for display of measurement results. This is a departure from some other Hewlett-Packard signal analyzers, which use a vector-scanned electrostatic-deflection CRT display. The raster-scan technology offers significant advantages for the HP 3561A. Probably the most important is compact size. With magnetic-deflection CRTs, large electron beam deflection angles are easily achieved. Because of this, magnetic-deflection CRTs can be much shorter than electrostatic-deflection CRTs for a given viewing area. The space saved by using a magnetic-deflection display helped fit the HP 3561A's extensive measurement capability into an oscilloscope-sized package. Another advantage of the raster-scan technology is low cost. The HP 3561A display, including drive electronics, costs about as much as an electrostatic CRT by itself.

Display quality was not sacrificed to gain these advantages. Because of their shorter electron beam path length, magnetic-deflection CRTs give bright, crisply focused displays. Vector scanning is limited in the number of traces that can be displayed before flickering becomes annoying. Raster scanning does not have this problem. This allows the HP 3561A to display measurement results in three-dimensional form with up to 60 traces of 400 points each (see Fig. 5 on page 10). That's the equivalent of over 24,000 vectors. Also, raster scanning consistently produces good-quality alphanumerics. Well annotated displays are an important part of the friendliness of the HP 3561A.

Display quality is further enhanced by scanning the raster vertically at a rate of 36,205 Hz with the vertical lines formed from left to right. Since most rapid trace variations occur vertically, this gives a smoother trace by reducing the jagged line appearance compared to a horizontally scanned raster.

The digital portion of the display contains a memory control circuit, a special sync circuit, and a processor inter-

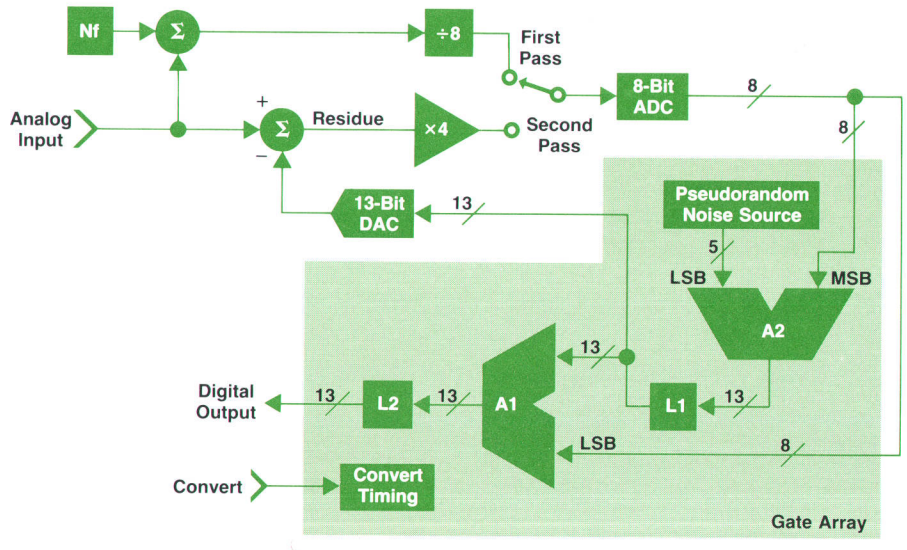


Fig. 4. *ADC of Fig. 3 with dithering circuitry added. This circuitry is incorporated as a gate array design to keep the added cost low.*

face. The display is bit mapped with a vertical resolution of 256 pixels and a horizontal resolution of 512 pixels. The pixel bit map is stored in two sets of eight 16K × 1 dynamic RAMs. One set is for full-intensity pixels and the other is for half-intensity pixels. Different intensities are used to give emphasis to specific areas on the CRT. To allow the display to be synchronized to the power line frequency, a special sync circuit is used which does not require phase-locked loops or analog circuitry. The circuit free-runs during display updates and synchronizes during retrace intervals when minor timing variations are not noticeable. The addition of the synchronization circuit requires only one more flip-flop.

The processor interface allows quick generation of display graphics. It includes alphanumeric generation capability and simple line-drawing hardware that can draw vertical, horizontal, 45°, and dotted lines at 1,400,000 pixels per second. It also supports several block operations such as set, clear, and complement at 11,300,000 pixels per second. These hardware features significantly reduce the amount of software overhead required for a bit-mapped display.

Power Supply

Requirements for the HP 3561A's power supply were that it be lightweight and compact for instrument portability and that it provide eight different outputs, four of which are isolated floating supplies. A switching supply design seemed to be the natural choice for the job. The potential disadvantage of using this type of supply is that it can be a source of EMI (electromagnetic interference). For this reason, an out-of-band switching frequency of 128 kHz synchronized to the system clock is used.

A half-bridge, off-line circuit topology was chosen, using 500-volt VMOS switching transistors (Fig. 5). Control loop circuitry regulates the main 5V logic supply. The output voltage is normally under control of a voltage feedback loop, but during overcurrent conditions a separate loop takes control to provide foldback current limiting. Both loops have their own loop shaping networks. All other supplies are derived from other windings on the same transformer and thus are slave preregulated by the regulation of the 5V supply. Each of these other supplies is then further regulated by a linear three-terminal regulator. This provides additional ripple rejection and supply isolation and the preregulation maximizes the efficiency of the linear regulators.

The power supply is protected by four separate protection circuits. The four fault-detection circuits are floating, nonfloating, and 5V overvoltage comparators and a primary-overcurrent-sense circuit. All protection circuit outputs are monitored by a latching shutdown circuit that turns off the gate drive to the power VMOS transistors and indicates on one of four LEDs which fault condition occurred.

Packaging and Shielding

The HP 3561A package is similar to that used for portable oscilloscopes. The circuit boards sit in a vertical card cage and plug into a motherboard. Since the display and the switching power supply were identified as potential EMI sources, they are both placed behind a shield that runs the length of the instrument.

Suppression of residual spurious signals by more than −80 dB relative to full scale required shielding of both the display and input amplifier. Electromagnetic field frequencies of concern are the display's vertical scan rate of 36,205 Hz and its harmonics. Electric fields caused by the high accelerating potentials for the CRT are attenuated by a grounded screen that covers the CRT face. High-impedance circuits in the HP 3561A's input amplifier are covered by a grounded aluminum shield can.

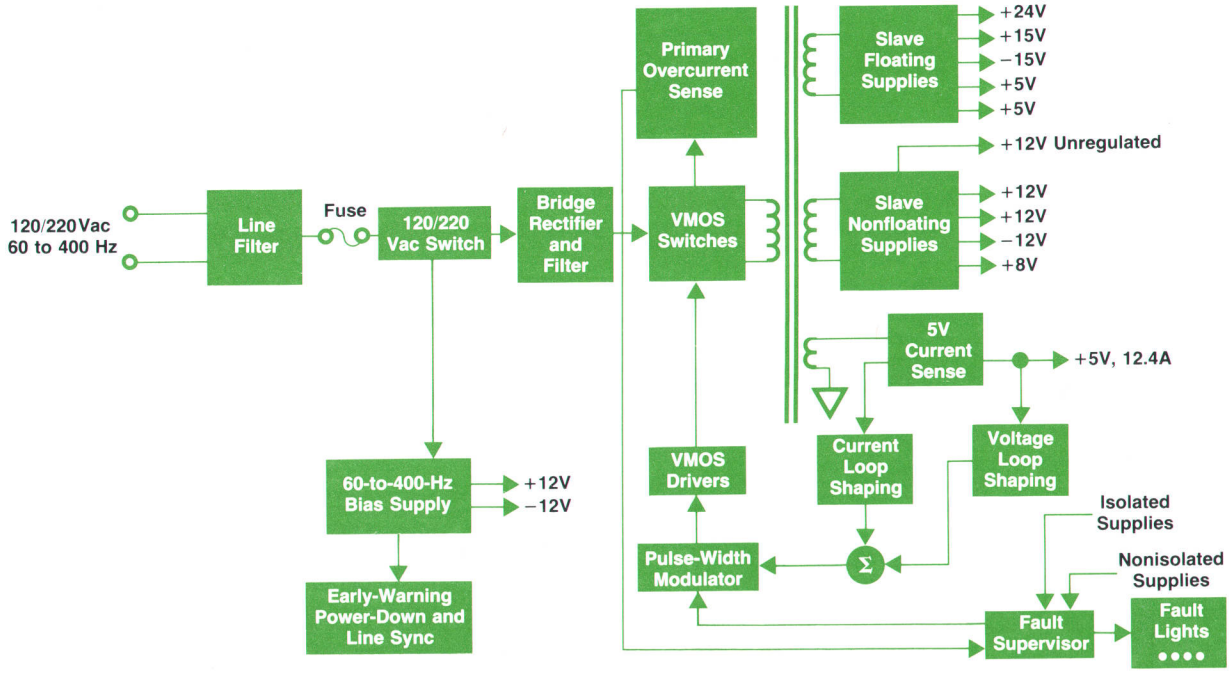


Fig. 5. *Block diagram of the HP 3561A's power supply.*

Varying magnetic fields are generated by the display's magnetic elements (flyback transformer and yoke) and by other ac high-current carrying conductors. These fields can generate currents in the instrument frame, causing local field radiation near the input amplifier. The initial field levels near the input amplifier were small (milligauss levels), but the performance specifications required additional attenuation of more than 40 dB. Shield materials have lower permeabilities at these low field levels, making the task more difficult. To simplify the problem, field radiation is reduced at the flyback transformer by a shorted turn of aluminum wrapped around the transformer. Circulating currents in this loop create opposing fields (Lenz's law), to help cancel some of the radiation, which in turn reduces field levels at the input amplifier. Further attenuation is achieved by enclosing the input amplifier board in a MuMetal™ shield. The CRT assembly (including yoke) is enclosed in another MuMetal shield, both to reduce its radiation and to shield the CRT from external fields that can cause display distortion. Fields from the power supply's magnetic components have frequencies at 128 kHz and its harmonics and thus are above the HP 3561A's measurement baseband of 100 kHz (one of the primary reasons for switching the power supply at 128 kHz). Aliased spurious supply radiation problems are eliminated by synchronizing the supply to the system clock, placing the aliased products at dc. Careful attention to grounding techniques also played an important part in the total solution. The resulting spurious signal levels are less than −80 dB relative to full scale on the HP 3561A's most sensitive range (−51 dBVrms).

