HEWLETT-PACKARD
JOURNAL

December 1993

HEWLETT
PACKARD

HEWLETT-PACKARD
JOURNAL

Articles

Departments

In this Issue

In 1966, shortly after I joined the staff of the HP Journal, we published an article on the HP 8405A RF vector voltmeter. It was one of the first instruments that could measure the phase of a signal as well as its amplitude, and engineers were excited about it. The display consisted of two analog meters. You made measurements one point at a time and plotted them manually on graph paper. Advanced for its time, the vector voltmeter would be useless for analyzing the time-varying signals and complex modulation types that are common today in communications, video, data storage, radar, sonar, and medical and industrial ultrasound. However, the same technologies that make these signals possible have made the vector voltmeter's descendants equal to the challenge, and no doubt as exciting to today's engineers as the vector voltmeter was in its day. The new HP vector signal analyzers whose theory and applications are described in the article on page 6 not only make the traditional measurements of frequency, power, distortion, and noise on time-varying and complex signals, but also offer new analysis types based on digital signal processing, such as amplitude, frequency, and phase demodulation, digital modulation analysis, correlation, coherence, and vector spectrum analysis. (See page 12 for applications of the demodulation capabilities.) The HP 89410A baseband vector signal analyzer is the foundation of the family, acting either as a standalone 10-MHz analyzer or as the user interface, final frequency converter, signal processor, and display for higher-frequency members of the family. Its design, which features a large-scale dithered analog-to-digital converter, is the subject of the article on page 31. The HP 89410A and the HP 89440A RF section make up the 1.8-GHz HP 89440A RF vector signal analyzer. The design and calibration of the RF section are discussed in the article on page 47. The HP 89411A down-converter (see page 58), with the HP 89410A and a higher-frequency HP spectrum analyzer, extends the family's vector signal analysis capabilities to frequencies above 1.8 GHz. The baseband and 1.8-GHz analyzers both have built-in sources for stimulus-response measurements. Source types provided are CW, random noise, periodic chirp, and arbitrary. Behind all of the measurement capabilities of these analyzers is a powerful three-processor architecture and a firmware system that's described in the article on page 17.

Optical spectrum analysis is the measurement of the optical power in a light beam as a function of wavelength. It's especially important now in the telecommunications industry, where high-performance fiber-optic systems are prevalent. Spectral measurements are essential for characterizing the components of these systems, such as laser sources, fibers, optical amplifiers, and receivers, and verifying their performance in the system. The HP 71450A optical spectrum analyzer makes optical spectrum measurements over a wavelength range of 600 to 1700 nanometers, covering all of the widely used fiber-optic bands. The HP 71451A optical spectrum analyzer makes spectrum measurements and offers four additional measurement modes for other types of measurements. The article on page 60 introduces these analyzers, describes their user interface, and demonstrates the capabilities of several of the downloadable application programs that are available, including programs for light-emitting diodes, Fabry-Perot lasers, and distributed feedback lasers. The analyzers acquire both high dynamic range and high sensitivity from a double-pass monochromator design (page 68). The monochromator, which is the wavelength-selective element of the analyzer, is based on a rotating diffraction grating. The grating is driven by a direct-drive positioning system that provides both high resolution and high speed (see page 75). At the output of the monochromator the light is coupled into a fiber. This design provides significant advantages, but it's not trivial to keep the light beam accurately aligned with the output fiber as the diffraction grating rotates. The article on page 80 describes a two-axis micropositioner that addresses this problem.

The HP vector signal analyzers featured in this issue are only one of many types of digital signal analyzers manufactured by the Hewlett-Packard Lake Stevens Instrument Division. This division's newer analyzers, including the vector signal analyzers, store data in a standard data format that allows Lake Stevens analyzers to exchange data with each other and with applications software (see page 85). Utility programs shipped with all Lake Stevens analyzers make it possible to convert data between the standard format and other formats, and to edit, display, and plot data stored in the standard format.

New digital cellular telephone technologies, developed to increase the number of users that can share a given frequency band, offer prime examples of the time-varying signals and complex modulations that HP vector signal analyzers are designed to measure. In the article on page 90, Dave Whipple of the Hewlett-Packard Spokane Division describes one of these technologies, a code division multiple access (CDMA) system standardized by the Telecommunications Industry Association for North American cellular applications. CDMA is a type of modulation in which all channels in a frequency band use the entire band and are separated by means of specialized codes.

Unlike North American cellular CDMA, the Digital European Cordless Telecommunications standard, or DECT, is not a standard implementation of a type of modulation, but a standard protocol for the radio portion of cordless communication links (DECT modulation is actually time division multiple access, or TDMA). DECT is defined by the European Telecommunications Standards Institute. The DECT standard defines channel frequencies and data packet formats and spells out the tests that cordless equipment must pass to be certified as conforming to the standard. On page 98, Mark Elo of the Hewlett-Packard Queensferry Microwave Division describes the DECT standard and a new downloadable program for HP 8590 E-Series spectrum analyzers that gives the analyzers measurement capabilities for testing to the DECT standard.

December is our annual index issue. The 1993 index begins on page 107.

R.P. Dolan
Editor

Cover

By adding a third axis (color) to the traditional spectrum analyzer display, the HP 89410A and 89440A vector signal analyzers can reveal the frequency content of a rapidly-changing signal in a particularly informative way. This spectrogram display represents more than 300 spectrum measurements covering the first 20 milliseconds of the turn-on transient of a marine-band handheld transmitter. The horizontal axis is frequency, but each spectrum measurement has been compressed to fit in one line of the display, with power levels shown as different colors. This allows a single screen to show vastly more information and reveals phenomena that would be difficult to spot otherwise. For example, transient distortion sidebands can be seen clearly in this measurement although they are present for only a few milliseconds just after turn-on. The sidebands gradually disappear toward the bottom of the screen. It can also be seen that these sidebands temporarily disappear each time the signal changes direction.

What's Ahead

The February issue will have twelve articles covering the design of the HP DeskJet 1200C color office printer.

Vector Signal Analyzers for Difficult Measurements on Time-Varying and Complex Modulated Signals

Called vector analyzers for their ability to quadrature detect an input signal and measure its magnitude and phase, these new analyzers offer conventional spectrum analysis capabilities along with a full set of measurements based on digital signal processing. The three-processor architecture includes a frequency selective front end and a digital IF section.

by Kenneth J. Blue, Robert T. Cutler, Dennis P. O'Brien, Douglas R. Wagner, and Benjamin R. Zarlingo

Swept spectrum analyzers are a fundamental tool for designers working in all aspects of electronics at frequencies from HF (high frequency) through microwave. They are powerful and accurate tools for measuring basic signal properties such as power, frequency, distortion, and noise. They have also been pressed into service to measure a group of more complex and dynamic signal properties that are often grouped together as modulation and sidebands (intentional amplitude, frequency, and phase modulation) or phase noise (generally unintentional or undesirable).

Recent trends have conspired to move many of today's signals beyond the measurement reach of traditional spectrum analyzers. These difficult signals generally fall into two groups:
- Time-varying. Burst, pulsed, gated, or time-division multiplexed signals whose measured properties change during a measurement sweep.
- Complex modulated. Signals with modulation that cannot be described in terms of simple AM, FM, and PM. Examples include the multiple varieties of QAM, QPSK, and PSK. To complicate the measurement task further, these complex modulated signals are often time-varying as well.

For design engineers in many different applications areas, dealing with these complex and challenging signals is now the rule rather than the exception. Examples include mainstream applications such as video, data storage, radar, sonar, and medical and industrial ultrasound.

By far the largest application area for these dynamic signals is communications. A staggering proliferation of wireless technologies is underway, bringing both new uses and vast numbers of new users to the limited RF spectrum that we all must share. The only way to accommodate these new demands within the existing frequency spectrum is to use it more efficiently, and this is the primary force driving the increased use of complex and time-varying signals.

A telling example is the transition from the current analog cellular telephone technology to the new digital methods in Europe (GSM), Japan (PDC), and North America (NADC). The capacity of the current analog system has been exceeded

in many areas, and the new technologies will support at least two to three times the number of users in a given frequency band. While they differ in some respects, all three of these new standards involve signals that are at once digital, time-varying, and complex modulated. They pose a distinct challenge to traditional signal analyzers, which are optimized for steady-state and simply modulated signals.

Several types of new tools have already been developed to address different portions of these application requirements:
- Digital Oscilloscopes. Oscilloscopes are excellent tools for capturing and viewing almost any complex or time-varying signal. However, they are optimized for viewing signals in the time domain and have insufficient digitizing resolution and accuracy for precision frequency, power, distortion, and noise measurements.
- Peak Power Analyzers. Fast-reacting power meters with data storage and displays can track the power component of rapidly changing signals. They solve one part of the signal measurement problem when frequency selective measurements are not required.
- Modulation Domain Analyzers. Also called frequency and time interval analyzers, they measure the frequency behavior of dynamic signals using fast zero-dead-time counter technology. Their analysis is limited to frequency and phase. They cannot measure amplitude or distortion, and cannot separate multiple signals.
- Spectrum Analyzers with Sweep Gating. Where time-varying signals repeat consistently and where a trigger signal is available, some spectrum analyzers can perform time-gated analysis. The analyzer sweeps selectively, synchronized with the trigger signal, and gradually builds up a measurement from many sweep segments.

Vector Signal Analyzers

The analog-to-digital conversion and digital signal processing technologies that have made these "problem" signals possible have also made possible a new generation of measurement solutions. HP's new vector signal analyzers (see Fig. 1) represent a two-pronged approach to dealing with today's time-varying and complex modulated signals and the systems

Fig. 1. The HP 89410A vector signal analyzer (center) is a 10-MHz baseband analyzer. It acts as the user interface, final digital IF, signal processor, and display section for the HP 894xxA family of analyzers. The HP 89440A RF section (bottom left) extends the analyzer's capabilities up to 1.8 GHz. The HP 89411A (right) is used with a higher-frequency analyzer to extend measurement coverage to frequency ranges above 1.8 GHz.

that use them. Traditional measurements—precision measurements of frequency, power, distortion, and noise—can be made as simply on time-varying signals as they are on steady-state signals. This allows designers to use the insight and design skills they have developed for simpler signals on more complex signals. Going beyond the traditional measurements, digital signal processing of precision sampled signals makes possible a variety of new analysis types including vector AM, FM, and PM demodulation, digital modulation analysis, correlation, coherence, and vector spectrum analysis. These new measurements are ideally suited for testing throughout the block diagrams of today's advanced communication and measurement systems.

The HP 894xxA vector signal analyzers are all based on a common measurement engine, the HP 89410A 10-MHz baseband analyzer. The HP 89410A acts as the user interface, final digital IF (intermediate frequency) section, signal processor, and display section for the entire family of analyzers. The HP 89440A and HP 89411A use analog RF (radio frequency) hardware to down-convert higher-frequency bands into the information bandwidth of the HP 89410A. The HP 89440A RF section extends the vector signal analysis capabilities up to 1.8 GHz. The HP 89411A is used in conjunction with a higher-frequency analyzer to extend the measurement coverage even higher in frequency. This design approach has made it possible to develop a core set of measurement features in the HP 89410A and make them immediately available for use in different frequency ranges. This provides measurement capabilities at RF employing digital signal processing (DSP) techniques that were previously only possible at baseband frequency ranges.

The HP 89410A and HP 89440A include a standard dc-to-10-MHz signal source. The HP 89440A also offers an optional

RF source. Both the baseband and the RF sources feature multiple source types for flexibility in circuit stimulation. Besides CW, random noise, and periodic chirp sources, the analyzers include an arbitrary source. Users can create their own signals or capture signals for later playback.

Signal Flow

The high-level measurement processing in the vector signal analyzers is shown in the signal flow diagram, Fig. 2. The HP 89410A has an alias-protected analog input bandwidth of dc to 10 MHz. The analog input signal flows into a hybrid dithered ADC (analog-to-digital converter), which samples the signal at a rate of 25.6 MHz. The digital data stream coming out of the ADC is routed to the digital local oscillator (LO). The digital LO performs frequency translation and quadrature detects the digitized input signal, resulting in a complex data stream consisting of real and imaginary parts. This full-rate data is then input to the digital decimating filters. These filters perform binary decimation (divide-by-2 sample rate reduction) in addition to providing image rejection. The output of these digital filters represents a bandlimited digital version of the analog input signal in the time domain. This digital data stream is then captured in a sample RAM (random-access memory). The sample RAM is a circular FIFO (first in, first out) buffer that collects individual data samples into blocks to be manipulated by the DSP (digital signal processor).

The ADC, digital LO, and filter sections are key contributions in the HP 894xxA analyzers. These blocks allow the use of powerful DSP techniques on systems and signals that have wider information bandwidths than previous fast Fourier transform (FFT) analyzers have allowed, as explained later.
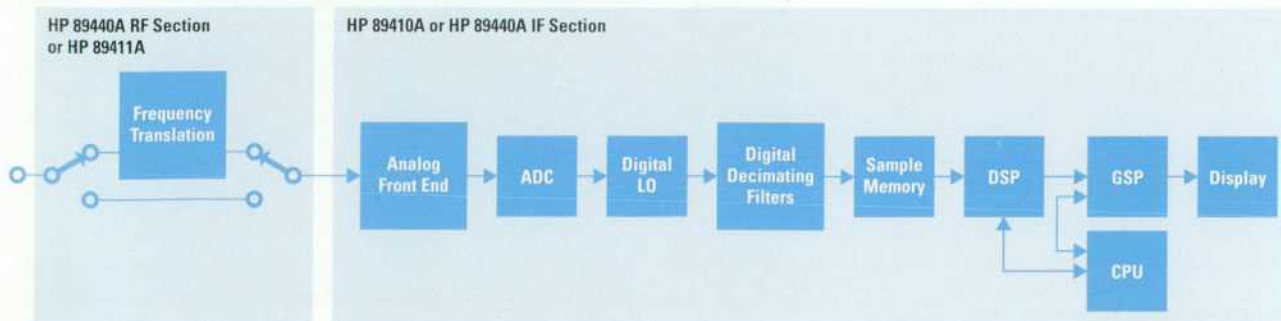
Fig. 2. Signal flow architecture of the HP 894xxA vector signal analyzers. (ADC = analog-to-digital converter. LO = local oscillator. DSP = digital signal processor. GSP = graphics system processor. CPU = central processing unit.)

The fundamental data used to produce all measurement results, whether in the frequency domain or the time domain, is the time data collected in the sample RAM. This can be contrasted with swept spectrum analyzers, which produce results directly in the frequency domain using swept filter and detection techniques. Since the input signal is processed and captured in this fundamental time-domain form, all aspects of the signal at that moment in time are captured—amplitude, frequency, and phase. The application of DSP algorithms to this time data can characterize and allow the user insight into all of these different views of the signal.

Multiple Processors

The HP 894xxA uses a set of three processors, each optimized for specific computational and control tasks, to perform the measurement postprocessing on the sample RAM time data. These processors are under the control of a real-time multitasking operating system (pSOS*) to provide overlapped operation and maximize throughput. The main CPU, which handles system management and processes user input, is a Motorola 68EC030.

The digital signal processor used in the HP 894xxA family is the Motorola DSP96002 floating-point dual-port processor. This processor performs all the block-oriented mathematical operations required by the DSP algorithms. A typical sequence of DSP operations to produce a spectrum result from the time data might include the following: input scaling, resampling, time-domain corrections, windowing, FFT, frequency-domain corrections, and display scaling. The HP 894xxA DSP architecture provides many more features than just the computation of FFTs. It has a full set of predefined measurement results including time, power spectrum, power spectral density, autocorrelation, crosscorrelation, cross spectrum, coherence, frequency response, and others. The HP 894xxA also supports a user-defined math capability, which allows the user to specify custom algorithms to be executed by the DSP.

The GSP (graphics system processor) used in the HP 894xxA family is the Texas Instruments TMS34020 processor. The GSP and color display provide display capabilities that exploit the measurement features of the HP 894xxA. These include multiple trace displays for simultaneous insight into different result domains, such as time and frequency. Other features include the use of waterfall and spectrogram displays to project measurement results into a third dimension on the display—history over time. This is useful for data presentation since the measurement and display throughput of the HP 894xxA can often reach or exceed 60 updates per second.

A flexible firmware architecture was required to realize the large feature set of the HP 894xxA. A core set of fundamental low-level DSP routines was developed to be used by all the various measurement modes and features within the analyzers. At a higher level, an extensible measurement architecture was designed to use these low-level DSP routines to produce the different measurement results. This architecture is described in detail in the article on page 17.

FFT Use in Previous Analyzers

Most spectrum analyzers do not use FFT processing to produce frequency-domain measurements. The majority of spectrum analyzers employ swept intermediate frequency (IF) hardware filter techniques to measure signal amplitude characteristics in the frequency domain. These analyzers have analog resolution bandwidth filters with electrical characteristics that limit their sweep rate to approximately one-half the square of the resolution bandwidth. Since these filters require a finite settling time before they can accurately represent the amplitude of a signal passing through the IF, the signal of interest must be static (nonvariant) over time. This produces two constraints for the user: the measurement speed for narrow resolution bandwidths is slow and the signals must not vary over time. These limitations can be overcome with the application of digital signal processing.

The HP 894xxA vector signal analyzers represent the second generation of HP signal analyzers that employ digital signal processing at RF. The HP 3588A[1] spectrum analyzer and HP 3589A network analyzer were pioneers in using digital filtering and FFT processing techniques in conjunction with traditional swept analyzer hardware. These predecessors used an all-digital final IF section to support higher-speed swept measurements while also supporting an FFT mode for faster narrowband measurements. These instruments laid the groundwork for the measurement architecture and approach of the HP 894xxA family of vector signal analyzers.

The application of the FFT in signal analyzers has been restricted in the past because of one of its fundamental algorithmic relationships—the frequency bandwidth of an FFT result is directly related to the sample rate of the input data. This gives rise to two design challenges in incorporating the FFT into higher-frequency analyzers. Because of the sampling rate constraints of high-dynamic-range ADCs, the frequency coverage (information bandwidth) of previous FFT analyzers has been limited, usually to less than 100 kHz. This has restricted their application to measurements of phenomena with a low information bandwidth (e.g., rotating machinery, servo loops, acoustics, and mechanical vibrations). Another prevalent limitation in these previous analyzers is also a result of the FFT fundamentals and hardware limitations. If the selections of sampling rates for time-domain data are limited going into the FFT algorithm, the user's ability to select an arbitrary frequency span for analysis is also limited to predefined and fixed analysis spans (e.g., 100 kHz, 50 kHz, 20 kHz, 10 kHz). A user who requires a span between one of these fixed values is forced to select either a larger bandwidth, thereby sacrificing signal isolation, or a smaller bandwidth, thereby sacrificing insight into adjacent spectral activity.

FFT Advantages

So why bother designing an FFT-based analyzer? The answers are speed and information. In a swept analyzer, the filter must be swept to a frequency and settled before a result is obtained at that single frequency. The filter is then swept to the next frequency, and so on. The FFT algorithm emulates a parallel bank of filters that can settle and measure simultaneously. For comparable-resolution filters, the FFT measurement can be much faster than the swept filter technique. The second part of the answer is information. Since the time data is the fundamental data type in the HP 894xxA signal flow, all characteristics of the signal are preserved for subsequent analysis. The FFT algorithm is a particularly useful

way to characterize amplitude and phase at all frequencies simultaneously. This makes the FFT technique useful for measuring time-variant or transient signals.

Another advantage of the FFT is that it provides true-rms detection. This is useful in making band power measurements. Traditional analyzers usually employ a peak detection mechanism to ensure that signals are "measured" whenever the resolution bin width is less than a display bin width. However, the result of this operation is that the noise floor is biased. The FFT algorithm has no such effect and thus allows accurate simultaneous measurements on signals and noise.

FFT Design Challenges

The fixed frequency span constraint was a significant issue in the design and development of the HP 894xxA family. Traditional swept spectrum analyzer users are not accustomed to being limited to predefined analysis spans. With a swept IF analyzer, the user has almost infinite control over setting the measurement span. Another significant issue was the fixed relationship of resolution bandwidth to the span and window filter function present in FFT analyzers. Swept spectrum analyzers have the freedom to select a span arbitrarily, and then to change the resolution bandwidth, changing the frequency resolution.

These challenges are overcome in the HP 894xxA while still using the FFT algorithm as the basis for all frequency-domain results. The fixed frequency span limitation is overcome by using the DSP to alter the sample rate of the time data input to the FFT algorithm. By changing the sample rate in software, the user is given back the ability to control the span setting arbitrarily without sacrificing the advantages of FFT processing (see "Frequency Selective Analysis" below). The fixed resolution bandwidth/span/window limitation is overcome by another DSP technique. Using frequency-domain interpolation (zero padding) in the time domain before the FFT, the effective filter bandwidth of a window can be changed without changing the span of the measurement. This restores the ability of the user to control resolution bandwidth independently of span (within limits).

These characteristics combine to provide a traditional spectrum analyzer look and feel in what is fundamentally an FFT analyzer. The HP 894xxA's scalar measurement mode is a prime example of this transparency. This measurement mode provides the greatest independence of span and resolution bandwidth of all the measurement modes. When set for a span and a resolution bandwidth that allow coverage by a single FFT, this measurement mode will set all of the frequency translating LOs (both analog and digital) to fixed values and process the incoming time data with a single FFT to provide the entire measurement result at once. If the user widens the span or reduces the resolution bandwidth so that the measurement cannot be realized by a single FFT, the scalar measurement mode will revert to a stepped mode of operation. In this mode, the frequency translating LOs are stepped to the beginning band within the sweep. Once settled, the instrument collects a time record and transforms it to produce the spectral result for that segment of the sweep. This segment result can be run through a software peak detector running in the DSP if the resolution bandwidth or bin width of the FFT is narrower than the display

bin width. This detected segment of the sweep is then processed for a partial update of the display. The measurement then steps the LOs to the next frequency span segment and repeats the process until the entire measurement span is covered (end of sweep). Since the display is updated as the LOs and FFTs are tracked across the span, the appearance is the same as that of a traditional swept spectrum analyzer.

Frequency Selective Analysis

Frequency selective analysis is a term used to describe the vector signal analyzer's ability to apply a bandlimiting filter to the measured signal to limit the signal's information bandwidth. In one sense this capability is not unique. Many oscilloscopes have a bandwidth limiting filter. Also, most traditional spectrum analyzers have a set of resolution bandwidth filters from which to choose. These filters limit the information bandwidth of the signal applied to the detector. What makes the vector signal analyzer unique is the combination of infinite adjustability of the filter bandwidth with nearly ideal filter characteristics. What are ideal characteristics? In this application an ideal filter would have a frequency response corresponding to a RECT function. That is, it would have linear phase, zero group delay, no amplitude ripple, infinite stop band attenuation, and a 1:1 shape factor. Obviously the filter described cannot be realized. However, using digital filters, resampling techniques, and time-domain corrections, a filtered signal path can be created that very nearly meets these goals. The composite filter and time-domain corrections typically result in a filter with a 1.3:1 shape factor, passband ripple less than 0.1 dB, zero group delay, phase linearity of ±1 degree, and stop-band attenuation of 111 dB. This filter is optimized for isolating signals, and not for time-domain characteristics such as overshoot and ringing.

Why does one need infinitely adjustable filters? There are two answers to this question. The first answer is based on the types of measurements that can be made with a vector signal analyzer, such as modulation analysis. The vector signal analyzer has built-in demodulation capability, and like any receiver, the fidelity of the demodulated signal will be degraded if signals other than the one to be demodulated are present. In most receiver systems the IF bandwidth is made as narrow as possible to provide an optimum amount of selectivity. Since the vector signal analyzer is a general-purpose tool and not a receiver optimized for a particular type of signal, the optimum bandwidth for the filter cannot be determined beforehand. Many traditional spectrum analyzers also have demodulators. In these instruments the resolution bandwidth filters serve to limit the information bandwidth. However, with only a finite number of resolution bandwidth filters to choose from, it's possible that the user is left with a choice between a filter that is either too narrow to pass the signal or too wide to reject another signal in close proximity. With the vector signal analyzer's infinitely adjustable bandwidth, an optimal information bandwidth can be set for any class of signal or measurement. In the vector signal analyzer the terms span and information bandwidth mean roughly the same thing. The only distinction between the two is that the information bandwidth corresponds to the 3-dB bandwidth, which is 12 to 17 percent wider than the span.

The second reason for wanting variable information bandwidth has to do with selecting a span, or equivalently, a

sample rate. With swept analyzers there aren't any limits on the selection of span. The user is free to choose any start-stop frequency pair. This has not been true for instruments using FFTs to compute the spectrum. In these instruments only a fixed number of spans can be selected. This limitation is based on the instrument's ability to change the rate at which the data is sampled. In most FFT analyzers, such as the HP 35665A dynamic signal analyzer, the sample rate is changed after the signal has been digitized by the ADC. This is done through the use of digital decimating filters. These filters, which are implemented in hardware, halve the sample rate by first bandlimiting the data and then discarding every other sample of the resulting oversampled data. For example, the ADC in the HP 89410A operates at a 25.6-MHz rate to produce a 10-MHz maximum span. To reduce the span to 5 MHz, the signal is passed through a digital filter which reduces the bandwidth of the signal from 10 MHz to 5 MHz. Then, every other sample is discarded to produce a 12.8-MHz sample rate. By cascading several decimating filters, the sample rate and span can be changed by $1/(2^N)$ where N is the number of decimating filters used. For more detail on how these filters are implemented in the vector signal analyzer, refer to the article on page 31.

The decimating filters allow the sample rate and span to be changed by powers of two. To obtain an arbitrary span, the sample rate must be made infinitely adjustable. This is done by means of a resampling or interpolation filter, which follows the decimation filters. A brief description of the resampling algorithm and how it can be used to change the sample rate is given in "The Resampling Process" at right.

Time-Domain Corrections

When it comes to calibrations and corrections, previous generations of FFT analyzers have mostly ignored the time data. This has occurred because it is much easier to correct the frequency spectrum data using multiplication than it is to correct the time data through convolution. In the vector signal analyzer the accuracy of the time data is very important. Not only is it the basis for all of the demodulation measurements, but it is also used directly for measurements such as instantaneous power as a function of time. Correcting the time data is the last step in creating a nearly ideal bandlimiting signal path.

While the digital filters and resampling algorithms are responsible for establishing the arbitrary bandwidth (sample rate and span), the time-domain corrections determine the final passband characteristics of the signal path. Time-domain corrections would be unnecessary if the analog and digital signal paths could be made ideal. Unfortunately, achieving nearly ideal characteristics in the analog and digital filters is either impossible or impractical. For example, the hardware decimation filters are implemented as infinite impulse response (IIR) filters rather than as linear-phase finite impulse response (FIR) filters. The IIR filters were chosen over the FIR filters because they are more economical to implement given the requirements for speed, shape factor, stop-band rejection, and the number of stages of filtering required.

Time-domain corrections work as an equalization filter to compensate for passband imperfections. These imperfections come from many sources. The IF filters in the HP 89440A RF section, the analog anti-aliasing filter in front of

The Resampling Process

In the HP 894xxA vector signal analyzers, the ADC and digital filters produce a digital sequence $x_1[n]$, which is obtained by sampling the filtered input signal $x(t)$ at an effective sampling interval of T_1. The resampling algorithm produces a different sequence $x_2[m]$ that is identical to the sequence that would have been obtained had $x(t)$ been sampled at a periodic interval $T_2 \neq T_1$. In other words, resampling changes the sample rate from $1/T_1$ to $1/T_2$ after sampling has already occurred.

The basic concept behind resampling comes from standard sampling theory, which states that a signal can be reconstructed from its samples provided that the samples are spaced so as to prevent aliasing. Using this concept, the most direct approach to changing the sample rate would be to reconstruct the original signal $x(t)$ from a sampled version of the signal and then sample the reconstructed signal at the new sample rate.

However, it's not necessary to convert the signal back into its continuous form to create $x_2[m]$. In the following derivation, the sequence $x_1[n]$ is treated as a continuous signal and is represented as a series of weighted Dirac delta functions. The sequence $x_1[n]$ is described by

$$x_1(\tau) = \sum_{n=-\infty}^{\infty} x(\tau)\delta(\tau - nT_1). \qquad (1)$$

The original signal is reconstructed by passing $x_1(\tau)$ through a reconstruction filter. The filter is described by its impulse response $h(t)$. Assuming that $h(t)$ is an appropriate filter, then $x(t)$ is obtained by performing the convolution $x_1(t) \star h(t)$. To distinguish this result from the original $x(t)$, the reconstructed signal will be referred to as $\hat{x}(t)$. Performing the convolution,

$$\hat{x}(t) = \int_{-\infty}^{\infty} x_1(\tau)h(t-\tau)d\tau. \qquad (2)$$

Using equations 1 and 2,

$$\hat{x}(t) = \int_{-\infty}^{\infty} \left[\sum_{n=-\infty}^{\infty} x(\tau)\delta(\tau-nT_1) \right] h(t-\tau)d\tau$$

$$= \sum_{n=-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x(\tau)h(t-\tau)\delta(\tau-nT_1)d\tau \right].$$

With the order of summation and integration reversed, the sifting property of the Dirac delta function can be used to evaluate the integral.

$$\hat{x}(t) = \sum_{n=-\infty}^{\infty} x(nT_1)h(t-nT_1). \qquad (3)$$

Given that $h(\tau)$ is known for any value of τ, equation 3 can be used to calculate $\hat{x}(t)$ for any value of t from the samples $x(nT_1)$. Limiting the values of t to the sample points for the sequence $x_2[m]$ produces the desired result, which can be written in sequence form as:

$$x_2[m] = \sum_{n=-\infty}^{\infty} x_1[n]h(mT_2-nT_1). \qquad (4)$$

Robert T. Cutler
Development Engineer
Lake Stevens Instrument Division

the ADC, and the decimation and resampling filters all contribute to passband ripple and phase nonlinearities within the selected span. The time-domain correction or equalization filter must compensate for these imperfections without adding any of its own.

The design of the equalization filter has many constraints:

- The filter must compensate for amplitude unflatness and phase nonlinearity within the specified information bandwidth (span).
- The filter must not compensate beyond the span or the desirable effects of the previous filters, such as stop-band attenuation, will be diminished.
- The filter must have a minimum-length impulse response.
- The filter must be designed on-the-fly for the current instrument setup (e.g., span) using the calibration data stored in nonvolatile RAM.

The design of the equalization filter begins by extracting the appropriate information about the analog signal path from the calibration data. This calibration data, which is generated by the instrument during self-calibration and stored in nonvolatile memory, contains data for all possible instrument configurations. The data extracted for the filter design will be a function of the selected center frequency, span, input range (attenuator settings), coupling, and input impedance. Once extracted, the data is used to create a frequency-domain correction vector (curve).

Once the analog correction vector has been computed, it is modified to include the effects of the decimation filters and the resampling filter. While the frequency response of each individual filter is known by design, the combined response cannot be computed until after the user has selected a span because the span determines the number of stages of decimation as well as the resampling ratio. The composite correction vector serves as the basis for the design of the equalization filter that will be applied to the time data.

Fig. 3 shows a typical plot of the composite analog/digital correction vector for a 9.9-MHz span. The upper trace shows that the amplitude of the correction varies over 6.2 dB. The lower trace shows the amount of compensation needed to correct phase nonlinearity. Over the 9.9-MHz span, this corresponds to group delay distortion of over 600 ns.

Bandlimited AM, PM, and FM Demodulation

Advanced floating-point DSP power in the HP 894xxA has enabled the development of high-speed AM, PM, and FM demodulation algorithms capable of up to 60 display updates per second. The hardware digital local oscillator and decimating digital filters allow fully alias-protected, bandlimited demodulation with spans as small as 1 Hz, and as large as 7 MHz for the RF receiver mode and 10 MHz for baseband receiver modes. Additionally, the HP 89411A extends the RF frequencies that can be demodulated to well above 1.8 GHz. The wide frequency coverage, bandlimited analysis, high accuracy through time-domain calibration filters, and typical dynamic range above 70 dB offer new insight into many demodulation applications (see "Applications for Demodulation" on page 12).

The analog demodulation signal processing block in Fig. 1 on page 17 contains the AM, PM, and FM demodulation algorithms. This block precedes time averaging, the only difference between the analog demodulation and vector measurement modes. Thus, most of the time-waveform signal processing capabilities in vector measurements, such as spectrums, time gating, and averaging, are also available for demodulated time waveforms. For example, a PM spectrum
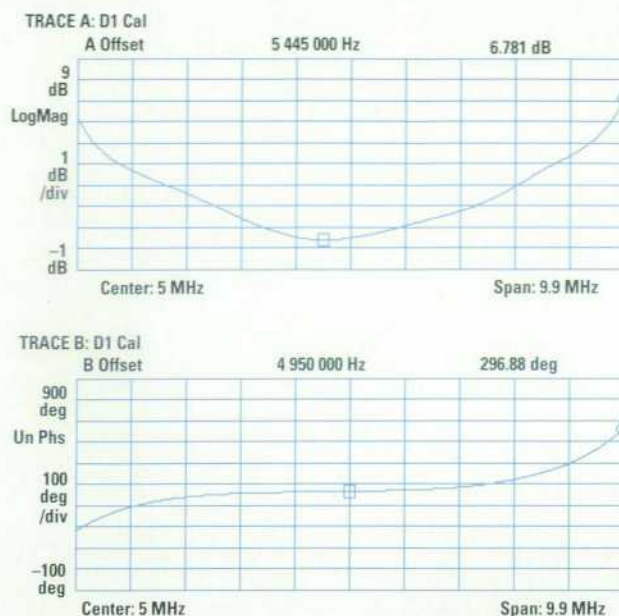
Fig. 3. The vector signal analyzers correct the basic time data for signal path imperfections. Shown here is a typical plot of the magnitude (Trace A) and phase (Trace B) of the composite analog/digital correction vector for a 9.9-MHz span.

can be averaged to generate a phase noise display. The measurement processes time data in block fashion with the resolution bandwidth determining the length of the time record.

Low-frequency information can be carried on a high-frequency sinusoidal signal, or carrier, by varying the carrier's amplitude and phase angle. The instantaneous frequency of a sinusoidal signal is given by the time derivative of its phase. Thus, the frequency modulation of a carrier can be computed by taking the time derivative of the phase modulation. These are the basic principles applied in the HP 894xxA demodulation algorithms, originally developed for the HP 3562A dynamic signal analyzer.[2]

A carrier without modulation can be expressed as:

$$Ce^{j(\omega_c t + \phi_c)} \tag{1}$$

with amplitude C, frequency ω_c, and phase offset ϕ_c. Consider a complex modulating waveform:

$$m(t) = [1 + a(t)]e^{j\phi(t)}$$

where $a(t)$ and $\phi(t)$ are real amplitude and phase modulating waveforms, respectively. Thus, a real modulated carrier can be written as:

$$x(t) = Cm(t)e^{j(\omega_c t + \phi_c)} + Cm^*(t)e^{-j(\omega_c t + \phi_c)}$$
$$= 2C[1 + a(t)] \cos[\omega_c t + \phi_c + \phi(t)].$$

If ω_c is large enough to prevent sidebands in the positive and negative frequency images of $m(t)$ from overlapping when summed in $x(t)$, the amplitude and phase modulating components in $m(t)$ can be unambiguously obtained by shifting the positive frequency image in $x(t)$ down to a frequency band near dc and filtering the complex time waveform with a digital low-pass filter to reject the other image.

Applications for Demodulation

The HP 894xxA vector signal analyzers provide digital signal processing (DSP) demodulation algorithms to extract AM, PM, and FM modulating signals. With the demodulator placed upstream in the signal processing flow, many of the vector signal analyzer's powerful analysis features can be used on demodulated wave-forms. A few of the many possible applications of vector signal analyzer demodulation are described here.

Phase Noise Measurements

The characterization of phase noise is an increasingly important requirement in modern communications systems. Traditionally this has been a very difficult and time-consuming measurement. The HP 894xxA vector signal analyzers greatly facilitate this measurement for all systems but those with the most demanding phase noise requirements. The power of these analyzers is in their ability to make very fast direct phase noise measurements in any domain relevant to the user. For instance, transmitter designers are most likely interested in measuring the noise spectral density around the carrier or the integrated band power in adjacent channels. Users interested in the recovery of digitally modulated information may be most concerned with the peak or rms phase deviation of the recovered vectors, which can be directly measured using the vector signal analyzers' PM demodulation capability. The HP 894xxA produces results quickly and easily in any of these domains. A typical measurement is shown in Fig. 1.

The vector signal analyzers are capable of mathematically locking to unlocked or drifting carriers using the AutoCarrier features (as long as the carrier is not digitally modulated), allowing fast and accurate averaged measurements even under these conditions. How fast is "fast"? For many measurement situations, users will make accurate measurements in seconds that had previously taken minutes or tens of minutes to complete. Measurement speed improvements of 10 to 1000 times can be expected. For more information refer to HP Product Note 89440A-2.

VCO Turn-on and LO settling

The vector signal analyzer's demodulation capabilities are powerful for a variety of VCO or local oscillator transient measurements. For example, the frequency trajectory of a VCO at turn-on can be evaluated using the FM demodulation feature. Similarly, an LO or phase-locked loop transient, such as that following a frequency change, can be directly measured for frequency or phase settling trajectories using the FM or PM demodulation features. For amplitude or power variations during the transient, the equivalent of a zero-span measurement is used, rather than the AM demodulation feature (see "Zero-Span Measurements" below). This is because the AM demodulator measures percent AM and the carrier power estimation is biased by the transient event itself.

When a transient is acquired in time-capture mode, it can be played back into the measurement as many times as is desired. Thus, one playback can be done with
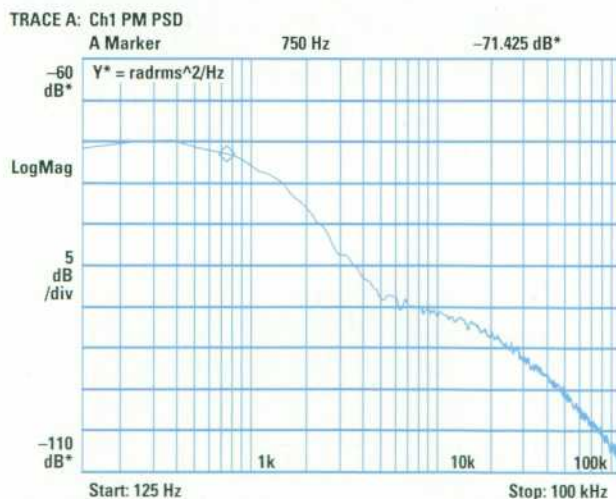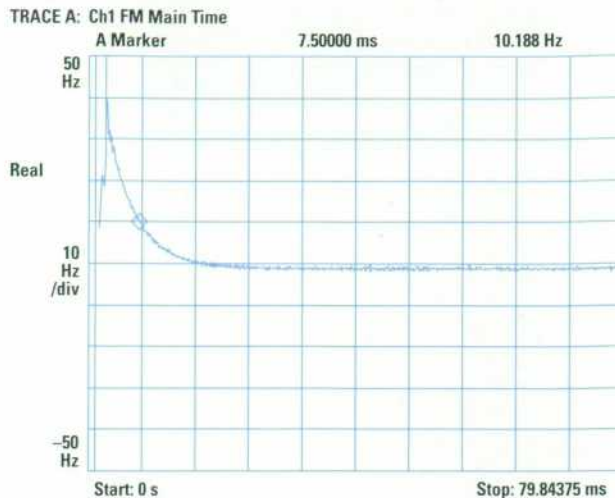

Fig. 2. An LO frequency settling measurement.

the analyzer set up to make the equivalent of a zero-span measurement, and another into the PM or FM demodulator. Displaying both results together provides a side-by-side comparison of instantaneous carrier turn-on power and phase or frequency during the transient.

Because of the vector signal analyzer's internal quadrature down-conversion, phase and frequency modulation are measured essentially continuously throughout the time record, without the cycle quantization limitations of counter-based modulation analyzers. This results in an extremely good combination of time resolution (of the transient event) and frequency or phase resolution for those demodulation types. An example of an LO frequency settling measurement is shown in Fig. 2. For more information refer to HP Product Note 89440A-5.

Zero-Span Measurements

Zero-span measurements with swept spectrum analyzers measure the AM envelope as a function of time. The vector signal analyzer also measures AM envelope, but with a different approach that offers distinct advantages over traditional swept technology.

In a swept spectrum analyzer, the center frequency is set near the carrier. The frequency span is set to zero to prevent the LO from sweeping. The sweep time is set to some nonzero value, limited by the resolution bandwidth filters and detector. In a vector signal analyzer, the RF signal is down-converted to IF and sampled. The signal is then down-converted in single-sideband fashion to dc and bandlimited. The bandlimiting is performed by digital filters, which perform the functional equivalent of resolution bandwidth filters for swept spectrum analyzers in zero-span mode. Thus, in a vector signal analyzer, the user sets the measurement span, not the resolution bandwidth. Setting the width of the main time record is equivalent to setting the sweep time. The complex time waveform is converted into an AM envelope by selecting either the linear or the logarithmic magnitude data format.

For clarification, the vector signal analyzer implements resolution bandwidth filters by way of the fast Fourier transform algorithm. The input signal bandwidth is first limited by digital filters set a little wider than the measurement span. In generating a spectrum display, the FFT can be thought of as a parallel bank of narrow resolution bandwidth filters that form a comb across the measurement span. However, no FFT is used if the time domain is displayed. Thus, the bandwidth for time displays is limited only by digital filters. The only effect of changing the resolution bandwidth on a time-domain display is to change its length.

Communication systems are requiring measurements of increasingly faster carrier ramps with zero-span measurements. The rate at which a zero-span measurement can handle a carrier ramp is limited by the rise time of the selected resolution bandwidth filter. As the resolution bandwidth is increased, the rise time decreases, allowing finer time resolution.

Fig. 1. A phase noise measurement made by an HP 894xxA vector signal analyzer.

Typically, swept spectrum analyzers have a maximum resolution bandwidth of 3 MHz. In contrast, the maximum span of the HP 894xxA vector signal analyzer is 7 MHz in the RF receiver mode and 10 MHz in the baseband receiver mode. Thus, a key advantage in the vector signal analyzer is that finer time resolution is possible.

The accuracy of swept spectrum analyzers is limited by the resolution bandwidth filters, logarithmic amplifier, and detector, all of which are implemented in analog circuitry for wide resolution bandwidths. The vector signal analyzer samples the signal directly at IF. Filtering, logarithmic conversion, and detection are all performed digitally in the vector signal analyzer, providing unmatched accuracy for AM envelope measurements.

Additionally, the vector signal analyzer offers signal processing that displays the AM modulating signal in units of AM modulation depth for nontransient amplitude-modulated carriers. This is done with signal processing algorithms that automatically remove the carrier amplitude offset and normalize to the carrier amplitude.

Ultrasound Transducer Analysis

Demodulation can provide useful insights into the behavior and performance of ultrasound transducers. The voltage on the transducer can be measured by the vector signal analyzer and FM demodulated to provide a frequency profile over time of a transmit pulse and echo. The demodulated FM time waveform will generally show four components. First, large FM noise deviations will dominate the time display where there is no signal, such as before excitation and before the echo is received. When the excitation arrives, the noise fluctuations turn into a flat signal with an FM deviation corresponding to the difference between the excitation frequency and the measurement center frequency. Upon removal of the excitation, the transducer begins to vibrate at its natural resonant frequency, dissipating stored energy. This induces an electrical signal, which the vector signal analyzer can measure and demodulate. Thus, the natural resonant frequency of the transducer can be measured.

Finally, the echo will produce transducer vibration, which can be detected in the demodulated FM time waveform. This is generally very weak and noisy in appearance. If a trigger is available, this measurement can be averaged to improve the signal-to-noise ratio.

The envelopes of the transmit and receive pulses can also be observed by selecting the linear or logarithmic display format in vector measurement mode. Bandwidth is controlled by the measurement span. The analyzer's high sensitivity and wide dynamic range allow surprisingly weak echoes to be observed on a logarithmic time display.

Automatic Carrier Frequency Determination for Broadband Transmissions

Automatic carrier frequency estimation is available for PM and FM demodulation. The carrier frequency is estimated independently over each time record. An additional feature in FM allows determining the carrier frequency more accurately when measuring signals with broadband modulation.

With averaging turned on, the FM demodulator will perform a weighted average on each of the carrier frequency estimates. This average can be displayed on the demodulated trace as a marker function. It is also used to compensate the FM result for the frequency offset arising from the difference between the carrier frequency and the measurement center frequency.

Best results are obtained by setting the measurement span slightly larger than the bandwidth of the modulated signal. The digital filters can easily reject alternate channels, but overlapping adjacent channels will degrade performance. One should select the minimum resolution bandwidth possible and select 3201 frequency points to increase the number of samples in the time record to the maximum allowable. More time samples will reduce the variance on each carrier frequency estimate.

Although not obvious, the carrier frequency of digitally modulated carriers as well as FM carriers can be determined using this technique. Data on a digitally modulated carrier must be random for the estimated carrier frequency to converge on the true carrier frequency. This condition is generally met in channels carrying normal data.

A typical application is to find the carrier frequency of a signal on a satellite channel. Satellites often carry a mixture of FM and digitally modulated carriers. The vector signal analyzer's digital filters can be chosen to select one channel of interest and reject others. The estimated carrier frequency can be used to verify that the proper transmitter frequency is being used for that assigned channel space.

Timothy L. Hillstrom
Douglas Wagner
Development Engineers
Lake Stevens Instrument Division

The digital receiver architecture of the HP 894xxA performs the equivalent of this single-sided shift and filter with a quadrature local oscillator and separate low-pass decimating filters to reject higher-order mixing components. The in-phase and quadrature-phase components are recombined in DSP memory and treated as a single-sided complex waveform.

The local oscillator can be represented as a complex sinusoid:

$$e^{-j(\omega_L t + \phi_L)} \qquad (2)$$

where ω_L must be close enough to ω_c that after multiplying x(t) by equation 2, the positive frequency image in x(t) is shifted near enough to dc to keep all sidebands of m(t) within the passband of the low-pass decimating filters. Under this assumption, the negative frequency image of x(t) is completely rejected after filtering, and the resulting complex time waveform is expressed as:

$$y(t) = Cm(t)e^{j[(\omega_c - \omega_L)t + \phi_c - \phi_L]}$$
$$= C[1 + a(t)]e^{j[(\omega_c - \omega_L)t + (\phi_c - \phi_L) + \phi(t)]}.$$

The amplitude modulating waveform, a(t), is recovered by first taking the magnitude of y(t):

$$|y(t)| = C[1 + a(t)], \qquad (3)$$

estimating the carrier amplitude with a weighted average on equation 3, removing the carrier amplitude from equation 3, and normalizing by the carrier amplitude:

$$a(t) = [|y(t)| - C]/C.$$

The result, a(t), is in units of amplitude modulation depth. For example, a maximum zero-to-peak amplitude of 0.5 for sinusoidal a(t) corresponds to a modulation index of 50% on the carrier.

Equation 3 represents the AM envelope and can be obtained in the vector measurement mode by selecting the linear-magnitude data format for the time waveform. This is useful for capturing transient events for which an average on equation 3 does not give a true estimate of the carrier amplitude.

The phase of y(t) includes the desired phase modulating waveform, $\phi(t)$, as well as a phase offset and ramp:

$$\angle y(t) = (\phi_c - \phi_L) + (\omega_c - \omega_L)t + \phi(t). \qquad (4)$$

Ideally, the local oscillator is equivalent to a coherent carrier, providing carrier-locked demodulation. The condition

of carrier lock is met when $\omega_L = \omega_c$ and $\phi_L = \phi_c$, in which case the phase of y(t) yields the desired phase modulating waveform:

$$\angle y_{coh}(t) = \phi(t).$$

The HP 894xxA provides a rear-panel 1, 2, 5, or 10-MHz reference input. The HP 894xxA's digital local oscillator can be set to the same frequency as a carrier synthesized from the reference. In this case, the digital local oscillator is frequency-locked and phase stable with respect to the carrier. The phase of the carrier is not available because the digital local oscillator is derived from frequency division and multiplication with respect to the carrier. This leaves a phase offset in the phase of y(t):

$$\angle y_{frq}(t) = \left(\phi_c - \phi_L\right) + \phi(t). \tag{5}$$

To obtain the desired phase modulating waveform for frequency-locked measurements, the phase offset $(\phi_c - \phi_L)$ is removed by computing a weighted average on equation 5.

A second PM demodulation mode is available when a frequency reference cannot be obtained. In this case, the carrier phase ramp, $(\omega_c - \omega_L)t$, requires different compensation to retrieve the desired phase modulating waveform. Time differentiating equation 4 eliminates the phase offset, giving:

$$\beta(t) = (\omega_c - \omega_L) + d\phi/dt. \tag{6}$$

Calculating a weighted average on equation 6 gives an estimate of the frequency offset, $\omega_c - \omega_L$. Removing the estimated frequency offset from equation 6 leaves the frequency modulating waveform. This is integrated with respect to time, providing the desired phase modulating waveform, $\phi(t)$, with all carrier components removed.

Frequency demodulation follows essentially the same steps as phase demodulation. A high-quality differentiator is applied to equation 5 to retrieve the frequency modulating waveform.

All carrier offset compensation can be turned off. In this case, PM demodulation is equivalent to selecting the phase data format on a time waveform in the vector measurement mode. Deactivating carrier offset compensation is useful for capturing transient events for which an average on equation 5 or 6 will not provide a true estimate of phase offset or frequency offset.

A frequency offset is independently estimated for each data record sent to the demodulation signal processing block. The estimated frequency offset summed with the HP 894xxA's local oscillator frequency (the center frequency of the measurement) is available for display on the demodulated data trace as the estimated carrier frequency.

The accuracy of the carrier compensation algorithms depends on how close a(t) and $\phi(t)$ approximate zero-mean functions over each data record sent to the demodulation signal processing block. Biases on the estimated carrier frequency become significant if a sideband amplitude is within several dB of the carrier and close enough to the carrier to result in fewer than approximately ten cycles over a data record.

With complex modulation, there can be unacceptable variance in the estimated carrier frequency. This variance can be reduced substantially by averaging the estimated carrier frequency. For FM demodulation only, averaging will activate an exponential average of each estimated frequency offset, and this averaged estimate will be used to compensate equation 6. An equivalent exponential average is performed on estimates of the carrier amplitude for averaging in AM demodulation.

The HP 89440A provides one channel of RF information for demodulation. Two channels can be independently demodulated in the baseband receiver mode. When the receiver is in RF mode, the second IF channel can measure baseband signals for comparison with a demodulated signal in the RF channel.

Simultaneous RF and Baseband Measurements

One of the more useful features of the HP 89440A vector signal analyzer is its ability to demodulate an RF signal on one channel while simultaneously measuring a baseband signal on the second. This feature can be used to isolate the signal causing disturbances in an LO, to measure the frequency response of a modulator, to study the loop characteristics of a phase-locked loop, or simply to measure the time delays between baseband and modulated RF signals. All of the two-channel time and frequency measurements that can be performed with the baseband analyzer, such as frequency response, correlation, coherence, or cross spectrum, can be performed for a demodulated RF signal and a baseband signal.

The concept behind these measurements is quite simple. Imagine comparing a signal that was used to modulate a carrier with the output of a demodulator operating on the modulated carrier. If the modulator and demodulator are ideal, then the two signals will be identical. If only the demodulator is ideal, then the two signals can be used to study the characteristics of the imperfect modulator. The vector signal analyzer can directly measure the first signal (on the baseband channel) and accurately demodulate the RF carrier to measure the second signal.

To measure the response of a modulator, the baseband source would be connected to the input of the modulator. For frequency response measurements, the source would be used to generate broadband signals such as a periodic chirp or random noise. For time-domain measurements, the arbitrary capability of the source might be used to generate a ramp or step. The source signal applied to the circuit under test would also be connected to the baseband input so that it could be measured as a reference signal. The RF channel would be used to demodulate the carrier at the modulator output. Using data from both channels, the response of the modulator to the source signal can be determined. The stimulus and response can be compared in the time domain or combined to compute a frequency response.

For determining the source of disturbances on an oscillator, the RF channel would be connected to the oscillator output and a demodulator selected. Typically a phase or frequency demodulator would be used. A probe would be attached to the baseband channel so that various signals in the circuit could be easily measured and compared to the demodulator output. For example, the probe might first be connected to the power supply line, and then moved to a nearby logic line. Depending on the type of disturbance, the user would either compare the signals directly in the time domain or use the

two signals to compute the crosscorrelation or coherence functions.

Internal to the instrument, both input channels are identically configured for RF/baseband measurements except that the RF channel has a demodulator added to the signal path and the digital LO for the baseband channel is configured for a center frequency of 0 Hz. This results in both channels having the same sample rate, which is necessary to allow cross-channel measurements such as frequency response. In the process of converting a complex signal into a real signal (with the same sample rate), the demodulator reduces the information bandwidth of the RF channel by half. The data for the second channel is also complex, but since the imaginary part of the waveform is zero (because of the frequency and phase of the digital LO), the signal is treated as a real signal and only half of the spectrum is displayed.

Time Selective Frequency Analysis

Often, today's spectrum analyzers are called upon to analyze signals that are not continuous. To avoid contamination by unwanted signals, pulsed or transient signals must be isolated in time before being converted to the frequency domain. The following few examples show how pervasive time-variant signals have become:

- In the United States, a frame of television video is broadcast as 525 lines in two interlaced fields. Lines 17 through 21 of each field are reserved for test signals, which may coexist with normal picture information.
- Time-division multiple access (TDMA) signals require pulsed modulation of an RF carrier. For the North American Digital Cellular (NADC) system, a 40-ms frame is composed of six 6.66-ms slots. A frame can carry three conversations, each using two slots. Base station transmission is not pulsed since power consumption is not a major concern. Mobile station transmission is pulsed to conserve power.
- Code-division multiple access (CDMA) communication systems combine digital modulation and spectrum spreading techniques to create broadband signals that are immune to noise. Power in a mobile CDMA phone is gated on and off in 1.25-ms bursts when the data rate is less than the full 9600 bits per second.

In addition to isolating valid information within a pulsed signal, time selectivity is critical to analyzing transients such as transmitter turn-on.

Often the user must think about a measurement problem in the time and frequency domains at the same time. As discussed earlier, the HP 894xxA vector signal analyzers' adjustable information bandwidth provides the user with control over the frequency-domain aspects of the measurement. With accurate hardware triggering and flexible time-record processing the user can address many of the time-domain issues. However, these two aspects of the measurement are not independent. Because of the nature of the FFT the extents and resolutions of the time-domain and frequency-domain data are interrelated. This is shown Fig. 4. Most FFT signal analyzers force the user to interact primarily in the frequency domain. The time record size is a fixed number of points (generally a power of two) and span is the only parameter under the user's control. This means that time record length (in seconds) is determined directly by the span. The user has no control over the instrument's resolution bandwidth aside from changing the span. Some of the newer FFT
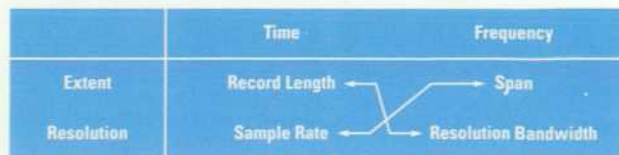

		Time	Frequency
Extent		Record Length	Span
Resolution		Sample Rate	Resolution Bandwidth

Fig. 4. Some measurements have both time-domain and frequency-domain aspects. Because of the nature of the fast Fourier transform, these two aspects are not independent. The extents and resolutions of the time and frequency data are interrelated as shown here.

signal analyzers give the user several options for the number of time points, but the time record length is still directly dependent on the span. While the HP 894xxA supports this traditional FFT mode it also allows the user to select the time record length or resolution bandwidth independently of the span. Time record length and resolution bandwidth are still related by:
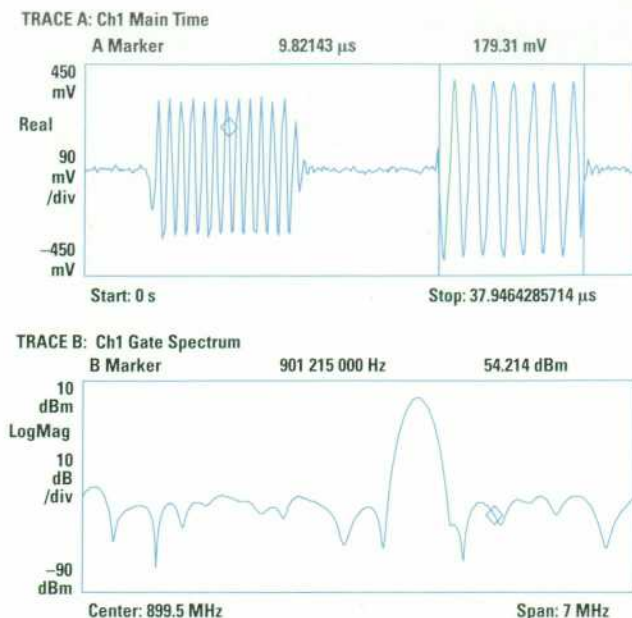
$$RBW = WBW/T,$$

where RBW is the resolution bandwidth in hertz, WBW is the noise bandwidth of the time window in bins, and T is the time record length in seconds. The number of time points can no longer be directly set. Instead, the maximum number of time points the analyzer can store imposes a limit on the choice of T (and on the minimum allowable resolution bandwidth) for the particular sample rate, which is determined by the span.

It is easy to see that to support variable resolution bandwidths, the HP 894xxA measurement architecture must be able to handle variable-sized time records. This affects all measurement and display-related signal processing, especially conversion between the time and frequency domains. Since the FFT is performed only on time records that are 2^N samples in size, the HP 894xxA's arbitrarily sized time records must be zero padded before performing the FFT. Zero padding merely adds zeros to the ends of the time record and changes the resolution of the frequency record. It does not alter the information content of the data. Zero padding is implemented through the application of the time window. The width of the time window is not constrained to be integer multiples of the sample time. Therefore, although the HP 894xxA vector signal analyzer operates on discrete time data, this imposes no quantizing effects on the resolution bandwidth. Based on its phase characteristics, the time window selection determines if zero padding is done at the end of the time record or if the time record is centered and zero padded at both ends. The effects of zero padding must be undone following the inverse FFT of the frequency data. In this case the time record must be resized and possibly shifted to remove all artificial zeros at the ends.

Often during a measurement, the user is not sure of what to look for or where the signal of interest lies. In this case the user may want to see a signal over a large segment of time to help isolate the portion to be analyzed. The HP 894xxA vector signal analyzer supports this through its time gating feature. The user sets a main data length, which defines the extent of the time data to be acquired. Once the main data has been acquired the user can gate out a region of the main data to be analyzed further. All subsequent time and frequency analysis will be performed on the gate region. Generally, the size and position of the gate region can be changed

TRACE A: Ch1 Main Time

A Marker   9.82143 µs   179.31 mV

```
450
mV

Real

90
mV
/div

-450
mV
```
Start: 0 s   Stop: 37.9464285714 µs

TRACE B: Ch1 Gate Spectrum

B Marker   901 215 000 Hz   54.214 dBm

```
10
dBm

LogMag

10
dB
/div

-90
dBm
```
Center: 899.5 MHz   Span: 7 MHz

Fig. 5. Time selective frequency analysis can display the spectrum of the second pulse of this two-pulse waveform.

and the data reanalyzed without retaking the main data. By simultaneously observing the main time, gate time, and gate spectrum the user gets a complete picture of the signal of interest including where it fits within a larger sequence of events. Fig. 5 shows how the HP vector signal analyzers interact with the user to perform time selective frequency analysis. The upper trace shows a two-pulse time waveform. A pair of vertical gate markers are placed over the right pulse. The frequency-domain representation of that gated pulse is shown in the lower trace of the display.

Certain measurement situations require that a large amount of time data be analyzed. If the signal is transient in nature, if it must be analyzed without any time gaps, or if it must reanalyzed several times the HP 894xxA's time capture feature can be used. In time capture mode, up to 10^6 samples (or half as many for each of two channels) can be acquired and stored in the sample RAM. This data can then be played back into the measurement as many times as necessary. During playback all of the measurement's features are available. For instance, the user can reanalyze the data with several resolution bandwidths or time windows.

Conclusions
Making the complex and powerful measurements associated with today's signals and systems is inevitably more difficult.

A primary goal for the vector signal analyzer project teams was to create analyzers that can handle the complex interactions between the frequency, time, and modulation domains by themselves, freeing the user to concentrate on the desired measurement results. This approach avoids alienating the users of traditional analyzers while providing the tools required for the demanding measurement needs of today and tomorrow.

References
1. K.C. Carlson, et al, "A 10-Hz-to-150-MHz Spectrum Analyzer with a Digital IF Section," *Hewlett-Packard Journal*, Vol. 42, no. 3., June 1991, pp. 44-60.
2. R.C. Blackham, et al, "Measurement Modes and Digital Demodulation for a Low-Frequency Analyzer," *Hewlett-Packard Journal*, Vol. 38, no. 1, January 1987, pp. 17-25.

pSOS is a trademark of Software Components Group, Inc.

A Firmware Architecture for Multiple High-Performance Measurements

The HP 894xxA vector signal analyzers perform fast, sophisticated measurements on complex waveforms. The firmware architecture provides access to multiple processors to meet the high-performance requirements while allowing individual measurements to share common features and protocol.

by Dennis P. O'Brien

The HP 894xxA vector signal analyzers offer a diverse set of advanced measurements. The measurement firmware architecture was designed with the two preeminent requirements of performance (i.e., speed of operation) and functional leverage across multiple measurement modes.

Measurement performance is measured in terms of both loop time and user command response time. Loop time is expressed as either display updates per second or real-time bandwidth, that is, the maximum frequency span at which the input signal can be processed without missing data. Of these two performance measures, loop time generally took precedence when design trade-offs were necessary. In Fig. 1, which shows the flow of signal processing for three of the HP 894xxA's four measurement modes, the loop time is the

Fig. 1. HP 894xxA vector signal analyzer measurement data flow.

time required to process the data from the sample RAM to the display.

Each measurement mode offers the user a variety of signal processing customizations in addition to numerous standard features. The standard features, common to all measurement modes, present a consistent look and feel to the user and make each measurement operate as a "good citizen" within the overall instrument firmware architecture. Fig. 1 shows that each of the HP 894xxA's measurements can use arbitrary span and time-domain corrections to condition the incoming time data, and that trace display conditioning operations such as user math, coordinate transformation, units conversion, scaling and pixelation are also used in common by all measurements.

Numerous additions to the standard features are required to meet the signal processing demands of individual measurements. In many cases this special signal processing is leveraged across several measurements—averaging, time gating, and time-to-frequency conversion are good examples. However, some additional signal processing is so specialized that it meets the needs of only a single measurement—analog and digital demodulation, for example.

A far bigger part of the measurement task than signal processing is control. It is by observing common conventions at its firmware interface that a measurement maintains its "good citizen" standing. Also, it is in observing common internal conventions that all measurements present a consistent interface to the user. Most of the control features of the HP 894xxA's measurement system are shared by all of its measurements. Some of the standard control features designed to provide consistency to the user are:
- Measurement pause/continue
- Measurement restart
- Single/continuous sweep modes
- SCPI* operation complete
- SCPI status register control
- Recomputation of results following parameter changes.

Some of the standard control features designed to provide a uniform interface to the rest of the firmware are:
- Parameter change handling
- Formalized interaction with the calibration process
- Formalized interaction with the autorange process
- Measurement activation and deactivation
- Measurement memory reallocation
- Premeasurement trace results information tracking
- Run-time measurement trace results information tracking.

To support a diverse set of measurements that share many common control and processing features and whose primary requirement is high performance, the measurement firmware architecture was shaped by these goals:
- Fully utilize the DSP96002 signal processor to provide the highest possible performance.
- Provide default control and signal processing implementations of basic features that are easily inherited by individual measurements.
- Make it easy to customize the default control and processing features for individual measurements.
- Implement the measurement architecture as a platform upon which future measurement designs can be built.

* SCPI = Standard Commands for Programmable Instruments.

Data Structure

All of the HP 894xxA measurements are block-oriented, that is, they operate on blocks or arrays of data that represent the signal over a given range rather than at just one point. The fundamental data structure is a vector, which embodies both a data array and a header structure. The data portion of the vector is generally represented as 32-bit floating-point numbers throughout the measurement and may be either real or complex. It is converted to integer format as part of the display processing. The header structure contains information about both the measurement setup that produces the data (center frequency, span, input range, etc.) and the data itself (such as whether or not it contains overloads).

Each measurement uses a collection of vectors. Fig. 2 shows the vector allocation for the HP 894xxA vector measurement mode. The timeData, freqData, and avgBufData vectors contain the measurement's raw data. These are the measurement's fundamental results. All user-selectable measurement results must be derived from these. The measData, dispData, and pixBufData vectors are trace-oriented, so there are four of each, or one per trace. The measData vectors are vector pointers that point to one of the mathMeasData vectors, which contain the final version of each user-selected measurement data result. There may be many more than four of these, since user math uses them as input operands and it is not constrained to use only displayed data. The mathMeasData, measData, dispData, and pixBufData vectors are collectively referred to as the trace vectors. As will be seen, several processes have access to the trace vectors. Therefore, a protection mechanism is required to restrict access to the measData vectors and all others that are farther downstream. Finally, there are a number of support vectors. Some of these contain vector constants that will not change during a measurement. Other support vectors are used for temporary storage (not shown in Fig. 2).

The timeData vector is the output of the time-domain correction filter block in Fig. 1 and is the input to the measurement-specific signal processing block. At the other end of the measurement loop, the mathMeasData vectors are the output of the measurement data and user math block.

The amount of memory for all required vectors varies from measurement to measurement. Not only does it vary between measurement modes, it may change with the setup of a given measurement. For example, changing the number of frequency points, the resolution bandwidth, or the average state can have considerable impact on the number or size of the vectors required. To make the best use of the available system RAM without implementing a sophisticated memory management scheme, all of the vectors are reallocated when a new measurement mode is selected. Their sizes are determined by user configuration of key parameter limits such as maximum frequency points. They are made large enough to support all permutations of the selected measurement mode within the limits set by the user. As mentioned earlier, some of the signal processing is shared by all measurements. Therefore, much of the vector allocation scheme is leveraged across all measurement modes.

The DSP96002 signal processor (hereafter referred to as the DSP) has a limited amount of high-speed RAM available. Operations on data blocks that reside in this RAM are up to seven times faster than if the data block is resident in the
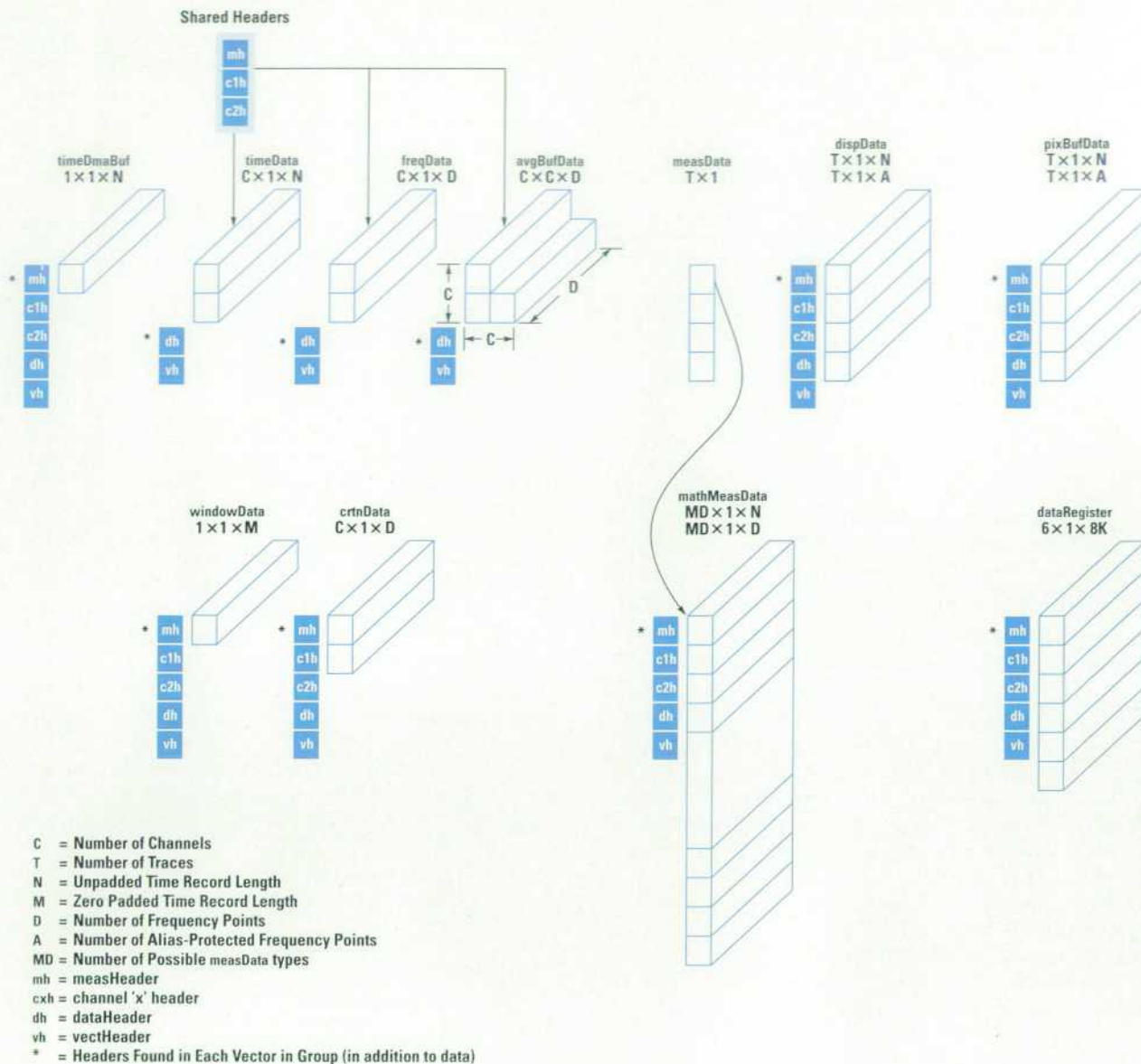
Shared Headers

C = Number of Channels
T = Number of Traces
N = Unpadded Time Record Length
M = Zero Padded Time Record Length
D = Number of Frequency Points
A = Number of Alias-Protected Frequency Points
MD = Number of Possible measData types
mh = measHeader
cxh = channel 'x' header
dh = dataHeader
vh = vectHeader
* = Headers Found in Each Vector in Group (in addition to data)

Fig. 2. HP 894xxA data vector architecture. The elongated blocks represent the data portions of the vectors.

system RAM. To take advantage of this improved performance the most critical vectors are placed in the DSP RAM. Unlike their system RAM counterparts, these vectors are reallocated to be exactly the required size whenever a parameter change is made that affects the vector makeup of the measurement. Reallocating and relocating vectors in response to parameter changes such as user selection of a new measurement result makes any DSP-resident vectors vulnerable to being lost. This is acceptable for vectors from mathMeasData downstream because they can be recomputed from the measurement's raw data. However, the loss of raw data would be a serious problem. To circumvent this problem the measurement system implements a scheme to preserve raw data across parameter changes. The scheme involves copying DSP-resident vectors to their system RAM counterparts before memory reallocation and then copying them back (if they continue to reside in the DSP memory) following reallocation. Much engineering effort was spent in

ensuring that the measurements make the best use of the fast DSP memory. This is a major contributor to the HP 894xxA's high 78-kHz real-time rate.

Process Structure

To achieve the highest possible performance, the measurement loop is split between the DSP and the main CPU. While many embedded systems have math processors that operate as slaves of the central processor, in the HP 894xxA the measurement loop runs primarily in the DSP. The DSP uses the main CPU as a slave to:

- Perform certain housekeeping functions via remote procedure calls such as run-time trace vector header information tracking
- Perform integer math on pixelated trace results
- Interface with the graphics system processor (GSP), which is the display processor.

Run-Time-Configurable Hardware Drivers

One of the difficult challenges facing any firmware team during a several-year project is turmoil in the software because of evolution of the hardware specification upon which the software must run. This hardware evolution is a natural process given the complexity of our systems and the demanding environmental and RFI testing they must pass. All too often the simplest of hardware changes can wreak havoc on the software interface to the hardware.

The high cost of prototypes and the evolutionary nature of complex hardware systems also can lead to the requirement that software be flexible enough to work with a mix of various hardware versions. Some of the factors contributing to software turmoil in hardware drivers are:

- Increased market understanding causing new expectations
- Technology changes
- Defective or marginal parts causing redesign
- Design flaws
- Interface requirements to other hardware in flux
- Software not available for complete turn-on
- Standards
- Printed circuit board layout constraints
- Manufacturing processes.

Early in the development of the HP 894xxA analyzers we realized that this would be a considerable problem since we were developing a new hardware platform. Given that the software would be required to deal with multiple revisions of each hardware board as well as various combinations of boards (Fig. 1), the following design criteria developed:

- Must be easy to accommodate multiple revisions.
- Must be easy to share code common to all revisions. This prevents having N versions of common algorithms.
- Must be easy to remove obsolete versions. If code for all versions is intermixed, pulling out obsolete code may be difficult and prone to introducing new problems.
- Must offer a clear, consistent interface from an external perspective. Clients of a hardware driver should not need to know or care which revision of the hardware is present.

C++ Hardware Drivers

Our solution was to develop a class hierarchy based on C++ classes. By providing a base class for a particular set of hardware and generating derived classes for particular hardware revisions, it became an easy task to support multiple hardware revisions.

A C++ base class is used to define the external interface of the hardware and to define what capabilities are required by derived classes of the base hardware class. The base class will often serve only as a definition and may have no useful code. In other cases the base class may have considerable code encompassing shared algorithms and advanced interface functions.

An example of a class structure is shown in Fig. 2. The base class AdcHw provides the external interface for the ADC hardware and the default implementation of much
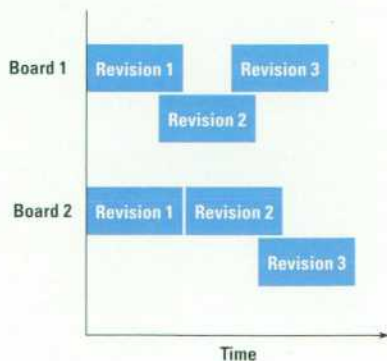

Fig. 1. The firmware must deal with multiple revisions of each hardware board and with different combinations of boards.

Fig. 2. Example of a class structure.

of the interface. The derived classes AdcHwR2, AdcHwR3, and AdcHwR4 provide overrides for unique aspects of revisions two, three, and four of the hardware. A portion of the AdcHw class is shown below:

```
class AdcHw
{
protected:
    int adcIndex;
    HwChannel channel;
    int revision = 0;
public:
    AdcHw(void);
    virtual void getAdcGa( char ** dataPtr, int * size);
    virtual void getRFAdcGa(char ** dataPtr, int * size);

    virtual void getBBAdcGa(char ** dataPtr, int * size);
    virtual void ditherCalSetting(int value);
    virtual void returnControlBits (void ** bitsPtr, int * count);
    virtual void setAdcIndex(int index);
    virtual void setControlWord(unsigned int control);
    virtual void setPreLoad(unsigned char newValue);
    virtual void setMuxControl(void);
    virtual void clearMuxControl(void);
    virtual int adcRev(void);
};

extern AdcHw * adcHwCh1;
extern AdcHw * adcHwCh2;
```

Member functions declared virtual can be optionally overridden by a derived class written for a different revision. The two variables adcHwCh1 and adcHwCh2 represent the external interface handle through which all other software interfaces with the ADC hardware.

Revision 4 of the ADC hardware required a new gate array program as well as a change in how control bits are configured. The implementation of a driver for this revision consisted of implementing several member functions to handle the incremental differences:

```
class AdcHwR4 : public AdcHw
{
public:
    AdcHwR4(); // class constructor
    virtual void returnControlBits(void ** bitsPtr, int *count);
    virtual void setControlWord(unsigned int control);
    virtual void getAdcGa(char ** dataPtr, int *size);
};
```

The two functions returnControlBits() and setControlWord() deal with the fact that the control bit interface changed and the function getAdcGa() deals with providing a unique gate array program. The software is now able to use version 4 ADC hardware without changing any other source code in the instrument. When a reference is made to adcHwCh1->setControlWord(), C++ will automatically arrange to invoke the version defined in the adcHwR4 class.

How do we determine which version of the class to use? Since multiple hardware drivers exist in the software, some mechanism needs to determine which driver to

use when the instrument powers up. In the HP 894xxA, this determination is done by reading an ID register on the hardware board of interest. This ID register is most often a hard-wired register that has three bits of version information. When a new board is made with a different hardware interface or capability, this version number is changed. The code to do this might be as simple as:

```
switch (adcRevision)
{
default:
case 0:
case 1:
    adcHwCh1 = new  AdcHw;
    adcHwCh2 = new  AdcHw;
    break;
case 2:
    adcHwCh1 = new  AdcHwR2;
    adcHwCh2 = new  AdcHwR2;
    break;
case 3:
    adcHwCh1 = new  AdcHwR3;
    adcHwCh2 = new  AdcHwR3;
    break;
case 4:
    adcHwCh1 = new  AdcHwR4;
    adcHwCh2 = new  AdcHwR4;
    break;
}
```

Summary

By implementing our hardware drivers as C++ classes, we were able to maintain a consistent software interface to the hardware while providing support for multiple versions of the hardware. Common code shared among all drivers is encapsulated in a base class, which itself knows nothing of the particular version in use, but instead relies on virtual functions to allow revision-specific drivers to override functionality seamlessly. The code is much cleaner because the software writer does not need to sprinkle conditional statements in the source code to accommodate various versions. In addition, it becomes an easy matter to remove obsolete versions of the hardware driver. Finally, the C++ compiler takes care of the mundane task of making sure the right function overrides are invoked for the installed version.

Glenn R. Engel
Development Engineer
Lake Stevens Instrument Division

For the measurement loop to run in the DSP with no run-time intervention or control from some higher authority in the main CPU, that loop has to be completely defined before being executed. Several options were open to the design team to address this problem. The approach chosen was to provide a minimum set of DSP-resident measurement loops, hereafter referred to as a measSequencer, which can be customized by individual measurements before run time. The customization is done through a compile process which will be described in a subsequent section. The default measSequencers provide every measurement with all of the control required to interact in a uniform way with all other entities in the instrument firmware architecture.

With this approach, a single designer was able to implement the measSequencers for all measurement modes. In addition to promoting a uniform interface for measurement loops of all measurement modes, this approach allowed other designers freedom to concentrate on the specific customizations required by their particular measurement mode. It is true that this is one trade-off that sacrificed run-time performance for ease of development. However, the overhead to support a generic measSequencer state machine is only about 300 μs, which is less than 1.5% of the loop time required for 78-kHz real-time operation at 1601 frequency points.

Fig. 3 shows the HP 894xxA measurement firmware architecture. The solid bubbles represent software modules. A module depicted with concentric rings is a process. Bubbles that are stacked indicate that multiple versions of that module exist to support the multiple measurement modes. Data stores are represented as pairs of parallel lines enclosing the stores' names.

Measurement Feature Leverage
There several ways to leverage standard features across measurement modes. Code can be copied as a means of reuse. The volatility of measurement firmware modules makes code developed in this manner very difficult to maintain. Dispatch tables, customized by the activation of a measurement mode, allow a foundational architecture to provide default functionality and yet flex to meet the custom needs of individual measurements. Dispatch tables, however, must be kept current manually. The HP 894xxA measurement firmware architecture uses C++ classes and inheritance to allow individual measurement modes to build on a foundational measurement. Several of the modules in Fig. 3 are implemented as C++ classes. A partial class hierarchy is:

```
MeasSeqCtlr
    HP89410MeasSeqCtlr
        VectorMeasSeqCtlr
        ScalarMeasSeqCtlr
        AnalogDemodMeasSeqCtlr
        DigitalDemodMeasSeqCtlr
MeasSeqGen
    HP89410MeasSeqGen
        BlockMeasSeqGen
            VectorMeasSeqGen
            ScalarMeasSeqGen
            AnalogDemodMeasSeqGen
            DigitalDemodMeasSeqGen
```

In the case of the MeasSeqCtlr class, a complete set of default actions is provided. Most of these actions are usable by all measurements. Where the default action is not what is required, the XXMeasSeqCtlr class (one of the classes derived from the MeasSeqCtlr class—"XX" stands for the name of the measurement mode) can overwrite the virtual function, replacing that default action with one of its own.

In the MeasSeqGen class, control of a complete compile sequence has been implemented. However, all of the measurement mode-specific customizations are the responsibility of a derived class.

Both the HP89410MeasSeqCtlr and HP89410MeasSeqGen classes provide customizations required by the HP 894xxA measurement system. Neither they, the MeasSeqCtlr class, nor the MeasSeqGen class provide sufficient functionality to implement a fully functional measurement. As such, they are abstract classes. That is, there are no instances of these classes in the system; there are only instances of their derived classes.
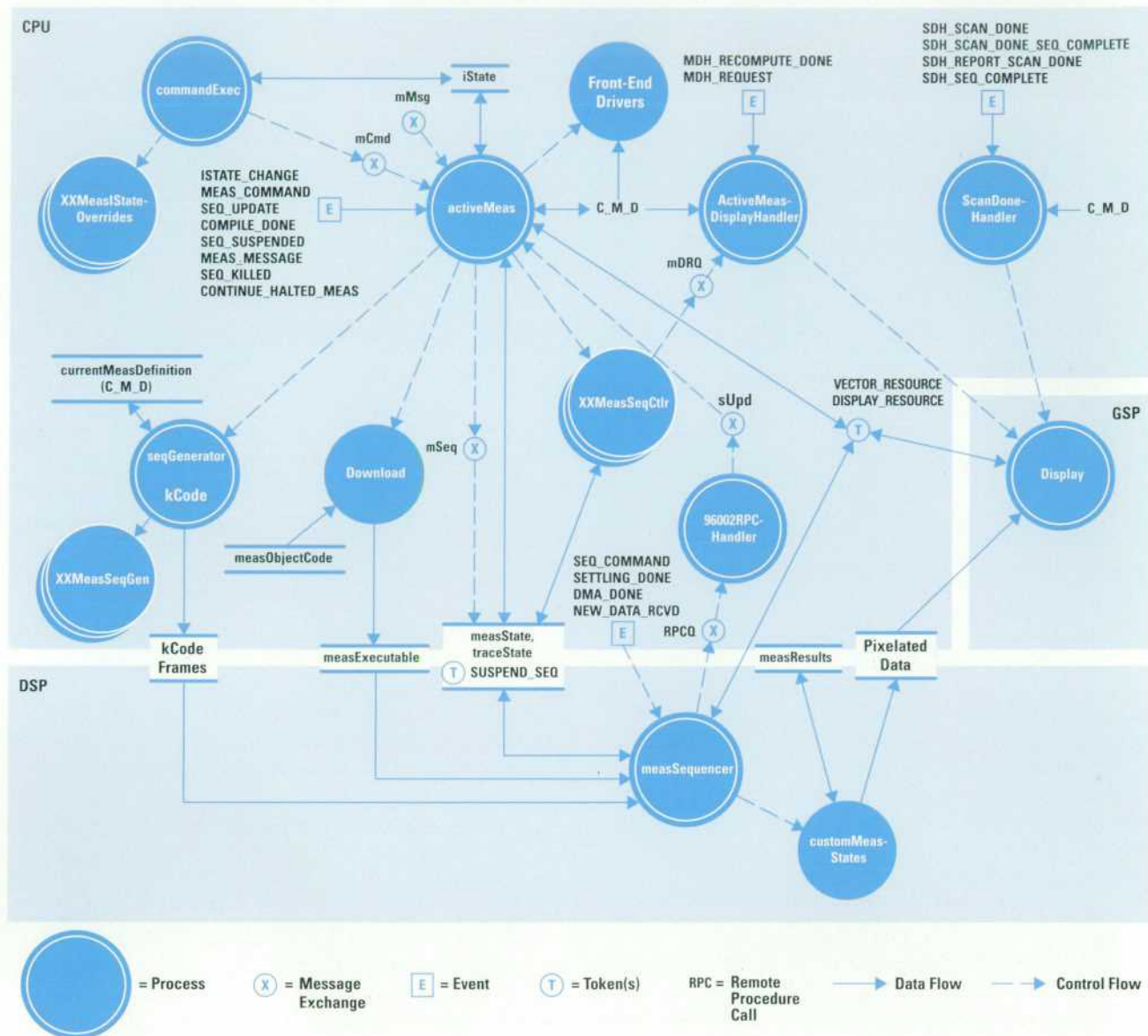
Fig. 3. Vector signal analyzer measurement firmware structure.

Customizations

The measSequencer is the sole process running in the DSP. It provides the default measurement loop control while the customMeasStates module executes customizations unique to the current active measurement. The exact set of customizations installed is determined by the current measurement mode's interpretation of the currentMeasDefinition data store (C_M_D in Fig. 3). CurrentMeasDefinition represents that portion of the instrument state (iState data store) that completely defines the measurement state. The seqGenerator controls the compile operation, which produces the composite measurement loop (i.e., measSequencer and customMeasStates). The XXMeasSeqGen modules ("XX" stands for the name of the measurement mode) interpret currentMeasDefinition according to the requirements of the current measurement mode, producing the measSequencer customizations executed by customMeasStates.

At the end of the measurement loop the measSequencer has, for each trace, produced pixelated data and tracked critical run-time information into both measState and traceState. At

this point, before beginning to process the next scan, the measSequencer signals the scanDoneHandler process by a remote procedure call. The scanDoneHandler then completes tracking header information from measState and traceState into the trace vectors' headers. When the vector headers are correct the scanDoneHandler computes any required marker values and further scales and formats the pixelated data. Finally, the data is transferred to the GSP for display. As the scanDoneHandler operates on its scan, the measSequencer is already processing the next scan. Since the measSequencer and the scanDoneHandler both access the trace vectors, a semaphore, VECTOR_RESOURCE, has been established to control access to these vectors.

In addition to the modules that generate a customized measurement loop and those that execute that loop, there are those that establish the overall interface to the measurement. The MeasSeqCtlr class provides default responses to measurement requests. It may handle a request itself or it may pass it along to the measSequencer. Examples of requests that it handles itself are those that cause the generation and
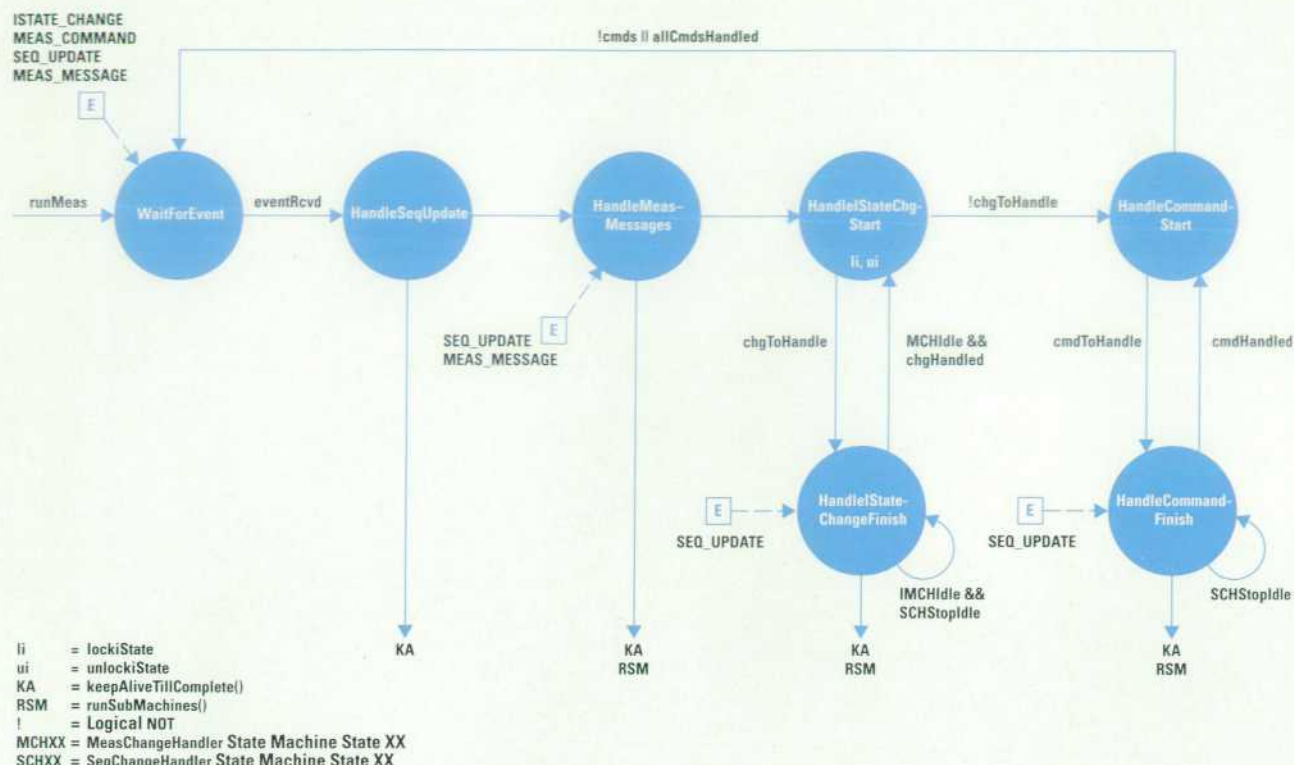
ISTATE_CHANGE
MEAS_COMMAND
SEQ_UPDATE
MEAS_MESSAGE

!cmds || allCmdsHandled

runMeas → WaitForEvent → eventRcvd → HandleSeqUpdate → HandleMeas-Messages → HandleIStateChg-Start → !chgToHandle → HandleCommand-Start

li, ui

SEQ_UPDATE
MEAS_MESSAGE

chgToHandle   MCHIdle && chgHandled   cmdToHandle   cmdHandled

SEQ_UPDATE   HandleIState-ChangeFinish   SEQ_UPDATE   HandleCommand-Finish

IMCHIdle && SCHStopIdle   SCHStopIdle

li = lockiState
ui = unlockiState
KA = keepAliveTillComplete()
RSM = runSubMachines()
! = Logical NOT
MCHXX = MeasChangeHandler State Machine State XX
SCHXX = SeqChangeHandler State Machine State XX

KA   KA RSM   KA RSM   KA RSM

Fig. 4. EventHandler state machine.

downloading of a new composite measurement loop. Start, pause, and continue are examples of requests that it directs to the measSequencer.

The XXMeasSeqCtlr module ("XX" stands for the name of the measurement mode) provides measurement-specific control when the default actions of the MeasSeqCtlr class are not sufficient. Likewise, the XXMeasIStateOverrides module provides measurement-specific filtering of user input when the default command interpretation is not sufficient.

Each specific measurement is defined by the default system operation plus the modifications made by the measurement-specific modules. To implement a new measurement mode, a measurement designer generates the three modules XXMeasSeqCtlr, XXMeasSeqGen, and XXMeasIStateOverrides.

Measurement Control

Each of the HP 894xxA measurement modes is implemented by single instances of XXMeasSeqCtlr and XXMeasSeqGen objects. Although there are four measurement modes, only a single mode is active at any time. This is referred to as the activeMeas.

The activeMeas must respond to many different types of requests. The most obvious are those that are the results of user actions. A change in the selected measurement results or coordinates must be handled immediately without affecting the raw data. A pause must be coordinated with the measSequencer to ensure a consistent set of valid trace vectors when the pause goes into effect. There are literally hundreds of softkeys that have an effect on a running measurement. While this is the source of the majority of measurement requests by far, there are internally generated requests as well. For example, the calibration manager requests permission of the activeMeas before performing the calibration. The

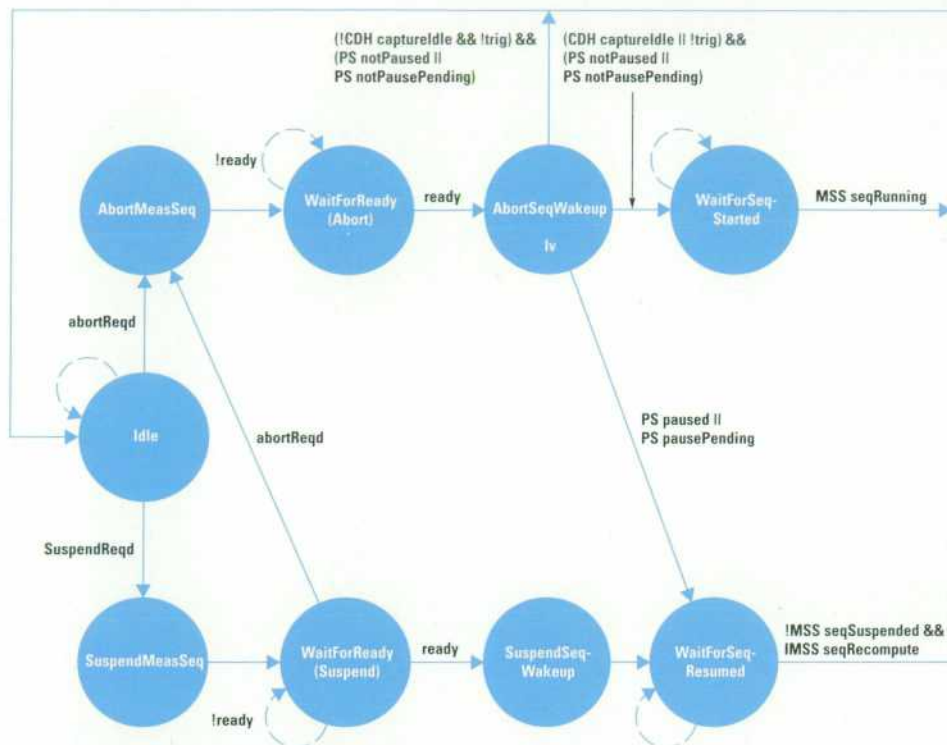activeMeas grants permission based on the its state and the state of the measSequencer.

The MeasSeqCtlr class must provide default actions to each request made of the activeMeas. To meet the goal of providing control of basic features which are easily inherited by individual measurements, these default actions must be useful for the majority of the XXMeasSeqCtlr derived classes.

At the core of the MeasSeqCtlr class are several state machines, which implement sets of basic actions. These state machines are not independent. Often their transitions rely on current states or state transitions of other machines. Three of the more central state machines are:

- EventHandler
- MeasChangeHandler
- SeqChangeHandler.

The activeMeas is normally a blocked process. It wakes up when one or more pSOS operating system events are received. Each event indicates that a different type of request has been made or that the measSequencer is informing the activeMeas of some important action. The EventHandler state machine (Fig. 4) addresses each event in a specified order. It makes sure that all requests represented by an event have been fully handled before addressing the next event. The EventHandler itself rarely fully handles a request. Instead, it generally asks one or more of the activeMeas's other state machines to handle it. A single event may represent many requests. This often happens when many iState changes are grouped together by the commandExec process. In an effort to improve user command response time, the EventHandler and the other activeMeas state machines are designed to handle multiple requests at the same time. Most requests require interaction with other processes before their handling is complete. If nothing else, the measSequencer generally must

States in the diagram:
- AbortMeasSeq
- WaitForReady (Abort)
- AbortSeqWakeup Iv
- WaitForSeq-Started
- Idle
- SuspendMeasSeq
- WaitForReady (Suspend)
- SuspendSeq-Wakeup
- WaitForSeq-Resumed

Transitions/labels:
!ready, ready, MSS seqRunning, abortReqd, SuspendReqd, PS paused || PS pausePending, !MSS seqSuspended && !MSS seqRecompute

(!CDH captureIdle && !trig) && (PS notPaused || PS notPausePending)
(CDH captureIdle || !trig) && (PS notPaused || PS notPausePending)

MSSXX = State XX of MeasSequencerStatus State Machine
SCHXX = State XX of SeqChangeHandler State Machine
DCHXX = State XX of DisplayChangeHandler State Machine
FCHXX = State XX of FrontEndChangeHandler State Machine
CDHXX = State XX of CaptureDataHandler State Machine
PSXX = State XX of PauseStatus State Machine
! = Logical NOT
ready = (MSS suspended || halted) &&
SCH idle &&
DCH idle &&
FCH ready

Fig. 5. MeasChangeHandler state machine.

be suspended as part of the request handling. In this case the EventHandler will block in a state other than WaitForEvent.

All of the activeMeas's other state machines run under the control of the EventHandler. The normal steps in handling an event are: (1) For the first request copy critical information from iState to currentMeasDefinition. (2) Request appropriate actions of other state machines. (3) Allow all state machines to run until they reach a quiescent state (i.e., no further state transitions). This is done in runSubMachines(). (4) If the request has not been completely handled then block. This is done in keepAliveTillComplete(). (5) When the request has been completely handled go to the next request. (6) When all requests for this event have been completely handled go to the next event. (7) When all events have been completely handled wait for more events in the WaitForEvent state.

Once started, the EventHandler will run as long as any other state machine is making state transitions. Once no machines can make any further advancement the EventHandler blocks, knowing that it will take an outside influence in the form of an event to allow any of the machines to advance further.

The MeasChangeHandler state machine (Fig. 5) is responsible for stopping the measSequencer, waiting until the current request has been handled and then restarting it. There are two options for this service: the measSequencer can be either suspended and resumed or aborted and restarted. The difference is that a suspension is coordinated with measSequencer

signal processing so that no data is lost, nor does raw measurement data become inconsistent with data in the trace vectors, while an abort happens immediately without regard for the validity of the vector data.

Finally, the SeqChangeHandler state machine (Fig. 6) is charged with the regeneration of a fully functional composite measurement loop. This may include requesting selection of a new measSequencer and a recompile of the customMeasStates by the XXMeasSeqGen. It must also compute many processing parameters such as the FFT window, the correction vectors, and the acquisition time record length. If any front-end hardware changes are required, then the SeqChangeHandler, in its DoSynchronizedChanges state, will coordinate with another state machine to make sure that the hardware setup is accomplished.