MuMetal is a trademark of Allegheny-Ludlum Steel Corporation.

Acknowledgments

Many people deserve recognition for their part in the development of the HP 3561A Analyzer. Early project management and product definition were done by Lynn Schmidt. Tom Rodine worked on the initial input amplifier design. Chuck Kingsford-Smith and Al Gee did the original work on the ADC and Ray Blackham designed the associated gate array and made other valuable contributions. Jerry Ringel designed the optional bubble memory board and did software design. Dean Payne worked on the LO. Paul Richer was the mechanical designer and helped with shielding solutions. Randy Eilert and Scott Lockhart were the industrial designers. Dick Bingham developed the product test system. Jerry Metz, John Ketchum, Dennis O'Brien, Jerry Weibel, Ramona Johnson, and Rob Mitchell were software designers. Larry Whatley, Mike Aken, Charlie Potter, and Nick Pendergrass deserve recognition for their fine leadership as project managers.

Instrument Software for Dynamic Signal Analysis

by Glenn R. Engel and Donald R. Hiller

THE SOFTWARE for a microprocessor-controlled instrument determines the instrument's usefulness and personality—how it interacts with the user and how easy it is for a new user to understand and use effectively. Because there are many variables in dynamic signal analysis measurements, the software for the HP 3561A Dynamic Signal Analyzer is particularly important. Some of the key features provided by the HP 3561A's operating system are:

- Friendly user interface
- Autocalibration
- Overlapped processing for quick display updates
- Spectral map display
- Direct output to HP-IB (IEEE 488) printers and plotters
- Nonvolatile memory storage
- Self-tests and service tests.

Architecture

At the heart of the HP 3561A's software is a 4K-byte multitasking operating system that provides process synchronization and resource control. There are five operating system tasks in the HP 3561A (see Fig. 1).

- I/O control
- Command processing
- Measurement/display
- Nonvolatile memory control
- Message handling.

The I/O control process interprets commands from the HP-IB and front panel. This process checks syntax and parameter limits and does immediate execution of simple commands. An example of a simple command would be setting the marker position. Commands that cannot be executed immediately such as buffer dumps and nonvolatile memory operations are passed on to the command process for completion. The command process then coordinates the operations necessary to complete the command. This includes locking necessary resources, communicating with other processes to handle parts of the command, and then unlocking resources after command completion.

The measurement/display process is responsible for all

signal processing functions and display updates. Signal processing functions include data collection, fast Fourier transforms of data (see box on page 20), amplitude corrections, phase corrections, averaging, and data scaling. The measurement/display process takes advantage of overlapped processing to increase data throughput. This process is shown in Fig. 2 and is discussed later.

The nonvolatile memory process is responsible for all nonvolatile memory functions such as saving and recalling traces and instrument states. With the optional bubble memory installed, the nonvolatile memory process becomes very important, because of the slower access time of bubble memory devices. By handling memory operations with a separate process, the HP 3561A can be making useful measurements while traces or states are being saved or recalled from the bubble memory.

The message process handles error messages and user warnings. This function is a separate process because messages need to be prioritized and most warnings have a display timeout. By using the message process, other processes do not have to worry about what happens to a message once the message process knows about it.

The operating system provides several functions to manipulate semaphores. The primary request and release functions provide a synchronization mechanism for resource allocation and are used to control counting semaphores for message communication between processes. When a process requires a resource, it calls the operating system's request function. If the resource is available, the process continues execution. If not, the operating system suspends the process and does not allow it to run until the resource is available. A resource becomes available when the process using it calls the operating system release function. When a resource is released, the operating system checks for any processes waiting for the released resource and schedules them for execution.

Friendly User Interface

The software for the HP 3561A's user interface was written to provide the user with as much feedback as possible concerning error conditions and instrument operation. For example, if the user enters an illegal value for a numeric input, the HP 3561A, instead of just beeping, responds by telling the user exactly what the minimum and maximum values can be. In addition, the software helps the user by giving suggested operation hints. For example, when the pseudorandom noise source is selected, the HP 3561A reminds the user that the uniform window is the most likely window selection for that noise source. If the uniform window is already chosen, no message is given. The friendly user interface extends to the HP-IB interface as well as front-panel operation. A diagnostic mode allows for easy debugging of HP-IB error conditions. This mode displays each character received from the HP-IB interface on the instrument's display and marks the exact location of syntax violations, which can save programmers considerable debugging time.

Autocalibration

Calibration in the HP 3561A is a three-step process. First, the software characterizes the amplitude and phase response of the instrument's analog front end over the entire operating range. Second, the software computes the front-end correction for the frequency range currently selected for measurements. This range, or span, can be a small portion of the HP 3561A's baseband of 100 kHz. Third, the software simulates the digital filters to determine the effects of the filtering operation on the magnitude and phase of the input signal. This results in one correction that represents any errors introduced by the HP 3561A between its input connector and the final digital representation of the input signal in RAM.

The key to providing high amplitude accuracy for the HP 3561A is directly related to its ability to measure its own magnitude and phase response properly. An accurate digital-state-to-analog-level (1-bit D-to-A) converter determines the accuracy of the internal calibration signal; it is driven by a pseudorandom noise generator. The spectrum of this calibration signal consists of a number of frequencies (determined by the length of the pseudorandom sequence) that are of nearly constant magnitude and whose magnitude and phase are theoretically exactly known. This allows rapid calibration cycles, since both magnitude and phase of all desired calibration frequencies can be determined with a single operation.

One problem of this autocalibration scheme is slew-rate limiting of the analog version of the signal, since the ideal signal has step changes between the two levels. If the slew-rate limiting is symmetrical, the effect is that of passing the signal through a low-pass filter with an impulse response that has the width of the slew time. For high slew rates (which are present in the HP 3561A's calibration circuit), this filtering causes negligible errors. On the other hand, if the slew rates are not symmetrical, the effect is that of adding a small signal that is the derivative of the pseudorandom sequence. This additional signal causes small but significant errors in the amplitudes of the pseudorandom noise frequencies and the resulting calibration constants. Fortunately, these errors are nearly identical and of the same phase when an amplitude-inverted pseudorandom sequence is passed through the same circuitry. This allows the error to be canceled by subtracting a measurement of the inverted sequence from a measurement of the noninverted sequence. Another way of looking at this situation is that the errors are even-order and can be cancelled by the same technique used in a differential transistor pair or a push-pull amplifier to reduce even-order

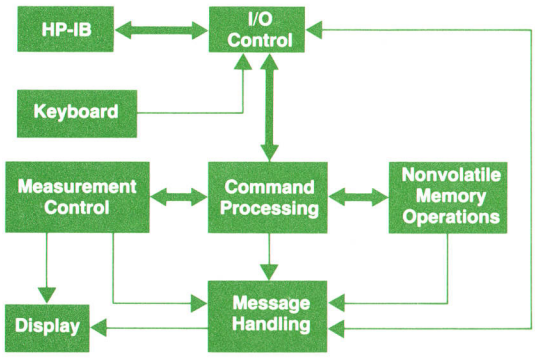


Fig. 1. *Operating tasks for the HP 3561A's software.*

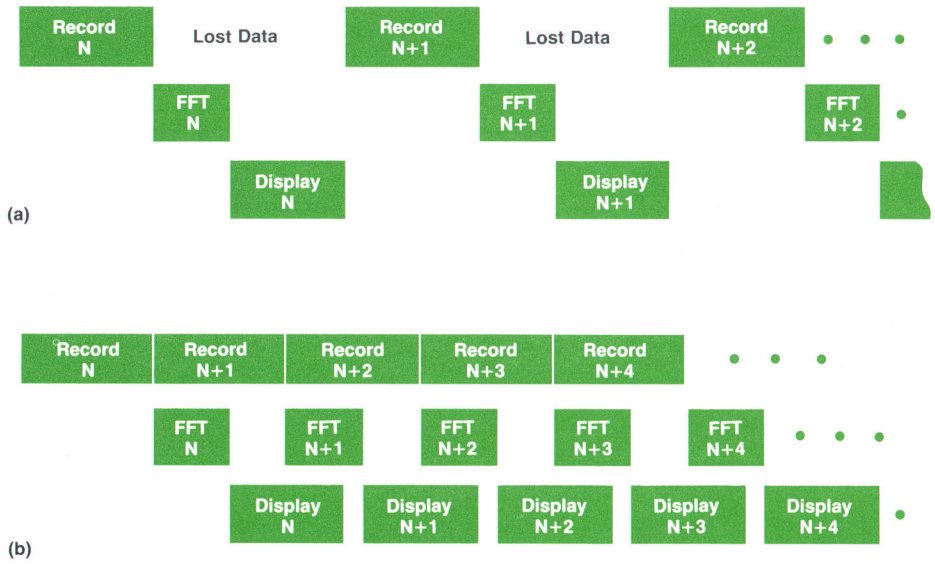


Fig. 2. *(a) The result of serial processing—data is lost while the processor is computing the FFT and updating the display. By using overlapped processing (b), records can be processed at a greater rate without losing data.*

distortion. Differences in propagation delays in the digital circuitry driving the analog conversion circuitry also cause the same kind of errors and so are also canceled.

By using pseudorandom noise autocalibration, the HP 3561A is able to correct for antialiasing filter rolloff well past the cutoff frequency of 100 kHz. However, because of possible aliasing problems and decreased dynamic range, the accuracy of the HP 3561A is not specified above 100 kHz. The resulting accuracy specification is ±0.15 dB from 0 to 100 kHz.

Overlapped Processing

Overlapped processing is possible in the HP 3561A thanks to a multiport RAM. This RAM is accessible from four different sources: the main processor, the slave FFT processor, and two DMA (direct memory access) channels. This allows the HP 3561A to almost triple its real-time rate by doing three operations simultaneously. Fig. 2 illustrates the possibilities. While the main processor is manipulating record N for display, the FFT processor is processing record N+1, and the DMA sequence is collecting record N+2.

Spectral Map Display

The software for the HP 3561A's spectral map display (see Fig. 5 on page 10) is quite simple. This display format provides a three-dimensional representation of the analyzed input signal versus time. As new spectra are placed into the map, they must be drawn as if they were behind the previous spectra. This operation is similar to viewing a distant range of mountains from an airplane. Mountain ranges that are behind others are only visible if they are higher than the ranges in front of them. That is, higher mountains can obscure lower mountains behind them. The hidden-line algorithm used by the HP 3561A uses exactly this idea. When a new spectrum is put on the display it is compared point by point with the previous spectrum to decide whether it is "visible." This process is simplified for a raster scan display because the two spectra only need to be compared at a limited number of points—once every horizontal pixel position. No interpolation is needed between points and only the largest value for each horizontal pixel position must be remembered for later spectra.