The SeqChangeHandler plays a critical role in measurement mode changes. In the StopSequenceCtlr state the activeMeas performs housekeeping required to deactivate itself. It then advances to StopIdle where control returns to the measurement manager, where a new measurement object is selected as the activeMeas. (The measurement manager is merely a repository for measurement objects and is not shown in Fig. 3). The new activeMeas then wakes up in the OverlayXXMeasSeqCtlr state where it initializes itself to operate as a fully functional, fully active measurement. It is here that available

Fig. 6. SeqChangeHandler state machine.

Diagram labels:

States: ForceRecompute (uv); Download-Executable; WaitForCompile; InitiateCompile; Idle; WaitForStopped (lv, uv); DoSynchronized-Changes; StopSequenceCtlr; StopIdle; OverlayXXMeas-SeqCtlr (li, ui, lv)

Transitions:
- !newExecutable && !trackCompile
- COMPILE_DONE && !newExecutable
- COMPILE_DONE && newExecutable
- newExecutable || trackCompile
- E / SEQ_KILLED
- E / COMPILE_DONE
- !measChgd
- measChgd
- seqStopped
- !newSeqCtlr
- !seqStopped
- newSeqCtlr
- !halted
- runMeas
- !runMeas
- E / CONTINUE_HALTED_MEAS

Legend:
```
li       = Lock istateResource
ui       = Unlock istateResource
lv       = Lock vectorResource
uv       = Unlock vectorResource
measChgd = MStateChg       ||
           recompute       ||
           newSeqCtlr      ||
           newExecutable   ||
           cfTrackReqd     ||
           trackCompile    ||
           measParmChg     ||
           syncFEChg
!        = Logical NOT
```

system RAM is reallocated according to the vector needs of this new activeMeas.

Each of the four individual measurement objects (modes) in the HP 894xxA measurement system has its own set of state machines. On a measurement mode change each state machine of the incoming measurement object must maintain continuity of its active state with its counterpart in the outgoing measurement object. With C++ it is easy for each measurement object to contain customized state machines and yet have equivalent machines in separate measurement objects share state variables. Consider this partial implementation of two state machine classes:

```
class StateMachine
{
    private:
      int crntState; // current state

    protected:
      int *crntStatePtr;

      void currentState(int newCurrent)
         { *crntStatePtr = newCurrent; }
      virtual void executeState();

      // add default StateMachine functionality here

    public:
      // constructor
      StateMachine();          { crntStatePtr = &crntState; }
      int currentState()       { return(*crntStatePtr); }
};
```

```
// possible states for instances of CustomStateMachine
typedef enum
{
    CUSTOM_STATE_0,
    CUSTOM_STATE_2,
    CUSTOM_STATE_3,
} CustomStates;

class CustomStateMachine : public StateMachine
{
    protected:
      // State variable – common to ALL instances of
      // CustomStateMachine
      static int currentState_;

    public:
      // constructor
      CustomStateMachine{crntStatePtr = &currentState_;
                         currentState(CUSTOM_STATE_0); }

      // add default CustomStateMachine functionality here
      virtual void customState0();
      virtual void customState1();
      virtual void customState2();
};
```

By accessing the current state through a pointer in the currentState() member function, CustomStateMachine can declare its own static current state variable and force its instances to use it by initializing the crntStatePtr to &currentState_ in its constructor. Since static member variables are shared by all instances of a class, classes derived from CustomStateMachine

always maintain synchronization of the current state in all of their instances. They may, however, overwrite the virtual state member functions and thus provide state machines customized to the needs of their measurement object owner.

XXMeasSeqCtlr

In describing the MeasSeqCtlr class and some of the state machines that it owns, actions specific to a measurement mode were mentioned. Computation of acquisition time record length is a good example. While the MeasSeqCtlr class provides a place for this to happen, it relegates the actual work to a derived class. While most actions of the MeasSeqCtlr class can be overwritten, very few are. Most of the default actions are well-suited for all of the HP 894xxA measurement modes. Most of the customization is limited to two areas: determining which iState changes affect the measurement and how they affect it, and computing acquisition-specific parameters and constant vectors.

XXMeasIStateOverrides

When the commandExec receives a change request it first validates the request. Then it puts the validated change in iState and informs the activeMeas of the change through a pSOS operating system event. The process of parameter validation must often be customized for a particular measurement. For example, the vector and scalar measurement modes impose different limits on the allowable resolution bandwidth. Custom parameter validation is done by the XXMeasIStateOverrides module. While conceptually separate, the XXMeasIStateOverrides module is actually implemented as part of the XXMeasSeqCtlr module.

Measurement Loop Customization

As mentioned earlier, there are two aspects to customizing the run-time measurement loop: measSequencer selection and generation of customMeasStates. MeasSequencer selection is simple, since there are only two measSequencers, one for averaged and one for nonaveraged measurements. Generation of customMeasStates is the majority of what the measSeqGen process does. Each measurement object has its own XXMeasSeqGen object. The XXMeasSeqGen instance associated with the activeMeas is known as the activeMeasSeqGen. The activeMeasSeqGen compile is broken down into five segments:
- Raw measurement (includes acquisition up to time gating)
- Time gate (includes FFT, frequency correction, and spectral average)
- Measurement data and user math calculation
- Coordinate transform and units conversion
- Scale, offset, and pixelation.

It is easy to see the correlation between compile segments and the flow of signal processing by comparing this list with Fig 1.

To minimize the user interface response time the compile process may be entered at any of the segments. However, once the compile process is started, all downstream segments must be recompiled. Generally, user change requests result in recompiling the minimum possible number of segments. For this to work the activeMeasSeqGen must observe two rules. First, the source operand vectors for any given segment must be completely defined by the previous segment and cannot be modified by this segment's compile.

Complete definition includes length and data memory address as well as header information. This is critical since the compiler relies heavily on the definition of source vector operands to produce a correct executable and to track and transform header information through math operations. For the very first segment, raw measurement, the activeMeas is responsible for initializing the timeData vector headers before starting the compile.

The second rule that must be observed is that a vector's data space must be allocated by the first segment that uses it. This is critical for vectors that reside in DSP RAM since their addresses are unknown until they are allocated at compile time. For system-RAM-resident vectors this operation does nothing since they are allocated when the measurement mode is switched to the current activeMeas.

All of the HP 894xxA XXMeasSeqGen objects inherit code generation for the user math step through the pixelation step. (Note that, while the user math and measurement data computations are in the same segment, they are implemented separately and can be overwritten and customized separately.) The measurement data code generation is inherited by all XXMeasSeqGens but the DigitalDemodMeasSeqGen class completely redefines it to support its unique set of measurement data result types. As might be expected, the raw measurement segment is overwritten and implemented differently for each of the measurements. The time gate segment is shared by the vector and analog demodulation measurements while it is overwritten and disabled for the scalar and digital demodulation measurements.

The end result of a compile is a set of *kCode* subroutines, each of which is composed of a group of kCode frames. Each frame represents an operation that can either perform block-oriented math on a vector or execute a remote procedure call. Those that perform vector math cause a DSP-resident interpreter to invoke highly tuned DSP-resident math routines. A compile segment can produce one or more of these subroutines. At run time the measSequencer invokes the kCode interpreter to execute a kCode subroutine at specific points in the measurement loop. In this way, customizations to the measurement loop signal processing are carried out.

kCode Compiler

The kCode compiler provides the XXMeasSeqGen designer access to all vector-oriented math operations as well as several other services that proved critical to efficient firmware development:
- Flexible kCode generation for vector math operations
- Complex/real vector type coercion (see below)
- DSP vector memory management
- Vector header information tracking
- Vector units tracking.

Only a fixed set of DSP-resident math routines are provided. Add, subtract, multiply, divide, conjugate, magnitude, phase, real, imaginary, square root, FFT, IFFT, natural logarithm, and exponential are provided, along with several routines tuned for specific measurement signal processing needs. For any given math operation the compiler generates one or more kCode frames. The kCode generated depends not only on the math operation but also on the operands. For example, computing the phase of a real vector requires filling the

destination vector with zeros while the phase of a complex vector is computed with an arc tangent operation.

Often the XXMeasSeqGen developer needs to take more control over the math operation, performing it only on selected elements of a vector. The compiler facilitates this with several levels of interface, each providing the user with a different level of control. The interface for the add operation is:

```
void v_add(Vector      *srcV1,
           Vector      *srcV2,
           Vector      *resV,
           long        *count)

void _v_add(Vector     *srcV1,
            Offset     *srcOffset1,
            Vecto      *srcV2,
            Offset     *srcOffset2,
            Vector     *resV,
            Offset     *resOffset,
            long       *count)

void __v_add(Vector    *srcV1,
             long      *src1IncP,
             Offset    *src1Offset,
             Vector    *srcV2,
             long      *src2IncP,
             Offset    *src2Offset,
             Vector    *resV,
             long      *resIncP,
             Offset    *resOffset,
             long      *count)
```

V_add() performs addition over the count elements of the vector. The user has no control other than count. _v_add() allows the user to begin the operation at an offset (potentially different for each operand). __v_add() allows the user to control the increment as well as the offset.

The compiler must often "coerce" the operands between real and complex numbers. The square root is a good example. The square root of a complex vector requires conversion to polar format, taking the square root of the magnitude, division of the phase by 2, and finally conversion back to rectangular coordinates. Since a real vector may contain negative values it must be coerced to a complex vector before undergoing a complex square root. The result, of course, is complex.

The compiler is responsible for allocating DSP-resident vectors. Since vector placement has a significant impact on measurement loop performance, the compiler must attempt to place as many vectors as possible in the high-speed DSP RAM. More important is that the most heavily used vectors are DSP-resident. The best job could be done by a two-pass compiler, but this would degrade user response time too much as well as add complexity to the design. A first come, first served approach was adopted with the enhancement of allowing the XXMeasSeqGen designer to force little-used vectors into the slower system RAM. The compiler groups the allocated vectors by segment so that at the onset of a compile, DSP memory belonging to all segments to be recompiled can be reclaimed.

A vector's headers contain information that is dependent on both the data and the math operations from which it is derived. For example, channel-specific information of a destination operand of a multiply is a combination of the channel information of both source operands. Another good example is the FFT, which requires that the data domain (which is part of the header information) be changed. For many of the trace vectors the XXMeasSeqGen designer can predict the correct header contents since the designer has complete control over all signal processing operations. However, for user math this is not true. The HP 894xxA kCode compiler takes a much safer approach, transforming the headers of destination operands based on the source operands and the operation. As part of each requested math operation the kCode compiler modifies the header of the destination operand according to a set of header tracking transformation rules.

The units of the vector data are critical to its interpretation, including units conversion and scaling for display. Like other vector header information, units are tracked by the compiler.

Measurement Loop Execution

The measSequencer state machine is shown in Fig. 7. The measurement loop itself begins in the WaitForData state and ends in LoopDone. Many of the states in this loop can be customized through kCode subroutines. The measSequencer remains in WaitForData until a complete scan has been acquired in the sample RAM. By the time Custom2 is exited the raw measurement data for this scan is available. The GetData, EnableDMA, and Custom1 states are designed to allow data for the first channel to be processed while data for the second channel is being transferred into timeDmaBuf (Fig. 2). Likewise, first-channel data for the next scan can be transferred during Custom2 and states that follow it. The measSequencer is not allowed to operate on trace vectors until it locks the VECTOR_RESOURCE semaphore in LockResources. The following two states, Custom3 and PostProcessMeasRes, perform all signal processing from measurement data on (see Fig. 1). Finally, in the Display state the measSequencer gives VECTOR_RESOURCE to the scanDone process via a remote procedure call. As described earlier, the scanDone process performs further processing to ready the trace data for display. To complete the synchronization with the measSequencer, scanDone will not unlock VECTOR_RESOURCE until trace data has been displayed.

One of the main differences between the averaged and non-averaged measSequencers is a direct transition from Custom2 to LoopDone, allowing trace-related processing to be bypassed in a fast average mode. Another is the transition from LoopDone to Halted to end an averaged measurement.

The main loop supports both free-run and triggered operation. An alternate path has been provided for single-sweep operation. By going to AwaitingArm from LoopDone the measSequencer is idled until the user arms the sweep.

The PreDataCltn1 state is provided to set up hardware before starting data acquisition. Generally the hardware has been set up properly by the activeMeas, but the scalar measurement must modify the LO frequency throughout the sweep. Before starting acquisition the StartDataCltn1 state waits for the hardware to settle.

Assault Handling

One of the most challenging aspects of measurement design is handling changes to the measurement setup in a graceful manner. It is not sufficient to recompile and restart a new measurement while destroying the existing measurement
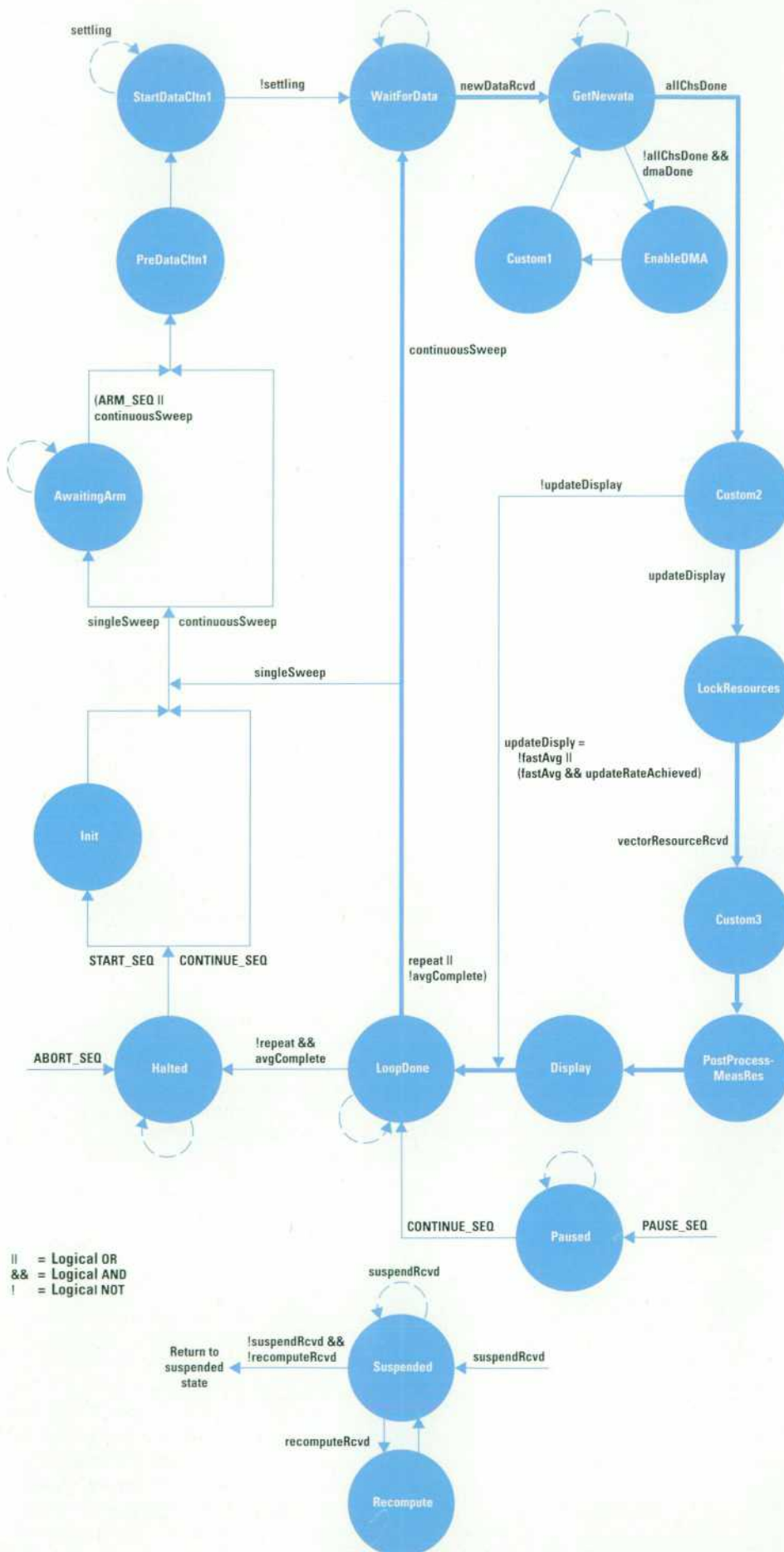
Fig. 7. MeasSequencer state machine. The heavier lines are the main loop.

II = Logical OR
&& = Logical AND
! = Logical NOT

Remote Debugging

One of the difficulties often encountered when developing software for embedded systems is a limited ability to debug the software. Frequently, this is because of lack of physical access to the CPU for traditional approaches such as logic analysis or processor emulation. Without full processor emulation with source-level debuggers and other productivity tools, the embedded system designer is often left with inefficient methods such as insertion of print statements, hardware tracing of memory accesses, or simply gestalt to isolate and fix complex software defects.

The software development team for the HP894xxA analyzers used a technique called remote debugging to aid in the development and debugging of the software. Remote debugging is the use of advanced debuggers running on a host computer to debug software that is running in an embedded system (or other computer) rather than on the host computer. Fig. 1 shows the concept.

Using the GNU C compiler (gcc) on a host workstation, an object code file is generated, which is then read by the GNU debugger (gdb). Using a communication line to the target hardware under development, gdb provides the workstation user debugging capabilities and insight into the running system.

The GNU debugger gdb provides capabilities such as stack backtraces, software breakpoints, printing C structures, disassembly, and other high-level features. It communicates with the target hardware by means of a simple ASCII protocol which is interpreted by a small kernel in the instrument itself.

Remote Debugging Protocol

The remote debugging protocol between gdb and the target instrument consists of nine basic commands:

Command	Function Response
g	Return the value of the CPU registers
G	Set the value of the CPU registers
mAA..AA,LLLL	Read LLLL bytes at address AA..AA
MAA..AA,LLLL:	Write LLLL bytes at address AA..AA
c	Resume at current address
cAA..AA	Continue at address AA..AA
s	Step one machine instruction
sAA..AA	Step one instruction from AA..AA
?	Return the current execution status

Each command is preceded by a header and has a checksum at the end to ensure data integrity. Since the debugger has detailed knowledge of the code running in the instrument, it can use the primitives above to acquire information for backtraces, printing variables, and change of execution control. For example, instead of providing a command in the protocol for setting a breakpoint, the debugger simply reads the value in memory where the breakpoint is to be placed and replaces the value with a trap instruction. When the breakpoint is reached, the debugger uses the write memory command to restore the original CPU instruction.

The use of this remote debugging capability in gdb allowed us to develop our firmware without the use of emulators. There are however, several classes of problems for which emulator or logic analysis still prove invaluable. For example, when a variable is being overwritten unexpectedly, hardware tools can be the quickest way of hunting down the problem. We've found that the use of a source-level debugger can significantly reduce the need for hardware solutions because the programmer is given a much better picture of the state of the system and how it got there.

In our environment we used RS-232 for the communication line. We wanted to be able to use this line for debugger communication as well as general input/output from printf() statements in the software. This was solved by using 8-bit ASCII on the RS-232 port and having all debugger traffic assert the most-significant data bit. RS-232 proved to be too slow for large data transfers such as the executable image so the HP-IB (IEEE 488, IEC 625) was used for high-speed data transfers.

In summary, remote debugging gave us the following advantages:
- Source-level debug. The use of source-level debugging contributed to a significant productivity gain.
- Lower development cost. The source-level debugger reduced our need for capital-intensive hardware debugging aids.
- Ability to debug during environmental testing. By simply plugging in an RS-232 connector, temperature-sensitive calibration problems could be quickly understood.
- Ability to debug an instrument without removing covers or boards. Hardware prototypes used by marketing or for documentation development could be interchanged with units used for software development with no changes.

Glenn R. Engel
Development Engineer
Lake Stevens Instrument Division

Fig. 1. Setup for remote debugging of embedded system software.

data. In most cases, a change must be reflected in the current measurement results. In the HP 894xxA this is particularly true since the measurement loop includes trace-related processing. These changes to the measSequencer are termed *assaults*, and the measSequencer and the activeMeas must both work closely together to handle them.

An assault to the measSequencer begins as a change request to the activeMeas. As part of its handling the activeMeas will likely modify measState and traceState settings used by the measSequencer. It may also recompute constant vectors such as those used in windowing and correcting the data. A compile often accompanies this change with a redefinition of the kCode subroutines and reallocation of DSP vectors. The measSequencer cannot be running when data critical to its proper operation is being changed. Therefore, one of the first steps of change handling is to suspend the measSequencer.

This is done by the activeMeas's MeasChangeHandler state machine (see Fig. 5). The measSequencer will hold off this assault until it has a consistent set of data in the raw and trace vectors. This means that it will not check for assaults from the time it leaves WaitForData until it enters LoopDone. The measSequencer begins to handle the assault by entering the suspended state and saving any raw DSP measurement vectors to their system RAM counterparts. It then signals the activeMeas that it is suspended. It will remain in this state until the activeMeas releases it after completely handling all changes to the measurement definition. If the change requires that existing data be interpreted under the new measurement setup the activeMeas will request the measSequencer to recompute the data. The recomputation uses the same kCode subroutines as the Custom3 and PostProcessMeasRes states. The extent of the recomputation is determined by the level of the recompile. Like the compile, as little recomputation as necessary is done. The measSequencer then signals the activeMeas that the recomputation has been completed. At this point, the activeMeas forces the new data to be redisplayed and then tells the measSequencer to resume processing. The measSequencer then returns to the state from which it was suspended and resumes operation. Finally, the measSequencer may, if directed by the activeMeas, be forced to restart acquisition (if hardware was changed, for example) or restart the entire measurement (for averaged measurement only).

Conclusions

Measurement design is a complicated process during which many decisions trading off performance for ease of development must be made. Many of the complications faced by the designer are compounded by the evolutionary nature of a measurement's feature definition. With careful design, a measurement architecture can be put in place that provides a foundation upon which multiple measurements can be implemented. Using a common measSequencer, the kCode compiler allows designers to customize the measurement loop quickly and easily with minimum attention to common features and protocol. Likewise, by using the object-oriented features of the C++ compiler, designers can develop classes that support the measurement loop by concentrating only on incremental changes to a foundational feature set. Although performance does suffer slightly, this approach yields rich dividends in development time by allowing multiple measurement designers to inherit a complete set of foundational features.

Baseband Vector Signal Analyzer Hardware Design

The HP 89410A combines superior front-end linearity and high-speed data conversion with powerful digital signal processing to provide advanced measurement capabilities. Extensive calibration, flexible triggering, and arbitrary source types provide the accuracy and versatility needed to make the sophisticated measurements required for complex signal analysis at RF information bandwidths.

by Manfred Bartz, Keith A. Bayern, Joseph R. Diederichs, and David F. Kelley

The HP 89410A vector signal analyzer provides an array of new capabilities to meet the emerging measurement requirements of complex signals that require simultaneous analysis in the time, frequency, and modulation domains. It makes measurements with resolution bandwidths as low as one millihertz and frequency spans as wide as 10 MHz to accommodate the wide information bandwidths of complex and frequency-agile communications signals. Its user interface is familiar to users of traditional swept spectrum analyzers, making it easy to use. With an optional second channel, a versatile source with arbitrary source types, flexible triggering, a built-in disk drive, and a variety of software and hardware options, the HP 89410A provides the upgrade paths necessary to accommodate the most sophisticated user needs.

The HP 89410A measurement engine is based upon powerful digital signal processing technologies that provide its speed and flexibility. The key elements of the hardware that support the HP 89410A's high performance include:
- An exceptionally linear front end with high input impedance capability, input protection, autoranging, and anti-alias protection.
- A high-speed, 25.6-MSa/s (million samples per second), wide-dynamic-range analog-to-digital converter (ADC) that employs a proprietary implementation of large-scale dithering to provide superior linearity
- A custom 25.6-MSa/s digital local oscillator and decimating filter chipset that provide up to 23 bits of effective resolution
- Powerful floating-point signal processing using the Motorola DSP96002, which delivers up to 48-MFLOP (million floating-point operations per second) peak performance
- Dedicated display processing using the Texas Instruments TMS34020 graphics system processor, providing up to 60 display updates per second to a color display
- A versatile signal source that provides sine, chirp, random noise, and arbitrary signal types
- Full calibration to provide superior accuracy and signal fidelity
- Flexible triggering with both pretrigger and post-trigger delay and arm delay
- A backplane with slots to accommodate hardware upgrade options.

The block diagram, Fig. 1, depicts the data flow path of the HP 89410A baseband vector signal analyzer. (Note that the HP 89410A and the HP 89440A IF section are identical. All further discussion will refer to the HP 89410A but applies equally to the HP 89440A IF section.) The signal input at the analog front end is amplified and alias protected. Following the analog signal conditioning, the signal undergoes analog-to-digital conversion. The digital data, which is sampled at 25.6 MSa/s, is routed to a digital switch assembly and on to a proprietary digital LO and digital decimating filter chipset. The digital LO and decimating filters perform frequency selective band translation. The translated data is then stored in a sample RAM with a capacity of 32K samples per channel, which can be optionally increased to 512K. From the sample RAM, the captured data is bused to the DSP96002 processor and subjected to corrections, windows, FFT algorithms, and other math operations. The DSP also formats the data for viewing and writes the data to the display system. In the final step of the data path, the display processor, the TMS34020 GSP, constructs a trace from the formatted data and presents it to the user on an internal 7.5-inch color CRT. By using the high-speed components in the data path, the HP 89410A's signal processing hardware can process approximately 300 512-point complex spectra per second with display update rates reaching 60 per second.

The data flow path for the source follows a similar sequence leading from the digital processing to the output of the analog source. In contrast to the receiver path, the digital LO and filters operate in the reverse mode in the source path, performing the necessary operations of data interpolation and frequency translation. The interpolated data is converted by the source digital-to-analog converter (DAC) at 25.6 MSa/s. A reconstruction filter and further conditioning circuitry prepare the analog signal to be output on the HP 89410A's front panel. The digital switch assembly between the front-end ADC and the digital local oscillator (LO) and decimating filters allows the digital source to be connected to the receiver data path for diagnostic purposes. Full analog calibration is accomplished by placing the internal calibration source path between the analog source and the two input channels.
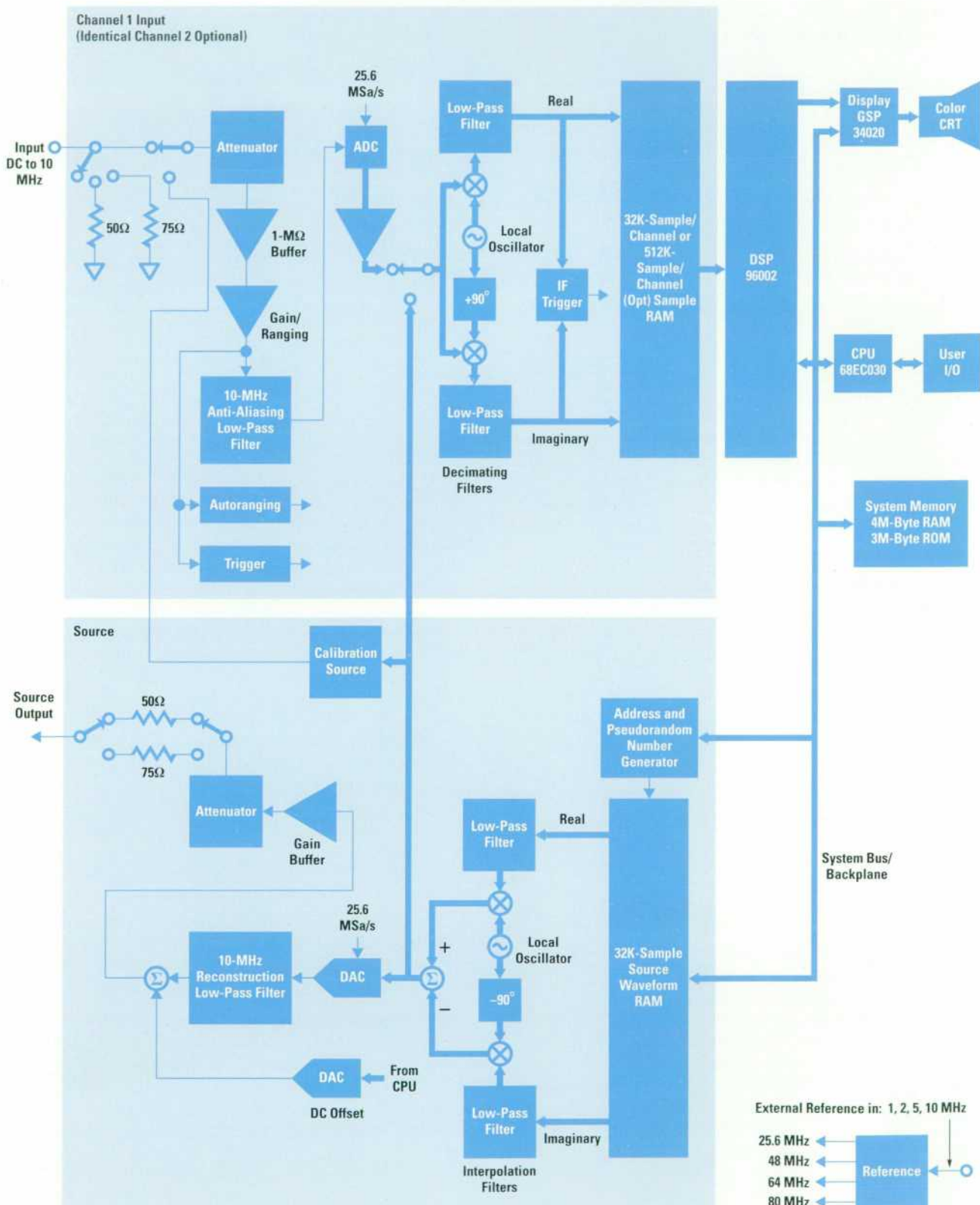
Fig. 1. Block diagram of the HP 89410A vector signal analyzer.

Analog Input

The analog input provides the interface between the signal to be measured and the instrument's analog-to-digital converter. Impedance matching, ranging, and anti-alias filtering must be accomplished without compromising signal fidelity. The input must be robust in the face of real-world signals

and inadvertent abuse. One input channel is standard, and a second is optional.

The input is single-ended with the low side at chassis ground potential. 50-ohm, 75-ohm, and 1-megohm terminations are provided. The specified return loss is 25 dB for the 50-ohm

termination and 20 dB for the 75-ohm. The 1-megohm input is specified at 2% accuracy and less than 80 pF of shunt capacitance.

Input ranges are provided in 2-dB steps. For the 50-ohm termination, input ranges extend from –30 dBm (10.0 mV peak) to +24 dBm (5.01V peak). The 75-ohm ranges extend over the same set of voltages as the 50-ohm ranges, but are numerically smaller by 1.761 dB. The 1-megohm ranges also start at –30 dBm (10.0 mV peak), but extend to +28 dBm (7.94V peak), dBm here implying a 50-ohm system reference.

A selectable anti-alias filter is provided. Alias protection is specified at 80 dB and is typically better than 95 dB.

Input Cable and RFI Suppression

The input BNC connector shell is dc isolated at the front panel (see Fig. 2). This is done to reduce spurious inputs caused by the instrument's switching power supply, display, and power line circuitry, preventing induced currents from flowing across the input cable shield. Dc isolating the connector shell necessitates ac bypassing it to chassis ground at the front panel to reduce radio frequency interference (RFI) emissions, which otherwise use the cable shield and connector shell as an exit path.

Ac bypassing is done with a small printed circuit board behind the front panel. Isolated areas around the connectors make contact with the connector shells via the connector

mounting hardware. These isolated areas are then bridged by surface mount capacitors to the ground areas of the board in contact with the front subpanel.

In the prototype, these isolated areas were bypassed to ground by a group of three surface mount capacitors, all on one side of the connectors. This design inadequately suppressed RFI in the 900-MHz range. These RFI emissions originated from the fast data buses in the instrument, which are strong sources of high-frequency energy. Network analyzer s_{11} measurements showed that this structure was indeed resonant at around 900 MHz and therefore was failing to suppress RFI there. The production revision connector bypass board has four surface mount capacitors placed evenly around and as close as possible to each connector. 0.047-µF capacitors are used, although the particular value doesn't seem to be very important. All emissions up to the test limit frequency of 1 GHz are effectively eliminated from the input connectors. Network analyzer measurements on the new design show that the resonance has been pushed out to beyond 2 GHz.

Each input cable is passed through five high-permeability ferrite toroids at the front-panel end. These serve multiple purposes, including forming a low-pass structure with the connector capacitors to minimize RFI, reducing spurious inputs, and minimizing the effects of measurement ground loops.
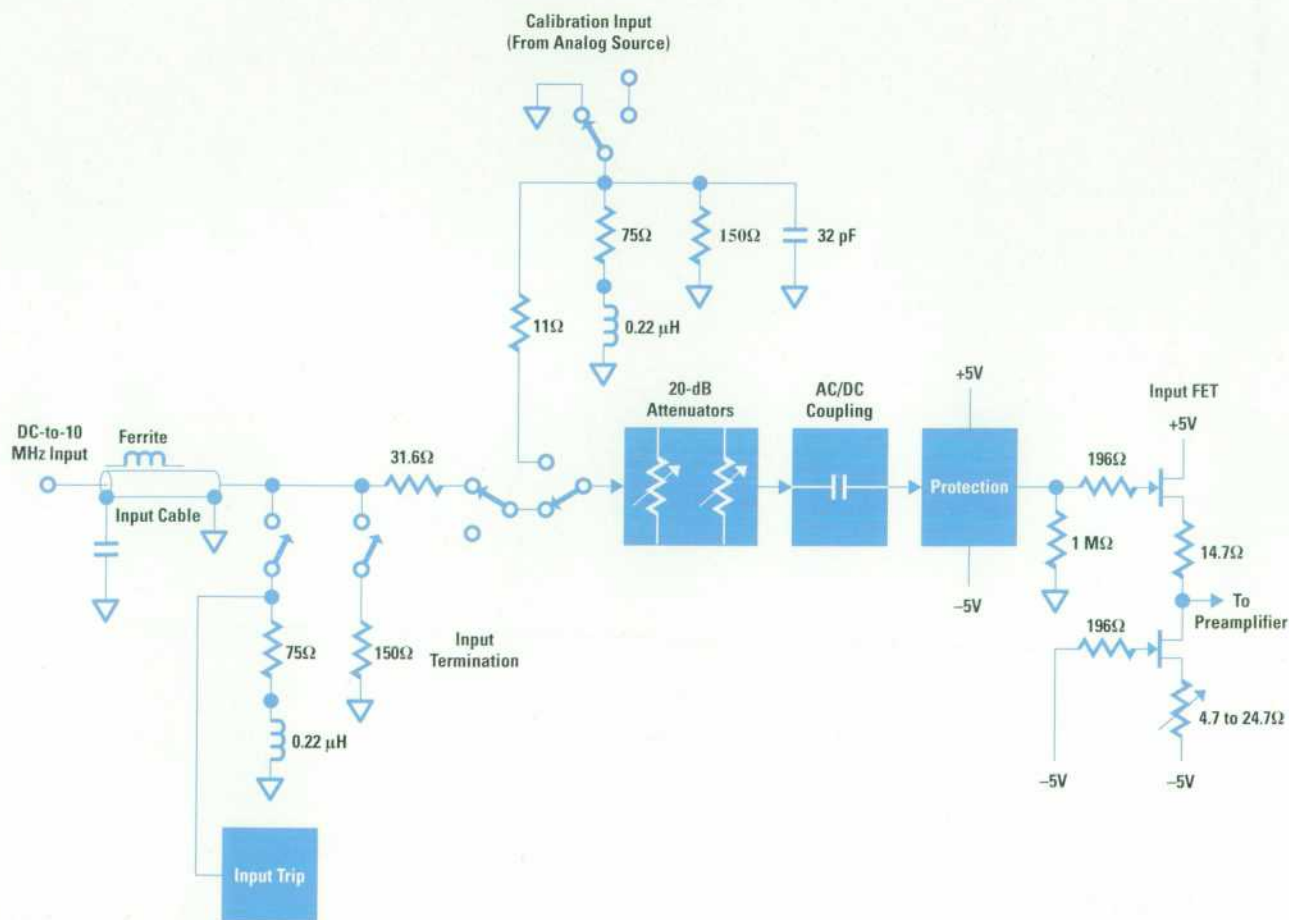
Fig. 2. Block diagram of the analog input front end.

Input Termination

The input is fundamentally a 1-megohm structure. 75-ohm and 50-ohm terminations are created by shunting the 1-megohm input with 75 ohms or 75 ohms in parallel with 150 ohms to make 50 ohms. However, by itself, this design produces unacceptable return loss, especially for 75 ohms, because of the 75-pF typical shunt capacitance across the 1-megohm input. At 10 MHz, 75 ohms in parallel with 75 pF gives only about a 15-dB return loss.

This return loss problem is solved with a small inductance in series with the termination resistor, which serves to tune out some of the effect of the shunt capacitance within the instrument's 10-MHz frequency range. 0.22 μH is used in series with the 75-ohm shunt termination. At 10 MHz, input return loss is improved to typically 23 dB for 75 ohms and better than 30 dB for 50 ohms.

Nothing is free of course, and one disadvantage of this tuning technique is worse return loss performance beyond 20 MHz, well beyond the instrument's frequency range. The other disadvantage, if this technique is pushed too far, is exaggerated frequency response differences among the 50-ohm, 75-ohm, and 1-megohm terminations. Since calibration does not account for the different terminations independently, these differences must be well-contained. With 75 pF of input shunt capacitance, 23 dB of 75-ohm return loss is achieved with only minimal (on the order of 0.01 dB) differences between termination frequency responses.

Input Trip

Active input trip circuitry detects excessive voltage on the 50-ohm and 75-ohm input terminations. Diode-capacitor peak detectors capture peak positive and negative voltages. Comparators then react, causing the logic on the input board to open the termination and protection relays and to alert the instrument CPU that the input is tripped. The trip point is nominally ±7.2V, or about +27 dBm, constrained by the dissipation capability of the terminations.

It is important that the detection diodes be reverse-biased for all normal input signal levels. The diodes only become forward-biased at signal levels approaching overload, preventing nonlinear conduction currents from being drawn from the input and causing distortion.

Calibration

The calibration signal is introduced into the input after the protection relay. This is less than optimum in the sense that, ideally, the calibration signal would replace the input signal for as much of the input circuitry as possible. With this approach, the input terminations are not included in the calibrated input circuitry.

The principal advantage of this approach is that the input terminations maintain their state relative to the input connector during calibration. This minimizes disruptions to the system connected to the instrument input. Also, the protection circuitry associated with the input terminations remains in effect during calibration. In the contrasting design in which the input terminations are included in the calibration, a temporary termination must be switched across the input to maintain the load on the system being measured. This is virtually impossible to do without a transient disruption of

the load, to which some systems are sensitive. Also, protection of the temporary termination is a problem that may require duplicating the trip circuitry.

Since the input terminations are not calibrated, the calibration termination must mimic the input accurately. Fortunately, this can be done very well up to 10 MHz. The resistor and return loss tuning inductor structure of the input termination is copied exactly, with a small empirically determined lumped capacitance added to model the capacitance of the input up to the calibration signal entry point. A transmission line detuning series resistor immediately following the calibration termination is also empirically determined, and models the series resistor following the input termination. These resistors are the subject of the following section.

Wide Bandwidth in a High-Impedance System

One of the best ways to achieve good linearity and distortion performance over a given bandwidth is to design so that the bandwidth of interest is a relatively small part of the overall bandwidth of the system. Distortion often increases dramatically at frequencies near the bandwidth limit of a circuit. Thus, a low-distortion design may reasonably employ circuits with ten times the intended application bandwidth.

Such is the case with the HP 89410A input. Designed to be used to 10 MHz, some of its constituent blocks have bandwidths on the order of 100 MHz or more. However, with some signal runs on the input board on the order of six inches in length, transmission line effects cannot be ignored.

The normal solution is to use a doubly terminated transmission line design. This design calls for carefully controlled transmission line impedances matched on both the source and receiving ends. However, two factors preclude the use of doubly terminated transmissions lines in this application. One is the 6-dB signal loss at the source-to-line divider, and the other is the inability to source the required current to drive matched 50-ohm or even 100-ohm transmission lines and maintain distortion performance.

Instead, a hybrid structure is used. Long signal runs have a small series back-match resistor at the driving end. Some runs also have capacitively coupled termination resistors at the receiving end. These circuit elements control the peaking in frequency response that would otherwise occur, but do not cause a large signal loss or draw large currents that would adversely affect distortion performance.

Attenuators and Input FET

The first elements in the ranging architecture of the input are a pair of 20-dB attenuators. These are designed for a 1-megohm system (presenting 1 megohm to the preceding circuitry when loaded with 1 megohm).

The attenuators transition from resistive dividers to capacitive dividers in the tens-of-kilohertz region. Each attenuator requires a tunable-capacitor flatness adjustment. The first attenuator also has an input capacitance adjustment to balance the input capacitance for the attenuated and nonattenuated settings.

Ac coupling is provided by a 0.1-μF capacitor, giving a low-frequency roll-off at nominally 1.6 Hz. The ac coupled and dc coupled paths are carefully balanced for capacitance to

ground. Without this careful balance, a frequency response difference would exist between the two paths at high frequencies, unaccounted for by calibration.

The input FET is the only discrete amplifier stage in the input. A source follower with current-source bias is constructed from a matched pair of JFETs in a single package. I_{DSS} (drain current at zero gate-to-source voltage) matching between the two devices to ±5% requires an adjustment to the current source to obtain 0Vdc across the stage. Signal gain is nominally about –1 dB for the stage.

The FET device used, a special high-transconductance, high-I_{DSS} type, is capable of maintaining distortion performance well in excess of 90 dB at 10 MHz with the –12 dBm signal levels found at this point in the circuit.

The source follower configuration is susceptible to oscillations if driven from a reactive source impedance. Gate-stopper resistors (small-value resistors placed right at the FET gate) ensure that the stage is stable for any source impedance presented at the input. 196 ohms was chosen because models showed it absolutely guaranteed stability while negligibly affecting the overall noise figure for the input, which is dominated by the preamplifier.

Preamplifier

Wide-bandwidth current-feedback operational amplifiers are used for all of the remaining signal amplification tasks in the input circuit (see Fig. 3). These very high-speed, high-fidelity devices allow amplification stages to be built with very wide bandwidth, relatively independent of the stage gain, because of their current-feedback topology.

This gain-bandwidth independence is exploited in the variable-gain preamp stage, which can be set to gains of 13.4 dB or 3.4 dB, thus giving the functionality of a 10-dB pad in the circuit for ranging. For the small input ranges, it is important to achieve as much signal gain in this stage as possible to have the best noise performance. The 13.4-dB gain setting is used, providing about a 17-dB noise figure for the stage. This dominates the sensitivity performance of the input. Large gain leads to a quiet preamp stage because the dominant noise mechanism for these amplifiers is inverting input current noise, which generates an output noise for the stage that is independent of gain setting. Referred back to the input, this

fixed output noise becomes a smaller equivalent input noise with larger gain.

Signal Return Ground

A special signal return ground is used for the input cable through the preamp stages. This ground is connected back to the main chassis ground at only one point: the point where the input cable shield comes onboard. This is required for reasons similar to those prompting the input connectors to be dc isolated at the front panel, that is, currents generated by other mechanisms in the instrument flow across the board ground planes in this area, generating small but significant spurious voltage drops that cannot be allowed to add to the input signal. Bringing all connections to the signal return ground back to a single chassis ground connection point prevents other currents from flowing across the signal return ground, generating spurious voltages.

After the preamp, the input signal is large enough not to be adversely effected by spurious ground currents, and the board ground planes are used for signal return. The inclusion of this separate signal return ground results in approximately a 10-dB reduction in the level of switching power supply related spurious signals, adding enough specification margin to guarantee producibility.

DC Offset

A 12-bit digital-to-analog converter sums dc into the dc offset stage, which is a high-speed current-feedback operational amplifier. This is done under control of the instrument CPU and autozero software routines, and compensates for dc offset all the way to the analog-to-digital converter. Care is taken not to introduce noise from the digital side and to maintain frequency response flatness in the summing stage.

Switchable Gain Amplifiers

The remainder of the ranging is accomplished in four amplification stages together called the switchable gain amplifiers. These stages have gains of 8 dB, 6 dB, 4 dB, and 2 dB, and can each be switched into the circuit or bypassed for a 0-dB gain for the stage. Again, high-speed current-feedback operational amplifiers are employed.

To maintain distortion performance, these stages are only used one at a time. Using one or both 20-dB attenuators, the
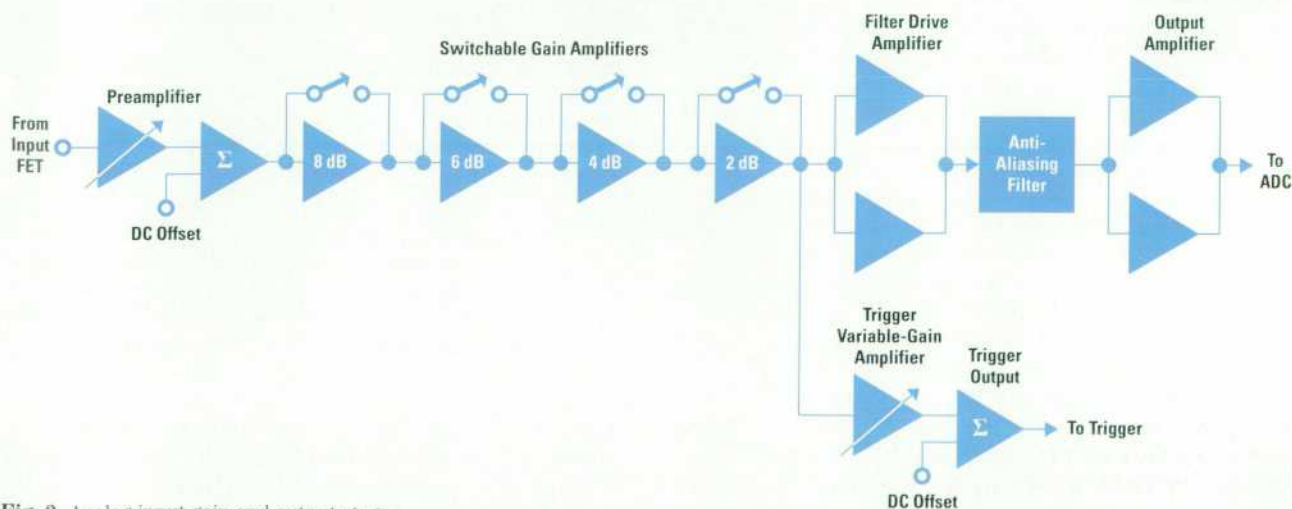
Fig. 3. Analog input gain and output stages.

10-dB variable-gain preamp, and one or none of the 8-dB, 6-dB, 4-dB, and 2-dB stages, ranging in 2-dB steps is achieved from −30 dBm to +28 dBm.

Trigger Output

At the output of the switchable gain amplifiers, ranging is complete. At this point, an input signal at full scale on any range is at a fixed size of about −5 dBm with as much of the input signal's bandwidth preserved as possible.

From this point, the input signal takes two paths. One path leads to the trigger board and the other to the analog-to-digital converter. The trigger circuitry demands a relatively large signal since high-speed comparators with fixed thresholds are used. Fortunately, signal fidelity demands are relaxed here.

The trigger signal first passes through another variable-gain stage very similar to the preamp. The gain is set according to the dither mode used by the analog-to-digital converter. For the half-scale dither mode used by the standalone HP 89410A, the higher 12-dB gain setting of the trigger variable-gain amplifier stage is used, raising the trigger signal level to +7 dBm. For the quarter-scale dither mode used by the HP 89440A RF vector signal analyzer, the lower 6-dB gain setting is used.

The final trigger output buffer serves to further isolate the input circuitry from the harsh environment of the trigger board and sums in the output of an 8-bit digital-to-analog converter for dc offset adjustment on the trigger signal. Depending on the range, on the order of 50 MHz of 3-dB bandwidth is preserved from the input connector to the trigger output.

The trigger board provides adjustable-level triggering, over-range detection, and half-range detection. The half-range detection circuitry in particular must be isolated from the input circuits, since for signals above approximately 6 dB below the range, the half-range comparators generate square waves at the input signal frequency, providing a strong source of odd harmonic energy, which could cause distortion if allowed to couple back into the input.

Filter Drivers and Anti-Alias Filter

High-speed current-feedback operational amplifiers in an inverting parallel configuration drive the anti-alias filter. The parallel amplifiers' outputs are summed into the filter input through 100-ohm resistors, presenting a 50-ohm source impedance to the filter.

The anti-alias filter is a nine-pole, eight-zero elliptical design. The filter corner is at approximately 10.2 MHz. The stop-band edge is at approximately 15.4 MHz. (Note that 15.6 MHz is the lowest frequency that can alias back into the 10-MHz band given the 25.6-MHz sample rate of the analog-to-digital conversion.) The overall input frequency response, including the effects of the filter passband, is flat within less than 1 dB, which is easily corrected by calibration. Stop-band attenuation is typically better than 95 dB for all frequencies beyond the stop-band edge.