HP-IB Plots and Printer Dumps

HP-IB plots and printer dumps are supported on the HP 3561A to allow the user to produce a hard-copy representation of the display. Printer dumps are achieved by simply reading the bit-mapped display memory and sending this information to a printer that supports raster-type graphics.

Plots are obtained by sending HP-GL plotter commands to a pen plotter. A high-quality copy of the HP 3561A display is obtained in this manner. Plotting the spectral map is more complex than generating the raster-display spectral map. When plotting the spectral map, the HP 3561A must keep a copy in its memory of the largest values plotted to aid in hidden-line removal. These values are kept with enough resolution to provide correct hidden-line removal down to the width of the plotting pen. HP-IB plots also allow the user to annotate specific marker values to show test results or merely to provide extra documentation.

Acknowledgments

We would like to acknowledge the engineers who contributed significantly to the design and success of the HP 3561A software. John Ketchum was responsible for much of the measurement software, Jerry Weibel and Ramona Johnson designed the HP-IB and user-interface software, Dennis O'Brien implemented the nonvolatile memory software, Jerry Ringel was responsible for implementing the full-octave and ⅓-octave software, Dave Larson implemented the trace math and HP-IB plot functions, Rob Mitchell implemented the autorange and time-averaging functions, and Glen Purdy was responsible for the self-tests and service tests.

FFT Implementation

One of the benefits of using an analyzer based on a fast Fourier transform (FFT) is measurement speed. The HP 3561A Dynamic Signal Analyzer uses the FFT algorithm to analyze signals with the equivalent of a parallel bank of 400 narrowband filters. This simultaneous spectral analysis technique is inherently much faster than using a swept-frequency spectrum analyzer that must cover the same frequency span serially, one frequency at a time. Although not yet practical at RF and microwave frequencies because of the fast FFT computation rate required, current technology does allow its use at low frequencies. This speed improvement can be put to good use in many of the HP 3561A's applications. The typical display update rate of five to ten spectra per second makes circuit adjustments easier. A noisy signal can be averaged 1000 times in less than a minute in the fast averaging mode.

These fast measurements require a fast FFT computation. In the HP 3561A, the FFT computation is performed by a TMS320 microprocessor. This processor has a reduced instruction set designed for digital signal processing applications. The processor and its support chips occupy only 30 square inches of printed circuit board area. It computes a 512-point complex FFT in 35 ms and a 1024-point real FFT in 42 ms, including the window multiplications.

Speed

Fast execution times are obtained by tailoring the FFT algorithm to the internal architecture of the TMS320. A modified radix-8, Sande-Tukey twiddle-factor algorithm[1] is used. With a slight rearrangement of the multiplications, this algorithm fits nicely into the TMS320's sum-of-products architecture. A radix-8 algorithm was chosen for several reasons. With this radix, bit-reversing is replaced by digit-reversing, that is, groups of bits are reversed at one time, which is easier and only needs to be done for every eighth data address. Also, a radix-8 algorithm requires fewer of the slower I/O operations to the instrument's main RAM. To reduce the effect of slow RAM accesses further, the algorithm performs arithmetic operations using the on-chip cache RAM while waiting for the I/O operations.

Noise and Accuracy

Noise performance of the FFT algorithm is equally important for the HP 3561A. A numerical algorithm is only as accurate as the number of bits used in the calculations. Inaccuracies in calculating the FFT can degrade absolute amplitude accuracy and linearity and generate phantom spectral components. These are all lumped into the category of "noise." In the HP 3561A, the FFT noise performance can be no better than the RAM data bus width, which is 16 bits. Given a perfect algorithm (no errors), the maximum dynamic range would be approximately 96 dB (6 dB per bit). This is more than adequate to ensure the instrument dynamic range specification of 80 dB, including all scaling overheads and specification margins.

The challenge is keeping a 16-bit FFT algorithm accurate to 16 bits at its output. The algorithm reads data words from main RAM, performs arithmetic operations, rounds the results and stores them back in main RAM. This occurs once for each FFT level, and there are four or five levels per transform. If a 16-bit data point is multiplied by a 16-bit coefficient, the result is 31 bits long. If this result is then summed with eight others just like it, three more bits are added and the result could overflow the TMS320's 32-bit accumulator. Therefore, prescaling is necessary.

The HP 3561A's FFT algorithm uses a simplified block scaling technique. It is implemented by saving the magnitude of the largest data point observed for an FFT level. The next level chooses to divide by one, two, four, or eight based on the previous level's maximum. This scaling method gains much of the performance improvement of a true block floating-point algorithm without the added complexity or speed reductions. With this type of scaling, the instrument noise floor drops with the level of the input signal.

Another important consideration is the nature of the noise floor. A noise floor that is uniform versus frequency, or white, is preferred. This means that the average power of the error in each of the output frequency bins is the same. Most of the bins are the result of successive multiply-accumulate operations, where many bits are discarded in the rounding operation. It is fairly easy to keep the error in these bins consistent. The problem bins are the ones that are the result of an accumulate only. In these bins, very few bits are thrown away during rounding. The error has only a few possible values, and there is an expected error, or bias, in the result. For example, if two 16-bit values are added and the 17-bit result is rounded back to 16 bits, the error is either zero or ½ LSB (least-significant bit). There is a bias of ¼ LSB. When these biased results are summed in following levels, the bias error adds up in a linear fashion.

Generally, FFT noise floors have spur-like bumps located on the output bins with large expected errors. A user might be tempted to interpret these as spurious components in the input signal even though they are outside of the specified dynamic range. To solve this problem, two types of rounding offsets are used. One results in a positive bias, and the other results in a negative bias. By appropriately choosing between these when rounding, the bias errors cancel.

The FFT limits amplitude accuracy in yet another way: the windowing operation. The input data block is multiplied by a tapering function, or window, that forces the start and the end of the record to zero. This reduces the effects of leakage, which occurs when the signal is not periodic in the measurement interval. Windowing is often thought of in the frequency domain as the frequency response of the analysis filters. To make accurate amplitude measurements, the effective filter frequency response should be very flat over a frequency span of plus or minus half of a bin spacing. The HP 3561A's flattop window function is flat to ±0.003 dB, much flatter than those used in other FFT-based analyzers. This flatter window is needed to complement the improved amplitude measurement accuracy of the HP 3561A.

Circuitry

The TMS320 hardware looks like a slave processor to the HP 3561A's 68000 main system processor. The 68000 sets up a small table in main RAM that tells the TMS320 what type of FFT to perform, scaling restrictions, data buffer start addresses, and other setup information. Then the TMS320's reset line is set true and FFT computation begins. When finished, the TMS320 writes the scale factors in main RAM and interrupts the 68000. Main RAM is shared among several processors and DMA ports. A bus arbitrator establishes priority and grants RAM accesses.

Acknowledgments

Ron Potter designed the improved flattop window.

Reference

1. E.O. Brigham, *The Fast Fourier Transform*, Prentice-Hall, 1974.

Bryan C. Hoog
R&D Engineer
Lake Stevens Instrument Division

Index HEWLETT-PACKARD JOURNAL

Volume 35 January 1984 through December 1984

Hewlett-Packard Company, 3000 Hanover Street, Palo Alto, California 94304 U.S.A.
Hewlett-Packard Central Mailing Dept., Van Heuven Goedhartlaan 121, 1181 KK Amstelveen, The Netherlands
Hewlett-Packard (Canada) Ltd., 6877 Goreway Drive, Mississauga, Ontario L4V 1M8 Canada
Yokogawa-Hewlett-Packard Ltd., Suginami-ku, Tokyo 168 Japan

PART 1: Chronological Index

PART 2: Subject Index

PART 3: Model Number Index

97080A	HP-UX, HP 9000 Model 520, multiuser	Mar.
97089A	HP-UX, HP 9000 Models 530 and 540, multiuser	Mar.
97935A	Data Pack	Jan.
98760A	Color Monitor Assembly	May
98770A	High-Performance Color Display	May
98780A	Monochromatic Monitor Assembly	May

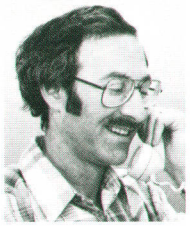
PART 4: Author Index

Authors

December 1984

4 — Dynamic Signal Analyzer

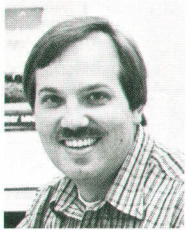
James S. Epstein



Jim Epstein has contributed to the firmware and digital hardware design of a number of HP products, among them the HP 3820A Total Station, the HP 3497A Data Acquistion/Control Unit, and the HP 3561A Dynamic Signal Analyzer. He was firmware project manager for the HP 3561A and is a coauthor of an HP Journal article about the HP 3497A, a coinventor for a patent related to a circuit for a distance measuring device, and a member of the Instrument Society of America. He studied electrical engineering at the Polytechnic Institute of New York (BS 1964) and the University of Missouri at Rolla (MS 1969) and worked for a major semiconductor manufacturer before joining HP in 1972. He also served two years in the U.S. Army, attaining the rank of first lieutenant. Jim was born in New York City. He is married, has three children, and lives in Loveland, Colorado. His outside activities include tennis, sailing, skiing, and swimming.

12 — DSA Hardware Design

Eric J. Wicklund



With HP since 1977, Eric Wicklund designed the input amplifier and switching power supply for the HP 3561A Analyzer. He also contributed to the development of the HP 3497A Data Acquistion/Control Unit and is currently a project manager for new sources. Born in Seattle, Washington, he received a BSEE degree in 1977 from the University of Washington and now lives in Woodinville, Washington. He is married, has three children, and is interested in photography, woodworking, and amateur radio (KR7A, 20-m CW). Eric just completed a two-year project of building his own home.

James S. Epstein

Author's biography appears elsewhere in this section.

Glenn R. Engel

Author's biography appears elsewhere in this section.

Donald R. Hiller

Author's biography appears elsewhere in this section.

Bryan C. Hoog

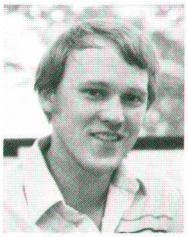
Bryan Hoog joined HP in 1978 after receiving a BSEE degree from the University of Missouri at Rolla. He designed the analog portion of the HP 3561A Analyzer's display and its FFT processor and 1/3-octave filter shapes. Born in St. Louis, Missouri, he now lives in Everett, Washington. His hobbies include photography, piano, hiking, tennis, soccer, and amateur archaeology.

Glen L. Purdy, Jr.

An R&D engineer with HP since 1979, Glen Purdy worked on the hardware and firmware for the HP 3561A Dynamic Signal Analyzer. He studied electrical engineering at the University of Michigan and received a BSEE degree in 1977 and an MSEE degree in 1979. Born in Eugene, Oregon, he now lives in Everett, Washington with his wife. He is a member of the IEEE and his interests include scuba diving, softball, and alpine skiing.

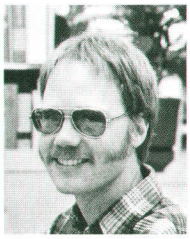
17 — DSA Software Design

Glenn R. Engel



Responsible for the operating system, display graphics, and spectral map display design for the HP 3561A Analyzer, Glenn Engel is a software designer at HP's Lake Stevens Instrument Division. With HP since 1979, he also contributed to the software for the phase noise measurement system in the HP 3047A Spectrum Analyzer. Born in Kansas City, Missouri, he attended Kansas State University and received BS degrees in electrical engineering and computer science in 1979. Living in Everett, Washington with his wife, he enjoys outdoor activities such as backpacking, skiing, and watching his wife ride in jumping events at horse shows. He also likes woodworking and communicating in code on 20 meters as an amateur radio operator (KC7SJ).

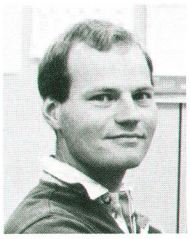
Donald R. Hiller



Born in Durango, Colorado, Don Hiller received a BS degree in engineering and applied science from the California Institute of Technology in 1975 and an MSEE degree from Stanford University in 1976. He then joined HP and has done R&D work on disc drives, fiber optics, digital signal processing, and A-to-D converters, including the ADC used in the HP 3561A Analyzer. He is a coauthor of three papers related to ADCs, fiber optics, and semiconductor measurements. Married and the father of a son, Don lives in Lake Stevens, Washington. He is active in his church and enjoys backpacking and hiking in the Cascade mountain ranges, riding his motorcycle, and reforesting the five acres around his home.

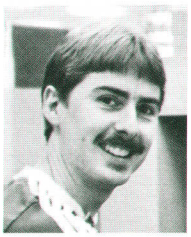
28 — Custom Digital Filters

Charles R. Panek



Born in Edison, New Jersey, Charlie Panek studied electrical engineering at the nearby Stevens Institute of Technology and received a BE degree in 1978. He then joined HP and while working full-time, continued his education to receive an MS degree from Colorado State University in 1981. At HP he worked briefly on 1/3-octave applications software before starting work on the custom digital filters used in the HP 3561A Analyzer. Charlie is currently working on digital signal analysis for a new product. Outside of work, he enjoys playing soccer, cooking, and touring on his motorcycle. He lives in Bothell, Washington.

Steven F. Kator



Steve Kator joined HP in Loveland, Colorado in 1979 and later moved to Everett, Washington with the Lake Stevens Instrument Division in 1983. He was responsible for designing the mixer multiplier and the rounding circuits used in the custom digital filters for the HP 3561A Analyzer and has contributed to several other IC development projects. Steve is an alumnus of Washington University (BSEE 1979) and Stanford University (MSEE 1983), and is a member of the IEEE. He recently left HP to do research in radio astronomy at the California Institute of Technology. He is newly married and lives in Pasadena, California. His interests include soccer, skiing, backpacking, beermaking, motorcycling, and performing early music.

Custom Digital Filters for Dynamic Signal Analysis

by Charles R. Panek and Steven F. Kator

DIGITAL FILTERING offers some powerful advantages when used with a fast Fourier transform (FFT), because it can perform frequency domain analysis in a narrow band around some arbitrary center frequency. To understand the need for a digital filter in dynamic signal analysis, let's first review the basics of FFT analysis and sampled systems.

Basic FFT Analysis

The fast Fourier transform is a special case of the discrete Fourier transform, a mathematical algorithm that translates a series of samples of a signal, evenly spaced in time, into a similar series of samples of the frequency spectrum of the signal. A finite set of these evenly spaced time samples is known as a time record. A time record with 1024 samples spaced 100 microseconds apart will be transformed into a frequency series or spectrum of 1024 evenly spaced samples between dc and the 10-kHz sampling frequency f_s. The frequency domain information for frequencies above $f_s/2$ is redundant, that is, the spectrum is essentially folded around $f_s/2$. This phenomenon determines the minimum rate at which a signal of a certain bandwidth can be sampled. Thus, provisions must be made to ensure that the sampled signal contains no frequency components of significance higher than $f_s/2$ so that the FFT results do not fold back on themselves, an effect called aliasing.

A very basic FFT-based signal analyzer (Fig. 1) consists of a sampler preceded by a low-pass antialiasing filter with a cutoff frequency less than $f_s/2$, and an FFT processor that operates on fixed-length time records. But what if the user has an interest in signals at very low frequencies, or in signals spaced very close together in frequency? With a sampling rate as described above, this analyzer would only have about 10 Hz of frequency resolution. How can this be improved? One possibility is to lengthen the time record. A 10,000-point time record would allow 1-Hz resolution. However, it would take 10 times as long to capture the data, and roughly 13 times as long to calculate the FFT.

Another possibility is to reduce the sampling frequency. This reduces the spacing of frequency components by a similar amount; a 1-kHz sampling frequency gives about 1-Hz resolution with a 1024-point transform. This would be an ideal solution, except that the cutoff frequency of the low-pass input filter must then be changed to ensure that the spectral content of the sampled signal is entirely below half the new sampling frequency. To allow for many different bandwidths, we would have to be able to switch between many analog low-pass filters, or perform a complicated retuning of a single filter.

There is another way to lower the sampling rate, however, and that is by reducing the bandwidth of the signal after it is sampled and then discarding samples at regular intervals. This process is known as sample rate decimation, and the low-pass filtering is performed by a digital decimation filter. To increase the resolution by a factor of five, we would pass the output of our sampler through the digital filter, filtering the data to one-fifth the original bandwidth, and then discard four out of every five samples. We can do this because the filtered signal is now oversampled, that is, it contains redundant information.

Digital filtering offers some other special advantages. The performance of a digital filter is a function of the coefficients used and the length of the data word. Once a digital filter design is completed, the bandwidth, flatness, noise, phase response, and other parameters of the filter are exactly repeatable from circuit to circuit. The adjustments and tolerance variances characteristic of analog filters are eliminated.

Zoom Operation

By cascading decimating digital filters, we can reduce the bandwidth and increase the frequency resolution by larger amounts. We still have an important restriction, however: the maximum frequency of our analysis continually drops lower. Suppose we wish to look at the modulation sidebands of a 4-kHz carrier frequency. If the modulation is very low in frequency, say 1 Hz, we will be unable to resolve the sidebands with the simple system of Fig. 1. Simply decimating the converter output by a factor of two will filter out the carrier, sidebands and all, assuming a 10-kHz sampling frequency.

The nature of the FFT limits us to frequency analysis of a time series whose frequency content extends from dc to half the sampling frequency. We can overcome this limitation by borrowing a technique from the designers of radio

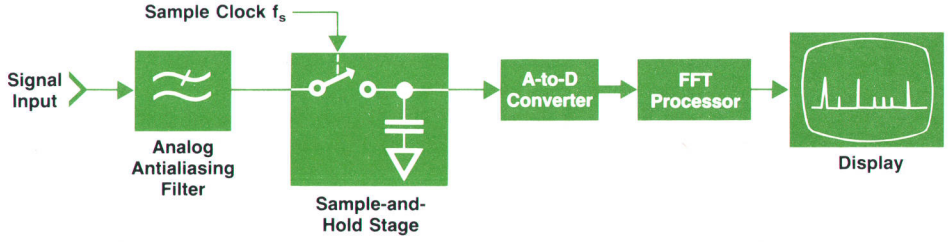


Fig. 1. *Block diagram of a basic fast Fourier transform signal analyzer.*

receivers and swept-frequency spectrum analyzers. That is, demodulation or analysis is performed in a fixed intermediate frequency (IF) band, and any frequency band of interest is mixed with a local oscillator (LO) signal to translate it to the IF band.

Here, we can consider our IF band to extend from dc up to half the sampling frequency, adjustable in bandwidth by the decimation rate of the digital filter. Thus, we can mix the 4-kHz carrier frequency with a 4-kHz local oscillator frequency and analyze the dc and low-frequency products using a decimated sample rate that will give us the desired resolution. In addition, we can perform quadrature mixing, that is, multiply the input by sine and cosine waves at the LO frequency. By processing the resulting complex numbers in a complex FFT, frequency data both above and below the LO frequency can be obtained without generating unwanted images. We can once again use the advantages of digital signal processing to do this mixing in a stable and accurate way. The output of the analog-to-digital converter (ADC) is fed into a pair of digital multipliers, along with sine and cosine wave sequences of the desired center frequency sampled at the same rate. The resulting digital products are fed into a pair of identical digital filters for sample rate reduction. This is referred to as "zoom" operation, for it allows the user to expand the view of a given portion of the input signal spectrum.

Architecture and Algorithm

We have arrived at the basic digital signal processing block diagram of the HP 3561A Dynamic Signal Analyzer (Fig. 2). A 100-kHz analog antialiasing filter precedes a sample-and-hold circuit and ADC running at a sample rate of 256 kHz. This sample rate gives nicely spaced frequency points when used with 2^N-sized time records (in this case, 1024 samples). The digital LO generates cosine and sine sequences, sampled at the ADC sample rate, which are fed into the mixer multipliers of the real and imaginary component digital filters, respectively, along with the ADC output. In baseband mode, where no frequency translation is desired, the cosine output of the LO is a constant, and only the real component filter is used. The 20-bit output of the filter chips is rounded to 16 bits by the control IC and fed into RAM, where it is accessed by the FFT processor.