The filter is driven from 50 ohms and is terminated in 250 ohms. To maintain signal fidelity, special high-linearity inductors are employed.

A filter bypass path can be selected if anti-alias filtering is not desired. This path is switched at both ends by high-isolation relays to prevent signal leakage in this path from decreasing the stop-band attenuation of the filtered path.

Output Amplifier

Like the filter driver, the output amplifier is a parallel configuration of a pair of high-speed current-feedback operational amplifiers. These together provide a 50-ohm output impedance driving the 50-ohm SMB cable path to the analog-to-digital converter. This path is terminated in 500 ohms on the analog-to-digital converter end. The 50-ohm back-match on the drive end absorbs the reflection from the near open-circuit termination and prevents standing waves. This termination system provides minimum dc loading, improving linearity and distortion performance.

Large-Scale-Dithered ADC

The ADC is one of the key enabling technologies for the development of a wide-information-bandwidth, wide-dynamic-range vector signal analyzer. As the critical block that bridges the analog and digital worlds, the ADC dominates several important instrument specifications such as the widest information bandwidth and the achievable noise and distortion dynamic range. Recent trends in commercial ADCs make it possible to use digital signal processing techniques at RF information bandwidths. However, the linearity limitations of commercial converters make them unsuitable for the precision measurements that many traditional analyzer customers require. Therefore, the ADCs for analyzer applications have customarily been developed in-house using proprietary techniques such as dithering to achieve superior linearity and signal-to-noise ratio (SNR) at higher sample rates.

To avoid the long development times associated with the design of a fully custom ADC, the HP 89410A converter design effort leveraged recent trends in commercially available high-speed ADCs through an informal parallel development effort with an external ADC vendor. This fast-track approach used the highest-speed commercially available 12-bit ADC under development as an embedded component in an HP large-scale-dithered circuit design being developed in parallel. The resulting converter has unprecedented linearity performance for high-speed converters, provides sample rates to 25.6 MSa/s, and achieves the necessary dynamic range for wideband vector signal analyzer instrumentation applications.

HP has considerable experience in the development of custom ADC architectures that use small-scale dithering. The benefits of small-scale dithering for achieving improved spurious performance are well-known. Small-scale dithering provides a means for reducing high-order spurious mechanisms. Dithering in conjunction with subsequent digital filtering makes it practical to extract signals far below the least-significant bit (LSB) step size of the converter.

Large-scale dithering not only reduces high-order distortion products but also seeks to achieve a significant improvement in the low-order distortion performance as well. This comes at the cost of some signal overhead associated with the level of dither signal applied. The HP 89410A ADC is the first known practical application by HP of large-scale dithering as
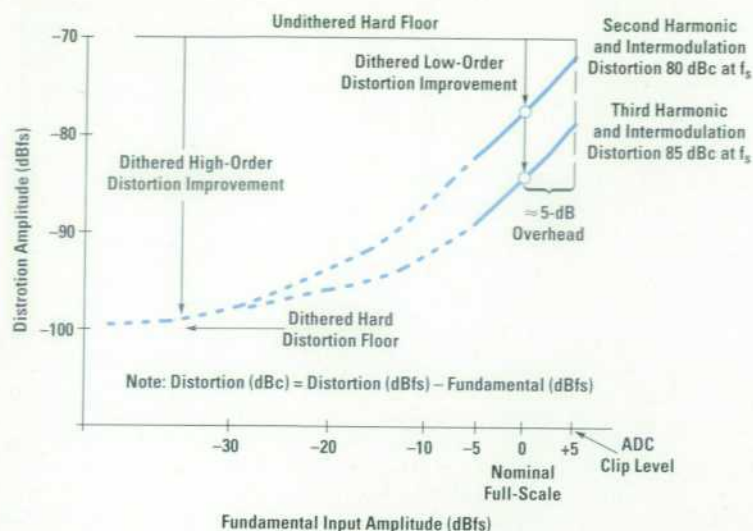
Fig. 4. ADC distortion dynamic range for half-scale dithered operation.

an external enhancement to an embedded ADC component. The goal was superior low-order linearity at high conversion rates.

The converter has two modes of operation: quarter-scale and half-scale dither levels relative to the converter's full-scale level. The two modes are optimized for the HP 89440A and HP 89410A, respectively. Fig. 4 depicts the resulting low-order and high-order distortion performance for the HP 89410A ADC at half-scale dithered operation. At nominal full scale, the HP 89410A ADC provides typically 80-dBc second-order and 85-dBc third-order distortion dynamic range with 5 dB of overhead. The high-order distortion associated with the hard floor, which is fixed for undithered converters, is significantly lowered, allowing "soft distortion" behavior at input signal levels other than nominal full scale (see "ADC Bits, Distortion, and Dynamic Range" on page 38 for a discussion of the hard floor and hard and soft distortion). Because of the soft distortion characteristic of the dithered converter, the performance measured in dBc remains nearly constant over 10 dB of input level range. This flexibility in the nominal operating point allowed better optimization of the overall instrument distortion performance. The unused overhead in the HP 89410A is available to users for higher-SNR measurements.

In contrast to small-scale dither, the application of large-scale dither poses some significant design challenges. The high-speed circuitry associated with dithering must exhibit extremely fast settling times. The large signal swings associated with the large-scale dither levels must be settled well within the 39.06-ns conversion time to avoid compromising the overall converter signal-to-noise ratio. To provide settling errors approximately 70 dB down requires the analog circuitry to have a settling time constant of 5 ns. Moreover, the close proximity of high-speed switching components to various precision linear devices and wideband amplifiers necessitated a careful RF layout. Particular attention was paid to grounding and placement to minimize high-speed reflections and achieve the required RF isolation.

ADC Block Diagram

Fig. 5 shows the block diagram of the large-scale dithered ADC. The dither signal consists of random noise, which must be uncorrelated with the input signal. A random 12-bit

sequence is generated by a pseudorandom number generator with a period of $2^{39}-1$ samples. At a sample rate of 25.6 MSa/s this yields a period of about six hours, ensuring that the periodicity of the dither is well below any measurement frequency of interest. Adjacent dither samples are designed to be highly uncorrelated to ensure that the power spectral density of the random noise is flat with frequency.

The sampled random noise sequence follows two signal paths in the block diagram. On the analog input side of the block diagram the sequence is converted to an analog random noise representation by the high-speed dither DAC, which runs at the sample rate of the converter. The analog representation of the dither signal is then combined with the analog input signal at a wideband summing junction implemented with very wideband, low-distortion operational amplifiers. The combined signal consisting of the input signal plus added random noise dither is applied to the track-and-hold input of the analog-to-digital conversion block.
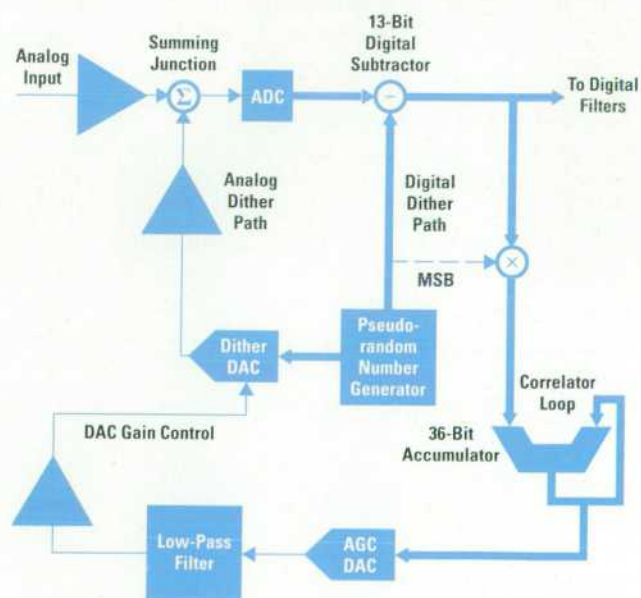
Fig. 5. Block diagram of the large-scale-dithered analog-to-digital converter.

ADC Bits, Distortion, and Dynamic Range

The number of output bits is often regarded as an accurate indication of the dynamic range of an analog-to-digital converter. In fact, this measure can be quite misleading. As ADCs trend toward more bits at higher speeds, static measures of converter performance such as integral and differential nonlinearity and the number of bits are giving way to dynamic measures such as signal-to-noise ratio and distortion dynamic range. Such dynamic measures often are far more useful in evaluating converter performance, particularly in applications where converters in conjunction with digital signal processing are replacing traditionally analog implementations. To help users better interpret the dynamic performance requirements for their applications the specifications of the HP 89410A and HP 89440A are written in terms of dynamic measures such as signal-to-noise ratio and distortion dynamic range instead of bits.

Often the 6-dB-per-bit rule is invoked in estimating the dynamic range potential of an ADC. For many converter architectures this may be an oversimplification. An obvious example is a sigma-delta converter, which may only use a single bit and by oversampling techniques achieve up to 18 effective bits of signal-to-noise dynamic range. Moreover, converters with large numbers of bits at the output may in fact suffer from inherent noise limitations that limit their performance to fewer effective bits.

A related misconception is to presume that a converter's ability to extract signals is limited by its resolution or LSB step size. For ADCs with dithered architectures followed by digital filtering the resolution is typically not limited to the converter step size. The dithering, which randomizes the quantization of the converter, works in conjunction with the inherent time averaging of digital filters to allow the extraction of signals far below the step size of the converter. For smaller effective resolution bandwidths of the digital filters, the resultant greater processing gains provide higher resolution limited only by the accuracy of the digital filters.

This concept is analogous to the reduction in noise power that occurs for narrower resolution bandwidths of analog IF filters. The digital filters in the HP 89410A and HP 89440A provide up to 23 bits of resolution corresponding to the narrowest resolution bandwidth of 1 mHz. This means the dithered ADC and digital filters can resolve signals to −140 dBc (dB relative to the fundamental amplitude). On the lowest input range setting of −30 dBm this corresponds to −170 dBm of sensitivity.

Similar arguments can be made for evaluating the distortion dynamic range of a converter merely on the basis of bits at the converter output. Often the actual linearity may be far worse than the 6-dB-per-bit rule would imply. By contrast, for dithered architectures the linearity typically far exceeds the linearity suggested by the number of bits (see "What is Dithering?" on page 44).

Hard Distortion Floor

Another distinction must be made between high-order and low-order distortion mechanisms, whose characterizations can be a source of confusion in dithered converters. Most undithered ADCs suffer from a hard floor limitation associated with the high-order distortion products generated by the staircase transfer function. This hard floor has the characteristic that as the fundamental input signal amplitude is lowered, the amplitudes of the distortion products remain relatively fixed. Most ADC distortion specifications are written in terms of dB spurious because the distortion performance is dominated by higher-order spurious mechanisms. These spurious products remain fixed in amplitude as the fundamental amplitude is lowered, thereby reducing the dynamic range in dBc.

This is in contrast to analog components and dithered converters, whose distortion is usually dominated by low-order mechanisms. Dithered converters exhibit low-order or "soft distortion" behavior because the reduction of high-order mechanisms significantly lowers the hard floor, allowing the low-order distortion mechanisms to dominate. As the signal amplitude is lowered, low-order distortion product amplitudes also decrease relative to full scale. The amplitude at which the distortion products no longer continue to drop and the dynamic range plateaus is called the hard distortion floor. This floor is often measured in terms of dBfs relative to the full-scale level of the converter and may often be referred to as the linearity of the converter. This is depicted in Fig. 1. Undithered, the HP 89410A hard floor would be around −68 dBfs. Dithering reduces the actual hard floor to approximately 95 to 100 dBfs.

Low-order distortion products are usually measured in dBc (dB relative to the fundamental signal amplitude). Because the distortion performance of the HP 89410A converter is dominated by low-order mechanisms the distortion is specified in dBc at nominal full scale.

Distortion order originally derives from the order of the term in the polynomial expansion of the transfer function. It also dictates the amplitude and frequency behavior of a particular distortion product. A dBc specification gives distortion performance relative to the carrier for a given amplitude. The traditional rule for determining the distortion amplitude behavior relative to the fundamental is 2 dB of second-order distortion amplitude reduction per dB of fundamental amplitude reduction, 3 dB per dB for third-order, and so on. First-order distortion changes by 1 dB per dB of reduction and therefore the dBc specification remains constant with changing fundamental amplitude. Because the dBc specification is dependent on the absolute level of the fundamental, many amplifiers and mixers are specified in terms of intercept point, which is the theoretical signal amplitude at which the fundamental and the particular distortion term are equal.

The order of a distortion product also dictates the frequency behavior of the distortion product in relation to the fundamental. A 1-Hz shift in fundamental frequency causes second-order distortion products to shift by 2 Hz, third-order by 3 Hz, and so on. Examples of third-order distortion include third-harmonic distortion or third-order intermodulation.
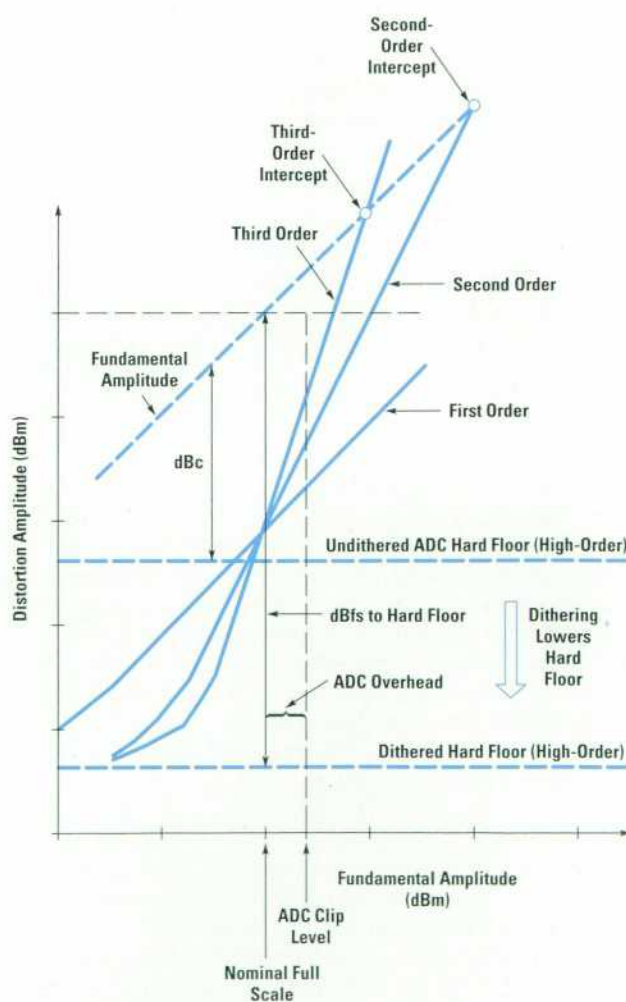
Fig. 1. Behavior of various orders of low-order distortion.
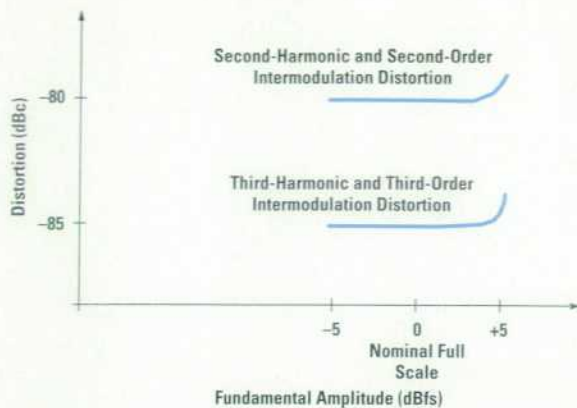
Fig. 2. Specified low-order distortion of the HP 89410A ADC.

Although it is generally true that both the frequency and the amplitude orders of behavior are consistent for a given distortion product, this is not always the case when a discontinuity is involved. An example is crossover distortion in a class B amplifier output stage; the second harmonic may not exhibit second-order amplitude behavior. Similar effects are seen in ADCs, which are fundamentally high-order with numerous discontinuities associated with the staircase transfer function. Dithering techniques seek to reduce or eliminate the effect of discontinuities.

Fig. 1 illustrates the difference between hard and soft distortion in terms of dBc, and shows the behavior of several orders of soft (low-order) distortion. The soft distortion characteristic for the HP 89410A ADC predominantly exhibits first-order amplitude behavior and is specified as shown in Fig. 2 for both second-harmonic and third-harmonic distortion and second-order and third-order intermodulation.

Manfred Bartz
Customer Support Engineering Manager
Lake Stevens Instrument Division

The digital path of the pseudorandom noise generator is routed to a high-speed 13-bit digital subtractor immediately following the analog-to-digital conversion. The subtractor consists of 4-bit pipelined adders with lookahead carry. It subtracts corresponding values of dither on a per-sample basis from the converted digital representation of the input signal plus dither. The appropriate number of delays in the digital path of the converter ensure that corresponding values of dither are correctly subtracted, thereby removing the dither and leaving only the digital representation of the original input signal.

In practice, it is a design challenge to ensure that the subtraction step is performed exactly so that no residual unsubtracted pseudorandom noise degrades the overall ADC signal-to-noise dynamic range. Several mechanisms can contribute to the converter noise, such as inaccurate subtraction of the dither, the settling time of the rapidly changing dither signal in analog components, and feedthrough of the dither signal. In addition, distortion errors in the dither DAC manifest themselves as dither subtraction errors which detract from the signal-to-noise ratio. Settling times associated with the analog components such as the dither DAC, the ADC track-and-hold circuit, and the wideband amplifiers must be minimized to ensure that their noise contributions are below the converter's own noise. Dither feedthrough is minimized by designing for high isolation. Fig. 6 shows a Pareto chart of the noise floor components in the large-scale-dithered ADC. The individual noise mechanisms were reduced to below the inherent noise of the vendor-supplied ADC component, resulting in a signal-to-noise ratio of 127 dBc/Hz. For half-scale dithered mode, the signal-to-noise ratio is 124 dBc/Hz, taking the additional 3 dB of dither overhead into account.

ADC Correlator

Small-scale dithered converters can often perform the subtraction step with minimal noise penalty because the dither signal levels are on the order of a few LSBs. Some small-scale dithered converters ignore the subtraction step entirely because any residual noise left by subtraction errors is negligible. For large-scale dither, the precision of the subtraction step becomes critical, because the large signal levels employed often exceed the magnitude of the input signal.

Accurate subtraction is achieved by using a digital correlator/accumulator as part of a low-frequency feedback loop to generate a feedback error signal. The feedback error signal is converted into an analog error signal by the automatic gain control (AGC) DAC. The analog error signal adjusts the amplitude of the dither signal at the output of the dither DAC to ensure exact subtraction.

The operation of the correlator can be described mathematically. At the output of the ADC, the digital representation of the kth sample of the input signal plus dither can be described as:

$$Aout_k = D_k(1 + e) + S_k,$$

where D_k is the kth dither value, S_k is the kth input signal, and e represents the dither gain error. Following the digital subtraction of the dither, the digital output signal consists of the digitized version of the input signal plus the residual dither errors:

$$Dout_k = D_k e + S_k.$$
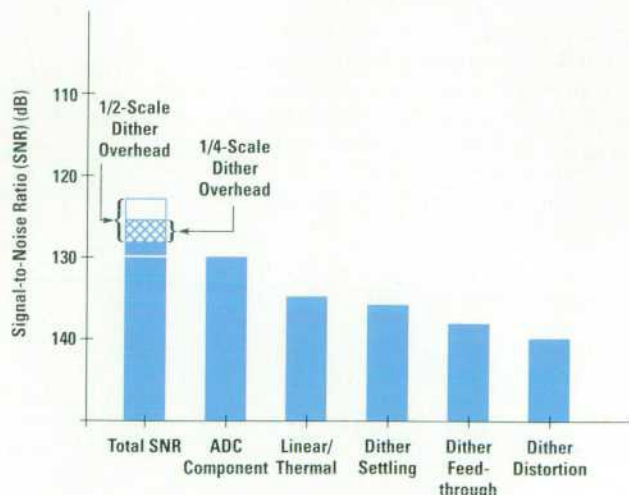
Fig. 6. Pareto chart of individual dither noise mechanisms.

(a)

(b)
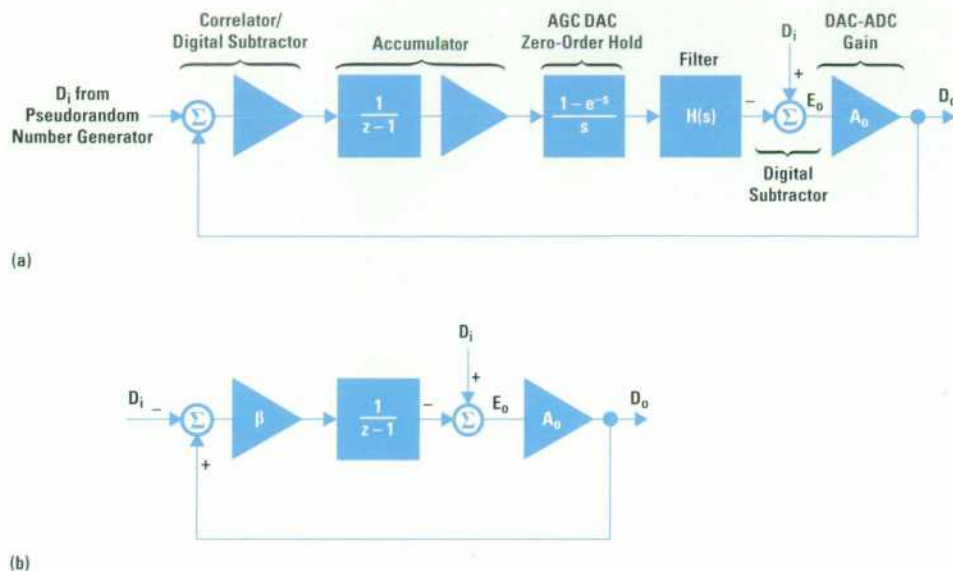
Fig. 7. (a) Simplified diagram of the mixed analog/digital correlator loop. (b) z-domain simplification.

The correlator multiplies each output sample by the corresponding dither sample, yielding:

$$\mathrm{Corr}_k = D_k^2 e + D_k S_k.$$

These values are consecutively added to the contents of the digital accumulator, yielding an accumulating average whose expected value is given by:

$$\mathrm{Acc} = \mathrm{E}\!\left[D_k^2 e + D_k S_k\right]$$
$$= \mathrm{E}\!\left[D_k^2 e\right] + \mathrm{E}\!\left[D_k S_k\right]$$

Because the dither signal is uncorrelated with the input signal, the expected value of their product is zero and the long-term average value of the accumulator is:

$$\mathrm{Acc} = \mathrm{E}\!\left[D_k^2 e\right]$$

Thus the expected value of the correlation samples is proportional to the dither gain error. Summation of these samples in the accumulator acts as a digital integrator. The feedback dither error signal consists of the 8 most-significant bits (MSB) in the accumulator. It is fed to an 8-bit AGC DAC whose analog output modulates the amplitude of the dither signal at the output of the high-speed dither DAC to drive the dither subtraction errors to zero.

Dither Gain Control Loop Analysis

The dither gain control loop is a mixed analog and digital control system. A simplified diagram is shown in Fig. 7a. Lumping the correlator and accumulator gains into a single feedback gain β and ignoring the s-domain effects of the zero-order hold associated with the AGC DAC and the loop shaping filter results in the simplified z-domain loop shown in Fig. 7b. The transfer function of the dither error derived from the simplified output error signal in the z domain is given by:

$$\frac{E_o(z)}{D_i(z)} = \frac{\beta(A_o - 1)}{z - (1 - A_o - \beta)}.$$

For the correlator loop the combined system gain of the dither DAC and ADC, A_o, is approximately 1 and the feedback gain β is $\ll 1$. For $A_o = 1$, the dither error goes to zero. For A_o approaching 1, the dither errors are proportional to β.

The bit size of the accumulator dominates the feedback gain constant β, which determines the time constant of the loop. The effect of the dither gain loop on the input signal is to produce a small amount of amplitude modulation. The magnitude and frequency of the AM sidebands are made negligibly small by choosing the accumulator size to be sufficiently large (36 bits). This corresponds to an approximate β value of $2^{(36-12)}$ for a 12-bit sample size, or modulation sidebands that are 144 dB down. In addition, one can estimate the loop bandwidth from the loop equation:

$$\mathrm{BW(3\text{-}dB)} = \frac{\beta f_s}{\pi}.$$

For $\beta = -144\ \mathrm{dB} = 10^{-144/20}$ the loop bandwidth is 0.5 Hz yielding a time constant of 2 seconds. These estimates are within an order of magnitude.

The long time constant associated with the dither error reduction loop would manifest itself in the system at power-up as a noise floor in the frequency domain that slowly lowers and settles out at the system noise floor as the dither gain loop drives the dither subtraction errors to zero. Although the decay is exponential, on a log amplitude scale in the frequency domain the settling time appears as a linear phenomenon in time. To overcome the slow settling at power-up an accumulator initialization step is implemented. During periodic calibrations the top 8 MSBs of the accumulator are stored in the instrument nonvolatile RAM. On a subsequent power-up, the stored value is preloaded into the MSB portion of the accumulator, thereby providing immediate ADC operation with optimal signal-to-noise ratio.

Hard Floor Mechanism

Although dithered converters substantially lower the hard floor distortion plateau, they eventually encounter a mechanism that limits the depth of the hard distortion floor. One such mechanism is unwanted coupling of the dithered ADC digital output to the analog input. This is explained heuristically by considering the MSB or sign bit of the digital output. This digital signal is synchronous with the fundamental input

signal frequency because it toggles in phase with the sign of the input signal. The MSB can be considered to be rich in distortion harmonics related to the input fundamental. These harmonics remain constant relative to the input signal level, and depending on the converter output-to-input isolation, can couple back into the input. Therefore, the reduction in distortion seen in dithered ADCs as the input signal is lowered is limited by this distortion feedback mechanism.

This mechanism was characterized with the converter operating in the HP 89410A with the help of the autocorrelation measurement feature of the instrument. Because this mechanism is correlated with the input but occurs at different time delays relative to the input, autocorrelation proved helpful in uncovering specific areas in the hardware where the mechanism was dominant. Correlation peaks of the instrument noise floor were observed at specific delay times corresponding to multiples of the sample clock delay following the ADC conversion. The most notable peaks were associated with delays corresponding to the output bus between the ADC output and the digital filters. To reduce the effect of this mechanism on the distortion hard floor in the ADC the digital outputs were changed from TTL to ECL, which has lower voltage thresholds.

Digital Implementation and Diagnostic Modes

The majority of the digital functionality of the large-scale-dithered ADC is implemented in two electronically programmable gate arrays. The gate array algorithms can be dynamically reloaded during instrument operation. This facility is used to provide the half-scale and quarter-scale dither modes of operation for the HP 89410A and HP 89440A, respectively. Flexibility in choosing different modes of operation allowed the dithered ADC to be tailored to the individual overhead and dynamic range requirements of each instrument. This capability is also used to provide factory service technicians a variety of ADC modes for diagnostic servicing of not only the dithered ADC but the instrument signal processing hardware as well. These include various test signal generators as well as dithered and undithered modes of operation.

Digital Local Oscillator and Decimation Filters

After signals are sampled by the analog-to-digital converter section, the digital signals are passed on to the digital local oscillator and decimation filter (LO/DF) block. The converted data stream is first routed through an intermediate assembly between the two ADC channels and the LO/DF block. This buffer/switch assembly provides a data multiplexer so that signal sources other than the analog front ends—digital inputs, for example—can drive the LO/DF section. It also allows further digital signal conditioning such as time gating to be performed on the data before it reaches the filters. For diagnostic purposes the digital source output can be routed through the buffer/switch assembly to drive the LO/DF section. This mode allows both the digital source and the LO/DF assembly to be thoroughly exercised during diagnostic tests.

The buffer/switch assembly also performs ECL-to-TTL signal level translation on the incoming ADC data. One of the design challenges encountered while building the HP 89410A involved the backplane signals and the method by which the ADC drives its data through the backplane and onto the digital LO/DF assembly. It was found that TTL drivers on the

ADC assembly combined with the capacitance of the backplane made for a noisy, high-current switching combination that fed noise back into the ADC signal-conversion circuitry. To eliminate this source of spurs, the ADC drivers were changed to quieter ECL buffers. The smaller noise margins of ECL logic required careful shielding between the ADC data bus and the surrounding TTL signals of the rest of the backplane to prevent crosstalk. The ECL buffer output and ECL receiver input terminations are matched to the impedance-controlled signal lines in the backplane to ensure optimum transmission line characteristics.

After the buffer/switch assembly the converted ADC signals from both input channels get passed on to the LO/DF assembly. This block provides two channels of frequency selective band analysis. Each channel provides one complex frequency shifter and two sets, real and imaginary, of digital decimation filters.

The digital LO is a proprietary Hewlett-Packard high-speed IC that consists of a precision quadrature local oscillator, two mixers, and two low-pass filters. It takes in real input data and produces low-pass-filtered complex output data. This operation is known as zoom mode. The local oscillator consists of a precision 40-bit phase accumulator followed by a sine/cosine generation circuit. The phase accumulator is designed to give decimal frequencies for typical sample rates with a minimum resolution of 25 μHz, although the HP 89410A requires a settable LO center frequency resolution of only 1 mHz. The sine/cosine generation hardware stores the first octant of sine and cosine in two 256-element ROMs. The ROM data is used with linear interpolation to generate two sinusoids offset by 90 degrees with worst-case spurs of –110 dBc and typical spurs of –125 dBc.

After mixing the local oscillator's complex sinusoid with the input data, the output is $y = xe^{-j\omega t}$, where x is the input data and $\omega = 2\pi f_{LO}$. This complex mixing allows the real input data to be frequency translated around dc for further baseband filtering and decimation. The input signal at frequency f_{in} is translated to $f_{out} = f_{in} - f_{LO}$.

The first set of low-pass filters is implemented in the digital LO. These filters can be bypassed for full-span (nondecimated) data or used when smaller spans are required. When the filters are used, the data is bandlimited to $f_s/4$ to prevent aliasing and the output is oversampled by two.

To provide user-selectable bandwidths, the down-converted complex baseband signal produced by the digital LO goes to two identical proprietary Hewlett-Packard digital decimation filter ICs, one for the real data and one for the imaginary data. The recursive decimation filters can be thought of as a cascaded chain of alternating digital low-pass filters and decimate-by-two blocks (see Fig. 8). With an ADC sample rate of 25.6 MHz, the available decimated output data rates are 25.6 MHz, 12.8 MHz, 6.4 MHz, ..., 3.052 Hz. These sample rates correspond to alias-free analysis spans of 10 MHz, 5 MHz, 2.5 MHz, ..., 1.192 Hz. By using a proprietary resampling algorithm implemented in the main DSP, this limited set of analysis spans can be expanded to provide any user-selected arbitrary span between 10 MHz and 1.0 Hz.

The filter chips support three different modes of operation: bypass, decimate, and interpolate. The digital source, described in a following section, uses the filter chips in the

Low-Pass Filter $f_s/4$ — 2× Decimate — Low-Pass Filter $f_s/8$ — ••• — 2× Decimate — Low-Pass Filter $f_s/2^{25}$

Local Oscillator 21 26 +90°

Input from ADC 24

Data Output Selection and Multiplexing

32 Real
32 Imaginary

26 Low-Pass Filter $f_s/4$ — 2× Decimate — Low-Pass Filter $f_s/8$ — ••• — 2× Decimate — Low-Pass Filter $f_s/2^{25}$
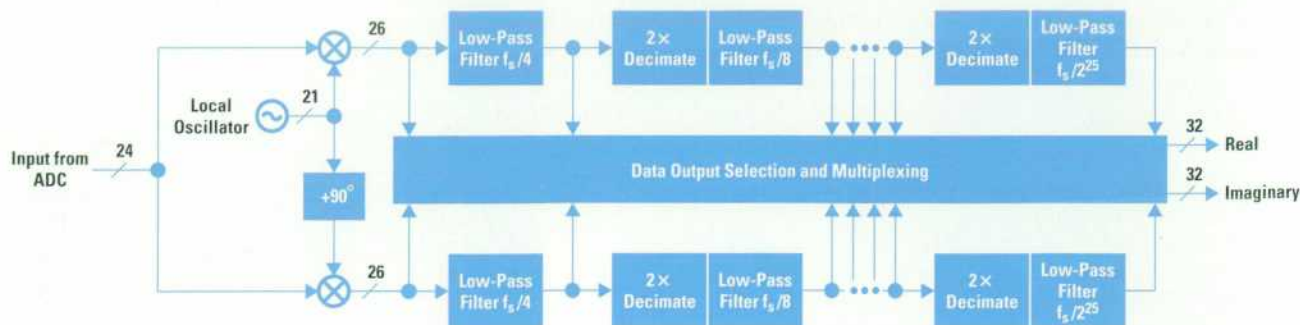
Fig. 8. Zoom and decimation filtering.

interpolate mode or "mooz" mode. The digital LO/DF assembly uses the filter chips in both the bypass and decimate modes. In bypass mode, the data is simply passed through the chip unchanged. This mode is used for full-span 10-MHz measurements and for the next smaller basic span, 5 MHz, because the first filter step is in the LO chip. For the 5-MHz span, the oversampled data passing through the decimation filter chips is decimated by two externally in hardware for the resultant output sample rate of 12.8 MHz. For spans smaller than 5 MHz, the digital filter chips are configured for the decimate mode. The data is processed through 23 passes of decimate/filter steps, each pass corresponding to one decimate/filter stage in Fig. 8. Each pass result is output from the digital filters, but only the pass of interest is deposited into the sample RAM. If desired, all of the basic spans of interest (except full span) can be gathered by the instrument at once, to be used in algorithms such as 1/N-octave analysis. The sequence of pass outputs is:

… 1 2 1 3 1 2 1 4 1 2 1 3 1 2 1 5 1 2 1 3 1 2 1 4
1 2 1 3 1 2 1 6 1 2 1 3 1 2 1 4 1 2 1 3 1 2 1 …,

where pass 1 refers to the output of the first cascaded decimate/filter section and pass 2 refers to the output of the second cascaded decimate/filter section. The sequence is structured so that for every pass k, there are two input points coming from the previous pass (k – 1). For example, between each pair of pass 2 outputs there are two pass 1 outputs, and between each pair of pass 3 outputs there are two pass 2 outputs. This sequence follows naturally from the architecture of the cascaded decimate/filter sections. A decimate-by-2 filter requires two input points for every output point. Therefore, two points from the (k – 1)th pass must be output before the kth pass can calculate one output point.

The low-pass filter is a compromise between optimum filter shape and calculation speed constraints. Although it has fairly good passband ripple characteristics, the filter's passband response droops to nearly –0.5 dB at the upper end of the analysis span. Similarly, while the filter stop band provides good rejection, the transition band of the filter, reflected about the Nyquist frequency, falls into the upper frequencies of the analysis span, limiting the worst-case alias protection to –86.7 dB. The worst-case passband ripple is about 0.2 dB and the worst-case stop-band rejection is better than 111 dB (all passes). Correction routines in the main DSP account for the passband roll-off anomaly and ensure very flat passband characteristics for the entire span. The IIR filter implementation enhances calculation performance while sacrificing time-domain performance as a result of the filter's sharp cutoff and nonlinear phase characteristics. Algorithms implemented in the main DSP provide the necessary overshoot and phase corrections for time-domain measurements.

The IF trigger circuit on the digital LO/DF assembly uses the complex digital data coming out of the decimation filters to allow frequency selective triggering. As opposed to traditional broadband triggering mechanisms like the analog input trigger, the IF trigger looks for energy present only in the frequency band selected. The IF trigger circuit looks at data coming out of the real and imaginary digital filters from either channel 1 or channel 2. A RAM is used to provide a complex-magnitude map that is compared to the incoming complex, band-selected data. The map can be visualized as a graph with the real part of the input data along the x axis and the imaginary part along the y axis. The magnitude of the complex input data is the square root of the sum of the squares of the real and imaginary parts, which describes a circle on the graph. For a given trigger value, the corresponding magnitude circle is calculated and placed in the map. The RAM is programmed to 0s for those complex pairs falling inside the trigger circle and to 1s for those values falling outside. When the incoming data's x-y location in the map moves from within the circle, where the RAM reads 0, to without, where the RAM reads 1, the magnitude of the data has crossed over the trigger point. When the IF trigger circuit detects the change of map values from 0 to 1, it generates a trigger to the instrument. With two bits in the map, two separate circles can be described, allowing additional trigger functionality such as hysteresis. The hysteresis algorithm requires that the input signal first fall below the programmed hysteresis level (smaller circle) before resetting the circuit to look for a new trigger (larger circle).

Sample RAM

The sample RAM assembly is an integral component of the measurement and triggering functions of the HP 89410A. In addition to its main function of capturing one or two channels of data from the digital LO/DF assembly, it provides the hardware necessary for pretrigger and post-trigger delays and block size accounting.

Two sample RAM options are available. The standard configuration provides storage for 32,768 64-bit samples (32-bit real, 32-bit imaginary) for each input channel in two-channel mode, and 65,536 64-bit samples in single-channel mode. The optional configuration provides storage for 512K 64-bit samples for each input channel in two-channel mode, and 1M 64-bit samples in single-channel mode. In single-channel

mode, the larger configuration provides about 41 ms worth of data at the full 25.6-MHz sample rate and about 191 hours worth of data at the lowest allowable sample rate of 3.052 Hz.

The main challenge in designing the sample RAMs was delivering high-speed data capture (>400 Mbytes/s) at a reasonable cost. The smaller sample RAM board is designed with high-speed SRAMs. The larger sample RAM required a different approach because of the size of the memory and the relatively high cost of high-speed static RAMs. The solution to this problem was found by using interleaved banks of video RAMs. The two-port video RAM has a high-speed serial-access register interfaced with a slower random-access memory. The data from the digital LO/DF assembly is loaded serially into the video RAM serial register (up to 512 samples) and is transferred with one operation into the random-access memory. The samples can then be accessed via the slower system bus by the DSP or system microprocessor.

8K samples per channel is sufficient data to support the display, which shows a maximum of 3200 lines of frequency data. The larger sample RAMs are used for the time capture measurement mode, in which the instrument captures a large amount of contiguous data for later processing.

Digital Signal Processor

The DSP assembly is designed around the Motorola DSP96002 IEEE floating-point digital signal processor. The DSP96002 provides two 32-bit memory ports and peak performance of 48 MFLOPs. The two memory ports are ideal for FFT calculations, the main function of the DSP. The DSP can become the bus master for the system, allowing it to access data from the sample RAM directly, perform the necessary data processing operations, and then place the results directly into the display assembly. By removing the CPU from the main data flow path, the instrument's throughput is significantly increased. On a typical fast average measurement, the DSP can transfer and process 300 512-point complex spectra (including corrections, FFTs, averaging and display formatting) per second.

Display

The display assembly is based on the Texas Instruments TMS34020 graphics system processor (GSP) and a 1M-byte bank of video RAMs. The GSP is a 32-bit high-speed general-purpose processor that is optimized for graphic display systems. The display system is designed to allow data points to be placed directly into display memory by the DSP. The GSP's program then processes these data points and constructs a trace and the surrounding annotation on the display independently, without outside processor intervention. This arrangement frees the DSP and the host CPU from the need to control the display operations directly. The result is up to 60 trace updates per second. Waterfall and spectrogram displays are limited only by the instrument's processing speed.

One of the challenges in designing the display assembly came from the analog video circuitry. The video RAMDAC and video buffers share a printed circuit board with the DSP and display blocks. The large data buses coming onto the board and the associated circuitry cause a lot of ground and power noise, which tends to show up on the video output and CRT. To reduce the effects of the digital noise, the analog

video subsection is placed on a split ground and power plane "island" and the video signals are routed to the backplane between sandwiched ground planes to provide shielding and a controlled 75-ohm trace impedance.

CPU

The HP 89410A central processing unit assembly is based on the Motorola MC68EC030 32-bit microprocessor and the MC68882 floating-point math coprocessor. The CPU handles many of the user interaction functions and the high-level control of the measurement system. The CPU assembly also provides system bus arbitration and 4M bytes of main memory, and controls the flexible disk, the HP-IB (IEEE 488, IEC 625), the serial port, and the keyboard. The HP 89410A's system code is contained in 3M bytes of in-circuit programmable flash EEPROM to allow easy firmware updates via the flexible disk drive.

Backplane

Another challenging aspect of the HP 89410A design was dealing with the large number of signals to be routed from assembly to assembly within the digital card nest. The card nest backplane motherboard connects eight digital logic assemblies and connects the display assembly to the internal CRT, the front ends to the LO/DF assembly, and the source to the digital source assembly. The backplane routes approximately 400 signals through five 150-pin connectors, three 200-pin connectors, and two 300-pin connectors. The backplane had to be carefully designed to carry all of these signals with minimal crosstalk and ground bounce and impedance-controlled ECL, video, and clock lines. The backplane is constructed as a ten-layer board. The top layer is a ground plane. The next layer provides the ECL, video, and system clock traces and is sandwiched between the top layer and another ground plane. The two ground planes, along with the 0.005-inch-wide traces, produce a 50-ohm characteristic impedance transmission line for the signals on this layer. The analog video signals and two ADC ECL data buses are further shielded from the TTL-level system clock lines that reside on this layer. The fourth layer provides the +5V power plane while the last six layers are used for the rest of the digital signals on the backplane. These layers are purposely separated from the ground and power planes to lower capacitance and thus increase the characteristic impedance of the signal lines. The higher impedance prevents ground bounce on the digital assemblies when a large number of outputs on the backplane are activated at the same time. The large buses such as the system address and data bus are placed on the component side of the backplane (bottom) to keep them as far away as possible from the ground planes. In addition to the extruded front-end analog card nest shielding, another shield is placed between the front end and the electrically noisy backplane.

Digital and Analog Source

The HP 89410A source is a DAC-driven, 50-ohm output impedance analog source with capabilities similar to many standalone sources. It provides a variety of output types including single-frequency sine, Gaussian distributed random noise, periodic chirp, and user-defined arbitrary waveforms. The source is used in calibrating the HP 89410A and in diagnosing the rest of the instrument.

What Is Dithering?

Dithering is a method for randomizing the quantization errors of an ADC by adding stimulus that is uncorrelated with the desired signal at the analog input of the converter. Fig. 1 depicts a basic block diagram of an externally dithered converter in which the dither signal is subtracted following the ADC conversion. There are various types of dithering, which can be differentiated in various ways, one being by the characteristics of the dither signal. Dither signals can be characterized on the basis of signal type (such as noise), amplitude (small-scale or large-scale), or frequency (narrowband or broadband), as shown in Fig. 2. Narrowband dither signals outside the information bandwidth can be removed following the conversion by digital filtering. Dither signals are also often characterized on the basis of the probability distribution of the dither amplitude (such as uniform or Gaussian for random noise).

Dithering a converter can provide significant improvements in the ability to extract signals below the resolution of the converter and in linearizing a converter's performance. The improvements vary with the degree and type of dither. Combined with the time-averaging property of subsequent digital filtering, dithering effectively removes or smooths the quantization noise inherent in the staircase transfer function of an ADC.

This can be illustrated as follows. Consider an ideal ADC with a staircase transfer function whose quantized ideal error transfer function is shown in Fig. 3a. The expected error transfer function resulting from dithering with a random noise signal can be computed by evaluating the expected transfer function with the weighted probability distribution of the dither. Fig. 3b depicts a uniformly distributed probability density function of random dither with exactly 1 LSB of dither amplitude.

The general equation for computing the expected value G for a transfer function with one random variable is given by:

$$G = \int_{-\infty}^{\infty} p(z)e(z)dz,$$

where e(z) represents the transfer function and p(z) represents the probability density function of the random variable z. The expected value computes the average of the transfer function weighted by the relative likelihood of z. For the case of dithering, in which the random variable z is added to the input signal of amplitude x, the resulting expected value of the transfer function appears as follows:

$$G(x) = \int_{-\infty}^{\infty} p(z)e(x + z)dz,$$

The function e(x+z) corresponds to the transfer function in the variable x added to the random variable z representing the dither.

This resembles a convolution, and hence we can illustrate the process graphically as the integration of the dither probability density function as it is moved past the ADC error transfer function as shown in Fig. 3c. For a dither amplitude of 1 LSB, the integration result is zero, yielding an ADC transfer function that is completely

Fig. 2. Types of dither in the frequency domain.
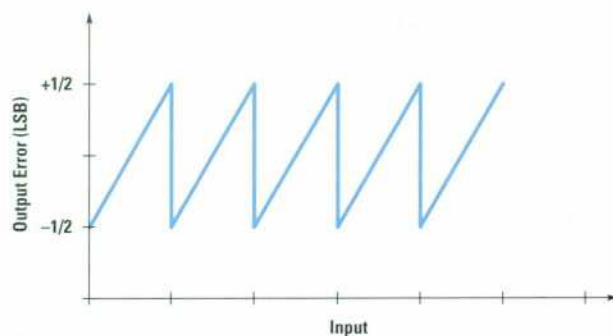
(a)
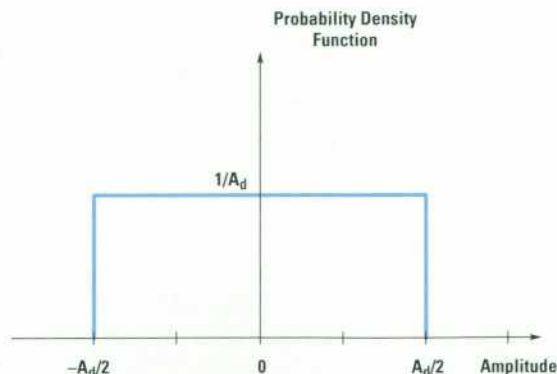
(b)
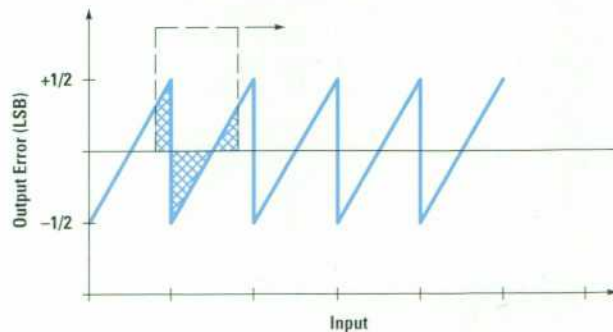
(c)

Fig. 3. (a) Error transfer function of an ideal ADC. (b) Probability density function of uniformly distributed random dither with 1-LSB amplitude. (c) Convolution of (a) with (b) yields zero.

Fig. 1. Block diagram of an externally dithered analog-to-digital converter.
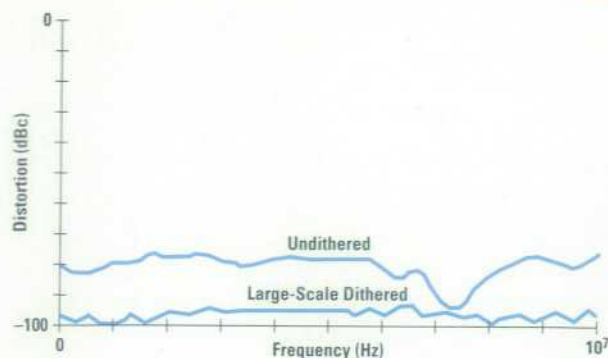
Fig. 4. Third-harmonic distortion of the same ADC with large-scale dither and no dither.

linear. The high-order distortion terms associated with the polynomial expansion of the quantized transfer function are effectively removed by the dithering. This result is true for dither amplitudes that are integral multiples of the ADC LSB.

Low-order distortion performance improvements can also be obtained by the application of large-scale dither. In practice the large-scale dither effectively smooths the inflections caused by the integral nonlinearity of the ADC transfer function. Although dithering a transfer characteristic with a single second-order or third-order distortion term does not improve that term, it does improve the low-order distortion of transfer functions whose polynomial expansions also contain higher-order terms, which is the case for real ADC transfer functions. The convolution generates lower-order terms from the higher-order terms, and these combine with the original low-order terms of the transfer function to make them smaller. This results in an overall improvement in low-order distortion. The degree of improvement is directly related to the magnitude of the dither signal employed. Larger-scale dither yields greater reduction of low-order distortion for a given signal amplitude.

Fig. 4 demonstrates this phenomenon with a measurement that compares the third-harmonic distortion performance of the HP 89410A converter for half-scale dither and undithered modes of operation. The third-harmonic distortion is significantly improved by the large-scale dithering.

Manfred Bartz
Customer Support Engineering Manager
Lake Stevens Instrument Division

The source is composed of two assemblies: a digital source assembly and an analog source assembly. The digital source assembly contains a 32K-sample complex (real and imaginary) waveform memory buffer, a pseudorandom noise generator, and the same digital LO/DF chipset that is used in the front-end receiver section. The digital source assembly creates a digital signal for the analog source assembly. The analog assembly contains a waveform DAC, a reconstruction filter, dc offset circuitry, an output amplifier, 10-dB step attenuators, and front-end calibrator circuitry.