It was determined on the basis of noise performance and data scaling that the digital filters should use 20-bit numbers to represent internal data. The A-to-D sample rate is specified at 256,000 samples per second. Hence, to allow serial processing of a decimated sample between each new ADC sample, a clock rate of 2 (samples) × 20 (bits) × 256 kHz, or 10.24 MHz, would be required.

Because of power limitations and other concerns, an alternative to strictly serial processing is used. Bits of the data word are processed in pairs of adjacent bits. Thus, the two least-significant bits (LSBs) are processed on the first clock cycle, the next two bits are processed on the second cycle, and so forth. This lowers the required clock rate to 5.12 MHz.

Mixer Multiplier

The multiplier used in the HP 3561A's signal processing circuitry uses a serial paired-bit shift-and-add scheme that operates on two's-complement words. To understand its operation, consider a basic serial/parallel unsigned multiplier as shown in Fig. 3.

The circuit is initialized by clearing all the D flip-flops except the one flip-flop at the rounding point, which is set, and then loading the multiplicand into the parallel latch. As each multiplier data bit is brought in (least-significant bit first), a new partial product is formed at the AND-gate outputs. The partial products are accumulated by the carry-save adders, and the result is presented serially at the output, with one new product bit for every new multiplier bit. The circuit is clocked until all the product bits are shifted out, and the input data is padded with zeros as necessary.

We may enhance this design to accommodate two's-complement numbers by adding a serial negator ahead of the most-significant AND gate in the serial input path. If the

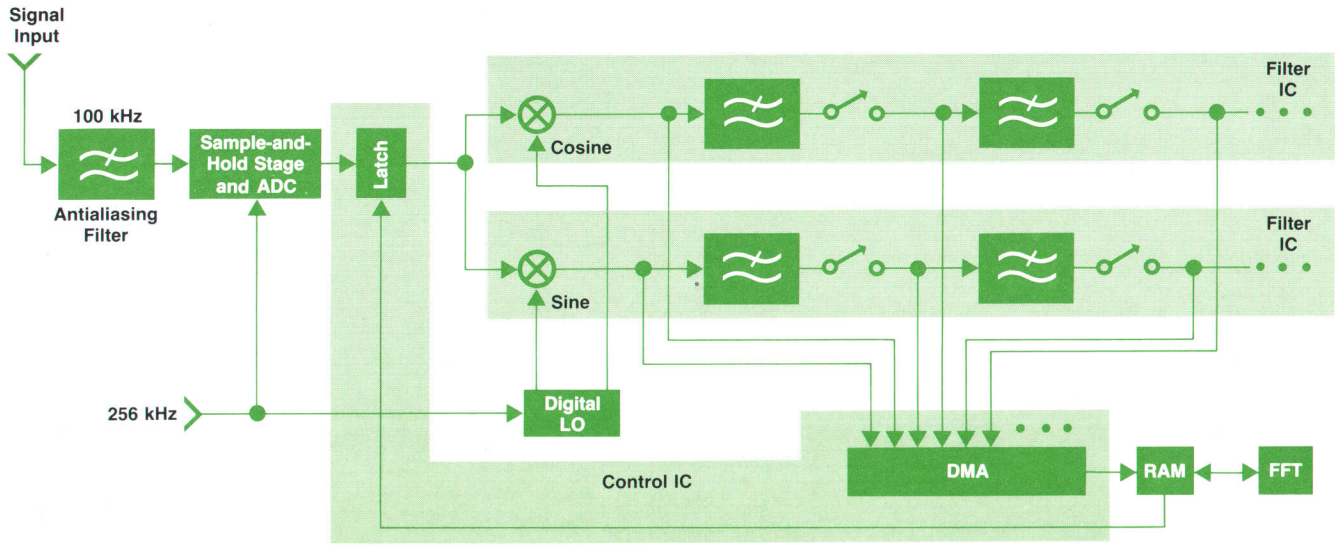


Fig. 2. *Basic digital signal processing block diagram of the HP 3561A Dynamic Signal Analyzer.*

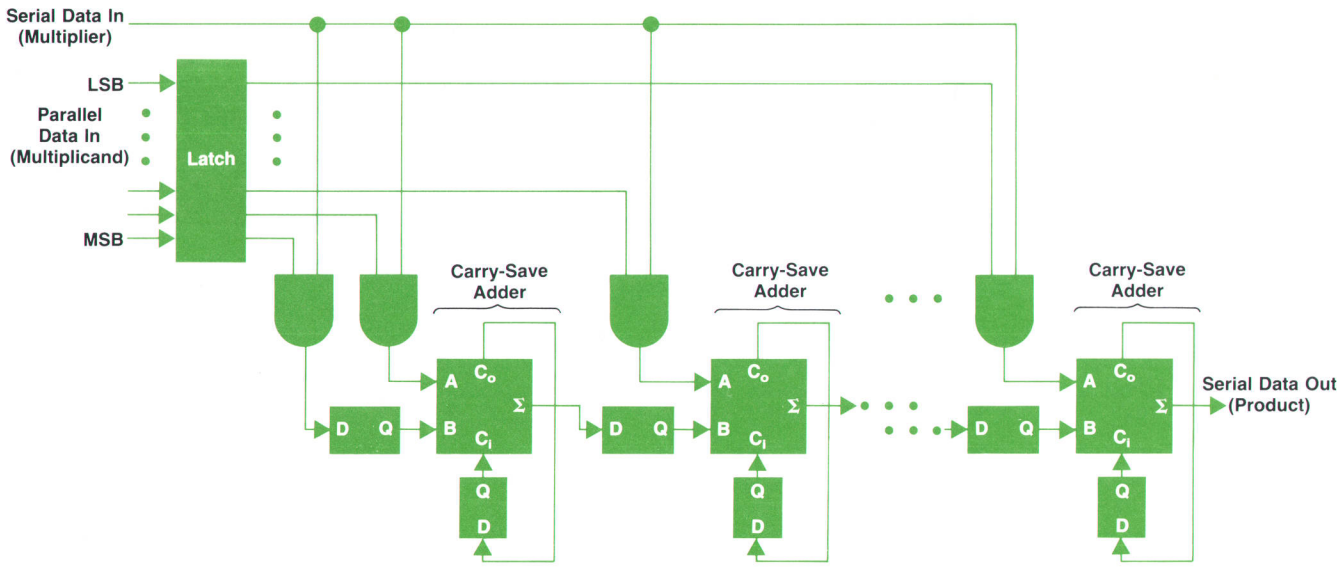


Fig. 3. *Basic serial-parallel unsigned multiplier.*

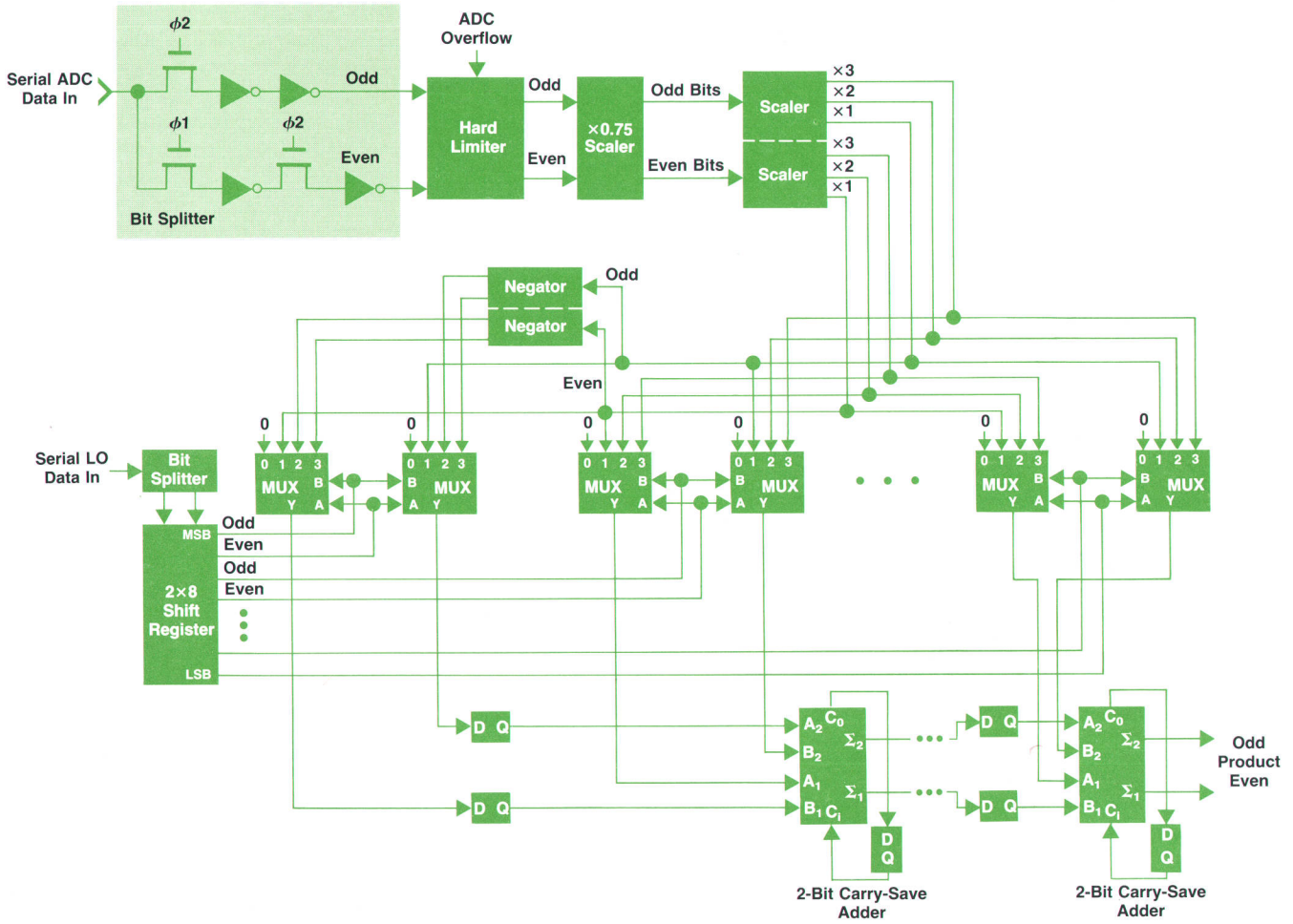


Fig. 4. *Two's-complement paired-bit mixer multiplier.*

multiplicand is negative, this has the same effect as subtracting the last partial product.