All source output signals are created digitally on the digital source assembly. They are generated via the source sample RAM, the pseudorandom noise generator, and the digital local oscillator IC, as shown in the block diagram, Fig. 1. The source signal flow is essentially the reverse of the front-end receiver section. The signal path starts with the source waveform memory, where the complex digital signals are stored to be fed into the real and imaginary digital interpolation filters. The address generator to the waveform RAM is either a linear counter or a pseudorandom noise source. The address counter is used with repetitive or single-block waveforms such as periodic chirp and user-defined source types. The pseudorandom noise addressing mechanism is used to generate noise outputs. The distribution of the random noise is determined by the waveform loaded into the source RAM. For example, if a Gaussian curve is loaded into the waveform memory, the random sampling of the curve by the pseudorandom noise address generator causes a Gaussian-distributed noise signal to be fed into the digital interpolation filters. The pseudorandom noise address generator is a maximal-length sequence that repeats itself approximately every 6 hours at the full output sample rate of 25.6 MHz.

The digital interpolation filters are used to increase the input (waveform) sample rate by a factor of 2^N, where N is programmable from 0 to 23. Since the output DAC sample rate is always 25.6 MHz, the digitized waveform store in the source RAM can have an effective sample rate as low as 3.05 Hz. This allows the HP 89410A to have extremely low-frequency waveforms and very large effective block sizes even though the waveform RAM is only 32K samples long. Also, since the DAC runs at one sample rate, only one reconstruction filter is needed.

The complex interpolated signals are routed into a digital LO chip where they are mixed from dc to any center frequency up to 10 MHz. By virtue of this complex mixing operation, the resultant output signal has a real two-sided spectrum around the positive and negative LO frequencies. In this operation, or "mooz" mode, the output is $y = \text{Re}\{xe^{j\omega t}\}$, where x is the input, Re represents the real part of the number in the braces, and $\omega = 2\pi f_{LO}$. This frequency translation causes the source's complex waveform signal at frequency f_{in} to appear at the output frequency $f_{out} = f_{in} + f_{LO}$.

The real digital source signal y is then converted into an analog waveform by a 12-bit DAC running at 25.6 MHz. The reconstruction filter follows the output of the DAC, limiting the bandwidth of the signal to 10 MHz while correcting for the $\sin(\omega t/2)/(\omega t/2)$ rolloff resulting from the zero-order hold effect of the DAC. The reconstructed signal is then summed with the output of a dc offset DAC and the result is buffered with a 20-dB-gain output amplifier. The amplified signal goes through the attenuator section to provide coarse control of the source level before being output on the front-panel source BNC.

Trigger
The HP 89410A provides very flexible triggering modes to support complex measurements. The HP 89410A can trigger from four sources: the external front-panel trigger input, either input channel, or the internal source block start signal. The trigger level and slope are user-definable. The trigger can be controlled by an external arm signal on the rear-panel arm BNC connector. Pretrigger and post-trigger delay and arm delay are available.

The HP 89410A has a sample rate of 25.6 MHz. With a 10-MHz input signal, up to 140.6 degrees of signal can be missed while waiting for the next sample time after a trigger condition is met. This amount of uncertainty would negate

the phase information that is so useful for many measurements. The solution is to measure the partial sample time. The time between the occurrence of the trigger conditions and the next sample point is measured and used to correct the sampled data. A pulse starts when the trigger conditions are fulfilled and ends when two sample points have been taken. The pulse charges a capacitor at a fast rate. When the pulse ends, the capacitor is discharged at a slower rate. The fast charge rate is about 731 times the slow charge rate. This effectively stretches the pulse by a factor of 731.

This stretched pulse gates a counter running at 6.4 MHz. When the stretched pulse ends, 10 bits of pulse length information is available for the software to use for corrections. The configuration gives a resolution of 214 picoseconds, or 0.77 degree of a 10-MHz input signal.

During a calibration cycle, the partial sample generator is calibrated by generating sine waves of known phase relationship to the sample clock and measuring the resulting partial sample values. This has shown the circuitry to have good linearity and repeatability.

Frequency Reference
The specifications of the HP 89410A require a clean, low-noise reference oscillator capable of locking to a customer-supplied signal or oven oscillator. Previous FFT analyzers don't share this requirement because their sample rates are much lower than their crystal oscillator clock frequencies. For example, the HP 35670A dynamic signal analyzer derives its 262-kHz sample rate from a 40-MHz crystal oscillator. This 150:1 ratio results in a 43-dB noise reduction. This approach has the effect of burying mediocre reference performance under the input noise floor.

Since the HP 89410A samples at a much higher rate than previous-generation analyzers, a different approach had to be taken. The design is modeled after the 80-MHz reference developed for the HP 3588A, a hybrid FFT-analog swept spectrum analyzer.[2] The key components of this design are a clean sampling phase detector and oscillator.[3] The HP 89410A reference is a crystal-controlled VCO running at 51.2 MHz tuned by a phase detector sampling at 400 kHz. The external lock signal is divided down by 25 so that a 10-MHz lock signal becomes 400 kHz. The sampling phase detector

filter is tuned to allow locking down to a 1-MHz lock input. The 51.2-MHz signal is divided by 2 and routed to all boards in the instrument in quadrature. The quadrature signal is developed by inverting and delaying by 10 nanoseconds. The 10-ns delay allows the data lines to settle before latching.

Another board locks a 16-MHz crystal oscillator to the reference and produces 48, 64, and 80 MHz. These signals are sent to other boards and used for CRT, DSP, CPU, and system bus clocks. Having all clocks in the instrument phase locked allowed the design team to locate and address fixed crosstalk-induced spurious signals rather than spurious inputs that drift and hide.

References
1. J.S. Epstein, et al, "Hardware Design for a Dynamic Signal Analyzer," *Hewlett-Packard Journal*, Vol. 35, no. 12, December 1984, pp. 12-17.
2. K.C. Carlson, et al, "A 10-Hz-to-150-MHz Spectrum Analyzer with a Digital IF Section," *Hewlett-Packard Journal*, Vol. 42, no. 3, June 1991, pp. 44-60.
3. T. Hillstrom, "Design Method Yields Low-Noise, Wide-Range Crystal Oscillators," *EDN*, March 17, 1988.
4. M. Bartz, "Large-Scale Dithering Enhances ADC Dynamic Range," *Microwaves & RF*, May 1993.

RF Vector Signal Analyzer Hardware Design

Based on the HP 89410A baseband vector signal analyzer, the HP 89440A RF vector signal analyzer extends the frequency range of both receiver and source to 1.8 GHz with a 7-MHz information bandwidth. All of the vector capabilities of the 10-MHz baseband instrument (up to a 7-MHz information bandwidth) can be translated to any frequency from 0 to 1.8 GHz.

by Robert T. Cutler, William J. Ginder, Timothy L. Hillstrom, Kevin L. Johnson, Roy L. Mason, and James Pietsch

The HP 89440A radio frequency (RF) vector signal analyzer consists of two components. The first is the HP 89440A intermediate frequency (IF) section, which is identical to the HP 89410A 10-MHz baseband vector signal analyzer described in the article on page 31. The second component is the HP 89440A RF section, which extends the vector signal analysis capabilities of the baseband analyzer to RF frequencies. This article concentrates on the RF section.

The 1.8-GHz RF section contains a triple-conversion heterodyned receiver, a source that mirrors the receiver, a local oscillator, frequency references, and dedicated processor control. The block diagram is shown in Fig. 1. Several aspects of the design differentiate it from traditional RF analyzers:

- The RF section maintains an 8-MHz IF bandwidth designed for good flatness. However, in a vector signal analyzer both amplitude and phase accuracy are very important, so sophisticated vector IF calibrations were developed.
- True-rms power detection and excellent amplitude accuracy combine for a powerful measurement capability.
- A local oscillator (LO) feedthrough cancellation circuit improves the LO feedthrough substantially, preserving dynamic range at low input frequencies.
- The RF local oscillator is not constrained to provide fine frequency resolution. (The baseband section provides that with its digital LOs). This degree of freedom had significant impact on design efficiency and development time.
- The RF source has all the capabilities of the baseband source, including sinusoid, pseudorandom, chirp, and arbitrary waveform source types, but translated to RF frequencies as high as 1.8 GHz. This provides an excellent value to users who need a flexible source at RF frequencies.
- Development time was a top priority. New methods were developed to maintain design flexibility and reduce risk. For example, receiver, source, and LO "plates" featuring RF coaxial vias were developed—a low-cost solution that also reduced technical risk by distributing and isolating critical blocks. Extensive RF and microwave simulations using the HP Microwave Design System (MDS) greatly improved turn-on rates and eliminated a prototype cycle.

Receiver

The HP 89440A RF section receiver is a triple-conversion down-converter with a 50-ohm input impedance. It translates signals with a maximum 8-MHz bandwidth between 2 MHz and 1.8 GHz to a final IF centered at 6 MHz with a bandwidth of 8 MHz. The 6-MHz IF output is connected directly to the input of the HP 89440A IF section. The IF section supplies a signal to calibrate the IF filters and amplifiers of the RF section receiver. The frequency response of the 2-MHz-to-1.8-GHz input is calibrated at the factory and stored in nonvolatile memory. The receiver is a conventional up/down converter with the first IF centered at 2.446 GHz followed by a down-conversion to an IF centered at 46 MHz. The final IF is centered at 6 MHz and provides gain and a buffered output to the IF section.

Conventional swept spectrum analyzers use a log detector, which produces a signal proportional to the logarithm of the power at the detector input. The input power is a function of the resolution bandwidth when noise is the dominant input, and is a function of the level of the input signal otherwise. In either case, log detectors have only 80 to 90 dB of usable dynamic range. Therefore, a variable-gain amplifier may precede the log detector to increase or decrease the input level so that it falls within the dynamic range of the detector. By contrast, there is no variable IF gain in the HP 89440A RF section receiver with the exception of a gain adjustment and a 1-dB gain step in the final 6-MHz IF. This 1-dB gain step is used with the 2-dB input range steps in the IF section receiver to compensate for temperature dependent gain variations in the three IFs of the RF section. The detector in the IF section is an analog-to-digital converter (ADC) with a dynamic range of 125 dB (the ratio of the full-scale input power to the noise power in a 1-Hz bandwidth). The signal-to-noise ratio (SNR) of the RF section is approximately 120 dB (the ratio of the maximum input for a 70-dB distortion-free dynamic range to the noise power in a 1-Hz bandwidth). The receiver's distortion-free dynamic range and noise figure are dominated by the performance of the first converter and first IF. Adding variable gain beyond these stages changes only the signal level but not the SNR. As a result, variable IF gain offers no improvement in dynamic range. The user only needs to control the input level to the first mixer, which is accomplished with an input step attenuator preceding the first mixer.
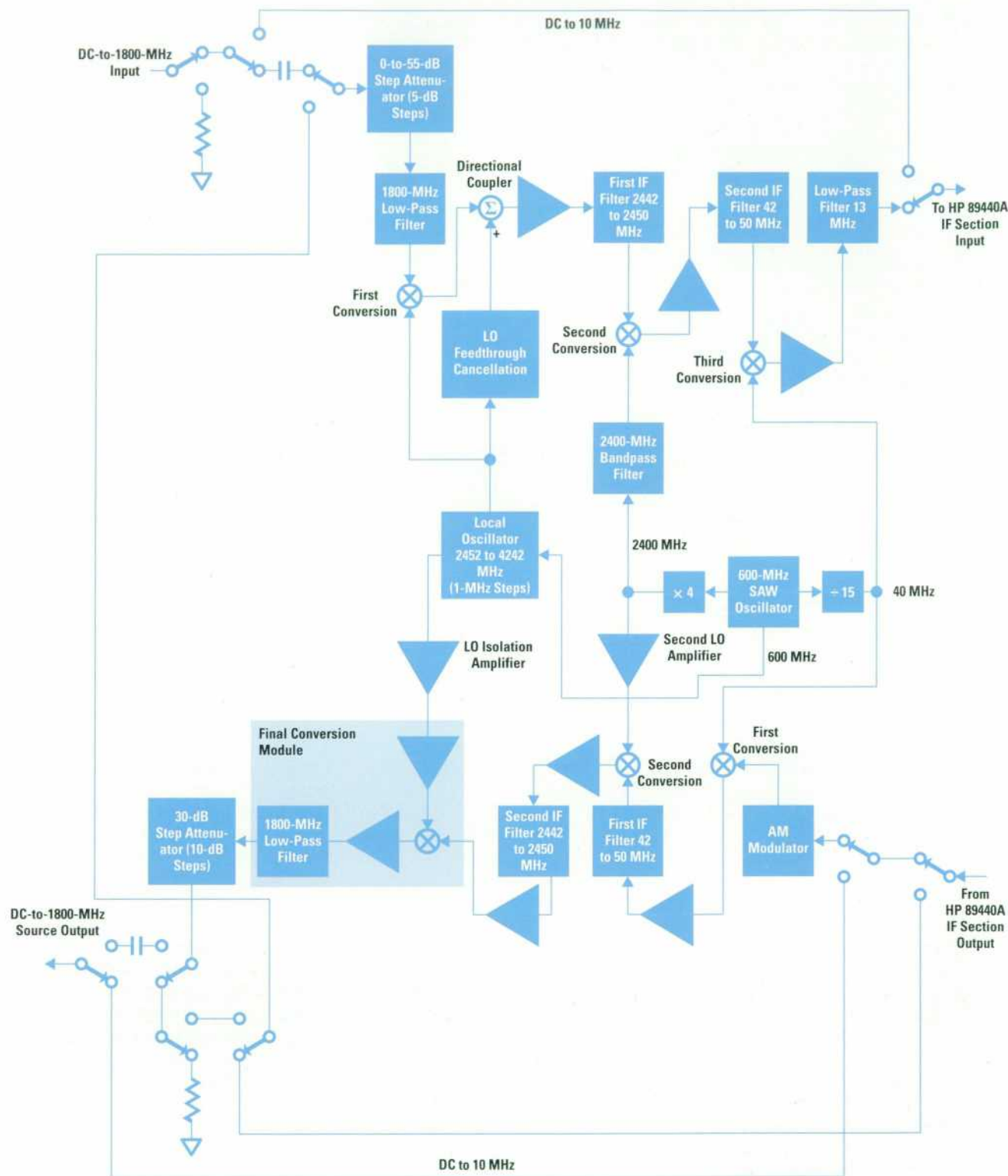
Fig. 1. General block diagram of the HP 89440A RF section.

The HP 89440A has one additional feature not found in traditional swept analyzers. A feed-forward LO feedthrough nulling circuit has been added to reduce the level of LO feedthrough in the first IF. Beyond the second converter, the second-IF filters remove the feedthrough term. Without this LO feedthrough nulling, LO feedthrough referred to the input could be 20 dB higher than a full-scale input signal. This

could result in residual responses and increased second-harmonic distortion at input frequencies below 15 MHz.

Cyanate Ester Printed Circuit Boards
The input attenuator and several other RF receiver, LO, and RF source boards that operate at frequencies beyond 1 GHz use cyanate ester printed circuit board material. Cyanate

ester was chosen in place of the standard glass epoxy printed circuit board (HP FR4) because of its lower loss tangent, which results in lower losses in the board. The dielectric constant of cyanate ester is 4.0 at 2 GHz while glass epoxy is typically specified at 4.5. All cyanate ester boards in the RF section are 0.030 inch thick rather than the standard 0.060 inch. The thinner printed circuit board material has two advantages. Many receiver, LO, and source boards in the RF section use surface mount parts with microstrip construction. With the thinner printed circuit board material, ground vias are shortened by 0.030 inch, reducing parasitic inductance in surface mount components needing a return to ground. In addition, microstrip transmission lines are narrower. A 50-ohm microstrip transmission line is nominally 0.060 inch wide on the 0.030-inch printed circuit board material while the same line on 0.060-inch printed circuit board is 0.110 inch wide. The disadvantage of 0.030-inch printed circuit board is less rigidity. However, all of the cyanate ester boards are mounted on "plates" and are well supported (see "Microwave Plate Assembly" on page 50).

Attenuator

The main signal path of the HP 89440A RF section starts with a step attenuator assembly that provides 0 to 55 dB of attenuation in 5-dB steps and is followed by the first converter. The step attenuator has an input for a calibration signal from the HP 89410A, a mode to terminate the user input during calibration, and a bypass mode to bypass the RF section. The bypass path connects the RF section receiver input connector directly to the IF section receiver input for frequencies below 2 MHz.

First Conversion

Following the input attenuator is the first converter. Input signals are converted to an IF centered at 2.446 GHz. The first mixer is a variant of the single-balanced design used in the HP 8590B spectrum analyzer. It is preceded by a 15-section low-pass filter with a cutoff frequency of 1.8 GHz. The input low-pass filter eliminates input image frequencies (4.89 GHz to 6.69 GHz) as well as spurious components (spurs) resulting from out-of-band inputs. The first-converter LO supplies a 20-dBm signal between 2.452 GHz and 4.242 GHz, which is attenuated by 3 dB at the LO port of the mixer to improve match. Following the mixer is a microstrip directional coupler, where the LO feedthrough cancellation signal is introduced. This is followed by a diplexer and a 4.5-GHz low-pass filter (not shown in Fig. 1) to eliminate mixer products and LO harmonics which can produce residual responses when mixed with the LO of the second converter. The entire mixer is built on Duroid board (Rogers Corporation) which has a dielectric constant of 2.33 ±0.05 and a loss tangent of 0.001 at 1 GHz. The board thickness and dielectric constant are tightly specified so that printed circuit board microwave filters, couplers, and transmission lines with repeatable performance can be produced. Early in the design it was recognized that skin effect losses in the input attenuator and the input low-pass filter preceding the first mixer would result in frequency dependent loss that is about 4 dB at 1.8 GHz. This unflatness can be calibrated and removed, but it results in a displayed noise floor that is unflat, and it reduces the effective dynamic range of the ADC by 4 dB. An amplitude equalizer (not shown in Fig. 1) was added

between the attenuator and the input low-pass filter to eliminate this effect. This has the added benefit of reducing the level of multiple tones in the first IF at low input frequencies. Multiple tones are present because the sum and difference products and the LO feedthrough are only separated by the input frequency and are not eliminated by the first IF filter if the input frequency is low.

The equalizer, the input low-pass filter, the 4.5 GHz low-pass filter, and the diplexer were designed using MDS. The mixer conversion loss was also simulated with MDS and the measured results were in excellent agreement with the simulated results. MDS eliminated at least one design turn of the printed circuit boards.

The diplexer (a traveling-wave directional filter) following the first mixer is a stripline design that terminates the first mixer IF output in 50 ohms at all frequencies. It has a bandpass frequency response that reduces the mixer sum product at the first IF amplifier. The diplexer is implemented with two 8.5-dB stripline couplers and two quarter-wavelength coupling arms. The insertion loss at 2.446 GHz is less than 1.5 dB and the 3-dB bandwidth is 150 MHz. The conversion loss from the input attenuator to the 2.446-GHz IF output of the first mixer board is typically 14 dB.

LO Feedthrough Cancellation

Without cancellation, LO feedthrough referred to the input is typically −10 dBm while the maximum input (0 dB input attenuator) is −30 dBm for 70 dB of distortion-free dynamic range. LO feedthrough cancellation reduces the LO feedthrough at the first IF amplifier by more than 20 dB. A sampled LO signal at the input of the LO port of the first mixer is amplified and split with a quadrature hybrid. At the output of the quadrature hybrid are two equal-magnitude signals that have a phase difference of 90 degrees. Each of these outputs drives the LO port of a double-balanced mixer configured as a current-controlled attenuator. The mixer IF ports are driven by separate current sources capable of sourcing or sinking up to 6 mA of current with 1-μA resolution. The signals from the mixer RF ports are summed by a Wilkinson combiner whose output is coupled into the IF port of the first mixer with a 14-dB microstrip coupler. During calibration of the RF section, the IF section adjusts each current source to produce a signal equal in magnitude and opposite in sign to that of the LO feedthrough at the IF port. The use of quadrature signals in the LO feedthrough cancellation circuitry reduces the interaction between the two current source controls so that ideally they are independent if the two signals are in perfect quadrature.

First IF

Following the first converter assembly are two cascaded low-noise GaAs MMIC (microwave monolithic integrated circuit) amplifiers which have a combined noise figure of 5 dB, a gain of 14 dB and an output third-order intercept* greater than 25 dBm. These amplifiers must be able to handle all mixer products leaving the first conversion assembly including LO feedthrough and both the sum and difference first-mixer products. Higher-order mixer products as well as

---

* Third-order intercept is the theoretical signal level at which the fundamental and the third-order distortion are equal.

Microwave Plate Assembly

At the inception of the HP 89440A RF vector signal analyzer project, it was known that the packaging technology used for the high-frequency portions of the analyzer circuitry would play an important role in the overall performance of the instrument. A number of packaging techniques were considered, including the conventional approach of independently packaged circuits connected with semirigid cable assemblies. Given the overall goals of shielding effectiveness, ease of assembly, and relatively low cost, it was decided to try something different. The high-frequency microwave plate assembly packaging scheme that evolved for the HP 89440A satisfies these goals, provides good flexibility for integrating other functionality, and provides a physical support structure for the individual modules. There are three such microwave plate assemblies in the HP 89440A: the RF source, the RF receiver, and the local oscillator.

The physical implementation of the packaging scheme is rather simple (see Fig. 1). Printed circuit assemblies are mounted with screws to the faces of a 0.375-inch-thick aluminum plate. Assemblies that require RF electrical interconnection on a given microwave plate assembly are mounted on opposite sides of the plate. The RF connection is supplied through a hole in the plate into which a conductor pin and an insulator are inserted. A 50-ohm characteristic impedance is maintained through the hole in the plate by sizing the diameters of the hole and the pin according to the dielectric constant of the insulator (PTFE). A microstrip-to-coaxial transition is formed as the conductor pin is soldered into a hole at the end of a microstrip on the printed circuit board. Good local ground contact at the transition is maintained by a multifingered ground ring that is seated in a shallow counterbore in the plate and is in contact with the printed circuit board ground plane at the perimeter of the hole. Two screws located near the hole secure the printed

circuit board to the plate and guarantee that the fingers of the ground ring are compressed. Good return loss is achieved with this interconnect solution for frequencies up to and beyond 6 GHz. A low-pass filter version of the conductor pin was developed that is physically interchangeable with the standard pin to suppress transmission of higher-frequency interfering signals (greater than 5 GHz) if required. Compared to the conventional approach, this method of RF interconnection reduces the number of expensive RF connectors and cable assemblies required. When shielding of the individual printed circuit assemblies is required, a shield can is screwed down over the assembly. The aluminum plate acts as the sixth wall of the shield can. A good RF seal is provided by conductive elastomer gasket material in the lip of the shield can. For some assemblies, shield-can resonances present a problem. Polyiron is fixed to the top of the shield can to suppress the resonances in those cases. Dc power is supplied to the shielded assemblies with standard screw-in feedthrough capacitors. The feedthroughs do an excellent job of preventing RF leakage into or out of the shielded assemblies.

Flexibility is provided by the packaging scheme to allow mounting of different types of devices onto the plates. For instance, several coupled coaxial resonator filters are used in the HP 89440A. These filters incorporate the plate into their design. The filter housing is screwed to the plate, which functions as the cover for the filter. Input and output coupling rods used in the filters are integrated with printed circuit board launches as described above to provide a connectorless interface between the filter and a printed circuit board. A significant cost saving was realized on these filters as a result of this design.

Additional functionality is included in the plate design in the form of features milled into the plate. An example is shielding compartments to isolate the individual sections of the step attenuators in the RF receiver and the RF source. The step attenuator printed circuit assemblies are screwed down to the shielding structures milled into the plate.

Certain aspects of the printed circuit board designs are optimized for use with the microwave plate packaging scheme. Coaxial-to-microstrip launch geometries are optimized for best return loss. The side of the printed circuit board toward the plate is mostly ground plane. Tin plating is used instead of SMOBC-HAL (solder mask on bare copper—hot air leveled) to ensure that the board fits flat onto the plate. For connection to other modules, printed circuit boards have bulkhead-mount SMA connectors soldered to the ground plane side. Each connector is pushed through a counterbored hole in the plate and its mounting nut is installed. Conductor pin geometry and screw hole locations were chosen so that a standard flange-mount SMA connector can be used to perform measurements on the individual printed circuit assemblies.

This high-frequency packaging approach is attractive for low-volume instruments. Unit cost is relatively low and minimal tooling charges are incurred.

Roy L. Mason
Development Engineer
Lake Stevens Instrument Division

Fig. 1. Microwave plate assembly packaging scheme.

LO harmonics are removed by the 4.5-GHz low-pass filter on the first conversion assembly. The sum and difference products are amplified by this board, and the third-order two-tone distortion products resulting from these tones appear as third-harmonic distortion to the user. At input frequencies above 60 MHz the diplexer reduces the sum tone by 6 dB at the input to the first IF amplifier, so the third-order distortion produced in the first IF is a concern only at low input frequencies. Following the first-IF amplifiers is a four-section coaxial filter with a 17-MHz bandwidth and an insertion loss less than 3 dB. This filter has a center frequency of 2.446 GHz and greater than 80 dB of rejection at the image frequency of 2.358 GHz.

Second Conversion and IF
Following the first IF filter is the second converter assembly which down-converts the 2.446-GHz IF to the second IF centered at 46 MHz. The second LO is a 2.4-GHz signal supplied by the reference. It is amplified to 10 dBm, filtered by a two-section combline filter to eliminate any sidebands at 50 MHz, and then amplified to 13 dBm before application to the second mixer LO port through a 3-dB attenuator. The 46-MHz IF output of the second mixer is amplified by a low-noise amplifier, resulting in a signal level equal to that of the input signal (−30 dBm full-scale). The noise figure at this point is nominally 23 dB. Following the second converter are three identical cascaded second-IF filter assemblies implemented

© Copr. 1949-1998 Hewlett-Packard Co.

with capacitively coupled resonators. The filter design is based on a Tchebychev filter with 0.1 dB of ripple. The filter is predistorted and has an insertion loss of 4 dB. The filter has a minimum rejection of 27 dB at 38 MHz (edge of the third-conversion image band), so the three cascaded stages have a minimum of 80 dB of rejection. So that each filter board will have a nominal gain of 0 dB a low-noise amplifier precedes each filter. Three identical filters were chosen over a design in which the entire filter resided on a single board. A single-board design would require shielding between sections of the filter and alignment of the filter would require a complex adjustment procedure. Using smaller identical boards, the alignment procedure is identical for each board and simplified by the fact that there are only four sections to adjust. Shielding is provided by the walls of the card nest in which the boards reside. The three cascaded filters have a 3-dB bandwidth of 8.5 MHz and a combined peak-to-peak ripple less than 1.2 dB.

Third Conversion

The final stage of the RF section receiver is the third-converter assembly which translates the 46-MHz second IF to the final IF centered at 6 MHz. A 40-MHz LO is provided by the reference at 3 dBm. Following the conversion are two wideband operational amplifiers, which provide the final gain and buffering before the signal is sent to the IF section. There is a 1-dB gain step that can be switched by the IF section and a manual gain adjustment to compensate for normal gain variance in the manufacturing process.

Local Oscillator

Since the IF section has the ability to tune with millihertz resolution to any frequency within its dc-to-10-MHz input bandwidth, the LO for the RF section does not have to replicate this function. The RF section down-converts the 1800-MHz input span to within the frequency range of the IF section and the digital LO tunes to the desired center frequency. In traditional spectrum analyzers, multiloop LOs are designed to realize millihertz resolution. Multiloop designs mean at least three phase-locked loops in the main LO. The step loop provides coarse frequency resolution over a wide range of frequencies. The interpolation loop tunes across a narrow span but with high frequency resolution. The sum loop combines the outputs of the two loops. In contrast, the HP 89440A RF section has a single-loop LO that tunes in 1-MHz steps. The elimination of the sum and interpolation loops means significant savings in complexity and reduced development risk within the LO section of the HP 89440A. This trade-off is not without consequence because a coarse LO resolution reduces the analysis bandwidth of the instrument. The IF bandwidth of the receiver chain supports an 8-MHz analysis bandwidth but an arbitrary center frequency at the input can only be placed within ±0.5 MHz of the IF center frequency because of the step size of the LO. Hence, for an arbitrary center frequency the 8-MHz IF bandwidth is reduced to a 7-MHz maximum analysis bandwidth.

The frequency range of the LO in the RF section basically starts at the first IF frequency and tunes to a frequency 1.8 GHz above that. A commercial YIG (yttrium iron garnet) oscillator was selected to cover 2.4 to 4.3 GHz. The output of the YIG oscillator goes to the LO distribution amplifier

and is also fed back to the synthesizer phase-locked loop (see Fig. 2).

Since the output frequencies involved are beyond the reach of programmable counters, the feedback path includes a down-conversion stage. The LO frequency range of 2.452 to 4.242 GHz is down-converted with one of three offsets (2.4, 3.0, or 3.6 GHz). The offset frequency is chosen to produce a down-converted signal between 42 and 642 MHz, which is within the range of the RF section programmable counter.

The value of the programmable counter (N) is chosen to divide the counter input frequency down to 1 MHz (N = 42 to 642). The output of the divider is phase-detected against a 1-MHz signal derived from the 40-MHz reference. This establishes the 1-MHz step size for the LO. The sign of the phase-locked loop must be switchable because the YIG oscillator tunes above and below the offset frequencies. The sign is switched by swapping the reference and feedback signals at the phase detector.

To compensate for the wide range of the loop gain (because N ranges from 42 to 642), a programmable gain block with 25-dB gain variation is added to the loop. Finally, a DAC-driven coarse tuning signal is used to steer the YIG oscillator into the lock range.

The LO is distributed across four printed circuit boards and one microwave plate assembly. The circuit boards contain the 600-MHz reference, the 40-MHz reference, the frequency counter, the phase detector, and the YIG driver. The circuits on the microwave plate are fabricated with cyanate ester printed circuit boards and include the frequency multipliers and the YIG down-conversion.

LO Distribution

The LO distribution amplifier is built around a packaged GaAs MMIC amplifier designed by HP's Microwave Technology Division. This amplifier has dual outputs which are used to supply the LO signal to both the RF receiver and the RF source sections.

LO Offsets

The three offset frequencies (2.4, 3.0, and 3.6 GHz) are generated by multiplying 600 MHz by integer values. The 600 MHz comes onto the LO microwave plate assembly and is split to provide signals for both the offset multipliers and the second-LO multipliers. P-i-n diode switches select the path to the activated offset multiplier and each path has its own final stage of amplification before the multiplier. Schottky diodes are used as the harmonic generating devices in each of the multipliers. At the output of each multiplier is a two-section coaxial filter to suppress the adjacent 600-MHz harmonic. The second LO (2.4 GHz) is generated in a similar manner.

YIG Down-Conversion

Down-conversion is implemented with a 7-dBm microwave mixer. Because of the low-level signals involved and the wide frequency span of the output IF (10 MHz to 700 MHz), the noise figure of the postconversion amplifier is important. The broadband noise within the 700-MHz span is sampled
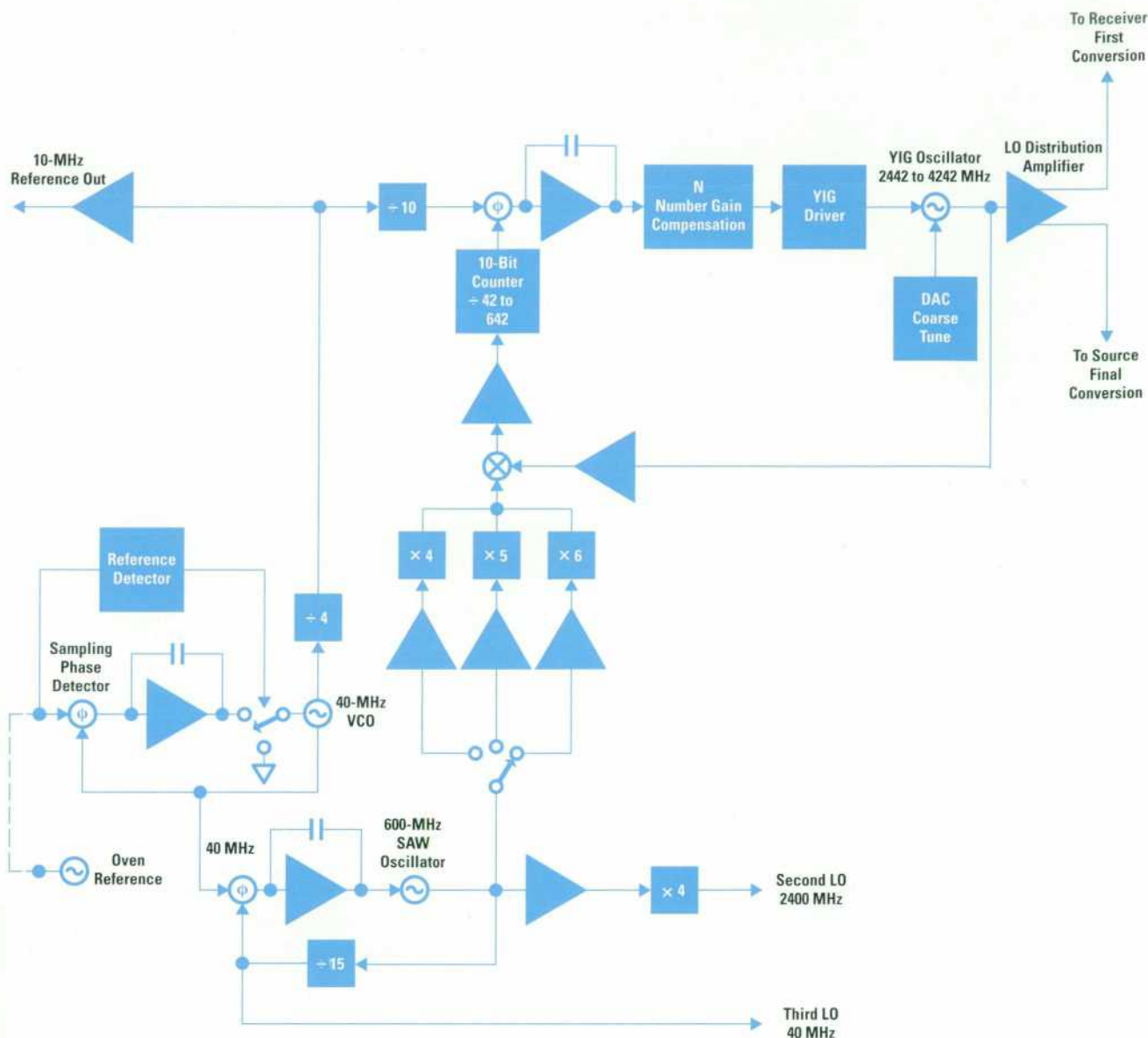
Fig. 2. HP 89440A RF section local oscillator.

down by the digital dividers (10log[700 MHz/1 MHz] = 28.5 dB noise gain) and can contribute to the phase noise pedestal.

10-Bit UHF Counter

A 10-bit counter is required to accommodate all integer divide numbers from 42 to 642. Since high-speed commercial counters are limited to 8 bits, a proprietary counter circuit was developed using ECL integrated circuits. This implementation allows the use of inexpensive commercial components to achieve maximum count frequencies of over 700 MHz, whereas standard design implementations only permit a 500-MHz maximum input frequency. This improvement was critical in realizing the design efficiency of the LO block diagram.

YIG Drivers and Tuning

The YIG oscillator has a main coil for coarse tuning and an FM coil for locking the phase-locked loop. Both coils are driven by voltage-to-current converters and high-current drivers.

The main coil is controlled by a 12-bit DAC, providing better than 1-MHz frequency resolution. Because of the extremely high gain of the main coil (20 GHz/A), noise filters are required. Ordinarily this would significantly affect LO switching time, so a speed-up circuit was designed to precharge the filter elements to their final values, greatly improving settling time.

The main coil tolerance is far too large for the available FM coil tuning range, so an automatic YIG tuning calibration is performed periodically, initiated by the instrument calibration timer. This improves the absolute accuracy of the main coil from ±400 MHz to ±2 MHz.

Frequency References

The 600-MHz reference is the source for the offset frequencies, the second LO, and the third LO (40 MHz). It is based on a phase-locked 600-MHz SAW (surface acoustic wave) oscillator. It is designed so that it does not contribute to the

LO system phase noise. The loop is locked to the 40-MHz reference.

The 40-MHz reference provides subharmonic (1-MHz, 2-MHz, 5-MHz, 10-MHz) reference locking for the instrument and acts as a cleanup loop for the user's external reference, if present. It contains a 40-MHz voltage-controlled crystal oscillator and a phase-locked loop with a sampling phase detector. In the absence of a user-provided external reference, the reference locks to the internal high-stability ovenized 10-MHz reference.

The 10-MHz ovenized reference is the widely used HP 10811, which has ultralow phase noise and extremely high temperature stability.

RF Source

The HP 89440A has an optional RF source to provide stimulus signals for a variety of test purposes. Output signals produced by the source are in the 2-MHz-to-1.8-GHz frequency range when the instrument is in the RF vector or demodulation modes. The HP 89440A is not a traditional swept analyzer, so the source provides several signal types in addition to the standard sine output to satisfy various measurement needs. Available HP 89440A source output signal types are sine, chirp, pseudorandom noise, and arbitrary (see "A Versatile Tracking and Arbitrary Source" on page 54). One or more of these source types are available at baseband in most modes using the RF source bypass path. Chirp or sine signal levels available from the RF source are +13 to –27 dBm.

As shown in Fig. 1, the source circuitry of the HP 89440A RF section is basically a frequency converter for the source output signal from the HP 89440A IF section. The circuitry that generates the IF section source output signal is described in the article on page 31. The source signal originating in the IF section is a chirp or noise signal centered at 6 MHz, a sine wave between 2.5 and 9.5 MHz, or an arbitrary signal. In the RF section, this source signal is fed to the first conversion assembly, which provides part of the signal switching functionality to bypass the RF section source or to route a calibration signal to the receiver. Amplitude modulation circuitry allows the signal to be modulated at a maximum frequency of about 1 MHz. This can be useful for impressing certain types of synchronization signals on the RF source output signal. The signal is then mixed with a 40-MHz fixed signal from the RF section reference to upconvert it to a 46-MHz IF center frequency. Following the first conversion assembly are two cascaded 46-MHz IF filter assemblies identical to those used in the RF receiver. To accommodate the wideband noise and chirp signals from the IF section source, the IF bandwidth of the entire RF source is approximately 8 MHz. This bandwidth also facilitates offsetting the frequency of sine waves up to 3.5 MHz from the tuned center frequency.

The 46-MHz IF signal is passed to the second conversion assembly where it is again upconverted, this time to a 2.446-GHz IF center frequency, by a mixer whose LO port is driven with 2.4 GHz. The fixed 2.4-GHz signal to drive the second mixer originates in the RF section reference and is amplified by the second-LO amplifier assembly located on the RF source assembly. Besides amplification of the 2.4-GHz signal, the second-LO amplifier assembly provides reverse isolation to prevent the source IF signals (particularly the 2.446-GHz IF) from leaking back through the 2.4-GHz LO distribution circuitry and into the receiver IF. If this were to happen, receiver sensitivity would be compromised when the source is functioning.

Following the second mixer are two IF amplifiers. The output of the last IF amplifier leaves the assembly and goes to the 2.446-GHz second-IF filter. This four-section coupled coaxial resonator bandpass filter has a 17-MHz bandwidth and is identical to the receiver IF filter. The filter is required to provide adequate rejection for the 2.4-GHz LO second-mixer feedthrough and the mixer lower sideband product centered at 2.354 GHz while maintaining reasonable insertion loss. Physically, the filter is optimized to take advantage of the microwave plate packaging scheme which also helped minimize its cost (see "Microwave Plate Assembly" on page 50). The filter is tuned by adjusting four self-locking tuning elements using a simple, noniterative tuning procedure. A tuning port on the filter housing aids in the procedure.

After the 2.446-GHz IF filter, additional IF amplification is provided by two amplifier stages on the IF gain assembly before the signal is applied to the final conversion module. An existing HP Microwave Technology Division design, the final conversion module was leveraged because its functionality is a good fit for the RF section source. The final conversion module uses a GaAs MMIC to mix the 2.446-GHz IF signal with a variable LO signal of 2.452 to 4.242 GHz to produce the baseband output signal, which is then applied to a thin-film low-pass filter. As in the case of the 2.4-GHz LO signal, sufficient isolation is needed in the variable-LO distribution path to the RF source to ensure that the 2.446-GHz source IF signal does not leak into the receiver IF. Some reverse isolation is afforded by a GaAs MMIC LO driver within the final conversion module, but this is insufficient by itself. Additional isolation is supplied by the isolation amplifier stage located between the final conversion module LO input and the main LO distribution amplifier output. A GaAs MMIC output amplifier brings the signal to the proper output level. Another thin-film low-pass filter follows the output amplifier. The GaAs devices inside the final conversion module require their dc supplies to power up and power down in a prescribed sequence. Since the RF section main power supply does not provide the proper sequencing, a local power supply assembly on the RF source accomplishes this along with voltage regulation and current limiting for all of the final conversion assembly dc supplies.

A step attenuator assembly is located at the output of the final conversion assembly. The step attenuator assembly applies 0, 10, 20, or 30 dB of attenuation to the output signal for amplitude control. The IF section provides fine amplitude control. An amplitude equalizer on the step attenuator assembly helps correct for roll-off in the final conversion assembly and the step attenuator assembly. A signal generated in the IF section can be routed through the RF source switching circuitry of the first conversion assembly and the step attenuator assembly to the RF receiver input for receiver IF calibration. The output of the RF source can also be routed to the receiver input for calibration and leveling of the RF source output. The RF source circuitry can be bypassed so that the IF section source output is available at the RF source output connector.

A Versatile Tracking and Arbitrary Source

The measurement power of the HP 89410A and 89440A vector signal analyzers is greatly enhanced by the inclusion of a versatile signal source. Available outputs are two tracking signals—pseudorandom noise and periodic chirp—and two independent source types—sine and arbitrary. The circuits that generate these signals are described in the accompanying article and in the article on page 31.

The pseudorandom noise is generated with a sequence length of $2^{39} - 1$ samples, which is about six hours long in the widest span and proportionately longer as the span is reduced. Its probability density function approximates a Gaussian probability density function to a width of at least ±3 standard deviations. These properties make it function like a true Gaussian random noise source for virtually all applications.

The periodic chirp is a swept sine wave that covers the span being measured by the receiver. Because it is generated as a swept sine, it has relatively constant amplitude when observed in the time domain, with the exception of bandlimiting filter ringing. As a result of being "flat" in the time domain, it is not absolutely flat when observed in the frequency domain.

The arbitrary source uses time-domain data from a user-selected stored data register and reconstructs the waveform with the original span and center frequency. The stored data can consist of measured waveforms, data generated on a computer and downloaded to the instrument, or the result of user math on either or both. The frequency of downloaded data is set by the headers that are loaded with it. This capability makes it simple to create, for example, an independent (not tracking) chirp source. To record the chirp, the receiver is connected to the tracking chirp source and the fixed span and center frequency are chosen. The time data is stored and is then available as an arbitrary source, at the chosen span, center frequency, and number of time points. It is worth noting that the arbitrary source data can be complex, that is, it can represent an I-Q (in-phase and quadrature) signal with the center frequency and span specified in the data register header information. Thus it is easy to obtain data communications waveforms and other hard-to-produce stimulus waveforms as long as the time-length constraints are acceptable. A number of these waveforms are included on a disk supplied with the instruments and can replace a number of external waveform generators.

Both the arbitrary source and the sine source can be placed anywhere within the 0-to-10-MHz HP 89410A frequency range and anywhere within 3 MHz (3.5 MHz

for sine) of the HP 89440A center frequency in vector and demodulation measurement modes. The chirp and random sources are designed to cover the frequency range being measured when the instrument is in the vector and demodulation modes. All four of these source types are generated in hardware—and software in the case of chirp and arbitrary—that to a great extent mirrors the signal processing in the receiver. Fig. 1 on page 48 illustrates this symmetry very clearly. The local oscillator (LO) frequencies and IF filter frequencies are all the same—the signal flow is simply in the opposite direction. This is also seen in the IF section (see Fig. 1 on page 32). The digital-to-analog converter (DAC) in the source performs the opposite function of the analog-to-digital converter (ADC) in the receiver. The 10-MHz reconstruction filter after the DAC suppresses alias components in a manner very much like the anti-alias filter preceding the ADC. One slight exception is that the source reconstruction filter contains additional peaking in the frequency domain to match the implicit $\sin(\omega t/2)/\omega t/2$ attenuation of the DAC. This difference arises because the aperture of the ADC is much narrower than the full sample width aperture of the DAC.

Before the DAC, the source is also a mirror image of the receiver. The decimating digital filter and LO ICs are designed to work in reverse, forming a complex interpolation filter. For the independent source types, the bandwidth and LO frequency of the digital source hardware are set independently from those of the receiver.

The source RAM contains the waveform to be output, but at baseband, before all the frequency translations of the upconversion chain. In the case of the sine source, this is simply a dc value—the upconversion shifts this 0-Hz signal to the frequency desired.

The chirp and arbitrary waveforms for the source RAM are computed in software. To implement arbitrary source spans given the fixed hardware sample rates, resampling must be performed, just as in the case of the receiver. The source resampling filter uses the same filter coefficients as the resampling filter in the receiver.

For computational simplicity, this filter has an alias-protected bandwidth of only one fourth of its incoming sample rate. It is designed to operate with incoming samples that are two times oversampled in the time domain, that is, with an extra sample point interpolated between each pair of stored data register samples.

The first conversion assembly and the two 46-MHz IF filter assemblies reside in the card nest structure within the RF section. The rest of the RF source circuitry resides on the source microwave plate assembly, which is a support, shielding and interconnection structure (see "Microwave Plate Assembly" on page 50). Two-sided, 0.030-inch-thick cyanate ester printed circuit board was used in the RF source because of its RF performance.

Controller

The control of the RF section is vested in a resident Motorola MC68HC11 microcontroller, which controls the hardware setup, communicates with the IF section via RS-232, stores HP 89440A calibration data in flash memory, and programmatically performs the YIG tuning calibration.

Calibration Contributions

The HP 89440A is one of the most accurate RF analyzers ever produced by Hewlett-Packard. At room temperature, at any frequency within the 2-to-1800-MHz measurement band, and at any level from 70 dB below full scale to full scale, HP 89440A level measurements are typically accurate within ±0.5 dB. Level accuracy is important because a vital application for the HP 89440A is true-rms power measurements on complex signals. Since the HP 89440A is a vector signal analyzer, the relative phase accuracy (deviation from linear

phase) is also important. To make accurate vector measurements and perform accurate demodulation of complex signals, the relative phase over narrow frequency spans must be accurate within a few tenths of a degree. This amplitude and phase accuracy is achieved through extensive self-calibration coupled with an extensive factory characterization.

The HP 89440A self-calibration routine calibrates both the IF section and the RF section. Since the calibration of the IF section is identical to the calibration of the HP 89410A as a separate instrument, only the HP 89440A RF section self-calibration is discussed here. Self-calibration can be set up to occur automatically at predetermined intervals to compensate for temperature drift. In addition to the calibration required for amplitude and phase accuracy, the HP 89440A self-calibration performs many other functions including source accuracy calibration, front-end dc offset compensation, trigger calibration, and first-LO feedthrough nulling.

Calibrator Hardware

Some parts of the HP 89440A self-calibration require a precise calibration signal. An internal calibration signal is generated by taking a single bit from the source RAM, reclocking the transitions, and clipping it to a precisely controlled amplitude. During the self-calibration the calibrator

In the receiver case, the oversampled data is generated by simply not performing the final decimation operation at the output of the digital filters. In the source case, this two times oversampled data is generated in software by inserting a sample point of value zero between each pair of stored data points. This simplistic interpolation leaves a huge alias component centered at the original sample rate, now half the new sample rate. This is the standard interpolation problem as seen in the frequency domain and is solved by additional filtering. The frequency response chosen for this filter is a raised cosine in the frequency domain. This interpolation filter is separate from the resampling filter, which follows it.

The source software also corrects the chirp and arbitrary waveforms for the frequency response errors (both amplitude and phase) of the digital and analog reconstruction filters. The analog correction is for the nominal frequency response (not individually measured) of the reconstruction filter in the HP 89410A; the filters in the HP 89440A RF section are not corrected for. The resampling filter also has amplitude errors which are corrected for.

The correction and oversampling are both performed in the frequency domain on overlapping blocks of data. This allows the use of frequency-domain correction data and makes the raised cosine oversampling filter easy to implement. The overlapping block approach conserves memory and removes size limitations from this portion of the signal processing. The source length limits are determined by the maximum stored data register length and the size of the source RAM.

The source RAM also places another interesting limit on the periodic chirp and arbitrary source waveforms. In spite of the arbitrary span capability of the resampling process, the length of one period of the source, in sample points, is limited to an integer number of source RAM samples. This is because the source RAM follows the resampling rather than precedes it. To assist the user in chirp measurements, the instrument defaults to a "chirp periodic" resolution of span choices when the chirp source is on (this is easily defeated if necessary). This gives receiver time records that are exactly the length of the source period for true periodic chirp measurements.