This design was adapted to work with bit pairs by splitting the serial input data stream into odd and even bits, and generating three two-bit-wide data streams representing the data multiplied by 1, 2, and 3. The parallel data (multiplicand) is also grouped into pairs. The AND gates of the basic circuit are replaced with pairs of 4:1 multiplexers, with the multiplicand bit-pairs driving the select lines and the $\times 1$, $\times 2$, and $\times 3$ multiplier bit-pairs driving the inputs.

The final design is shown in Fig. 4. The ADC and LO data words are brought in serially and converted to bit pairs. The LO word enters a two-bit-wide, serial-in, parallel-out shift register. The ADC word passes through a hard limiter that prevents overflowed ADC data from entering the filter, and a 0.75 scaler, which reduces the amplitude of the data so that overshoot in the filter will not result in overflows. This multiplier outputs the 20 most-significant bits (MSBs) of the product of a 16-bit LO word and a 13-bit ADC data word in ten clock cycles, which suits the requirements of the digital filter perfectly.

Cascaded, Pipelined, Second-Order Sections

There are two filters on the HP 3561A's custom digital filter chip. One, with a bandwidth of approximately 20% of the input sample rate, is used before reducing f_s by a factor of 2. The other filter, with a 7.8% bandwidth, is used before decimation by a factor of 5. Both of these filters are implemented as a cascade of second-order sections. Both filters also share the same digital circuitry. Filter selection is accomplished by an additional address bit to the on-chip ROMs that contain the filter coefficients. Each filter is an eighth-order elliptical design, with coefficients chosen to meet design criteria while minimizing the number of bits to represent them. Having fewer bits in the coefficients reduces the amount of digital hardware needed to implement the filters.

Pipelining of the circuitry is used extensively between the second-order sections of the filter. That is, data is delayed by one sample between each filter section. Thus, the entire filter processes several samples at one time, which increases throughput without increasing clock frequency.

There are ten channels of data storage associated with each filter section and the output of the filter. Using this memory, data can be recycled through the filter up to ten times after the sample rate has been reduced each time. By processing decimated data between samples from the ADC, the filter sample rate need only be twice that of the A-to-D converter.[1]

Peled-Liu Second-Order Sections

Each second-order filter section must implement a simple difference equation:

$$y_n = a_0x_n + a_1x_{n-1} + a_2x_{n-2} - b_1y_{n-1} - b_2y_{n-2} \qquad (1)$$

where x_n is the nth input word, y_n is the nth output word, and a_i and b_i are the filter coefficients.

To implement this equation with a minimum of hardware, a technique variously known as the Peled-Liu[2] or Princeton multiplier is used. A brief explanation of this technique and its adaptation to this design follows. If we assume that data is represented by two's-complement binary numbers of B bits bounded by 1 and -1, we can represent x_n of Equation 1 by:

$$x_n = -x_n^0 + \sum_{j=1}^{B-1} x_n^j\, 2^{-j} \qquad x_n^j = 0 \text{ or } 1 \qquad (2)$$

where x_n^j is the jth bit of the nth input word. A similar expression represents y_n of Equation 1. Using these expressions, the difference equation can be rewritten as:

$$\begin{aligned} y_n = a_0 &\left[\sum_{j=1}^{B-1} x_n^j 2^{-j} - x_n^0\right] + a_1\left[\sum_{j=1}^{B-1} x_{n-1}^j 2^{-j} - x_{n-1}^0\right] \\ &+ a_2\left[\sum_{j=1}^{B-1} x_{n-2}^j 2^{-j} - x_{n-2}^0\right] - b_1\left[\sum_{j=1}^{B-1} y_{n-1}^j 2^{-j} - y_{n-1}^0\right] \\ &- b_2\left[\sum_{j=1}^{B-1} y_{n-2}^j 2^{-j} - y_{n-2}^0\right] \end{aligned} \qquad (3)$$

We can define a function ψ with five binary arguments as:

$$\psi(x^1, x^2, x^3, x^4, x^5) = a_0x^1 + a_1x^2 + a_2x^3 - b_1x^4 - b_2x^5 \qquad (4)$$

where $x^k = 0$ or 1

Then Equation 2 can be rewritten as:

$$y_n = \sum_{j=1}^{B-1} 2^{-j}\,\psi(x_n^j, x_{n-1}^j, x_{n-2}^j, y_{n-1}^j, y_{n-2}^j) - \psi(x_n^0, x_{n-1}^0, x_{n-2}^0, y_{n-1}^0, y_{n-2}^0) \qquad (5)$$

The function ψ can have only 32 possible values, depend-

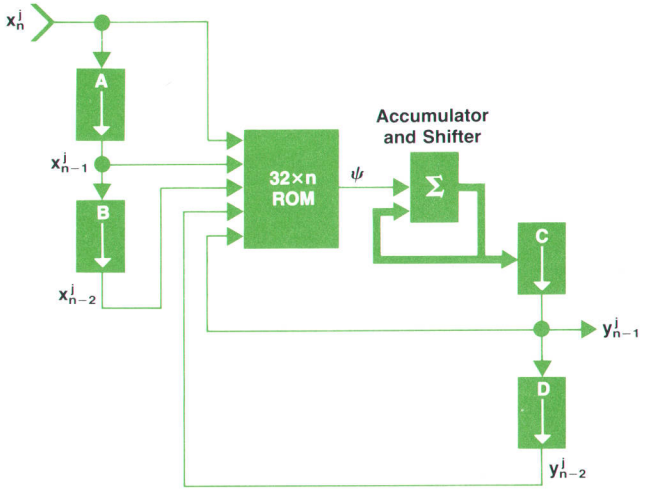


Fig. 5. *Basic Peled-Liu second-order filter. Blocks A, B, and D are serial-in, serial-out shift registers. Block C is a parallel-in, serial-out shift register.*

ing on its five binary arguments. Since the values of a_i and b_i are fixed in each filter section, this suggests implementing this function in ROM. The entire second-order section is shown in Fig. 5.

Shift registers A through D are all B bits long. The nth input word x_n is fed, least-significant-bit first, into the ROM and shift register A. The ROM is simultaneously addressed by the corresponding bits of the x_{n-1}, x_{n-2}, y_{n-1} and y_{n-2} words. The ψ value generated on each clock cycle is accumulated with the result of the previous operation, shifted right one bit. On the first cycle, the accumulator register is preloaded with a rounding value, which is added to assure minimum bias when the result is truncated back to B bits. On the final cycle of each filter operation, the ψ value corresponding to the sign bits of the data word is subtracted from the accumulator, and the accumulator value is parallel-loaded into shift register C.

Paired-Bit Implementation

To maintain a lower clock rate, as mentioned earlier, the design of Fig. 5 was adapted to process bit pairs, using two Princeton multipliers per second-order section. Equation 5 can be rewritten:

$$y_n = \sum_{j=1,3,\ldots}^{B-1} 2^{-j}\psi\,(x_n^j, x_{n-1}^j, x_{n-2}^j, y_{n-1}^j, y_{n-2}^j) + \sum_{j=2,4,\ldots}^{B-2} 2^{-j}\psi\,(x_n^j, x_{n-1}^j, x_{n-2}^j, y_{n-1}^j, y_{n-2}^j) - \psi(x_n^0, x_{n-1}^0, x_{n-2}^0, y_{n-1}^0, y_{n-2}^0) \quad (6)$$

where B is an even number.

This leads to a second-order section design as shown in

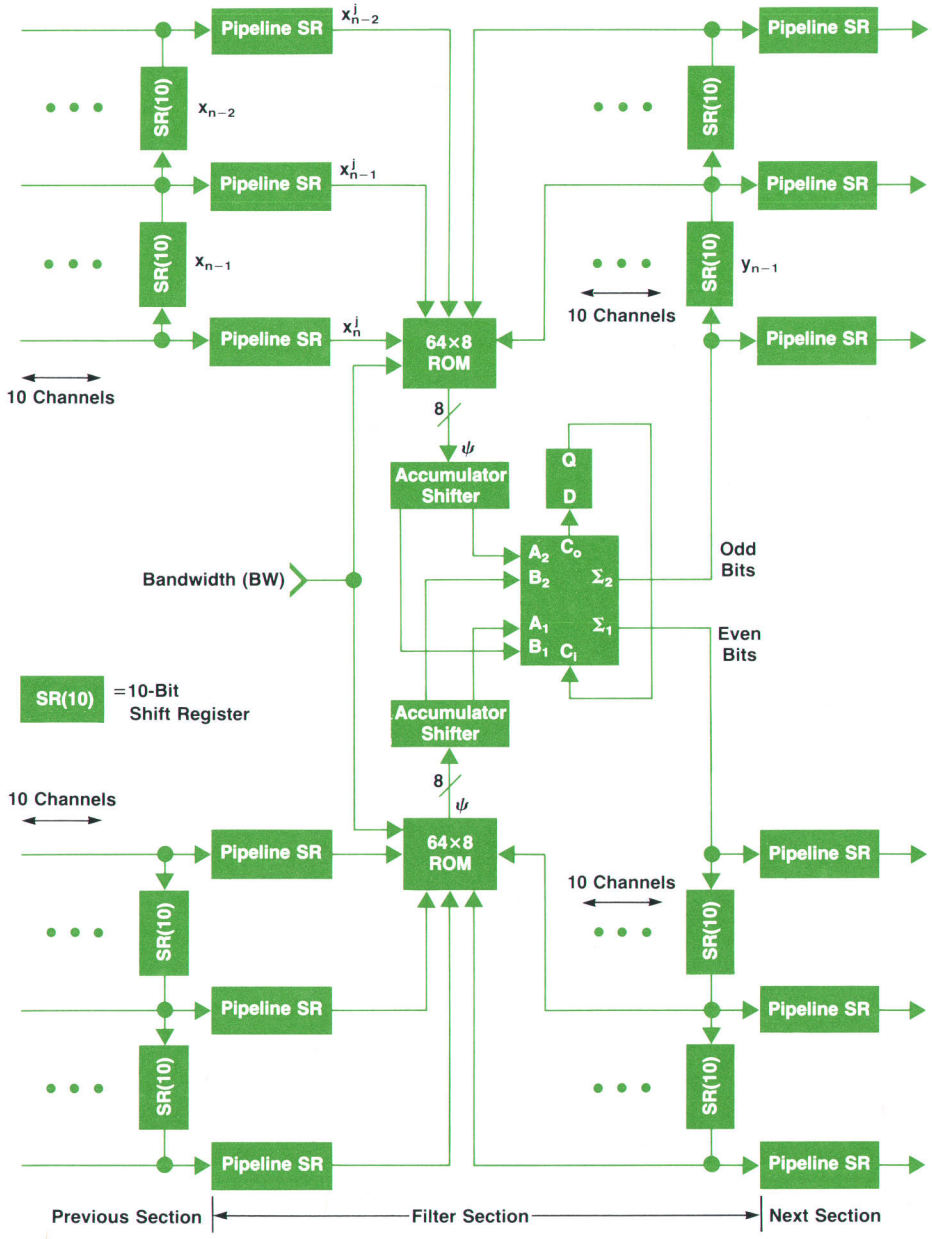


Fig. 6. *Basic circuit of Fig. 5 adapted to process bit pairs using two 10-channel memory banks.*

Fig. 6. Memory is organized as banks of 10-bit shift registers. One bank stores the even bits of the 20-bit data words, and the other stores the odd bits. Each bank consists of 10 channels of memory, one for each pass the data can make through the filter. Thus, a second-order section can filter one sample from the output of the mixer multiplier, and then store away its state and operate on a sample from a data sequence that has already passed through the filter and had its sample rate reduced.

Each Princeton multiplier generates a partial product, 20 bits long, after rounding. These numbers are then added, two bits at a time, in the carry/save adder. The resulting sum is then split into its odd and even bits, and fed into the appropriate channel of y_{n-1} memory.

The y_{n-1} memory for the first filter section serves double duty as the x_{n-1} memory for the second section. Pipeline registers are used for delaying and resynchronizing the data between sections.

The simplified block diagram for the complete filter is given in Fig. 7. Note that actually five second-order sections are used to implement an eighth-order filter. The first section implements a transfer function consisting of a scaling constant and two poles, and the last section implements a transfer function with only a pair of zeros. This implementation was chosen to minimize circuit area on the IC.

The last section of the filter is followed by an overflow detect and hard-limit circuit. Although the data is scaled down at the input so that internal overflows cannot occur, to preserve unity gain, the data is scaled up in the last section. If an overflow does occur there, the output data is set to the appropriate ± full-scale value.

Once data emerges from the hard limiter, it is stored in the 10-channel-by-20-bit feedback memory, where it is recirculated until new data is written into the same channels. Thus, sample-rate reduction of filtered data is accomplished by merely writing into feedback memory more often than reading from it.

Limit Cycles

Just as there are challenges that face the designer of an analog filter who wishes to maximize noise performance and dynamic range, the digital filter designer must cope with problems unique to digital signal processing. These include noise-like signals and dc offsets generated within the filter because of errors in rounding, and low-level self-oscillations in the absence of a signal.

The phenomenon of limit cycle oscillations occurs in many digital signal processing systems where there is a rounding operation within a feedback loop. As a simple example, consider the following difference equation:

$$y_n = x_n + 0.9\,y_{n-1} \tag{7}$$

If we set $y_{n-1}=0$ and $x_n=1$ when $n=0$, and $x_n=0$ when $n>0$, then the following sequence results:

$$\begin{aligned} y_0 &= 1 \\ y_1 &= 0.9 \\ y_2 &= 0.81 \\ y_3 &= 0.729 \\ y_4 &= 0.6561 \\ &\vdots \end{aligned}$$

As can be seen, the sequence decays exponentially toward zero. Now let's limit the precision of the arithmetic, and round each result to the nearest tenth. Then:

$$\begin{aligned} y_0 &= 1 \\ y_1 &= 0.9 \rightsquigarrow 0.9 \\ y_2 &= 0.81 \rightsquigarrow 0.8 \\ y_3 &= 0.72 \rightsquigarrow 0.7 \\ y_4 &= 0.63 \rightsquigarrow 0.6 \\ y_5 &= 0.54 \rightsquigarrow 0.5 \\ y_6 &= 0.45 \rightsquigarrow 0.5 \\ y_7 &= 0.45 \rightsquigarrow 0.5 \end{aligned}$$

This time the sequence "hangs up" at 0.5; even with no input there is an output. A dc output like this is known as a deadband. When there are two feedback terms in our difference equation as in Equation 1 (thus, implementing a pole pair), the output can decay into an ac oscillation known as a zero-input limit cycle. Bounds for the amplitudes of these limit cycles have been developed.[3] In general, they relate to the Q of the poles of the transfer function. As the Q of the poles increases, limit cycles of higher amplitude become more likely. Since sharp-cutoff filters tend to have high-Q poles, a digital filter designer can expect to have problems with limit cycles when trying to maximize performance.

When incorporating a digital filter in an FFT analyzer, the presence of limit-cycle oscillations is especially undesirable. Because these oscillations tend to be roughly sinusoidal near the resonant frequency of the pole pair, they will appear as a spurious signal in the frequency display of the analyzer when the input is removed or held at very low levels. To add to the confusion, the oscillations can have different values or go away altogether, depending on the trajectory the filter input takes as it decays.

Limit cycles are not readily predictable in complicated

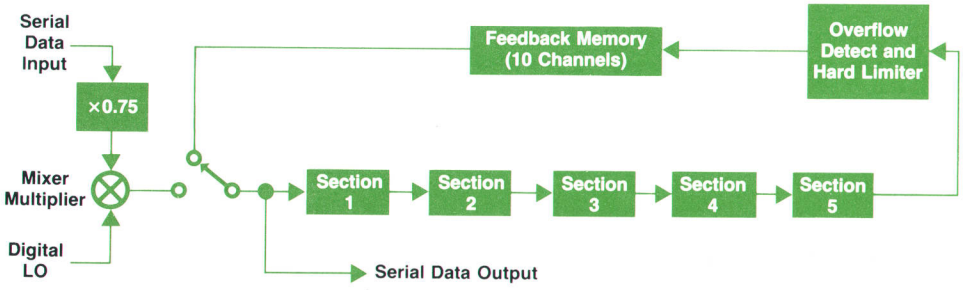


Fig. 7. *Simplified block diagram of complete digital filter.*

filters. For a basic second-order filter, it is a simple matter to find all the zero-input limit cycles by repeatedly allowing the filter output to decay (or settle into oscillation) after presetting the filter's registers to every possible combination of small values (one pair at a time). However, the results of such an experiment do not guarantee the limit cycle performance when the filter input is other than zero, or when the filter is cascaded with other second-order sections.

Deadbands and limit-cycle oscillations can be explained qualitatively as the rounding operation dominating the filter coefficients in determining the decay of the filter's output. A design tactic that has often proven successful is to inject a low-level dither signal at the rounding point. In the example shown above, subtracting a value of 0.05 from the number before rounding will allow the filter output to drop below 0.5. If we add 0.05 as often as we subtract it, the rounding operation will still yield a minimum dc offset.

This solution has its disadvantages, however. If a simple periodic dither signal is used, it too will appear as a spurious signal on the output of the FFT stage, albeit of predictable amplitude and frequency. As a matter of practice, such a dither signal often must be of an amplitude equal to or greater than the amplitude of the oscillations it must break up.

In the HP 3561A Dynamic Signal Analyzer, limit cycles are eliminated by using a dither signal derived from a long pseudorandom noise generator. This proves quite effective in eliminating deadbands and limit cycles altogether. The energy from this noise generator is not concentrated at one frequency, but is spread evenly over the entire spectrum when its sequence length is sufficiently long. Thus, instead of a spurious component at some frequency on the FFT display, the dither appears as an increase in the noise floor. In fact, the peak value of the pseudorandom noise dither at each rounding point in the filter is less than one LSB of the 20-bit data word. Therefore, not only are there no limit cycles or deadbands, but the dither generator is not visible at the output under zero input conditions.

Offsets

Offsets or errors as a result of rounding can result in significant measurement inaccuracies. Offsets in the filter output are most problematic during zoom measurements, because input signals at the same frequency as the local oscillator are downconverted to dc and are displayed at midscreen. A dc error in the filter output thus results in a spurious component at the center frequency on the display.

One would expect that since the filter data is 20 bits long, any reasonable rounding scheme should result in acceptably low distortion. However, several effects exacerbate the problem.

Since the HP 3561A filter is a cascade of five sections, and since the decimation algorithm entails further cascading the filter with itself up to ten times, any consistent offset in the section arithmetic will accumulate. Also, since rounding occurs after each multiply within the second-order sections, and because of the quantization of the coefficients, the results of certain coefficient multiplications have nonuniform statistical distributions. Hence, biases occur. Finally, the dc gain from some internal nodes to the output of the filter can be fairly high, particularly with the higher-Q, divide-by-5 coefficients.

Curing DC Offsets

As a first example, consider the problem of rounding 20 bits to 16 bits. If we simply truncate the unwanted bits, the magnitude of the error introduced will be distributed between 0 and −15 LSBs of the 20-bit word. If we assume that the value of the truncated bits is uniformly distributed, the error will have a mean of −7.5 LSBs (see Fig. 8a).

The traditional solution is to add a value of 8 LSBs (½ LSB of the 16-bit result), which yields a mean error of ½ LSB of the 20-bit word. This is quite good, but not good enough for our application.

Also consider the case of nonuniformly distributed values that are truncated. For example, suppose the first bit below the truncation point has a 75% probability of being zero. The truncation error will be distributed as in Fig. 8b, with a mean of −5.5 LSBs. If we do the standard procedure of adding 8 LSBs, an error of 2.5 LSBs still exists.

The solution employed in the HP 3561A digital filter consists of adding one of two predetermined values before truncation. The two values are calculated so that their mean added to the mean truncation error equals zero. The choice of which value to add is controlled by a pseudorandom register bit with a 50% duty cycle. Thus, on the average, the rounding error will be zero. This technique has two major advantages. It is flexible enough to deal with highly nonuniform distributions, and it allows the designer to adjust the variance in the rounding (noise power) to suit such needs as breaking up limit cycles.