The pseudorandom source data is generated entirely in hardware using the source RAM as a Gaussian probability density function lookup table, rather than as a source of actual sample points. Because of this, it is not corrected for frequency response errors other than the $\sin(\omega t/2)/\omega t/2$ correction included in the analog reconstruction filter. Resampling is not available, so the spans are limited to the factor-of-two choices implemented by the digital reconstruction filters. This is not a major problem—the source span is chosen to include the receiver span.

The chirp and arbitrary source types have a single-shot capability in addition to the normal repeating, periodic output mode. When this mode is chosen, the source is triggered by the measurement trigger. This gives only one source burst for each measurement, which can be positioned relative to the measurement by adjusting the trigger delay. This is functional in all but the source trigger mode (the measurement is triggered by the source in this mode) so the analyzer can be set up to make a single-burst measurement conditioned by an external or internal trigger. The source can be summed into an existing signal, such as a TV waveform, allowing a waveform to be inserted onto a particular TV line if the appropriate TV line trigger is input to the analyzer's external trigger port. One minor artifact is that there is a variable latency (from zero to one source RAM sample) between the trigger signal and the start of reading out the contents of the source RAM.

Another valuable, application-driven feature is that when the source is forced to the 0-to-10-MHz mode while the HP 89440A 2-to-1800-MHz receiver mode is engaged, the source outputs its waveform across half the vector span, which is equal to the demodulation span. With the second input channel option, this allows direct frequency response measurements of modulation systems, including phase-locked loops. In this latter application, the source can be used to inject an error waveform into a portion of the phase processing circuitry. A reference response can be measured at another place in the phase circuitry using the second input channel, which also operates at half span. The actual phase can then be measured with the RF receiver in phase demodulation mode with the instrument center frequency placed directly at the RF frequency being analyzed. This can also be applied to other modulation systems.

Don Hiller
Design Engineer
Lake Stevens Instrument Division

is internally connected to the input channel and the input remains terminated—no manual connections are required.

The calibrator output level is calibrated at the factory or field service center by comparing a known signal at the input (supplied by the factory or service test station) with the calibrator signal. The difference is stored in nonvolatile RAM.

IF Section Vector Calibration

The HP 89440A self-calibration routine generates complex-valued correction data for the IF section input channels over a dc-to-10-MHz frequency band. These corrections are valid for measurements made at the IF section's input connectors when the analyzer is in the dc-to-10-MHz baseband receiver mode. The correction data is combined with the known response of the digital filters to provide complete calibration of both amplitude and phase.

The correction data is computed from calibration data obtained by passing a calibration signal through the IF. The calibration signal is generated from a 256-bit binary sequence clocked at 25.6 MHz. This produces a comb spectrum with spectral lines spaced every 100 kHz, each with known energy level and phase relative to the source trigger. Only 24 of the spectral lines are actually used to model the frequency response of the IF. By limiting the number of frequencies measured, the amount of memory required to store the calibration data is reduced. Also, the binary sequence that generates the comb spectrum can be optimized to maximize the energy level at the calibration frequencies, thereby improving the SNR of the calibration measurement. The frequencies measured are not evenly spaced, but were chosen based on the characteristics of the filters to be characterized. To compute correction data, a spline routine is used to interpolate between the calibration points. The calibration data is complex, so the real data and the imaginary data are interpolated separately. Before interpolation, excess phase caused by delay is removed from the data. This improves the accuracy of the interpolation by reducing the order of the data. In other words, the spline routines are interpolating a lower-order curve. After interpolation the delay is reintroduced into the correction data.

When measuring the calibration signal, source triggering and time averaging are used to reduce the noise of the measurement. After the first measurement of the calibration signal is made, the bit sequence is inverted (ones become zeros and vice versa), and another time-averaged measurement is made. These results are combined and compared to the known spectrum of the calibration signal to produce the

vector corrections for the input channel at the specific frequencies generated by the calibration signal.

The vector corrections are measured for each input channel configuration that is likely to produce a different frequency response. Six primary calibrations are performed. Four secondary calibrations are performed on the IF section's 20-dB attenuator and 13-or-3-dB amplifier. The secondary calibrations are similar to the primary calibrations except that the measurements are only made at one frequency. The results of the six primary calibrations and the four secondary calibrations are stored in nonvolatile RAM and combined in a variety of ways to compute a correction vector for each input range.

IF Section DC Offset Calibration

Each IF section input channel has a dc offset DAC that is used to compensate for residual dc offsets in the analog input circuit and in the ADC. The autozero calibration measures the residual dc and adjusts the dc offset DAC to minimize the amount of dc offset. This calibration must be performed for each different configuration of the input channel's active elements. Twenty dc offset calibrations are performed for each channel.

IF Section Source Calibration

The IF section's dc-to-10-MHz source is calibrated by internally connecting the source to the previously calibrated input channel. The uncorrected source is programmed for a sine wave at 1.5 MHz and the level of the signal is measured. The resulting amplitude correction factor is applied to source levels entered by the operator.

The source has a dc offset DAC to provide user-selected dc offsets. The gain of the dc offset DAC is measured and this correction is applied to the source dc offset value entered by the operator. The source dc offset DAC is also used to compensate for any dc offsets in the source DAC and associated analog circuits. The calibration routine finds the setting of the source DAC that will produce zero volts dc at the output.

Trigger Calibration

Three separate calibrations are required for the IF section's trigger circuits. First, the trigger dc offset DAC is used to compensate for any residual dc offsets in the trigger circuits. The calibration routine must determine the correct settings for this DAC. Since the dc offset of the trigger cannot be measured directly, it must be inferred from another measurement. To determine the dc offset, the source is programmed for a 1-MHz sine wave and internally connected to the input channel. Input triggering is used to measure the amplitude of the sine wave at the trigger point. Two measurements are made, one using a positive trigger slope and one a negative trigger slope. If there is no dc offset in the trigger then the two measurements will have equal amplitudes but opposite polarities.

Each of the trigger types (external, channel, and source) has a different amount of delay relative to the input signal. In addition, there is a different delay for each receiver mode. The trigger calibration routine must measure these delays so the correct trigger point can be determined. To measure delay, the calibrator is programmed for an 800-kHz square wave and is internally connected to the input channel. The phase of the calibration signal is measured using each trigger

type and the corresponding delays are stored in nonvolatile RAM.

The partial-trigger delay counter allows the analyzer to determine the correct trigger point even if the trigger point occurs between ADC samples. The partial-trigger delay counter is based on an analog pulse stretcher that must be characterized before stable triggering can be achieved. The two characteristics that must be determined are the minimum number the counter will return for zero trigger delay relative to the sample clock, and the maximum number returned for a trigger with slightly less than one sample clock of delay. To ascertain these values, the internal source is programmed to generate a sine wave and this signal is internally connected to the input channel. The maximum and minimum numbers are read from the partial-trigger delay counter as the trigger point is moved relative to the sample clock. The trigger point is adjusted by changing the phase of the sine wave. The maximum and minimum numbers are used to find the coefficients of a first-order equation and this equation is used to compute the delay for any other counter value returned.

RF Section Vector Calibration

The RF section calibration provides amplitude and phase corrections that are valid for measurements made at the input connector of the RF section when the HP 89440A is in the 2-to-1800-MHz receiver mode. The HP 89440A self-calibration routine generates vector correction data for the analyzer's 7-MHz IF. The IF section's input channel is considered part of the IF for this calibration—the correction vectors previously calculated for the IF section's input channel are not used.

The RF section vector calibration is almost identical to the IF section vector calibration. The calibration signal is internally connected to the RF input with the RF section tuned to 6 MHz. The RF section LO adds an arbitrary phase rotation to the measured data and the self-calibration routine must determine the amount of phase rotation for each measurement and correct the data before the results are averaged. Since changes in the RF attenuator do not affect the frequency response of the IF, the vector calibration is only performed at the 10-dB attenuator setting (–20-dBm range).

The RF section attenuators are calibrated by using the internal calibrator and the IF section input channel to measure attenuation relative to the 10-dB attenuator setting. Only five attenuator combinations are measured and these measurements are combined in a variety of ways to produce gain corrections at all 12 attenuator settings.

LO Feedthrough Nulling

Two DACs in the RF section must be adjusted to minimize the amount of receiver first-LO feedthrough. Minimizing the LO feedthrough reduces low-frequency residuals and reduces distortion problems caused by having multiple tones in the IF at low frequencies. The LO feedthrough is minimized by a circuit that samples the LO and adds a small amount of the LO to the signal in the first IF. Both in-phase and quadrature components of the LO are added. The DACs are used to adjust the in-phase and quadrature components to null the LO feedthrough.

The calibration software adjusts the DACs to produce a minimal amount of LO feedthrough with the RF section tuned to 6 MHz. At this frequency the LO feedthrough term shows up at 12 MHz, requiring the anti-alias filter in the IF section to be disabled for the calibration measurement. Since the feedthrough term is the only signal present, aliasing is not a concern. To find the optimal DAC settings, the calibration routine uses what is commonly referred to as a golden section search.[1] The search is carried out first on one DAC and then on the other. However, since the in-phase and quadrature-phase signals used to cancel the feedthrough term are not in perfect quadrature, there is a certain amount of interaction between the two DACs. Several iterations are required to find the optimal DAC settings. To minimize the search time, the DAC settings obtained in the previous calibration are used as a starting point in the search.

RF Source Calibration

The RF source is calibrated by internally connecting the source to the RF input and measuring the uncorrected source level over the 1800-MHz frequency range. The resulting correction factors are stored in nonvolatile RAM and are applied to the user-selected source level.

RF Section Factory Calibration

To achieve a ±0.5-dB typical accuracy, the flatness of the RF receiver must be extensively characterized at the factory and at field service centers. This is accomplished by using the test setups shown in Fig. 3. In the first step (Fig. 3a), the tracking and flatness of the power splitter and cables are characterized by connecting a second power meter channel to the end of the test cable. The signal generator is programmed for each calibration frequency and the difference between the readings of the two power meter channels is stored. This step transfers the flatness of the second power meter channel to the test setup with very little degradation.

In the second setup (Fig. 3b), the test cable is connected to the RF input and the gain of the RF section is measured at each calibration frequency. The level of the input signal is measured by reading the power meter and correcting the reading using the stored results from the first step. The input to the IF section (which is always at the same frequency and at nearly the same level) is measured with the HP 89440A in dc-to-10-MHz receiver mode. The ratio of the input signal level and the IF signal level is the gain of the RF section. The absolute accuracy of the gain measurement is not important since only the flatness of the RF section is of interest. The actual gain of the RF section is calibrated by the self-calibration described previously.

The RF flatness is measured for all attenuator settings of the RF section and the results are stored in nonvolatile RAM within the RF unit. In all, over 1800 calibration points are stored in the RF section. At power-up, the IF section reads the RF section calibration data and uses this data to correct level measurements. Storing the RF calibration data in the RF section allows any IF section to operate with any RF section. An HP Instrument BASIC program, running in the HP 89440A, allows service centers or customers to perform this calibration easily.

(a)

(b)
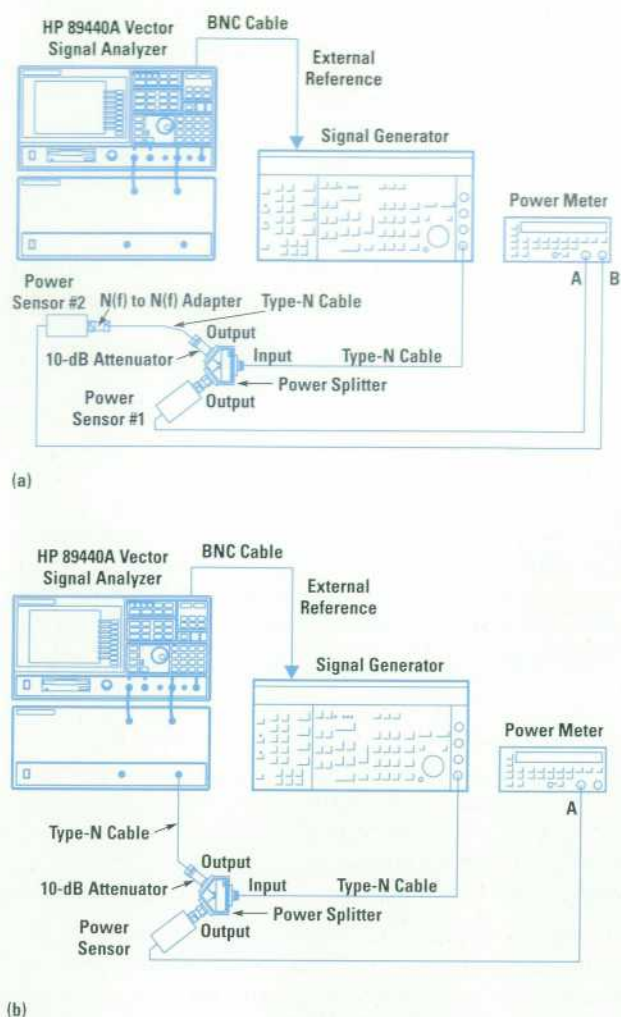
Fig. 3. Setups for characterizing the receiver flatness of the HP 89440A RF section. (a) Setup for characterizing the tracking and flatness of the power splitter and cables. (b) Setup for measuring the gain of the RF section as a function of frequency.

Performance Verification

The factory calibration, coupled with the automatic self-calibration, produces an instrument with performance that can be very difficult to verify. The amplitude accuracy and IF flatness of the HP 89440A are verified using a method identical to the factory RF calibration shown in Fig. 3b. Using this method, the HP 89440A's measured level accuracy at room temperature conditions is typically better than the measurement uncertainty.

To verify the vector performance of the HP 89440A, the deviation from linear phase (relative phase error) within one IF bandwidth must be measured. Ideally, we could calculate the worst-case relative phase error from the worst-case IF flatness because the IF correction data consists of complex vectors. The amplitude and phase corrections are not independent and any phase errors would have corresponding amplitude errors. In practice, however, the scalar RF flatness calibration data is also used to apply second-order corrections to the IF levels. This means that it is possible for

Vector Measurements beyond 1.8 GHz

The HP 89410A and HP 89440A vector signal analyzers provide unprecedented analysis capabilities for looking at complex signals. Driven by a need for more and larger spectrum requirements, many communication schemes are moving to higher frequencies. The HP 89411A provides a way to apply the powerful analysis capabilities of the HP 89410A to signals that lie above 1.8 GHz. It does not have the level of integration or the complete set of features provided by the HP 89440A, but it does allow a user to view and analyze signals above 1.8 GHz as they have probably never done before. The HP 89411A provides this capability by translating the auxiliary IF (intermediate frequency) output from one of a number of RF and microwave spectrum analyzers to a center frequency within the analysis range of the HP 89410A. It will directly translate the IF outputs of analyzers such as the HP 71000 Series and HP 8566A/B. Customers who already own one of these analyzers can now extend their instrument's capabilities by combining them with the measurement capabilities of the HP 89410A. In addition, an integrated solution can be constructed using features of the HP 89410A such as HP Instrument BASIC, HP-IB (IEEE 488, IEC 625) controller capability, and built-in firmware features describing external down-converters such as the HP 89411A.

The HP 89411A is fixed-frequency down-converter that translates a band of signals up to 7 MHz wide from a center frequency of 21.4 MHz to a band centered at 5.6 MHz. As the block diagram, Fig. 1, shows, it does this in two frequency conversion steps and provides conversion gain and image filtering for the signal band of interest. Its internal oscillators are phase-locked to an external 10-MHz reference frequency to allow high-quality magnitude and phase measurements on a variety of signals. The HP 89411A package has the same footprint as the HP 89410A so it can be stacked directly below it.

The HP 89411A's performance goals were aimed at making it appear largely invisible to the user. Its broadband noise, distortion, spurious, and phase noise performance are similar to the HP 89410A's. In a typical system the overall performance will be determined by the RF or microwave spectrum analyzer.

To support the HP 89411A as well as other down-converters, several firmware features are incorporated into the HP 89410A. A menu under the instrument mode key allows the user to define some of the attributes of an external receiver such as the HP 89411A. These attributes include the tuning range, IF bandwidth, display "mirroring", and enabling of HP-IB control. When the external receiver mode is selected the instrument's x axis and markers are labeled with the actual input frequency. Selecting the mirroring function instructs the HP 89410A to mirror or flip the spectrum display. This capability is provided to undo the mirroring that can occur as a result of the mixing scheme used in some RF and microwave spectrum analyzers. The HP-IB control capability is directly compatible with all HP 71000 Series spectrum analyzers and with the HP 8566A/B. When changing frequencies the user simply enters the new center frequency and span on the HP 89410A. The HP 89410A then checks to see that the parameters are not out of range for the defined external receiver setup, and then can optionally issue HP-IB commands to tune the RF or microwave analyzer to the desired frequency. The user can account for conversion gain differences in various setups by applying trace math to the
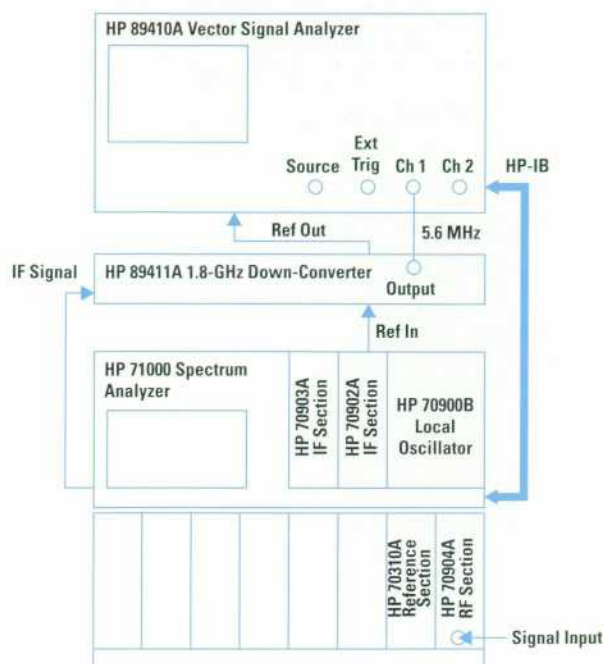
Fig. 2. Connection diagram for Modular Measurement System hookup with an HP 71000 Series spectrum analyzer.

measured results. If more complicated control functions are needed the HP 89410A can be configured as an HP-IB bus controller, and with an HP Instrument BASIC program, an integrated down-conversion system can be constructed. The instrument connections for such a system are shown in Fig. 2.

Joe Tarantino
Design Engineer
Lake Stevens Instrument Division
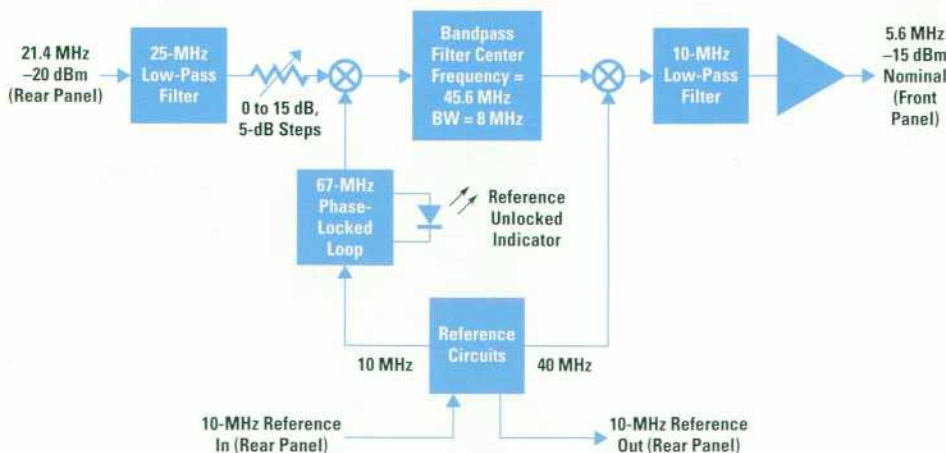
Fig. 1. Block diagram of the HP 89411A down-converter.

the analyzer's IF flatness to be significantly better than the measured phase accuracy would imply.

To measure deviation from linear phase, a test signal is needed that has three or more spectral components with known phase relationships between them. Since HP 89440A phase measurements have arbitrary delay and offset terms, the phase of a single tone and the phase relationship between two tones are arbitrary. However, the phase difference between two tones relative to the phase difference between one of these tones and a third tone is not arbitrary. For example, suppose the HP 89440A is used to measure an amplitude modulated carrier. The difference between the carrier phase and the phase of the upper AM sideband should have the same magnitude as the difference between the carrier phase and the phase of the lower AM sideband, but the opposite sign. If the source has no incidental PM, then the sum of the upper and lower sideband phase differences relative to the carrier phase is a measure of the analyzer's deviation from linear phase. A method similar to this is used at the factory to measure the deviation from linear phase of the HP 89440A's 7-MHz IF bandwidth.

Acknowledgments

Reference

1. W. Press, et al, *Numerical Recipes in C*, Cambridge University Press, 1988.

Optical Spectrum Analyzers with High Dynamic Range and Excellent Input Sensitivity

The diffraction-grating-based HP 71450A and 71451A optical spectrum analyzers provide the basic spectral measurement of optical power versus wavelength and advanced functions for measuring and characterizing LEDs, DFB Lasers, and Fabry-Perot lasers.

by David A. Bailey and James R. Stimple

The telecommunications industry is one of the most lively and interesting areas of the electronics industry today. The development of high-performance fiber-optic systems requires the ultimate performance of components such as laser sources, fibers, optical amplifiers, and receivers. Accurately measuring the performance of these components and confirming their operation in the system is essential to prove the design. The optical spectrum analyzer is one of the most valuable tools for making these measurements. The HP 71450A and 71451A optical spectrum analyzers are designed to make spectral measurements in the laboratory and in a production environment. The HP 71451A optical spectrum analyzer is shown in Fig. 1.

Both analyzers can make spectral measurements between 600 nm and 1700 nm on LEDs, Fabry-Perot lasers, distributed feedback lasers, and erbium-doped fiber amplifiers. These basic measurement capabilities are described later in this article. A new double-pass monochromator enables the analyzers to provide the high dynamic range of double-monochromator instruments (55 dB at 0.5 nm from the peak) and the sensitivity of single-monochromator instruments

(better than –90 dBm). Fig. 2 shows block diagrams of the HP 71450A and 71451A optical spectrum analyzers.

The HP 71451A optical spectrum analyzer offers measurement capabilities that go beyond basic optical spectral measurements by providing four measurement ports: monochromator input, photodetector input, monochromator output, and transimpedance amplifier input (see Fig. 2b). These ports allow five different modes of operation:

- Optical spectrum analyzer mode. This mode provides basic optical spectrum analysis with precise amplitude accuracy and less than 0.5 dB polarization sensitivity.
- Preselector mode. This mode allows front-panel output of light that passes through the monochromator. Wavelength division multiplexed channels, individual modes of Fabry-Perot lasers, and selected widths of LEDs or white light sources can be output on 62-μm fiber for further use or analysis.
- Stimulus-response mode. When broad spontaneous emission light is applied to the monochromator input, the monochromator output becomes a variable-wavelength source. The user can pass the light through a device or filter, and then
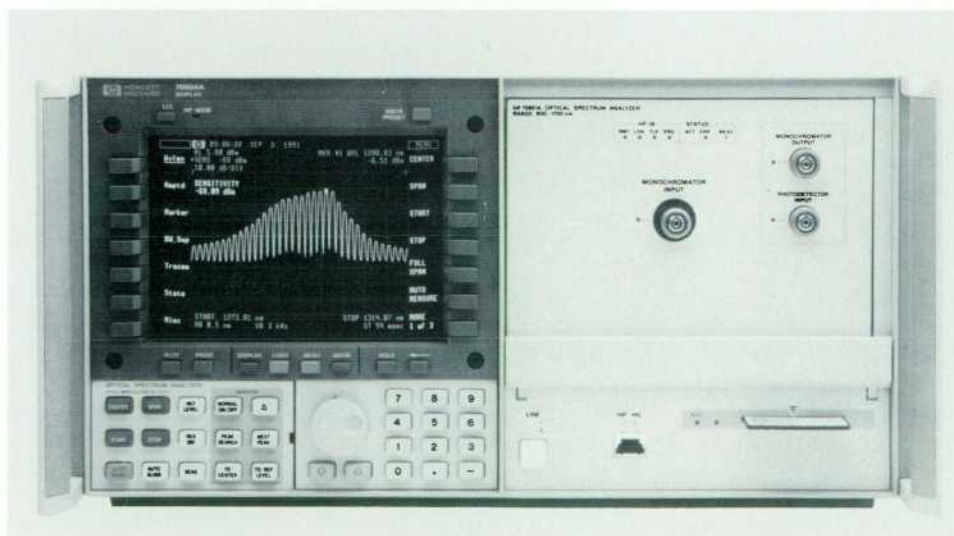
Fig. 1. The HP 71451A optical spectrum analyzer makes spectral and other measurements between 600 nm and 1700 nm.
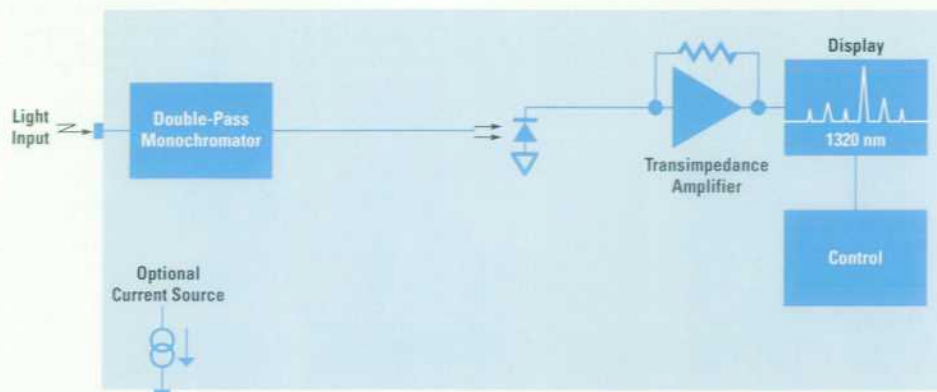
(a)
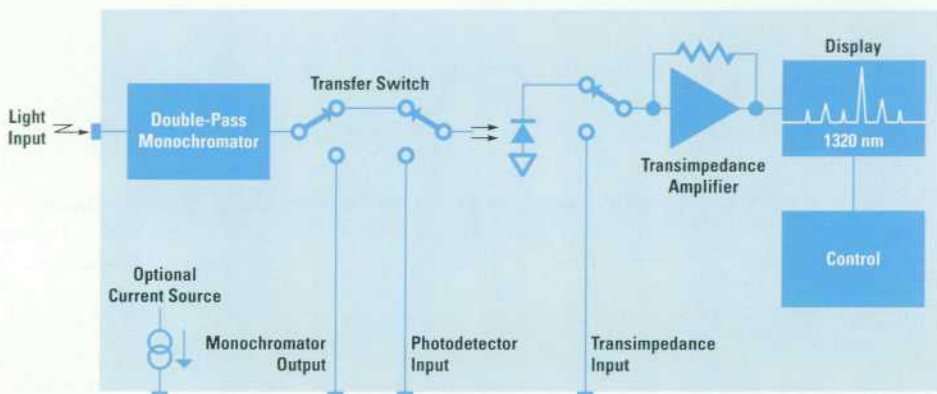
(b)

Fig. 2. (a) Block diagram of the HP 71450A optical spectrum analyzer. (b) Block diagram of the HP71451A, which extends the capabilities of the HP 71450A.

reinsert it into the photodetector input for analysis. Filters, fibers, amplifiers, isolators, switches, and other components can be characterized using this mode.

- Power meter mode. This mode offers direct access to the photodetector. In this mode, a trace of average power versus time is displayed, allowing the user to record any amplitude change over time or monitor amplitude while adjustments are made. Long-term drift can also be monitored.
- Photodetector mode. This mode is similar to the stimulus-response mode except that the device under test (DUT) is an optical-to-electrical component. By comparing the calibrated response of the internal photodiode with the measured response of the DUT (via the transimpedance input), responsivity versus wavelength can be calculated and displayed.

The rest of this article describes the user interface and the advanced measurement programs provided with the HP 71450A and 71451A optical spectrum analyzers. Other articles in this issue describe the design and implementation of the components in these analyzers.

User Interface

The user interface of the HP 71450A and 71451A optical spectrum analyzers is designed to have the same look and feel as HP's RF and microwave spectrum analyzers. The instrument functions are selected from the front panel via 14 softkeys and 15 hard keys. The only obvious difference between optical and RF and microwave instruments is that the signal information is displayed in wavelength (nanometers)

for optical instruments and frequency (Hz) for RF or microwave instruments. Displaying signal information as a function of wavelength has always been an optical tradition.

Menu Keys

Over 250 instrument functions are available from the 14 keys located on the sides of the instrument display (Fig. 3). These functions are grouped into seven measurement categories called firmkeys, which are always displayed on the left side of the display. Pressing a firmkey accesses a group
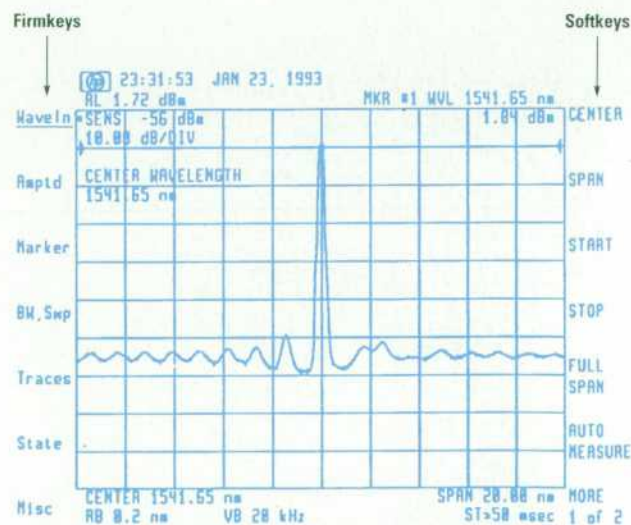
Fig. 3. An example of the softkeys and firmkeys that appear on the front panel of the HP 71450A and 71451A spectrum analyzers.

Optical Spectrum Analysis

Optical spectrum analysis is the measurement of optical power as a function of wavelength. Applications include testing laser and LED light sources for spectral purity and power distribution, and testing the transmission characteristics of optical devices.

The spectral width of a light source is an important parameter in fiber-optic communication systems because of chromatic dispersion, which occurs in the fiber and limits the modulation bandwidth of the system. The effect of chromatic dispersion can be seen in the time domain as pulse broadening of the digital information waveform. Since chromatic dispersion is a function of the spectral width of the light source, narrow spectral widths are desirable for high-speed communication systems.

Fig. 1 shows the spectrum of a Fabry-Perot laser. The laser is not purely monochromatic; its spectrum consists of a series of evenly spaced coherent spectral lines with an amplitude profile determined by the characteristics of the gain medium.

Optical spectrum analyzers can be divided into three categories: diffraction-grating-based and the Fabry-Perot and Michelson interferometer-based optical spectrum analyzers. Diffraction-grating-based optical spectrum analyzers are capable of measuring the spectra of lasers and LEDs. The resolution of these instruments is variable, typically ranging from 0.1 nm to 5 or 10 nm. Fabry-Perot-interferometer-based optical spectrum analyzers have a fixed, narrow resolution, typically specified in frequency, between 100 MHz and 10 GHz. This narrow resolution allows them

Fig. 2. Simplified block diagram of an optical spectrum analyzer.

to be used for measuring laser chirp, but can limit their measurement spans much more than the diffraction-grating-based optical spectrum analyzers. Michelson-interferometer-based optical spectrum analyzers, which are used for direct coherence-length measurements, display the spectrum by calculating the Fourier transform of measured interference patterns.

The HP 71450A and 71451A optical spectrum analyzers are diffraction-grating-based.

Basic System

A simplified block diagram of a grating-based optical spectrum analyzer is shown in Fig. 2. The incoming light passes through a tunable-wavelength optical filter (monochromator) which resolves the individual spectral components. The photodetector then converts the optical signal to an electrical current proportional to the incident optical power.

The current from the photodetector is converted to a voltage by the transimpedance amplifier and then digitized. Any remaining signal processing, such as applying correction factors, is performed digitally. The signal is then applied to the display as the vertical, or amplitude, data. A motor rotates the diffraction grating, tuning the wavelength of the optical filter. The angular position of the diffraction grating determines the horizontal location of the trace as it sweeps from left to right. A trace of optical power versus wavelength results. The displayed width of each mode of the laser is a function of the spectral resolution of the tunable-wavelength optical filter.

Fig. 1. Optical spectrum of a Fabry-Perot laser.

of functions called softkeys, which are displayed on the right side of the display. Pressing the MORE softkey displays an additional set of softkeys. These additional keys tend to be the keys used less often. This menu tree structure allows easy access to all of the instrument functions.

Hard Keys

To allow easy access to functions that are used most often, there are 15 hard keys on the front panel (see Fig. 4). These functions control sweep wavelengths, resolution, reference level, and marker positions. Three functions have been added that are not found on a microwave spectrum analyzer. These functions are automeasure (**AUTO MEAS** key), sensitivity (**SENS** key), and autoalign (**AUTO ALIGN** key).

Automeasure. When the **AUTO MEAS** key is pressed, the analyzer searches the full wavelength span and locates the largest detected signal. If a signal cannot be found, the sensitivity is increased and the search continues. Once a signal is found,

it is positioned on the screen by adjusting the center wavelength, sensitivity, and reference level. The signal width is also measured and the span is reduced so that most (or all) of the signal power is displayed. The amplitude scale is set to 10 dB/division.

The user can modify this operation by selecting a wavelength span and a final amplitude scale to be used upon completion of the automatic measurement routine. If multiple signals are present and the signal of interest is a lower-level signal, the user can position the marker on that signal and the automatic measurement routine will acquire the peak closest to the marker.

Sensitivity. The **SENS** key is used to adjust the sensitivity of the instrument. Normally the optical spectrum analyzer automatically selects the greatest sensitivity that does not affect the sweep speed. The sensitivity function allows the user to select the smallest signal amplitude to be displayed across
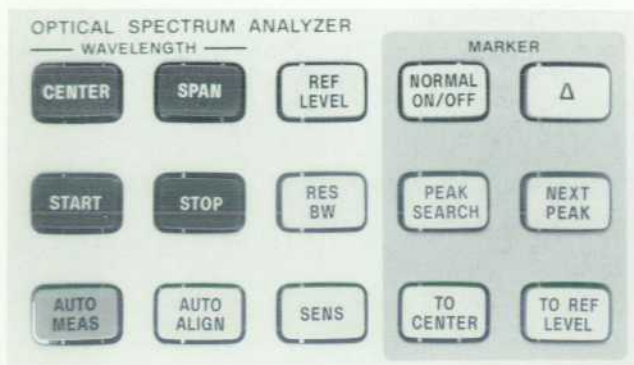
Fig. 4. A representation of the hard keys on the front panel of the HP 71450A and 71451A spectrum analyzers.

the current wavelength range. Increasing the sensitivity causes low-level signals to change the amplifier gain. These gain changes require pauses, which decrease the sweep speed. An increase in sensitivity may also require a narrower video bandwidth filter, which will also slow the sweep speed. Adjusting the reference level to the highest signal level to be measured and the sensitivity to the lowest signal level to be measured will optimize the sweep speed.

Autoalign. Alignment of the optical spectrum analyzer is easy to perform. When the **AUTO ALIGN** key is pressed, the optical spectrum analyzer automatically adjusts the mechanical position of the optical output fiber to ensure amplitude accuracy. No manual adjustments are necessary, and the optical spectrum analyzer can use an input signal of any wavelength for alignment.

Advanced Measurement Programs

The HP 71450A and 71451A optical spectrum analyzers provide the capability to download and execute custom programs, which are called advanced measurement programs. These programs provide one-button measurement solutions without using an external computer. The programs can be downloaded from a disk or memory card and stored in the analyzer's nonvolatile RAM. They can be run by pressing the USER firmkey and then selecting the displayed softkey. They can also be accessed remotely via the HP-IB interface. Three advanced measurement programs are supplied with the optical spectrum analyzers. These programs automatically measure the following light sources:

- Light-emitting diodes (LEDs)
- Fabry-Perot lasers
- Distributed feedback lasers (DFBs).

Light-Emitting Diodes

Light-emitting diodes produce light with a wide spectral width. When used in fiber-optic communication systems, they can be modulated at frequencies up to about 200 MHz. LEDs have the advantages of low temperature sensitivity and no sensitivity to back reflections. Additionally, the incoherent emitted light is not sensitive to optical interference from reflections.

A light-emitting diode generates light by spontaneous emission. This occurs when an electron in a high-energy conduction band changes to a low-energy valence band (Fig. 5).
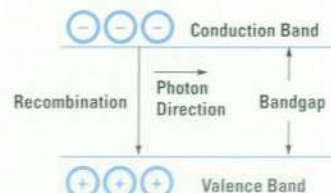
Fig. 5. Spontaneous emission. Most electrons move from the conduction band to the valence band during recombination.

The energy lost by the electron is released as a photon. For a given material, discrete energy levels represent the different orbital states of the electron. The energy of the released photon is equal to the energy lost by the electron, and the wavelength of the emitted photon is a function of its energy. As a result, the wavelength of the photon is determined by the material used to make the LED.

The spontaneous emission is caused by the recombination of electrons from the conduction band to the valence band. The difference in energy between the conduction band and the valence band is called the bandgap energy (E_g) and is expressed in units of either joules or electron volts (eV). The wavelength of the emitted photon is determined by the bandgap energy. The wavelength is expressed as:

$$\lambda = hc/E_g = 1.24 \ \mu m/E_g,$$

where h (Planck's constant) is equal to 6.62×10^{-34} Ws^2, c (speed of light) is 2.998×10^8 m/s, and E_g (bandgap energy of the material) is expressed in units of joules.

The conduction-band electrons are generated by a forward bias placed on the p-n junction of the diode (Fig. 6). The material on the n-layer side of the junction has immobile positive charges evenly distributed throughout the layer, with mobile negative charges, or electrons, responsible for electrical current flow. Conversely, the material on the p-layer side of the junction has immobile negative charges evenly distributed throughout the layer with mobile positively charged holes, actually locations of missing electrons, responsible for electrical current flow.

At the junction, the mobile electrons from the n layer and the mobile holes from the p layer recombine and produce

= Immobile Negative Charges

= Immobile Positive Charges
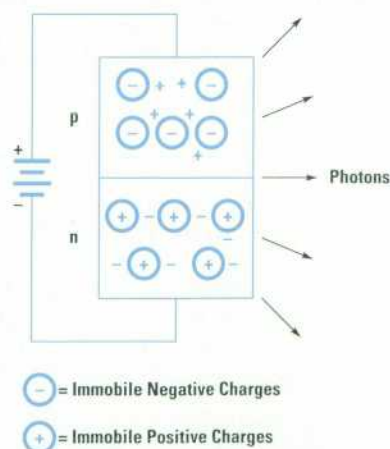
Fig. 6. Diagram of a forward biased p-n junction showing the location of immobile charges and mobile current charges.

photons. While LEDs in use today consist of multiple layers of semiconductor material, the light-generation process is the same.

The spectrum of a light-emitting diode results in a broad distribution of wavelengths centered about the wavelength calculated by the above equation. The spectral width is often specified at the half-power points of the spectrum, or FWHM (full width at half maximum) points. Typical values for full width at half maximum range from 20 nm to 80 nm for LEDs.

LED Advanced Measurement Program

The LED advanced measurement program automatically measures many LED parameters. Some parameters such as mean wavelength and spectral width are measured by two methods. One method takes into account the entire spectrum, while the other takes into account only a few points of the spectrum. All the measurements are made at the end of each sweep. Fig. 7 shows the measurement display provided by the HP 71450A and HP 71451A LED advanced measurement program. The following is an explanation of each of the values computed by the LED advanced measurement program.

- Total power. The total power parameter is the summation of the power at each trace point between two user-selected points, normalized by the ratio of the trace point spacing and resolution bandwidth. This normalization is required because the spectrum of the LED is continuous rather than containing discrete spectral components as a laser does.
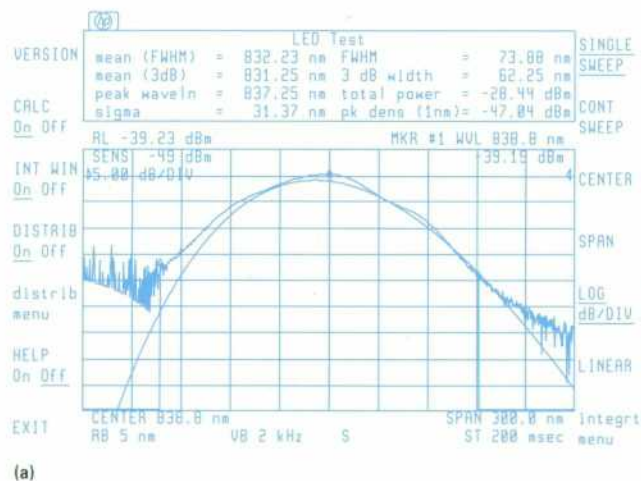
(a)
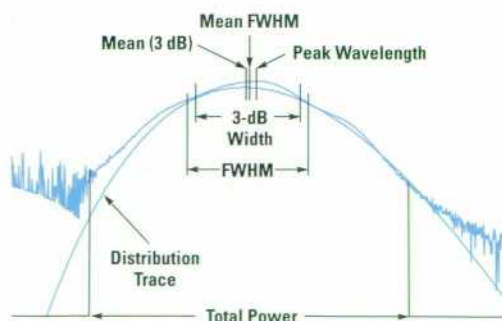
(b)

Fig. 7. (a) The spectrum of a light-emitting diode. (b) The parameters provided by the LED automatic measurement program.

The total power of an LED being tested is determined by the equation:

$$\text{Total Power} = P_o = \sum_{i=1}^{N} P_i \, (\text{TS/RBW}),$$

where RBW is the resolution bandwidth and TS is the trace point spacing.

- Mean FWHM. The mean wavelength of full width at half maximum points represents the center of mass of the trace points. The power and wavelength of each trace point are used to calculate the mean FWHM wavelength:

$$\text{Mean FWHM} = \overline{\lambda} = \sum_{i=1}^{N} P_i \, (\text{TS/RBW}) \, \lambda_i / P_o.$$

- Sigma. This measurement is an rms calculation of the spectral width of an LED based on a Gaussian distribution. The power and wavelength of each trace point are used to calculate sigma.

$$\text{Sigma} = \sigma = \sqrt{\sum_{i=1}^{N} P_i \, (\text{TS/RBW}) \left(\lambda_i - \overline{\lambda}\right)^2 / P_o}.$$

Sigma is also used to calculate the distribution trace (described below).

- FWHM (full width at half maximum). This measurement describes the spectral width of the half-power points of the LED, assuming a continuous Gaussian power distribution. The half-power points are those points where the power spectral density is one-half that of the peak amplitude.

$$\text{FWHM} = 2.355 \times \text{Sigma}.$$

- 3-dB width. This value is used to describe the spectral width of an LED based on the separation of the two wavelengths that each have a power spectral density equal to one-half the peak power spectral density. The 3-dB width is determined by finding the peak of the LED spectrum and dropping down 3 dB on each side.

- Mean (3 dB). This is the wavelength that is the average of the two wavelengths determined in the 3-dB width measurement.

- Peak wavelength. This is the wavelength at which the peak of the LED's spectrum occurs.

- Peak density (1 nm). This is the power spectral density (normalized to a 1-nm bandwidth) of the LED at the peak wavelength.

- Distribution trace. This is a trace that is based on the total power, the power distribution, and the mean wavelength of an LED. This trace has a Gaussian spectral distribution and represents a Gaussian approximation of the measured spectrum.

Fabry-Perot Lasers

Lasers are capable of producing high output power and directional light beams. When used in fiber-optic communication systems, semiconductor lasers can be modulated at rates up to about 10 GHz. However, lasers are sensitive to temperature and back reflections. Additionally, the coherent emitted light is sensitive to optical interference from reflections.
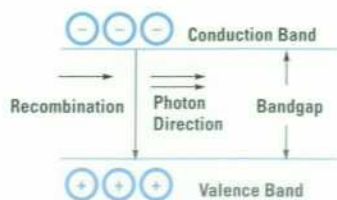
Fig. 8. Stimulated emission is the release of a photon because of an electron hole recombination triggered by another photon.

The design of the Fabry-Perot laser is simpler than the distributed feedback laser (described later). However, it is more susceptible to chromatic dispersion when used in fiber-optic systems because it has a wider spectral bandwidth. A Fabry-Perot laser differs from a light-emitting diode in that it generates light mainly by stimulated emission. Some of the photons are generated by spontaneous emission, as described for the LED, but the majority of the photons are generated by stimulated emission, where photons trigger additional electron-hole recombinations, resulting in additional photons as shown in Fig. 8. A stimulated photon travels in the same direction and has the same wavelength and phase as the photon that triggered its generation.

Stimulated emission can be thought of as the amplification of light (laser is an acronym for light amplification by stimulated emission of radiation). As one photon passes through the region of holes and conduction band electrons, additional photons are generated. If the material were long enough, enough photons might be generated to produce a significant amount of power at a single wavelength.

An easier way to build up power is to place a reflective mirror at each end of the region just described so that the photons travel back and forth between the mirrors, building up the number of photons with each trip. These mirrors form a resonator, which is a requirement for laser operation.

Laser operation has two additional requirements. One requirement is that for stimulated emission to occur, a greater number of conduction-band electrons than valence-band electrons must be present. This is called a population inversion. It is achieved by forcing a high current density in the active layer of the diode structure. The second requirement is that the gain exceed the losses from absorption and radiation. Part of the radiation loss is the amount of light released at the laser output. As the current increases, the gain increases. The current for which stimulated emissions occur is the threshold current of the laser.

The resonator is often just highly reflective, cleaved surfaces on the edges of the diode (Fig. 9). As the light reflects between the mirrors, the photons of a given wavelength must be in phase to add constructively. The resonator acts as a Fabry-Perot interferometer because only light for which the resonator spacing is an integral number of half wavelengths will add constructively. As a result, the spectrum of a Fabry-Perot laser contains multiple discrete-wavelength components.

The possible wavelengths produced by the resonator are given by:

$$f_{res} = mc/2ln,$$
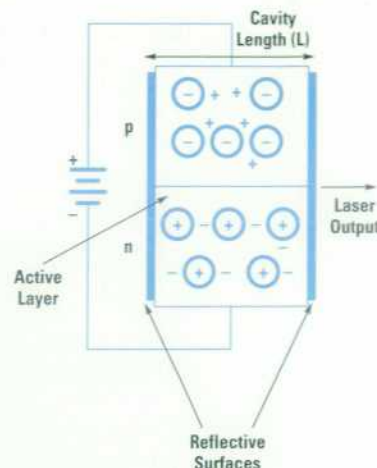
Fig. 9. The reflective surfaces at the edges of the laser diode act as a Fabry-Perot type resonator.

where m = integer, c = speed of light, l = length of cavity, and n = refractive index of cavity.

The actual output power at each of these wavelengths is determined by the laser gain and mirror reflectivity at that wavelength. As with the LED, the center wavelength can be determined from the bandgap energy. The separation between the different wavelengths (mode spacing) can be determined from the separation of the mirrors as follows:

$$\text{Mode Spacing} = c/2ln \text{ (Hz) or } \lambda^2/2ln \text{ (nm).}$$

Fabry-Perot Laser Advanced Measurement Program

The Fabry-Perot laser advanced measurement program automatically measures the parameters of the Fabry-Perot laser at the end of each sweep. All of the measurements are based upon the detected trace peaks of the laser (see Fig. 10). What defines a peak is controlled by the peak excursion function. The peak excursion value (in dB) can be set by the user and is used to determine which trace peaks are accepted as discrete spectral responses.

A peaks function is supplied in the measurement program to verify that a proper peak excursion value is being used. The peaks function, when enabled, displays a vertical line from the bottom of the grid to each counted spectral component of the signal (see Fig. 10a).

A distribution trace function is also supplied with the program. This function displays a trace that is based on the total power, individual wavelengths, mean wavelength, and mode spacing of the laser. This trace can be a Gaussian, Lorentzian, or envelope spectral distribution that represent continuous approximations to the real discrete spectrum.