To see how this so-called "dithered rounding" is applied, return to the first example above (Fig. 8a). If we alternately add zero and 15 LSBs to the value before truncation, there will be an average error of zero in the operation. Likewise in the second example (Fig. 8b), we may alternately add zero and 11 LSBs before truncating, again achieving zero mean error. Clearly, this has profound effects on the rounding noise, but the average noise power will generally be less than four times that of simple truncation. The use of random alternation is important, because the alternative use of periodically determined rounding is equivalent to

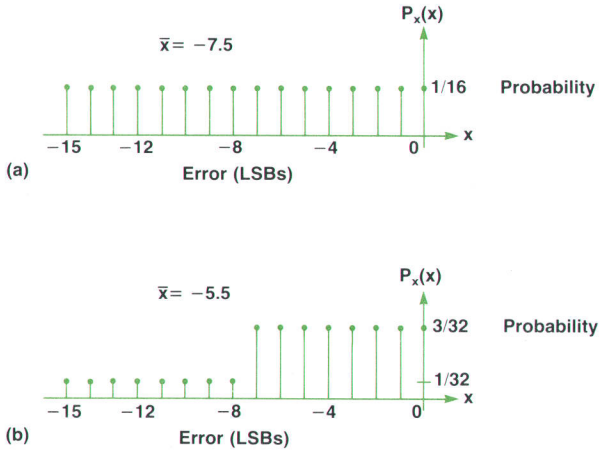


Fig. 8. *Example of probability mass function of truncation error for (a) uniformly distributed values and (b) nonuniformly distributed values of LSBs when truncating a 20-bit word to 16 bits.*

modulating the signal, which is not normally desirable in a digital filter.

Mask Design Process

The digital decimation filters used in the HP 3561A are implemented as a set of three custom NMOS integrated circuits, fabricated at HP's Loveland Technology Center. Two identical filter chips, containing 30,000 transistors each, perform simultaneous filtering of the real and imaginary data. A control chip containing about 10,000 transistors provides correct sequencing of control signals for the filter ICs and provides an interface between them and the main instrument processor and the memory used by the FFT processor.

Designing integrated circuits of this complexity requires a methodology that is logical and well thought out, and that leaves no room for errors. Once the filter algorithms and topology are determined, and the digital design completed, the process of translating schematics into silicon begins. This process starts with preliminary layouts of critical portions of the circuitry, both to help prove feasibility of the design, and to gain insight into the area requirements of the circuit. A large portion of the circuitry in the digital filter consists of shift register memory. A simple shift register cell layout was optimized for size, and, along with a few other cells, this critical cell was used to develop an architectural plan showing the sizes and locations of the major blocks of circuitry and the power and signal routing between these blocks.

The artwork for the chips is built in a hierarchical fashion. The chip is divided into large functional blocks, (multiplier, filter arithmetic, memory, et cetera). These blocks are rectangular and their horizontal and vertical dimensions match. The large blocks are themselves composed of smaller, regularly shaped blocks of medium complexity (control circuitry, parallel arithmetic, input latches, et cetera). This hierarchy continues down to the most basic blocks of circuitry, called leaf cells (e.g., shift registers and flip-flops).

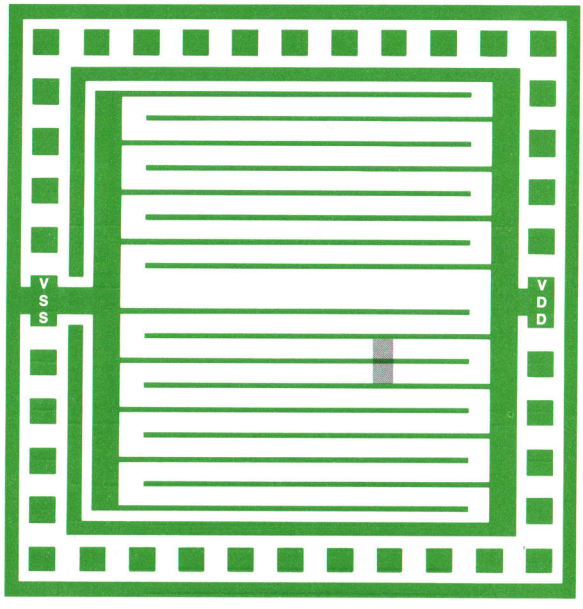
A keystone of this design method is the use of a standard metal pattern. The metal layer on the IC is used to distribute power and route signals over long distances. Power distribution on the chip is done using a metal pattern than looks like a pair of interdigitated combs (Fig. 9a). One set of fingers is connected to the 5V supply, the other to ground. Thus, the ground and 5V lines are spaced at regular intervals, alternating across the chip. Another element of the standard metal pattern is that room for a few metal signal lines is left between the power lines.

All leaf cells are built using the standard metal pattern (Fig. 9b). Every cell must be the same height, and includes half of a 5V line at the top and bottom of the cell, and a ground line across the middle of the cell. Between the power lines is room for up to four metal signal bus lines. Circuitry is laid out underneath these metal lines, using either the metal bus lines or portions of the underlying polysilicon layer for interconnect. A half-height cell can also be used for smaller circuit blocks. Thus, chip area that might otherwise be used just for busing signals from one block to another can also contain circuitry. Care must be exercised in layout, of course, to ensure that noise-sensitive nodes are not placed under clock or data lines.

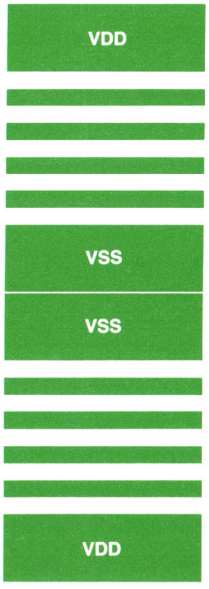
Production Considerations

It is not sufficient that an IC design be functional in an instrument under nominal conditions. Design goals for the chip must be developed early that will permit reliable operation given variances in ambient temperature, supply voltages, fabrication processing, and the performance of external circuitry.

Testing. Given these criteria, production testing of these integrated circuits is a two-part process. Each IC must be thoroughly checked for correct functional operation of all of its many thousands of transistors, then this operation must be confirmed under test conditions that attempt to simulate worst-case conditions of operation in the instrument.



(a)



(b)

Fig. 9. *(a) Simplified layout of top metal pattern for digital filter chip. (b) Standard metal pattern used in designing cells for the digital filter chip (see shaded area in (a)).*

Complete functional testing of large, complex chips presents a special challenge, especially with a circuit like the digital filter, whose data inputs are all serial. Determining a sequence of ones and zeros on these inputs that is guaranteed to excite all of the many thousands of nodes on the chip is a virtually impossible task, given the length of time a complete simulation of the entire chip would take.

Because most of the internal data path of the digital filter is also serial, direct accessibility of internal nodes is not the problem it might otherwise be. If any node in the serial data path is stuck at a one or zero, this condition will propagate an error to the output. Other, nonserial portions of the circuit, such as ROMs, parallel adders, and control circuits, are tested by operating the chip in each of its modes, and exciting the filter with signals of large dynamic range for as long a test as is feasible. In addition, a special test mode allows use of any of the 10 channels of data memory without passing data through the previous channels first, thus speeding up the memory test and helping to isolate any problems to a specific channel.

Test Program Generation. Rather than manually generate a set of stimulus and response vectors to test the chips and then code these vectors into test system programming language one clock cycle at a time, a software system was developed that automatically generates test system language code when given a stimulus sequence. This system generates response vectors using the same software simulator that was used to model the chips during the design process. Other software then converts these vectors into serial data streams used by the chips and generates digital test code for the wafer and packaged parts test system used in production.

Packaging, Cooling

Dense, high-speed chips like the digital filter and control chips require special consideration for packaging and cooling. A 64-pin, grid-array package was chosen for both parts. This choice not only provides the needed number of pins for the control chip, but offers a low thermal resistance between the substrate and the top surface of the package.

A cast aluminum heat sink that fits over the chips on the printed circuit board in the HP 3561A Analyzer promotes cool, reliable operation even under conditions of elevated ambient temperature. The heat sink mounts to the board using a spring loading scheme that keeps the heat sink's mating surface flush and in firm contact with the surface of the package without unduly stressing the brittle ceramic package.

Acknowledgments

Andy Purcell did the circuit design and much of the layout design of the control chip. Wayne Gravelle brought his many years of experience in layout to bear in executing the filter chip. His concepts of hierarchical design made an overwhelming task manageable. Dick Roberts of the University of Colorado gave valuable assistance in the design of the filter architecture. Lynn Schmidt provided much guidance during the early stages of the project. Charlie Potter did the groundwork for the control chip and managed the development of both ICs, keeping the road ahead clear.

Special thanks go to all the folks at the Loveland Technology Center. In particular, thanks to Larry Chesler and Doug Bartlett for their support and assistance during the design phase, to Wayne Gisel and Hal Cook for helping to bring these two designs into production, and last, but certainly not least, kudos to Sandy Shertzer and her crew in the computer-aided artwork department for their many long hours and patience with the last-minute changes of engineers.

The HP Design Aids group, along with the customer service group at the Loveland Technology Center, developed a set of software tools that helped to close the design loop on these complicated ICs.

References

1. L.A. Schmidt, "Designing Programmable Digital Filters for LSI Implementation," *Hewlett-Packard Journal*, Vol. 29, no. 13, September 1978.
2. A. Peled and B. Liu, "A New Hardware Realization of Digital Filters," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-22, December 1974, pp 456-462.
3. S.R. Parker, *Limit Cycles and Correlated Noise in Digital Filters*, Point Lobos Press, 1979.

Hewlett-Packard Company, 3000 Hanover Street, Palo Alto, California 94304

Bulk Rate
U.S. Postage
Paid
Hewlett-Packard
Company

HEWLETT-PACKARD JOURNAL

DECEMBER 1984 Volume 35 • Number 12

Technical Information from the Laboratories of Hewlett-Packard Company

Hewlett-Packard Company, 3000 Hanover Street
Palo Alto, California 94304 U.S.A.
Hewlett-Packard Central Mailing Department
Van Heuven Goedhartlaan 121
1181 KK Amstelveen, The Netherlands
Yokogawa-Hewlett-Packard Ltd., Suginami-Ku Tokyo 168 Japan
Hewlett-Packard (Canada) Ltd.
6877 Goreway Drive, Mississauga, Ontario L4V 1M8 Canada

CHANGE OF ADDRESS: To subscribe, change your address, or delete your name from our mailing list, send your request to Hewlett-Packard Journal, 3000 Hanover Street, Palo Alto, CA 94304 U.S.A. Include your old address label, if any. Allow 60 days.

5953-8530