The other parameters computed by the Fabry-Perot measurement program include:
• Total power. This is the summation of the power in each of the displayed spectral components, or modes, that satisfy the peak excursion criteria.

$$\text{Total Power} = P_o = \sum_{i=1}^{N} P_i,$$

Peak Indicators

(a)

Mean Wavelength
Peak Wavelength
Peak Amplitude
Full Width at Half Maximum
Distribution Trace
Mode Spacing
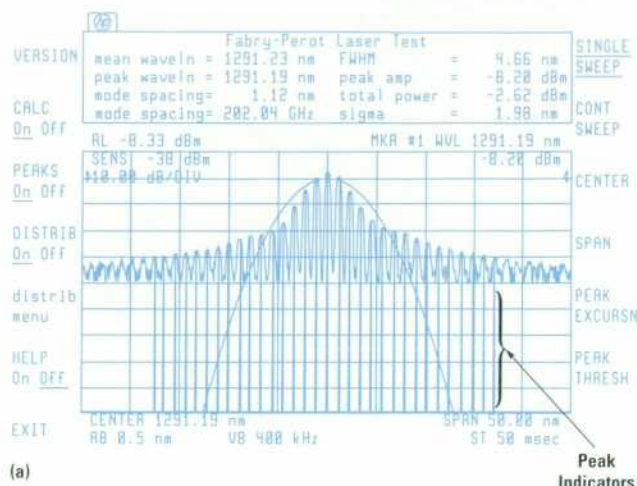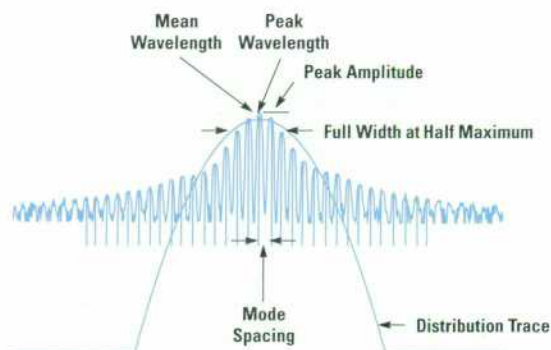
(b)

Fig. 10. (a) The spectrum of a Fabry-Perot laser. (b) The parameters provided by the Fabry-Perot laser measurement program.

- Mean wavelength. This parameter represents the center of mass of the spectral components onscreen. The power and wavelength of each spectral component are used to calculate the mean wavelength.

$$\text{Mean Wavelength} = \bar{\lambda} = \sum_{i=1}^{N} P_i\lambda_i/P_o.$$

- Sigma. This is an rms calculation of the spectral width of the Fabry-Perot laser based on a Gaussian distribution.

$$\text{Sigma} = \sqrt{\sum_{i=1}^{N} P_i\left(\lambda_i - \bar{\lambda}\right)^2/P_o}.$$

- FWHM (full width at half maximum). This parameter describes the spectral width of the half-power points of the Fabry-Perot laser, assuming a continuous, Gaussian power distribution. The half-power points are those where the power spectral density is one-half that of the peak amplitude.

$$\text{FWHM} = 2.355 \times \text{Sigma}.$$

- Mode spacing. This is the average wavelength spacing between the individual spectral components of the Fabry-Perot laser.
- Peak amplitude. The power level of the peak spectral component of the Fabry-Perot laser.
- Peak wavelength. This is the wavelength at which the peak spectral component of the Fabry-Perot laser occurs.
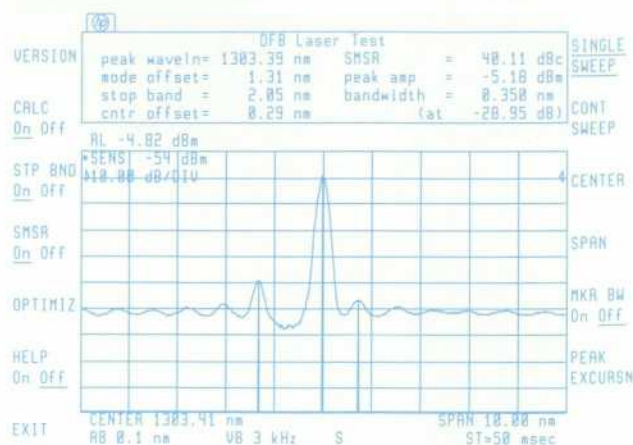
Fig. 11. The spectrum of a distributed feedback laser.

Distributed Feedback Lasers

Distributed feedback (DFB) lasers are similar to Fabry-Perot lasers, except that all but one of their spectral components are significantly reduced (see Fig. 11). Because its spectrum has only one line, the spectral width of a distributed feedback laser is much less than that of a Fabry-Perot laser. This greatly reduces the effect of chromatic dispersion in fiber-optic systems, allowing for greater transmission bandwidths.

The distributed feedback laser uses a grating, which is a series of corrugated ridges, just above the active layer of the semiconductor (see Fig. 12). Rather than using just the two reflecting surfaces at the ends of the diode, as a Fabry-Perot laser does, the distributed feedback laser uses each ridge of the corrugation as a reflective surface. At the resonant wavelength, all reflections from the different ridges add in phase. Because of the much smaller spacings between the resonator elements compared to the Fabry-Perot laser, the possible resonant wavelengths are much farther apart in wavelength, and only one resonant wavelength is in the region of laser gain. This results in the single laser wavelength.

The ends of the p-n diode still act as a resonator in the DFB laser, producing lower-amplitude side modes. Ideally, the dimensions of the reflective surfaces are selected so that the end reflections add in phase with the grating reflections. In this case, the main mode will occur at a wavelength halfway between the two adjacent side modes; any deviation is called a center offset. Center offset is measured as the difference between the main-mode wavelength and the average wavelength of the two adjacent side modes.

The amplitude of the largest side mode is typically between 30 and 50 dB lower than the main spectral output of the laser. Because side modes are so close to the main mode (typically between 0.5 nm and 1 nm) the dynamic range of an optical spectrum analyzer determines its ability to measure them. Dynamic range is specified at offsets of 0.5 nm and 1.0 nm from a large response. The HP 71450 and 71451A optical spectrum analyzers specify a dynamic range of –55 dBc at offsets of 0.5 nm and greater, and –60 dBc at offsets of 1.0 nm and greater. This indicates the amplitude level of side modes that can be detected at the given offsets.
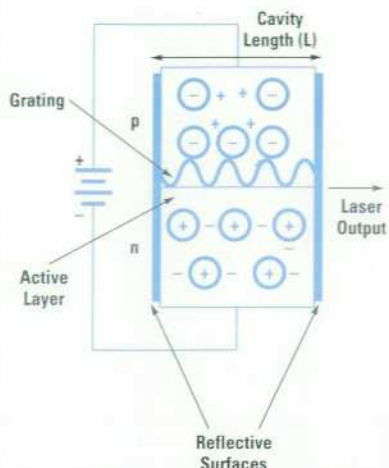
Fig. 12. A p-n junction representation of a distributed feedback laser which uses a series of reflecting ridges to reduce the amplitude of all but one of the spectral components of the laser.

Fig. 13. The parameters provided by the distributed feedback laser automatic program.

Distributed Feedback Laser Advanced Measurement Program

The distributed feedback laser advanced measurement program automatically measures the parameters of the distributed feedback laser at the end of each sweep. Like the Fabry-Perot laser advanced measurement program, all of the measurements are based upon the detected modes of the laser, or trace peaks. What defines a peak is controlled by the peak excursion function.

A stop band display function is supplied to verify that a proper peak excursion value is being used to determine the correct stop band modes. The stop band function, when enabled, displays a vertical line from the bottom of the grid to each of the selected side modes shown in Figs. 11 and 13.

A side-mode suppression ratio display function is also supplied to verify that a proper peak excursion value is being used to determine the correct side mode.

The other parameters provided by the DFB measurement program include:

- Peak wavelength. This is the wavelength at which the main spectral component of the DFB laser occurs.
- Mode offset. This is the wavelength separation (in nanometers) between the main spectral component and the largest side mode.
- Peak amplitude. This is the power level of the main spectral component of the DFB laser.
- Stop band. This is the wavelength spacing between the upper and lower side modes adjacent to the main mode.
- Center offset. This parameter indicates how well the main mode is centered in the stop band. This value equals the wavelength of the main spectral component minus the mean of the upper and lower stop band component wavelengths.
- Bandwidth. This parameter provides a measurement of the displayed bandwidth of the main spectral component of the DFB laser. The amplitude level, relative to the peak, that is used to measure the bandwidth can be set by the user. The default amplitude level used is –20 dBc. Because of the narrow line width of lasers, the result of this measurement for an unmodulated laser is strictly dependent upon the resolution bandwidth filter of the optical spectrum analyzer. With modulation applied, the resultant waveform is a convolution of the analyzer's filter and the modulated laser's spectrum, causing the measured bandwidth to increase. The combination of the modulated and unmodulated readings can be used to determine the bandwidth of the modulated laser and the presence of chirp.

Conclusion

The HP 71450A and 71451A are HP's first optical spectrum analyzers. By leveraging our expertise in user interfaces from the RF and microwave products, we were able to provide a product that has the familiar look and feel that users expect from HP.

Acknowledgments

A Double-Pass Monochromator for Wavelength Selection in an Optical Spectrum Analyzer

The wavelength-selection scheme used in the HP 71450A and HP 71451A optical spectrum analyzers propagates the light from the device under test twice through the refraction and diffraction elements in the monochromator.

by Kenneth R. Wildnauer and Zoltan Azary

For many users of spectral analysis instruments, measurement speed is of primary concern, and having a display of an optical spectrum in real time is highly desirable. Many users interested in the purity of their source are also interested in being able to detect low-level signals that are very close in wavelength to the primary signal. The ratio of the power of these low-level signals to the main signal can be easily smaller than 10^{-4} (–40 dBc) at offsets less than one nanometer away. The ability of an instrument to resolve or display these signals will be referred to as close-in dynamic range in this article.

Higher transmission rates, better transmission quality, and the longer transmission distances of today's fiber-optic transmission systems have created the need to measure and analyze these low-level optical signals. To measure low-level optical signals an instrument must be efficient and sensitive. To perform these measurements quickly is an added challenge. Also, many times the polarization state (i.e., the orientation of the electric field) of the input signal is either not known or variable. Hence, the instrument measurements should be relatively insensitive to changes of the input polarization state.

Because many applications find it useful to filter an input signal optically, an instrument that can produce optical output should also have variable optical bandwidth. With the need for more precise optical measurements, the instrument should be repeatable and accurate and be able to make these measurements in a standard instrument environment—without the need for an optics table. Finally, it would be convenient if the instrument that measures these low-level signals could be small and rugged enough so that it can be moved around without any special care.

The HP 71450A and 71451A optical spectrum analyzers provide the features mentioned above by using a specially developed wavelength-selection scheme—the double-pass monochromator. A block diagram of these analyzers is shown in Fig. 1. This article describes the operation and performance of the double-pass monochromator and the operation and characteristics of the components in the data acquisition and processing system in Fig. 1.
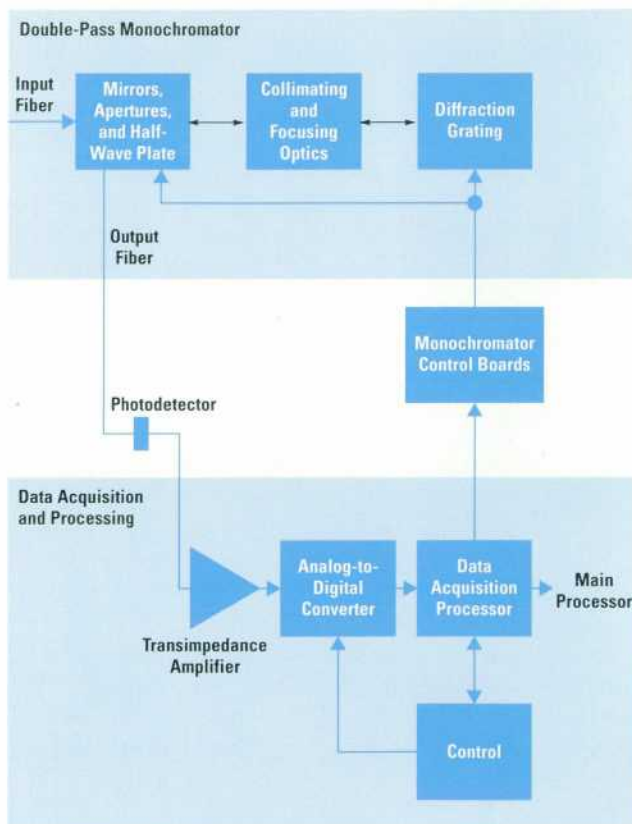
Fig. 1. Block diagram of the major components in the HP 71450A and 71451A optical spectrum analyzers.

Double-Pass Monochromator

A double-pass-monochromator-based design was chosen for the HP 71450A and HP 71451A optical spectrum analyzers rather than a spectrometer-based design for two reasons. First, a single photodetector (which is what the monochromator uses) has an inherent advantage over the detector array of the spectrometer for close-in dynamic range measurements. The detector array also costs more than a single detector. Second, a monochromator puts less demand on the

Masking Wheel* Input Aperture Wheel Slit* Half-Wave Plate Collimating Lens

Output Fiber

Other Data Acquisition and Processing Components Detector Transimpedance Amplifier

Diffraction Grating
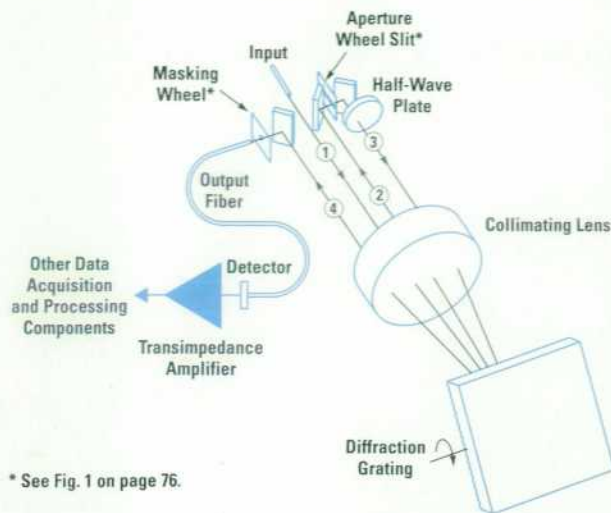
* See Fig. 1 on page 76.

Fig. 2. The elements of the double-pass monochromator showing the light beam as it makes the two passes through the optical components.

optical system for imaging a large span of spatially dispersed wavelengths. The double-pass monochromator configuration further increases the close-in dynamic range of the instrument, essentially obtaining the range of two cascaded monochromators.

A refractive optical system was chosen to reduce the size requirements of the monochromator so as to minimize the overall instrument size. Careful design of the refractive elements has reduced the inherent chromatic aberrations associated with such systems to a tolerable level. The inherent disadvantage of slower measurement speed because of scanning with the diffraction grating is reduced with a direct-drive system and the properties of the double-pass configuration.

Operation. The propagation of light through the system starts with the light entering the monochromator from the device under test (see Fig. 2). This light is then relayed by the input connector assembly (① in Fig. 2). The user end of this connector assembly is a flat-polish physical contact with interchangeable adapters to allow connection to standard fiber interfaces. The monochromator end is an angled interface to air. Both interfaces minimize reflections to the user end. The light then propagates towards the lens and is collimated for illumination of the diffraction grating. The diffraction grating is operated in very nearly a Littrow condition* and can be rotated to the desired wavelength. The light is dispersed by the diffraction grating and returns through the same lens to be reflected by the first plane mirror and imaged onto one of the apertures on the rotatable aperture wheel ②. By rotating the aperture wheel, different aperture widths (slits) and hence resolution bandwidths can be selected by the user. Once the light enters and leaves the aperture slit the first pass of the double-pass monochromator is effectively complete.

The second pass starts when the light exits the aperture slit and is reflected by the second plane mirror and propagates through an achromatic half-wave plate ③. The half-wave

plate is oriented so that it causes a 90-degree rotation of the s and p polarization components as defined with respect to the lines on the diffraction grating.** The beam is again collimated by the same lens and again illuminates the same diffraction grating. However, because of the orientation of the first and second mirrors with respect to the dispersion direction of the diffraction grating, on the second pass through the system the light is not dispersed any further by the diffraction grating but is collapsed or recombined, creating a filtered replica of the input signal. This recombined beam ④ is then imaged by the lens onto a fiber after reflection from a third plane mirror near the fiber. This fiber, which is called the output fiber in Fig. 2, is a piece of multimode fiber that carries the light to the photodetector for conversion into photocurrent for analysis and display. Another function of this output fiber is to act as a second aperture in the system. As an option, this light can be directed to the front panel of the instrument providing an optical output for the user. In front of this output fiber, there is also a mask wheel that is coaxial with the aperture slit wheel and controlled by the same motor. To measure the dark current of the photodetector, this mask wheel can be rotated such that the output signal is blocked.

This double-pass monochromator system provides the high sensitivity typically found in a single-pass monochromator system and the high dynamic range typically found in a double monochromator. Also, because of the half-wave plate, it provides excellent polarization insensitivity.

Performance. Two main factors affect the time it takes to make a swept measurement with an optical spectrum analyzer. The first is the ability to move the diffraction grating quickly, and the second is the signal-to-noise ratio. The instrument settings and the power of the user's input beam will affect which of these two factors limits the measurement speed. For relatively medium to high power levels, which are greater than microwatts (>−60 dBm), the measurement speed is usually limited by motor speed. For measurements in which the user is concerned about power levels less than this, the measurement speed is usually limited by signal-to-noise ratio. The ability to move the grating quickly, precisely, and reliably in the HP 71450A and 71451A analyzers is provided by a direct-drive system. This direct-drive system offers significant improvement in measurement speed for cases in which the measurement is limited by motor speed. This system is described in the article on page 75.

To minimize the measurement time for cases in which the signal-to-noise ratio is the limiting factor, it would be advantageous to maximize the signal-to-noise ratio by maximizing the signal and minimizing the noise. For a given input signal level the only way to maximize the signal without amplification is to minimize the loss through the system. This is achieved through the use of highly efficient and low-loss optics to collect as much of the user's input signal as possible.

The two main components of noise are noise resulting from scattered light and electrical noise from the photodetector and its components. To minimize the scattered and stray light in the HP 71450A and 71451A analyzers, careful attention

* See "Diffraction Grating" on page 70 for a description of the Littrow condition.

** See "Polarization Sensitivity" on page 71 for a description of s and p polarization.

Diffraction Grating

A diffraction grating is made up of an array of equidistant parallel slits (in the case of a transmissive grating) or reflectors (in the case of a reflective grating). The spacing of the slits or reflectors is on the order of the wavelength of the light for which the grating is intended to be used. The HP 71450A and 71451A optical spectrum analyzers use a reflective grating.

The basic operation of a diffraction grating begins when light that strikes the reflective lines of the grating is diffracted. For a given wavelength there will be a certain angle at which the diffracted wavelets will be exactly one wavelength out of phase with one another and will add constructively in a parallel wavefront (see Fig. 1).

The light of a given wavelength leaves the grating at a specific angle, and light of other wavelengths leaves the grating at other angles. Although it is based on a different principle, the diffraction grating spatially separates the wavelengths of light that strike it much the way a prism separates the wavelengths of light that pass through it.

Fig. 2. The diffraction grating in the 71450A and HP 71451A optical spectrum analyzers operates in a configuration called a Littrow condition in which the wavelength of interest travels back along the path of the incident beam.

The general equation for a diffraction grating is:

$$n\lambda = 2d(\sin \alpha + \sin \beta),$$

where λ is the wavelength of the light, d is the spacing of the lines on the grating, α is the angle of the incident light relative to the grating normal, β is the angle at which light of wavelength λ leaves the grating, and n is an integer that is called the order of the spectrum.

When the wavelets are each one wavelength out of phase the spectrum is called a first-order spectrum. At another angle where the wavelets are all exactly two wavelengths out of phase and will also add constructively, the spectrum is called a second-order spectrum. Higher-order spectra may also be present.

In the HP 71450A and 71451A optical spectrum analyzers the diffraction grating is operated in a special configuration called the Littrow condition. In this arrangement, the wavelength of interest leaves the diffraction grating and goes directly back along the path of the incident beam (see Fig. 2). Thus, in the grating equation $\alpha = \beta \ (= \theta)$ (see Fig. 2) and the equation can be simplified to:

$$n\lambda = 2d \sin \theta.$$

Fig. 1. The basic operation of a diffraction grating. Incoming light strikes the reflective grating producing diffracted wavelets that are exactly one wavelength out of phase with one another and add constructively in a parallel wavefront.

was paid to coating the optics. Also, careful analysis and attention was paid to the absorption and scattering of light because of reflections from the input beam and the dispersed light from the grating reflecting off the internal surfaces of the monochromator. The effect of this stray light is reduced through background light subtraction (described later).

The noise from the photodetector and its associated electronics can be lowered by reducing the electrical bandwidth (video bandwidth) of the detection system. There is of course a penalty in measurement time because of the need to allow the video bandwidth filters to settle.

One way to minimize the noise power from the photodetector without reducing the bandwidth is to use the smallest detector possible. An important feature of the double-pass configuration is the ability to use a relatively small detector, because on the second pass the beam is not dispersed but

recombined. In fact, the detector need only be as large as the output image plus any wander or movement of the output image resulting from rotation of the diffraction grating during measurement. In a symmetrical system with no mechanical errors, there should be no movement of the output image as the grating rotates. Because complete elimination of all mechanical errors is not possible, either the detector must be larger so that it does not miss the image when the grating rotates and the image moves, or there must be some means for the detector to track the movement of the output image. To keep the detector small, we have chosen to provide a means for tracking the output image movement in our analyzers. This tracking mechanism is described in the article on page 80. This smaller detector generates less inherent noise and thus allows the use of a wider video bandwidth with the benefit of a faster measurement or sweep speed.

Another benefit of the small detector at the output is the ability to have a small output aperture. A small output aperture coupled with the second pass through the system increases the close-in dynamic range. This is true whether the second pass is set up to further disperse the light or to recombine the light (collapse the dispersion). Besides allowing a small detector, there are several other advantages to recombining the light on the second pass rather than further dispersing the light. First, the output aperture does not affect the optical or resolution bandwidth of the system. This is solely determined by the first pass. Second, the inherent time dispersion of the first pass is canceled. Because of the wavelength dispersion of the first pass, there has to be a corresponding time dispersion. This can be seen by noting that the grating is tilted with respect to the wavefront illuminating the grating. Hence different parts of the wavefront see different path lengths and different time delays. By recombining the dispersed light on the second pass, this time dispersion is canceled because of the inverse path lengths across the wavefront of the beam on the second pass with respect to that of the first pass. Finally, recombining light on the second pass results in an optical output that can be conveniently provided on an optical fiber. In the configuration mentioned above, this optical output has an optical resolution bandwidth that is variable and selectable by the user and not limited by the output aperture. Thus, the full set of resolution bandwidths provided by the instrument are available. This allows the user to use the monochromator as a tunable, variable-bandwidth optical filter or preselector.

The efficiency of the diffraction grating in the HP 71450A and 71451A optical spectrum analyzers is inherently dependent on the polarization state of the light illuminating it. The grating is also the optical element with the largest loss in the system and the largest change in loss with input polarization. Because the input polarization state can be different for different users and can change during a measurement, it is desirable to compensate for the polarization dependent efficiency. This is accomplished by rotating the s and p polarization components (See "Polarization Sensitivity," this page) by 90 degrees between the first and second passes. A rotation of 90 degrees will cause what was the s component on the first reflection off the grating (first pass) to become the p component on the second reflection off the grating (second pass). This same exchange happens for the p polarization. Thus the efficiency of the grating is always the product of the s and p efficiencies of the grating for any input polarization. This compensation scheme also has the benefit of maintaining the same degree of polarization of the light. Therefore, the light at the optical output has the same degree of polarization as at the input beam, but not the same polarization. The 90-degree rotation is accomplished by the achromatic half-wave plate.

Data Acquisition and Processing

After the incoming light beam has been optically filtered by the double-pass monochromator, the components in the data acquisition and processing section of the HP 71450A and HP 71451A optical spectrum analyzers are responsible for detecting the incoming light and converting it to an electrical signal, which is then converted to digital format for processing and display. Fig. 3 shows components included in this section of the optical spectrum analyzers.

Polarization Sensitivity

Polarization sensitivity results because the reflection loss of the diffraction grating is a function of the polarization angle of the light that strikes it. As the polarization angle of the light varies, so does the loss in the monochromator. Polarized light can be divided into two components. The component parallel to the direction of the lines on the diffraction grating is often labeled p polarization, and the component perpendicular to the direction of the lines on the diffraction grating is often labeled s polarization. The loss at the diffraction grating differs for the two different polarizations and each loss varies with wavelength. At each wavelength, the loss of p polarized light and the loss of s polarized light represent the minimum and maximum losses possible for linearly polarized light. At some wavelengths, the loss experienced by p polarized light is greater than that of s polarized light, while at other wavelengths, the situation is reversed. This polarization sensitivity results in an amplitude uncertainty for measurements of polarized light and is specified as polarization dependence.

To reduce polarization sensitivity in the HP 71450A and 71451A optical spectrum analyzers, a half-wave plate is located in the path of the optical signal between the first and second pass in the double-pass monochromator (see Fig. 2 in accompanying article). This half-wave plate rotates the components of polarization by 90 degrees. The result is that the component of polarization that receives the maximum attenuation on the first pass will receive the minimum attenuation on the second pass, and vice versa.

Photodiode. As mentioned above, because of the monochromator's ability to produce a small output image it is possible to use a photodiode with a correspondingly small active area for electrical detection. This photodiode has the advantage of having a very large shunt resistance (on the order of 1 GΩ), which allows detection of very low light levels. Sensitivity at the photodiode is around −90 dBm when using a 100-MΩ transimpedance with a 10-Hz video bandwidth.

The photodiode output current is directly proportional to input optical power. Therefore, when converting the photodiode current into decibel units, the base-10 logarithm is multiplied by 10 instead of 20 as would be done with electrical signals. All electrical gains referred to in this article are computed this way. For example, an amplifier with a voltage gain of 100 is referred to as having a gain of 20 dB, not 40 dB as is usually done. This convention causes all electrical gains to be viewed in terms of optical power levels.

Transimpedance Amplifier. The transimpedance amplifier is an FET-input monolithic IC that has very low input bias current (less than 1 pA maximum at room temperature) and very low input noise (600 nV rms, 10 Hz to 10 kHz). The transimpedance amplifier's gain can be selected from six discrete feedback resistor settings between 1 kΩ and 100 MΩ in decade increments. A 100-MΩ feedback resistor is permanently connected to the amplifier. Five additional (lower) gains are available by connecting other resistors in parallel with the first one. Shielded reed relays are used for this purpose and are located on the the low-impedance (output) side of the feedback path. By locating the relays here, the effects of relay leakage currents are minimized and less expensive relays can be used.

Analog-to-Digital Converter. The data acquisition subsystem has a total dynamic range of 100 dB, 90 dB of which is available at any one time. A 16-bit ADC running at 27 kHz provides 25 dB of dynamic range with about 0.02 dB of resolution at the low end. By switching in a 25-dB amplifier (an electrical
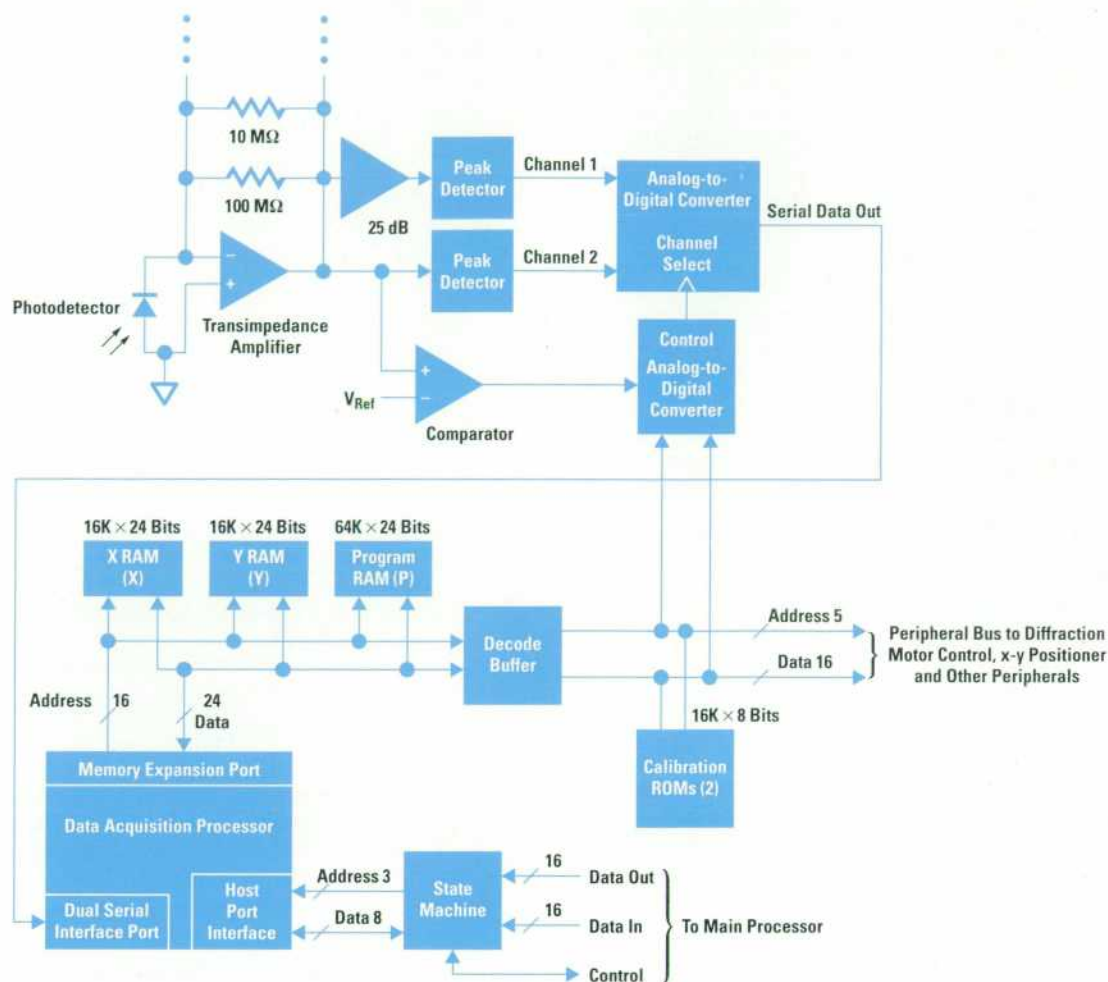
Fig. 3. A block diagram of the components in the data acquisition section of the optical spectrum analyzer.

voltage gain of 320:1), the effective input range of the ADC is increased to 50 dB. A comparator decides whether the ADC should convert the output of the 25-dB amplifier or the unamplified signal. If the output from the transimpedance amplifier is <30 mV the amplified signal is used.

The range of transimpedance gains available (1 kΩ to 100 MΩ) adds an additional 50 dB to the spectrum analyzer's dynamic range. The firmware takes advantage of this by offering a mode of operation in which transimpedance gains are automatically switched during a measurement to keep the measured signal within range. This mode is called autoranging. To guarantee adequate overlap between ranges, only 90 dB of the dynamic range is available during any one measurement.

Detection and Sampling. The data acquisition hardware is capable of either sampling or peak-detecting the transimpedance amplifier's output. Peak detection is useful for catching narrow signals when the measurement span is wide compared to the resolution bandwidth. The firmware digitally extends peak detection to capture the peak of many ADC readings taken for each trace data point.

When the hardware is set to sample detection, the firmware passes the ADC readings through a single-pole IIR (infinite input response) digital filter. By filtering ADC readings,

system noise is reduced and lower optical power levels can be measured.

Grating Angle Measurement. When a swept-wavelength measurement is taken, the motor control hardware is instructed to move the diffraction grating from the start wavelength angle to the stop wavelength angle at a constant angular velocity. This includes a certain amount of oversweep so that the motor reaches full velocity before the start wavelength and doesn't start decelerating until after the stop wavelength has been reached.

During a measurement, the ADC samples at a constant frequency and the data taken must be correlated with the wavelength. Two factors make a simple time correlation of ADC data with wavelength impractical. First, although the command input to the grating motor controller (the desired diffraction grating angle) can be made to follow an accurate and predictable velocity profile, the problem is that to the degree of accuracy we require, which is less than an arc-second, the actual trajectory followed by the grating will exhibit substantial and unpredictable variations from the desired trajectory. Second, the relationship between the diffraction grating angle and the wavelength is not linear. These two factors make a simple time correlation of ADC data with the wavelength impossible.

A simple solution is used to overcome these problems in the HP 71450A and 71451A. When the data acquisition system takes a sample, it also sends a signal to the diffraction grating motor control hardware. This signal causes the actual angle of the grating to be stored. The firmware can then read the angle and determine the wavelength when the ADC conversion is finished.

The diffraction grating motor system is described in the article on page 75.

Data Representation. The digital signal processor processes data as 24-bit or 48-bit signed fixed-point fractions. The leftmost bit (bit 23) is the sign bit and the radix point is assumed to be just to the left of the most-significant data bit (bit 22). With 24-bit data, the least-significant bit is equal to 2^{-23} and the range of numbers that can be represented is -1 to $1 - 2^{-23}$.

Amplitude data is represented by a 48-bit signed fraction. The maximum ADC reading at the lowest-gain transimpedance without using the 25-dB gain block is scaled to 0.5. Readings taken from the 25-dB gain block or with other transimpedances are scaled accordingly. This data representation has a dynamic range of about 120 dB with 0.02 dB of resolution at the low end.

Zeroing and Chopping. Mostly because of drift in the analog hardware such as op-amp input offset voltages and currents, the ADC reading corresponding to zero optical input power changes with time. To compensate for this, the firmware periodically remeasures the zero-input ADC reading. It does this by rotating the aperture wheel to a point halfway between two apertures (slits). This blocks light from making the second pass through the monochromator. It also blocks nearly all stray light from entering the output fiber because the output mask wheel has no hole at this position. Normally, this operation is carried out during retrace, and the time spent rezeroing varies with instrument settings (from less than 200 ms to several seconds).

The optical spectrum analyzer's dynamic range can be limited by stray light in the monochromator when large signals are present at other wavelengths. Stray light from high-level signals can enter the monochromator output fiber and mask a low-level signal. It turns out that stray light in the monochromator is uniformly distributed over any given small area. One of these areas is the small area surrounding the output beam. Because of this fact the firmware can displace the output fiber away from the output beam and get an accurate estimate of the stray light level at the output beam. The stray light level can then be subtracted from measurements of the output beam amplitude to increase overall dynamic range. This is the background light subtraction technique mentioned earlier. More details about this technique are given in the article on page 80.

In our implementation, the two-axis micropositioner, which positions the output fiber, is chopped between its normal position and a displaced position. The ADC readings from these two fiber locations are subtracted to arrive at a measurement that is much less sensitive to stray light levels inside the monochromator. Chopping runs at a 20-Hz rate and is automatically enabled whenever at least one chop cycle can be executed per trace bucket.

Digital Data Processing. The majority of processing performed on ADC data occurs inside the ADC interrupt service routines running in the data acquisition processor. Amplitude-versus-wavelength (flatness) corrections and conversion of linear amplitudes to dBm are the only two operations that do not occur during interrupt servicing. All measurement data sent to the main processor is in dBm units. The main processor converts the data back to linear units when a linear amplitude display is requested.

The exact processing performed on ADC data during interrupt servicing varies depending on instrument settings. Most of this variation is accommodated by providing different interrupt service routines to process ADC data for different modes of operation. This reduces processing time because the interrupt service routine does not have to spend time figuring out what mode the instrument is in. Some instrument settings are examined during interrupt servicing. For example, examining whether autoranging is enabled. These tests do not incur a large performance penalty.

The interrupt service routine also updates the output fiber micropositioner DACs as necessary during a measurement. This causes the output fiber to track the output beam's movement as the diffraction grating rotates.

Digital Signal Processing. The data acquisition section is controlled by a Motorola DSP56000 digital signal processor. It runs with a 20-MHz clock and typically executes instructions at a 5- to 10-MHz rate (two-word instructions require 200 ns to execute). Many DSP56000 instructions are capable of performing an ALU operation and moving two words of data to or from different memory spaces in one instruction cycle. The processor also has a memory expansion port, a host interface port, which is used to communicate with the main processor, and a dual serial interface port, which is used to receive ADC data.

The DSP56000 has access to four different memory spaces, one for program instructions (P), two for data storage (X and Y), and one for long word data storage (L). The L memory space is not actually separate, but consists of the X and Y spaces concatenated. The processor also has a small amount of internal program and data memory. Accesses to internal data memory are faster than for external memory so most of the data used by interrupt service routines is located there.

The DSP56000 is provided with a full complement (64K words*) of external program RAM and 16K words of X and Y data RAM. I/O is memory mapped at the upper end of Y memory and the DSP56000 provides low-overhead instructions to access this memory area. 16K bytes of each X and Y memory space provides access to two ROMs containing calibration data. One ROM is used for the monochromator and the other for the data acquisition printed circuit board.

To avoid excessive bus loading, the DSP56000's memory expansion port is buffered to form a peripheral bus to which the calibration ROMs and all peripheral devices are connected. Peripherals connected to this bus include the diffraction grating motor controller, the aperture wheel motor controller, the fiber micropositioner DACs, the current source control registers, and the ADC interface registers.

* One word is 24 bits wide.

The DSP56000's host interface bus is connected to the main processor through two 16-bit buses and a state machine. The state machine allows the main processor to write to the DSP56000's host input registers by placing 8 bits of data and a 3-bit register number onto a 16-bit control bus (only 11 of the 16 bits are used). When the DSP56000 has data to send to the main processor, the state machine reads three 8-bit host registers and formats the 24-bit result as two 16-bit data transfers over the main processor's input bus.

The DSP56000's firmware is written in C for most control functions and assembly language for time-critical data processing functions. It is stored in compressed format in the main processor's ROM. During power-up, the main processor decompresses and downloads the firmware into the DSP56000.

Main Processor. The main processor's printed circuit board is identical to the processor board used in the HP 70900B local oscillator module. Using this printed circuit board allowed us to take advantage of the thousands of engineering hours that have gone into the the HP 70900 modular spectrum analyzer firmware.

The main processor board contains an MC68020 CPU running at 20 MHz, and an MC68881 floating-point coprocessor. The board also accepts a plug-in daughter board which is currently loaded with 1M byte of ROM and 1M byte of RAM.

The main processor board has two unidirectional 16-bit buses (one output, one input), which are used to interface to other printed circuit boards in the HP 70900B local oscillator module. During the investigation phase of this project, we determined that we could build a state machine to interface these two buses with the DSP56000's host interface. Since the HP 70900 controlled its measurement hardware

over these two buses, we were able to substitute the optical spectrum analyzer's hardware without major changes in the HP 70900 firmware. This arrangement allows the DSP56000 to handle hardware control and data acquisition tasks, and the HP 70900 firmware to provide high-level user interface functions.

Originally, nearly all of the HP 70900 firmware access to external hardware went through a firmware module called the local oscillator slave. By replacing the LO slave module with our own slave module, we were able to use the HP 70900 firmware with surprisingly little alteration. Some additional work was required to add wavelength units and other spectrum-analyzer-specific functions to the user interface.

The HP 70900B local oscillator module and the HP 71450A and 71451A optical spectrum analyzers are now shipped with identical main controller boards and identical firmware. During power-on, the firmware checks to see which external hardware is present and configures itself as either a local oscillator or an optical spectrum analyzer. This has the additional advantage in that features added to the HP 70900 firmware are available in both the local oscillator and the optical spectrum analyzers.

A High-Resolution Direct-Drive Diffraction Grating Rotation System

Creating a high-resolution, high-speed positioning system that can provide
over two million data points per revolution of the diffraction grating
required a design that is much different from the gear-reduction
positioning systems typically used in optical spectrum analyzers.

by Joseph N. West and J. Douglas Knight

The wavelength tuning of the double-pass monochromator used in the HP 71450A and 71451A optical spectrum analyzers is controlled by the angular position of the diffraction grating. After the input beam is collimated by the lens, it strikes the diffraction grating, where each wavelength is dispersed at a different angle. For each angle of the diffraction grating a corresponding wavelength is passed back through the optics and focused on the center of the first-pass aperture (slit). The width of the slit determines the resolution bandwidth of the wavelengths that pass through the remainder of the system to the detector. Rotating the grating causes the dispersed wavelengths to sweep across the slit, making the monochromator act as a tunable filter. Fig. 1 shows the housing that contains the optical and electromechanical components that make up the double-pass monochromator assembly.

The angular resolution requirements for the grating positioning system in the monochromator can be determined by calculating the relationship between the angular position of the diffraction grating relative to the collimated light and the spatial dispersion of the light at the resolution bandwidth slit.

Using the grating equation* for a scanning monochromator like the one used in the HP 71450A and 71451A optical spectrum analyzers, roughly 600 microradians of diffraction grating rotation per nanometer of optical dispersion (at 1300 nm) can be calculated. To represent narrow signals, it is desirable to have at least sixteen data points across the narrowest resolution bandwidth of the instrument (0.08 nm). This translates to 200 data points per nanometer of dispersion. Dividing 600 microradians by 200 points gives an angular resolution requirement of about three microradians (0.00017 degree) per point, or about 2,100,000 data points per revolution of the diffraction grating.

Conventional Methods

The traditional approach to building such a high-resolution positioning system is to use large amounts of gear reduction (see Fig. 2). This is the approach used in most older optical spectrum analyzers. In these systems there are commonly two stages of gear reduction. The first stage might consist of a planetary gearhead with a reduction of about 20:1 which would be followed by a worm drive with an additional reduction of about 30:1 for a total gear reduction of about

600:1. The advantage of this approach is that it reduces the resolution requirements of the primary feedback device, often an optical encoder. It is possible to get away with using a fairly low-technology, 1,000-line TTL-output encoder. Looking at every zero crossing from the two quadrature channels of this encoder gives a resolution of 4,000 counts per revolution of the encoder which combined with the 600:1 gear reduction provides sufficient resolution to position a diffraction grating.

However, the gear reduction approach has several drawbacks. One of these is speed. With a 600:1 gear reduction, the diffraction grating is rotating at only 1/600 the speed of the motor. Similarly, the acceleration of the diffraction grating is only 1/600 the acceleration of the motor. Moving the diffraction grating at any significant speed requires that the motor and gear train be accelerated to very high speeds. Reversing the direction of the grating requires decelerating the motor and gear train and accelerating them to high speeds in the opposite direction. In the past this speed penalty came to be accepted as inevitable because no alternative was generally available for such a high-resolution system.

A second drawback to the gear reduction approach is backlash. Backlash is the slop exhibited by a gear train when the direction of rotation changes and the gears change from contact on one gear face to contact on the opposite gear face. A number of techniques can be applied to minimize backlash in a system. These techniques generally consist of a method for compliantly loading the gear train to ensure that the same gear faces remain in contact regardless of the direction of rotation. Fig. 3 shows a simple spring used to reduce backlash in a system that rotates over a small angle. Obviously, in a full-rotation system a more complicated scheme is required. These antibacklash techniques help, but do not eliminate backlash. Even though the same gear faces remain in contact, as the gear train reverses direction lubricants are smeared in an opposite direction and the bearings that support the gears are deflected in an opposite direction. For a high-resolution system, some degree of backlash will still be evident.

A third drawback of geared systems is susceptibility to errors caused by wear or by changes in environmental conditions. In a typical gear-reduction system, the angular position of the drive motor is monitored with an optical encoder. The angular position of the diffraction grating is not measured

* See "Diffraction Grating" on page 70.
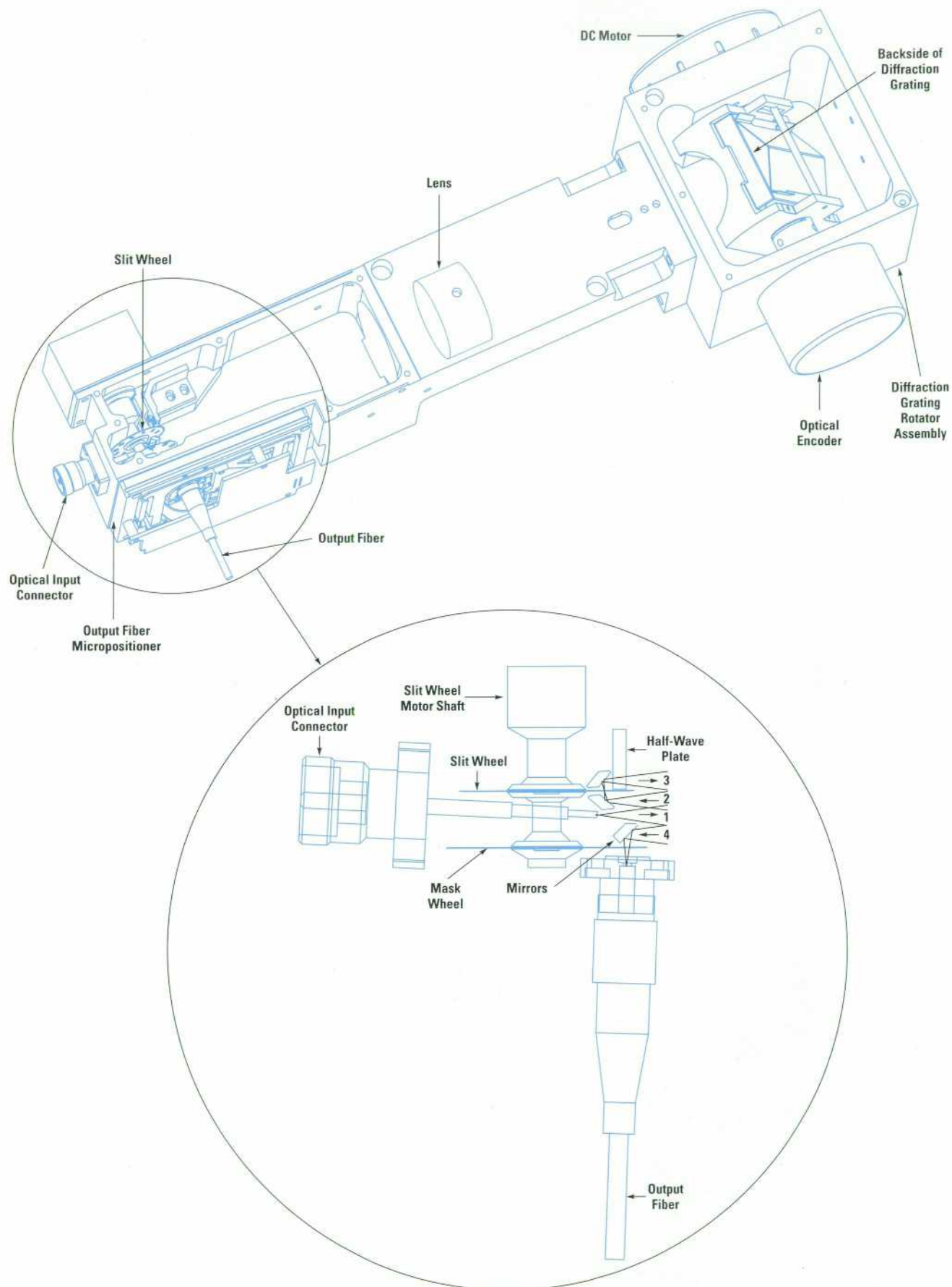
Fig. 1. The housing for the optical and electromechanical components in the HP 71450A and 71451A optical spectrum analyzers. The dimensions for the housing are 80 mm (3.15 in) high, 150 mm (6 in) wide, and 400 mm (15.75 in) long.
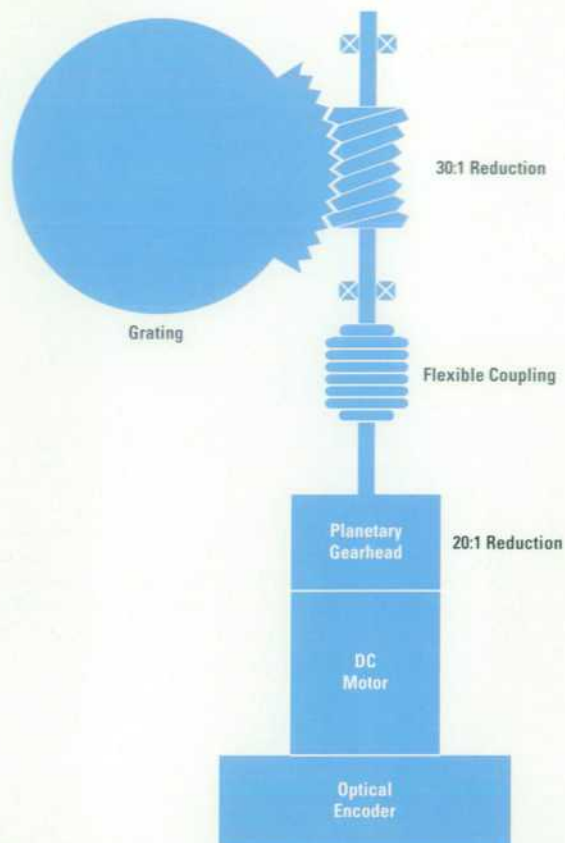
Fig. 2. A worm drive grating rotation system typically used in optical spectrum analyzers.

directly, but is inferred from the motor position and the gear ratio. As gears wear or expand and contract with temperature, or as lubricant viscosities increase over time, the actual position of the diffraction grating relative to the motor position will change. Periodic recalibration is needed to correct these errors.

Direct-Drive Grating Rotation System

Based on the drawbacks mentioned above, the decision was made to use a direct-drive, direct-readout grating rotation system in HP's double-pass scanning monochromator. Fig. 4 shows this system. A drive motor and a rotary optical encoder are directly attached to the shaft that holds the diffraction grating. This system has the following advantages:

- Direct measurement of the angle of the diffraction grating
- No errors from backlash or "wind-up" deflections in the gear train
- No susceptibility to wear and much less sensitivity to environmental changes
- Speed, compactness, and ruggedness.

Implementing a high-resolution direct-drive, direct-readout system placed some stringent requirements on the components used and the design process. Two things were required. The first was a high-torque motor for fast starting and stopping and high-speed scanning. The motor we chose is a frameless, brushless dc motor. The permanent magnet rotor uses strong rare-earth magnets and mounts directly to the diffraction grating shaft. The stator mounts to the fixed outer housing where heat can be dissipated in a controlled manner. The motor is brushless so there will be no debris
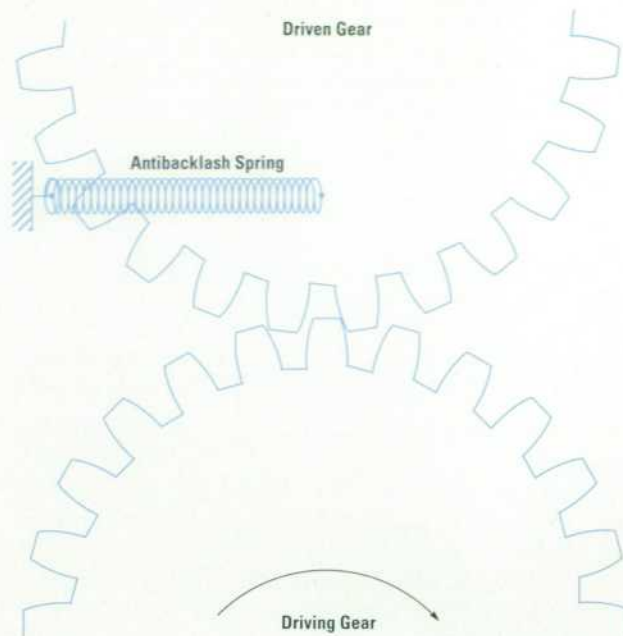
Fig. 3. Backlash is the slop exhibited by a gear train when the direction of rotation changes and the gears change from contact on one gear face to contact on the opposite gear face. A simple spring can be used to reduce backlash in a system that rotates over a small angle.
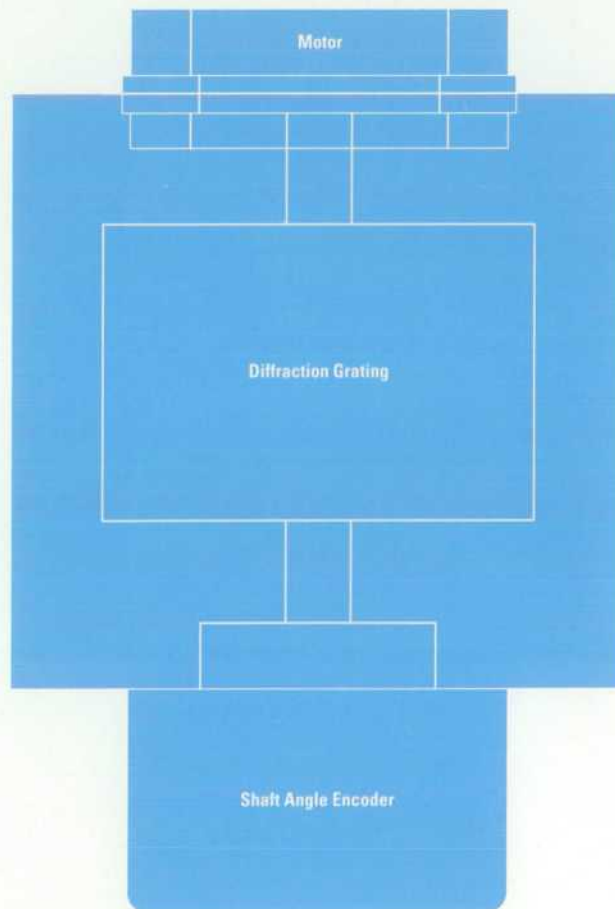
Fig. 4. The direct-drive, direct-readout grating rotation system used in the double-pass monochromator.

generated by brushes wearing and no maintenance (brush replacement) over the life of the instrument. In addition, brushless dc motors have an advantage over conventional brush-type motors in that there is no friction because of rubbing between the commutator and the brushes. Friction is a problem in the control of high-resolution systems.

A second, and more difficult requirement is a high-resolution, optical-encoder-based measurement system that is able to resolve directly more than two million points per revolution of the diffraction grating. Building such a system involved searching for the latest in optical encoder technology and then applying considerable design effort to accomplish the necessary resolution goals.

The encoder used in our system is a sine wave output incremental rotary optical encoder with 9,000 lines on the rotating disk. In an incremental encoder a single light source, typically a light-emitting diode, shines a beam of light through a rotating disk that contains radial slits which alternately transmit or block the light. The light passing through the slits is detected by two sets of photodetectors that convert the light into electrical signals. Before hitting the photodetectors, the light also passes through a phase plate containing two additional patterns of slits. These two patterns are offset slightly relative to one another so that the signals received by the two sets of photodetectors are 90 degrees out of phase. The quadrature relationship of the signals makes it possible for the user to know the direction of rotation of the encoder by looking at which channel is leading by 90 degrees and which channel is lagging. There is also a third channel that provides an index pulse once per revolution for determining absolute position. The outputs of the main A and B channels are very close to sinusoidal (see Fig. 5).

Each zero crossing of the A and B channels increments or decrements a position counter to provide coarse position information, depending on the relative phase of the A and B channels. To increase the resolution beyond the usual 4× linecount value, a process called interpolation is used. Because the signals are sinusoidal, there is additional analog information between zero crossings which can be extracted. Commercial circuits are available that perform the interpolation function, but they typically have interpolation ratios that are convenient numbers in the base 10 number system, such as 5× or 50×. For a digital control scheme, it is more convenient to have an interpolation ratio that is an integer power of two. The HP 71450A and 71451A optical spectrum analyzers use an interpolation ratio of 64:1 (actually 256:1 for data acquisition, but the two lowest-order bits are not used for
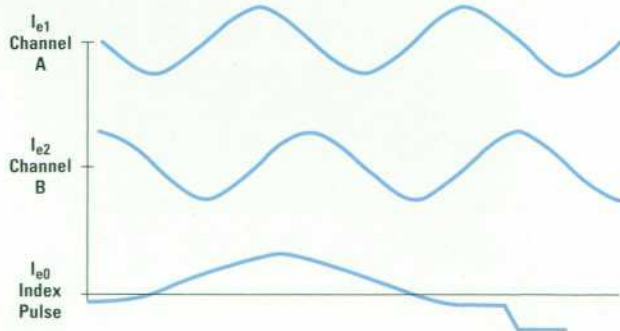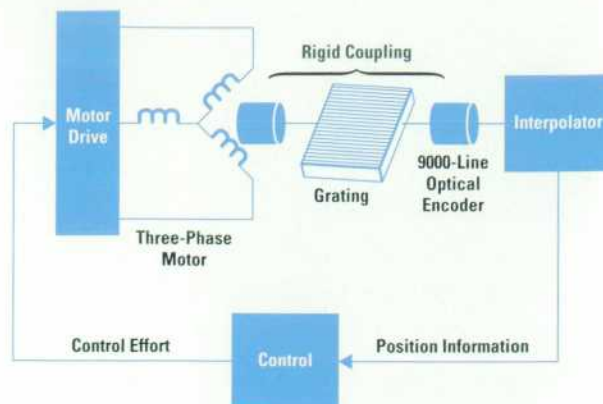
Fig. 6. A block diagram of the closed loop diffraction grating system.

control). This gives a resolution of $9000 \times 4 \times 64 = 2,304,000$ counts per revolution determined directly from the encoder.

Interpolation in this design is achieved by amplifying the two sinusoidal outputs of the encoder until the minimum and maximum values are just within the range of an analog-to-digital converter (ADC). The outputs of the ADC are then a digital representation of the sine and cosine signals. The ratio of the two digitized outputs is the tangent of the angle. By looking up the ratio in an arctangent table, the angle that is the interpolated fractional position between the sine and cosine zero crossings can be found. The accuracy of the interpolation is dependent upon the degree of distortion in the sine and cosine signals, the phase angle between them, and the number of resolvable bits in the analog-to-digital conversion process.

One of the problems encountered in controlling high-resolution systems such as this is the change in behavior of the friction in the system as the system changes from moving to a fixed position. This happens when the grating is either tuned to a fixed wavelength in zero span or it is momentarily stopped while changing directions at the beginning or end of a sweep. The degree of resolution is so fine that the difference between the static case and the dynamic case becomes readily apparent.

While the system is in motion, that is, servoing to a moving target position, the friction is a nicely behaved linear damping term. The rotational inertia of the system interacting with the motor winding resistance forms a simple pole. There is also a pole at zero frequency since a constant input voltage to the motor gives a steadily increasing angular position. This is a fairly simple system to close a servo loop around (see Fig. 6). Fig. 7a shows the open-loop frequency response of the system while in motion. The noisy measurement at frequencies below 1 Hz is because of a signal-to-noise ratio problem in this particular measurement and not an indication of the actual low-frequency response.

When the system is servoing to a fixed target position, static friction will lock the pieces together. The elastic behavior of the pieces then changes the system behavior to that of a spring-mass system with a complex pole pair. Fig. 7b shows the measured loop characteristics of the system when it is servoing to a fixed target position. Again, the entire open-loop response is shown, not just the rotor characteristics.

Fig. 5. Electrical outputs from the optical shaft encoder.
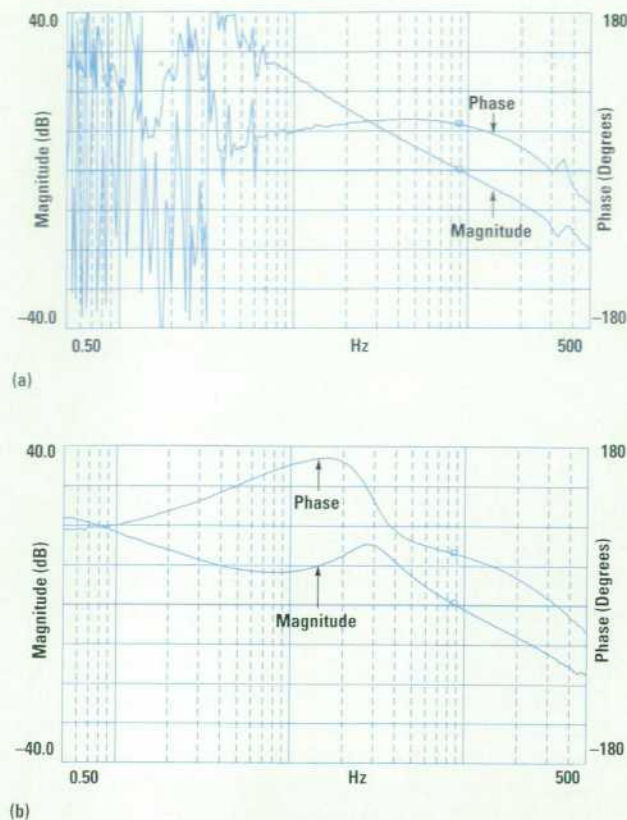
Fig. 7. Frequency response of the diffraction grating rotator (a) when the target position is moving and (b) when the target position is fixed.

The spring-mass resonance in the fixed position is readily apparent from the peak in the magnitude response at 30 Hz.

The torque required to break the system loose from a fixed position to rotation is not a well-known value. The best that can be done is to specify the maximum breakaway torque of the mechanical elements and then design the control system so that it is able to deal with the system breaking free at some lower torque. The amount of breakaway torque will vary depending on the position of the system, the current environmental conditions, and other variables.

At very low rotation rates, the system rapidly jumps between servoing to a moving target position or to a fixed target position. If the loop is not compensated to take this change of behavior into account, the result can be a system that is stable when moving but that oscillates when it servos at a fixed position.

In our design we found a single set of loop compensation values that provide a stable response for either operating mode, ensuring that the system works well under all conditions.

Conclusion

Direct-drive technology has been applied with great success in a number of industrial applications ranging from phonograph turntables to industrial robots and military gun turrets. Applying these techniques to an optical spectrum analyzer produces a system that provides fast, accurate, and reliable rotation of the diffraction grating, and with regard to motion control, brings the latest technology to optical spectrum analysis.

A Two-Axis Micropositioner for Optical Fiber Alignment

A positioning system with submicron resolution is used to keep the output fiber accurately aligned with the light coming out of the monochromator during movement of the diffraction grating.

by J. Douglas Knight and Joseph N. West

The double-pass monochromator design used in the HP 71450A and 71451A optical spectrum analyzers offers a number of performance advantages over competing monochromator designs. Several of these performance advantages come from the secondary filtering effects of the optical fiber used at the output of the second pass of the monochromator. The limited cross-sectional area and limited numerical aperture of the fiber help reject stray light, giving good dynamic range performance and good spurious response rejection. Coupling the light into fiber also allows the use of a small, low-noise photodetector which results in excellent sensitivity even with rapid sweep speeds and minimal video filtering. In addition, going into fiber allows the instrument to have an optical monochromator output that offers both fixed-wavelength and swept-wavelength modes of operation with a full range of resolution bandwidths selectable by the user. These advantages are significant, but designing a system to keep the output fiber accurately aligned with the light coming out of the monochromator during sweeps proved to be a considerable design challenge.

The Positioning Problem

Ideally, in a perfectly symmetric double-pass monochromator, the spot of light at the output of the second pass would not move. However, in reality, as the diffraction grating rotates and the instrument sweeps in wavelength, the spot of light at the output of the monochromator does move slightly in two dimensions. It is therefore necessary to track the moving spot with the output fiber to capture the light completely and to realize the desired filtering effects.

In the dispersion direction (y-axis in our implementation) the movement of the output spot is the result of asymmetry in the system. The second pass is farther off the axis of the lens than the first pass. This asymmetry is necessary to avoid picking up light from the first pass with the photodetector. The movement in the y-axis is predicted by theory and is consistent from unit to unit.

In the nondispersion direction (x-axis), the movement of the output spot is the result of manufacturing tolerances that cause the lines of the diffraction grating not to be perfectly parallel to the axis of rotation of the diffraction grating (see Fig. 1). Parallelism and perpendicularity of critical parts in the grating rotator assembly are aligned to very close tolerances. Even so, tens of micrometers of beam movement still result. If the lines of the grating are not perfectly parallel to

the axis of rotation, the lines of the grating will precess* about the axis of rotation as the grating rotates, causing the output spot to move in the x-axis. For a given monochromator this movement is very repeatable and tracking is possible with a precise positioning device.

Tracking the Output Spot

Once it was determined that a micropositioning device was necessary to track the output spot, other advantages of having such a device were envisioned. One advantage is noise and stray light cancellation. Most systems that attempt to do noise cancellation chop the optical signal with an aperture that alternately passes or blocks the light to the detector. When the light to the detector is blocked it is possible to measure the electrical noise of the detection system which can be subtracted from the reading obtained when the aperture passes light to the detector.

When the aperture passes light to the detector, the output of the detection system represents signal + stray light + electrical noise. Subtracting electrical noise leaves signal + stray light.

* Precession is the type of motion experienced by a gyroscope or top when a force is applied at right angles to the axis of rotation.
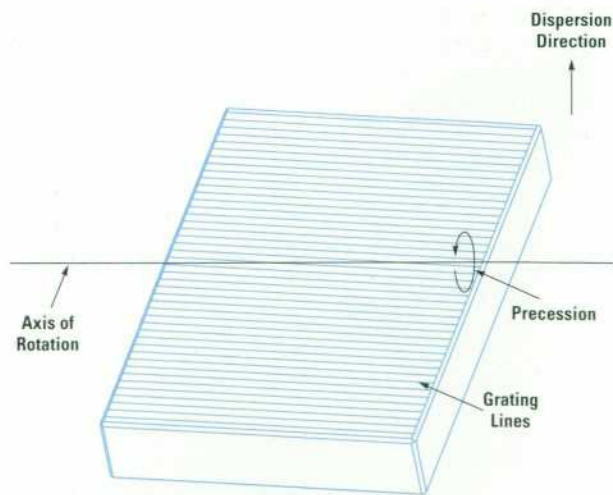
Fig. 1. A diffraction grating in which the lines of the grating are not perfectly parallel to the grating's axis of rotation. This slight misalignment will cause movement of the output spot in the x-axis.

With a fast and accurate micropositioner at the output of the monochromator it is possible to perform another kind of optical chopping to remove both electrical noise and stray light. If we assume that the stray light is relatively uniform in the region around the output spot as it might be in the case of scatter from optical components and diffuse reflections from the inside of the monochromator cavity, physically moving the output fiber laterally away from the output beam would allow a measurement to be made of stray light + electrical noise. Alternately moving the output fiber into and out of the output beam allows the stray light + electrical noise term to be subtracted from the signal + stray light + electrical noise term leaving only the signal value. The digital signal processor described in the article on page 68 controls the micropositioner and performs the subtraction of stray light and electrical noise from the measurement. This mode is activated automatically in the HP 71450A and 71451A analyzers when the user requests a very sensitive setting that results in a sweep time greater than 40 seconds.

Having an electrically actuated micropositioner at the output of the monochromator also eliminates the need for the user to make manual adjustments to the second-pass aperture (output fiber) relative to the first-pass aperture (slit) to maintain signal symmetry or to adjust the optical output of the HP 71451A for fiber-in/fiber-out measurements.* Most optical spectrum analyzers that have double monochromators require the user to adjust the optical output for maximum signal strength at a given wavelength with manual micrometers. With the HP optical spectrum analyzers, output coupling is automatically maintained over the entire wavelength range. If the instrument is dropped or experiences significant changes in temperature, there is an **AUTO ALIGN** key on the front panel of the instrument that the user can push to initiate an alignment routine to ensure that optimum output coupling is reestablished.

Implementation

To couple light into the output fiber efficiently, the fiber must be accurately held in the output focal plane of the monochromator while it is aligned to the spot within fractions of a micrometer. The output fiber must be able to track the movement of the spot smoothly as the instrument sweeps. Thus, the micropositioner must have the following characteristics:

- Submicrometer resolution
- Smooth operation (no roughness from rolling elements)
- No friction from sliding members
- No screw or gear backlash
- Compact size
- Immunity to changes in orientation relative to gravity
- Some insensitivity to vibration and environmental disturbances
- Negligible movement in the z-axis while moving in the x-axis and the y-axis
- Approximately ±175 μm of travel in each axis
- Remote activation
- Fast response time.

The micropositioner design used in our double-pass monochromator consists of a two-axis planar flexure plate actuated by voice coil linear actuators with strain gauges mounted on the flexing beams to sense the deflections of

those beams and therefore the displacement of the fiber (Fig. 2a). The strain gauges and the linear actuator associated with each axis together form a closed-loop position servo system (Fig. 2b). The displacement of each axis of the micropositioner is determined by the target position, which comes from the calibration ROMs in the data acquisition unit described in the article on page 68. These ROMs contain factory calibrated x-y position values (described below) that are correlated with the angular position of the diffraction grating. To select a particular target position, the data acquisition unit computes the wavelength associated with the current position of the diffraction grating and uses this wavelength to index into the calibration ROMs to retrieve the x-y pair associated with the wavelength of interest. Each x-y pair is sent to the appropriate 12-bit digital-to-analog converter where it is converted to a voltage and applied as reference input to the micropositioner servo system shown in Fig. 2b. The voltage value from the converter is proportional to the displacement each axis must make to keep the output fiber aligned with the light coming out of the monochromator during movement of the diffraction grating.

Flexure Plate. The flexure plate used in the micropositioner consists of a frame-within-a-frame design (see Fig. 3). The outer frame mounts rigidly to the monochromator housing shown in Fig. 1 on page 76. A four-bar (parallelogram) flexure linkage connects the inner frame to the outer frame and guides the inner frame in x-axis motion. A second four-bar flexure linkage connects the fiber mounting collar** to the inner frame and guides the fiber mounting collar in y-axis motion. Actually the inner frame and the fiber mounting collar each move in an arc, but for small displacements the motion is close to linear. During calibration at the factory the x and y locations of the micropositioner are recorded for each angular grating position. It is the ability of the positioner to repeat to these preset positions that is important and not the ability to reach absolute locations in Cartesian space.

Flexures of various types have been used in designs for many years. The term flexure refers to an assembly composed of one or more flexible hinges made of an elastic material, which is typically a metal that deflects within its elastic range. Flexure assemblies are sometimes constructed from strips of metal such as beryllium copper or phosphor bronze. Flexures can also be made by machining material away from a solid metal plate in such a manner as to leave webs in appropriate locations to act as flexure joints.

The micropositioner in the double-pass monochromator is a monolithic flexure of this second type. Two types of flexure joints are used in the micropositioner. The first type is called a transverse circular flexure joint. This type of joint consists of the web that remains when two holes are drilled and reamed near one another. This type of joint localizes the bending to a small region, which gives good stiffness perpendicular to the plane of motion as well as good torsional stiffness. Because the bending is localized, this type of joint sees relatively high stresses for a given angular displacement. The spring constant for a right circular flexure hinge can be expressed by the equation:[1]

$$k_B = \frac{M}{\theta} \approx \frac{2Ebt^{5/2}}{9\pi R^{1/2}}$$

---

* The HP 71450A doesn't provide customer access to the monochromator output fiber.

** The fiber mounting collar is where the monochromator's output fiber is mounted.
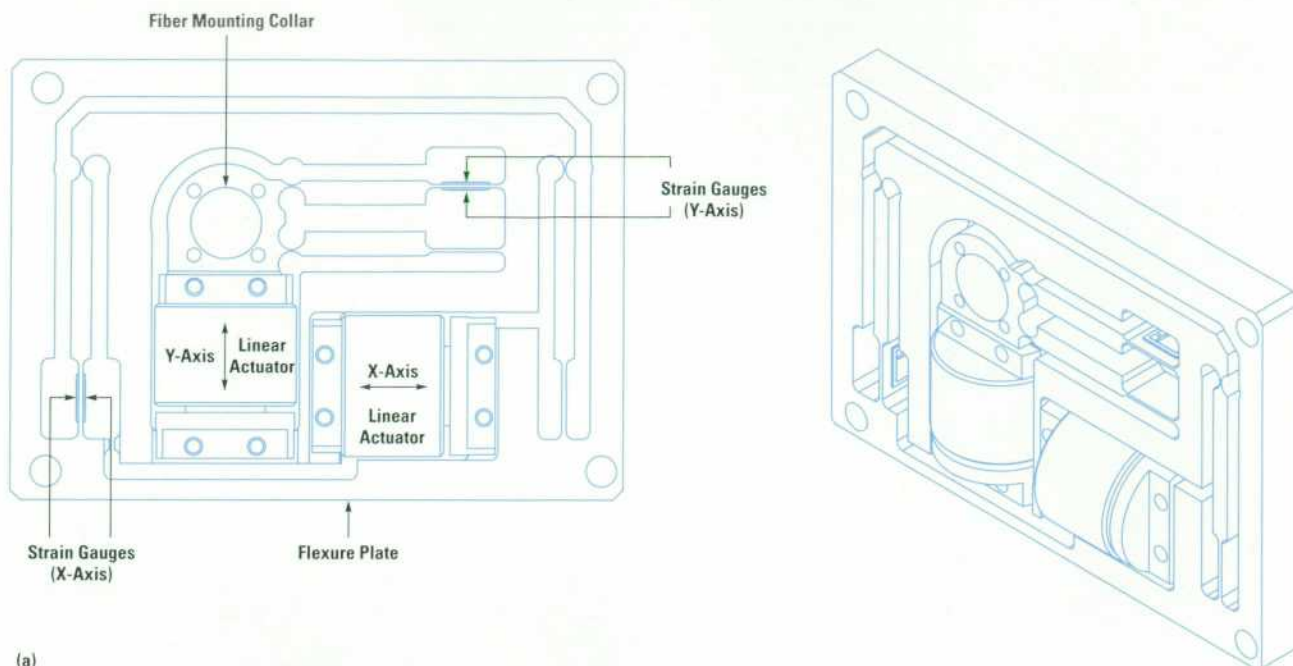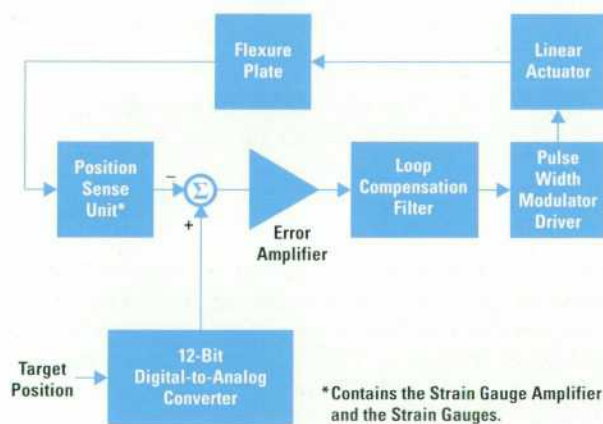
Fig. 2. (a) Top and side views of the x-y micropositioner used in the double-pass monochromator. (b) The servo loop schematic for one axis of the positioner.

where:

> M is the moment required to bend the flexure through an angle θ
> E is the elastic modulus of the flexure material
> b is the width of the flexure joint
> t is the minimum thickness of the flexure
> R is the radius of the right circular cutouts.

Fig. 4 shows these parameters.

The maximum stress in the right circular hinge can be expressed by the equation :

$$\sigma_{max} \approx \frac{4Et^{1/2}\theta}{3\pi R^{1/2}}$$

The second type of flexure joint used in the micropositioner is a thin rectangular beam flexure similar to the strips used in nonmonolithic flexures. This type of flexure joint distributes the bending over a longer distance and was chosen because it is suitable for mounting strain gauges to sense deflection. Standard rectangular beam equations can be used for calculations associated with this type of flexure joint.

A discussion of flexures would not be complete without mentioning the topic of fatigue. The designer must constantly calculate and recalculate the fatigue life of the flexure joints as the design is modified to ensure that fatigue failures will not occur even after many years of continuous use.

Linear Actuator. Each axis of the micropositioner is actuated by a voice coil linear actuator. The linear actuator consists of a moving coil in a fixed permanent magnet structure. It is very much like a voice coil used in a loudspeaker driver, but performance is optimized for force rather than frequency response. Since there is no contact between the coil and the magnet structure, the actuator does not contribute any roughness or sliding friction to the operation of the positioner.
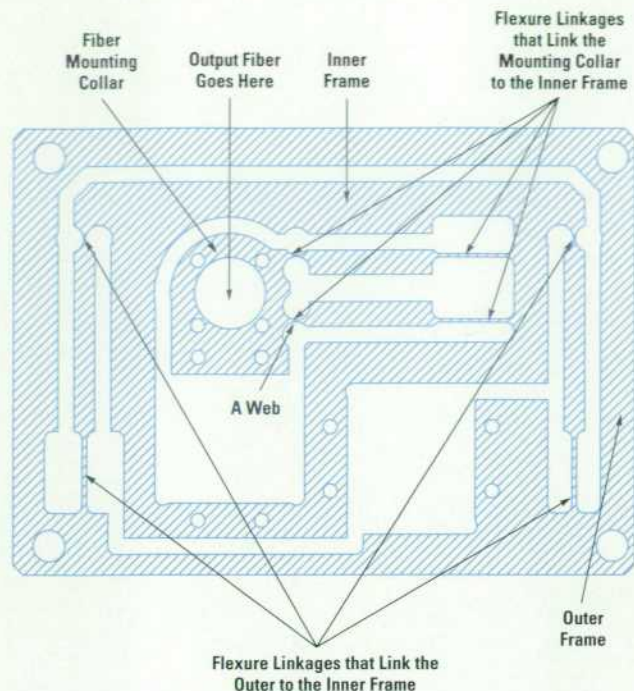
Fig. 3. The flexure plate used in the micropositioner.

Strain Gauges

Strain Gauges. The feedback sensors used in the closed-loop position servos are strain gauges. Strain gauges were invented in the 1930s and are used primarily in force transducers or load cells. They are essentially resistive elements whose resistance changes with strain as they are stretched or compressed. They are bonded to the flexing beams of the positioner with special low-creep adhesive. As the beams bend, the outer surfaces of each beam experience tension and compression. Since the strain gauges are tightly bonded in place, each gauge experiences essentially the same strain as the surface upon which it is mounted.

The signals produced by a strain gauge are rather small. For the geometry of the micropositioner used in the optical spectrum analyzer, the change in resistance for each strain gauge is about 0.5 milliohm (out of 350 ohms) for a one-micrometer position change in the flexure. To detect small changes such as this, balanced bridge circuits are the usual
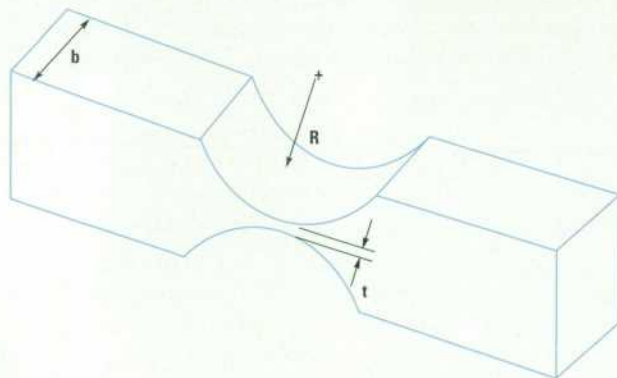
Fig. 4. The locations on a transverse circular flexure joint for the parameters used in the equation to compute the spring constant for the circular flexure.
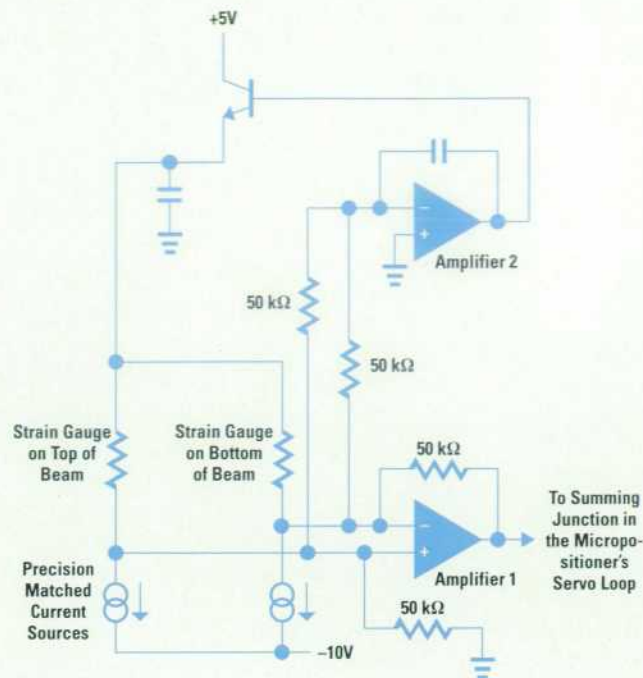
Fig. 5. The strain gauge amplifier used in the micropositioner. There is one amplifier for each axis.

choice. In the case here, it is desirable to have a linear relation between change in position and change in resistance. To achieve this, the usual Wheatstone bridge circuit must have two of its resistor elements replaced with matched current sources (see Fig. 5).

The change in voltage that results from the current flowing in the strain gauge resistors must be amplified to usable levels. The current is limited to around 5 mA to limit heating of the flexure. The output from the bridge circuit is around 4.8 microvolts per micrometer of position shift. This signal is amplified to 0.7 millivolts per micrometer for use in the position servo loop.

An amplifier used in an application such as the double-pass monochromator must have good input characteristics, including low input offset voltage and low input offset drift. Good common-mode rejection is also desirable. The amplifier used in our design (an OP 77) has very good input characteristics, but the common-mode rejection is not adequate. The amplifier alone has a common-mode rejection ratio of one microvolt per volt or 120 dB. This can be degraded to about 48 dB by the 0.1% resistors used in the rest of the circuit. The common-mode voltage in the original breadboard was around 3.25 volts. This could produce an output offset of up to ±13 mV, which is the equivalent of 19 micrometers of movement in the flexure. Since submicrometer resolution was needed, this was clearly a problem.

The scheme used in our monochromator is to drive the strain gauges in such a way that the common-mode voltage is sensed and servoed to zero volts. In Fig. 5 amplifier 1 is the strain gauge amplifier, while amplifier 2 is used to sense the common-mode voltage and drive the strain gauge voltage source up or down until the common-mode voltage is zero at the sense point.
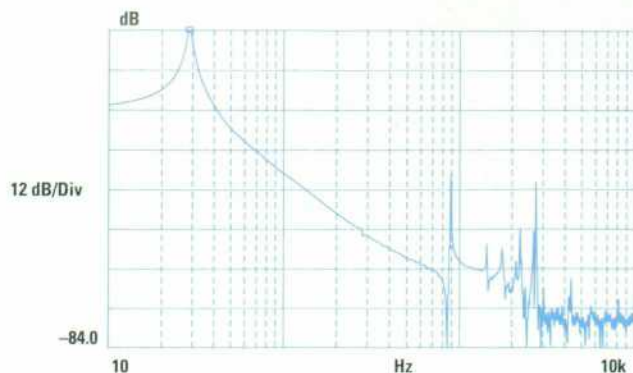
Fig. 6. The flexure frequency response from the actuator drive to the strain gauge output.

Performance

The flexure frequency response from the actuator drive (the input) to the strain gauge amplifier output is shown in Fig. 6. This is for the x-axis; the y-axis is very similar. The response is dominated by the resonance at 29 Hz. This resonance is the natural response of the spring-mass system formed by the flexure beam spring characteristic and the mass of all the moving parts supported by the flexure. The resonance introduces a 180-degree phase shift at frequencies beyond resonance.

We wanted to be able to chop the optical beam using the flexure at around a 20-Hz rate. To do this with a square-wave position characteristic requires that the control loop have a bandwidth that is several times the chop rate. The loop bandwidth must include the third harmonic of the chop rate, and the fifth harmonic or higher is desirable if the settling time is to be reasonably short. In addition, the 29-Hz resonance is a manifestation of a sensitivity to vibration. If the loop bandwidth includes the resonance then the loop will damp the resonance.

Closing a feedback servo loop around a resonance in which the resulting open-loop gain at resonance is greater than one is usually avoided because oscillation is the typical (and undesired) result. In our case the design required that we close the loop without oscillation. If the 29-Hz resonance were the only complex pole pair in the flexure, then the design of the feedback would have been straightforward. However, there were also mechanical resonances at 950 Hz and

at 2.2 kHz (2.7 kHz in the y-axis) that put an upper limit on what the loop bandwidth could be. In addition, there was a delay mechanism because of the construction of the voice coil actuator that added phase shift to the flexure response, constraining the amount of phase margin that could be obtained. Phase margin has a large impact on the transient response of the system. (55 to 60 degrees of phase margin turns out to give a very nicely behaved transient response, while 45 degrees is about as little as can be allowed if there is any concern over the transient response.) Since we were interested in a good transient response, the available phase margin became the determining factor in setting the loop bandwidth at around 100 Hz.

The target position for the servo loop is taken from the output of a 12-bit DAC. This is a settability of one part in 4096. However, only a portion of the DAC range is used to drive the flexure. The remainder of the range is used to compensate for the mechanical and electrical tolerances that determine the strain gauge amplifier output for the undriven rest position of the flexure.

Conclusion

The micropositioner was one of a number of key components needed to be able to build a double-pass scanning monochromator. As described above, it provides a means for translating the output optical fiber in a plane perpendicular to the output light beam to track the output light beam during rotation of the diffraction grating. It is able to move quickly and accurately over the necessary range of motion and has proven to be a valuable asset in achieving our performance goals.

Acknowledgments

Reference

1. J. M. Paros and L. Weisbord, "How to Design Flexure Hinges," *Machine Design*, November 15, 1965, pp. 151-156.

A Standard Data Format for Instrument Data Interchange

This standard format allows many HP analyzers to exchange data with each other and with applications software. Utilities provide data conversion, editing, viewing, and plotting and a function library provides access to SDF data from programs.

by Michael L. Hall

The Standard Data Format (SDF) is a record-based binary data file format that is used to store data from a variety of analyzers manufactured by the Hewlett-Packard Lake Stevens Instrument Division. These analyzers range from portable acoustic analyzers and low-frequency FFT analyzers to RF vector signal analyzers (see Fig. 1). The SDF file format is flexible enough to contain multiple channels of data, multiple data results in a single file, multiple scans of a data result (waterfall), and deep capture of contiguous time data. The HP 894xxA vector signal analyzer described in the article on page 6 uses SDF to store trace data (single results), time capture data (up to one million time samples), and waterfall data.

Included with each instrument that saves data files is the Standard Data Format utilities, a set of MS-DOS® programs that make it possible to convert data from one format to another, edit SDF records and data, graphically view data, and plot (single or batch) data from SDF files. Fig. 2 lists the SDF utilities.

Interchangeability

Storing data in SDF format allows many instruments and applications to interchange measurement data, time capture data, and waterfall or map data (see Fig. 1).

Since the amount of memory varies from instrument to instrument, there are restrictions on the amount of data that each instrument can import from another source. For example, the HP 894xxA vector signal analyzer with Option AY9 is restricted to one million samples of time capture data. Other instruments, depending upon the amount of memory purchased, have other restrictions.

The HP 894xxA contains a 3.5-inch flexible disk drive that supports both MS-DOS and HP Logical Interchange Format (LIF) file systems. SDF data stored on a DOS file system can be directly interchanged between this analyzer and other instruments in the first column of Fig. 1 by exchanging disks.

Other instruments either do not store data directly in SDF format or do not support MS-DOS flexible disks. The SDF utilities include programs to translate instrument data from other formats to SDF (see "Instrument Translators" in Fig. 2).

Interface	
LIF	Logical Interchange Format filer
LIFDIAG	LIF diagnostics
DOWNLOAD	Download HP 3560A or HP 3569A file via RS-232
HPIB63	Transfer HP 3562A or HP 3563A traces via HP-IB

Instrument Translators	
63TOSDF	HP 3562A or HP 3563A to SDF
SDFTO63	SDF to HP 3562A or HP 3563A
660TOSDF	HP 35660A to SDF
60TOSDF	HP 3560A to SDF
69TOSDF	HP 3569A to SDF
88TOSDF	HP 3588A to SDF
89TOSDF	HP 3589A to SDF

Application Converters	
SDFTOASC	SDF to ASCII format
SDFTOML	SDF to MATLAB matrix format
SDFTOMX	SDF to MATRIXx matrix format
SDFTO58	SDF to Data Set 58

Examining Files	
VIEWDATA	Graphically view SDF data
REPEAT	Repetitively execute another program (e.g., batch plot)
SDFPRINT	Textually view SDF headers

Changing Files	
ASCTOSDF	ASCII data to SDF
SDFEDIT	Change SDF headers
SDFTOSDF	Split SDF file

Program Interface	
SDFUTIL	SDF libraries
FILTERSDF	MATLAB filter time capture files

Fig. 2. The Standard Data Format utilities.

Direct Exchange (Supports MS-DOS)	Exchange through Translator	Supported Application/Data Formats
HP 89410A	HP 3562A	Spreadsheets
HP 89440A	HP 3563A	General ASCII Data
HP 35665A	HP 35660A	MATLAB
HP 35670A	HP 3560A	MATRIXx
HP 3566A	HP 3569A	Data Set 58
HP 3567A	HP 3588A	Libraries
	HP 3589A	
	HP 3587S	

Fig. 1. Standard Data Format supported instruments and applications.

Foreign File Systems

Some instruments do not support MS-DOS as a file system, but do support LIF, including the HP 3562A, HP 3563A, HP 35660A, HP 3588A, and HP 3589A analyzers. To support data interchange with these instruments, the SDF utility LIF can read and write on a LIF disk in the computer's internal disk drive or an HP-IB connected external disk drive. The utility LIF can:

- Identify HP-IB card location and any connected external disk drives
- Copy files to and from LIF disks using optional wild-card file names
- Delete LIF files
- List a LIF directory
- Initialize a LIF disk.

In addition, the SDF utility LIFDIAG performs LIF disk back-ups even when a disk is damaged. It includes the ability to read, modify, and write individual disk sectors (256-byte blocks).

The HP 3560A and HP 3569A analyzers are battery-powered portable instruments that contain a nonvolatile RAM disk and have an RS-232 interface. The SDF utility DOWNLOAD can transfer a file from the instrument to the computer via RS-232. In addition, the extended data transfer utilities for the HP 3569A contain FILE69, a filer that can bidirectionally transfer groups of files via RS-232.

SDF Format

SDF files contain binary records that describe various attributes of the data (see Fig. 3). All records contain a combination of the following types of data: 8-bit, 16-bit, and 32-bit integer, 32-bit and 64-bit floating-point, and null-terminated strings (C-style strings). Different processors can interpret multibyte numbers in two ways: either with the most significant byte (MSB) of the number appearing first (lowest address in memory) or with the MSB last. Numeric quantities are stored in an SDF file with the MSB first.

The first two bytes in an SDF file contain a format descriptor to identify this file as being in SDF format. Following the format descriptor are one or more records in the following format:

Field	Size
recordType	16-bit integer
recordSize	32-bit integer
record dependent	record dependent

The recordType defines the type of record and the contents of the record dependent section. The recordSize is the size of the record in bytes and is used to find the next record in the file. All records of the same type must be contiguous in the file (see Fig. 3). The SDF_FILE_HDR is the first record in the file and the SDF_MEAS_HDR is the second record in the file.

The SDF_FILE_HDR record contains a pointer to the first record of each of the other record types (except SDF_MEAS_HDR) and a count of how many records of each type are in the file. The SDF_FILE_HDR also contains information identifying the source of the data in the file and the date and time the data was created.

Fig. 3. Standard Data Format file structure.

The SDF_MEAS_HDR record contains information identifying the measurement that created this data, including the frequency and average parameters.

The SDF_DATA_HDR record contains information identifying each type of data that is contained in the file, including the data's type, length, format, and number of logical channels (rows and columns) of the result. Most instruments store only one logical channel of one data result in a given file.

The SDF_VECTOR_HDR record contains information relating a logical channel of a given data result to a physical input

channel (e.g., spectrum data) or a pair of channels (e.g., frequency response data). There must be one of these records for each logical channel (row and column) referenced in each SDF_DATA_HDR.

The SDF_CHANNEL_HDR record contains information identifying each physical data input channel in the instrument that is referenced by any data in the file. Most instruments have only one or two channels. This record includes information on how the input is set up (range, coupling, input impedance), identifies the filters applied to the input data, and indicates whether the input overloaded while acquiring data.

The SDF_HDR (Unique) records describe information that is unique to a particular instrument and not used by other instruments. For the HP 894xxA, there is a unique record associated with the data that describes the format of the displayed trace (coordinates and scaling) at the time the data is saved. The display is restored to this state when the data is recalled and displayed within the HP 894xxA.

The SDF_SCAN_STRUCT record describes data results that are organized in scans. Scans are multiple measurements of the same result separated in time. Both waterfalls or maps and time capture results have scans. This record contains the number of scans and the data for the scan variable. Each scan has a value that describes its location in time; it may be a time offset from the start of the measurement, a shaft speed (r/min) value (the HP 35670A has a tachometer input), or just a scan number.

The SDF_SCAN_BIG record describes data results that are organized in scans (similar to SDF_SCAN_STRUCT). This record describes the type of scan orientation and contains the number of scans used for long (>32767 scans) scan-oriented data files. This record can serve as a replacement for the SDF_SCAN_STRUCT.

The SDF_SCAN_VAR record describes one scan variable for data results that are organized in scans. Each scan has a value that describes its location in time. It may be a time offset from the start of the measurement, a r/min value (the HP 35670A has a tachometer input), or just a scan number. There can be multiple SDF_SCAN_VAR records in a file, so that multiple attributes can be described (e.g., time and r/min for each scan).

The SDF_COMMENT_HDR record describes any free-form text that the user may want to associate with the data file. It may contain information describing the particular test setup with which the data was acquired. It may contain anything the user wants. This record has a variable length so any amount of text can be included in the data file.

The SDF_HDR (x) record contains the x-axis data for any data result that contains arbitrarily spaced x data (i.e., not linearly or logarithmically spaced). For the HP 3566A, for example, swept sine frequency response data is arbitrarily spaced.

The SDF_HDR (y) record contains all the y-axis data for all data results in the file. Information in the SDF_DATA_HDR and the SDF_SCAN_STRUCT determine the location of the y-axis data within this record. The data is in the same order as the SDF_DATA_HDR records in the file.

Expandability

SDF is not a static standard. Since each record has a specified size, it is possible to add fields to the end of a record and increase its record size. This mechanism has been used to enhance the standard twice in the past. The first revision of SDF was released with the first release of the HP 3566A and HP 3567A analyzers. The second revision of SDF was released at the time of the HP 35665A analyzer, later releases of the HP 3566A and HP 3567A analyzers, and the release of the SDF utilities. It included support for time capture files, better support for waterfall (scan-based) data, and better accuracy for frequency parameters. The third release of SDF includes support for multiple scan variables, comment records, and larger data records.

Exchanging Data with Applications

Any particular instrument is not always capable of performing the analysis desired by the user. To address this need, the SDF utilities contain programs to convert SDF data files to other formats that are easily imported into computer-based analysis applications.

The most universal format used by many applications is ASCII numbers. Spreadsheet programs can import ASCII numbers into columns or rows of a spreadsheet. The SDF utility SDFTOASC provides the capability to convert any portion of the data in an SDF file to a flexible ASCII format. The user can specify any C-style printf format specifier.

PC-MATLAB from The MathWorks, Inc., is a software package for general digital signal processing and filtering. An optional signal processing toolkit provides the ability to perform digital filtering and FFT operations on time data. MATLAB's basic data type is a matrix. The SDF utility SDFTOML converts any portion of an SDF data file to MATLAB matrix format, allowing complex data to be imported directly into MATLAB.

MATRIXx, a product of Integrated Systems, Inc., is a software package for control system analysis. It is similar to MATLAB in that the elementary data type is a matrix. The SDF utility SDFTOMX converts any portion of an SDF data file to MATRIXx matrix format.

Data Set 58 is the universal file format for mechanical applications. The SDF utility SDFTO58 converts any portion of an SDF data file to Data Set 58 format as a matrix.

Additional third-party converters are available to convert SDF data to other formats.

The SDF utility REPEAT makes it easier to deal with a group of data files by automating batches of operations on SDF files, such as file conversions.

Examining SDF Files

Sometimes the user does not want to postprocess measurement data, but wants to verify what data is in a file, or wants to make a graphical hard copy of the data with the appropriate annotation. The SDF utility VIEWDATA allows the user to view up to three traces of data simultaneously in either a

stacked or overlaid format. Full x and y annotation is provided, just as if the trace were being displayed by the instrument. The user has a full choice of coordinate systems to display the data: linear, log, or dB magnitude, real part, imaginary part, wrapped or unwrapped phase, Nichols (dB versus phase), or polar (real versus imaginary). Full marker (and offset marker) functionality is provided with the arrow keys or the mouse used to control movement. Overlaid text and imported HP-GL graphics can be placed over the trace. Hard copy is provided to any HP-GL plotter or PCL printer. With the SDF utility REPEAT, VIEWDATA can be used to batch plot multiple SDF files. Although many instruments can plot results directly to a hard-copy device, the batch plot capability of VIEWDATA allows the user to avoid tying up expensive instrument time doing a slow operation such as plotting.

There are many pieces of information in an SDF file in addition to the circumstances of the data acquisition in the SDF data headers. The SDF utility SDFPRINT allows the user to see this information in a textual format.

Changing SDF Files

Up to this point, I have only described a one-way flow of data being created by an instrument and being transferred to another analysis tool. The SDF utilities provide tools for creating and modifying SDF files so they can be transferred back into an instrument. You can use an analysis program to create time data to load into the HP 894xxA analyzer's arbitrary source or to create a user-defined input filter to be used with the HP 894xxA's digital demodulation measurement. The SDF utility ASCTOSDF creates a time-domain or frequency-domain SDF file from ASCII data. To customize an SDF file, the SDF utility SDFEDIT can be used to change any header field in an SDF file. Some instruments create an SDF file that has many results in it (e.g., the HP 3567A), but most instruments will only use the first result in an SDF file (e.g., the HP 894xxA). The SDF utility SDFTOSDF can be used to split a multiresult SDF file into separate SDF files so that an instrument has access to all results. Some old instruments only support a certain number of frequency points in data results (e.g., the HP 3562A supports only 801 points). The SDF utility SDFTOSDF can be used to respace data to any number of frequency points.

Accessing SDF from a Program

Some users want direct access to SDF data from a programming language. The SDF library provides that interface. Library functions allow access to all standard headers in an SDF file and allow reading and writing of the SDF data. The SDF library scales the SDF data and converts the data to the format the user requests (16-bit or 32-bit integer or 32-bit or 64-bit floating-point). The SDF library is also available as a Microsoft® Windows dynamic link library, providing access to Windows-based programs such as Microsoft Visual BASIC. MATLAB has the capability to call user-written functions (MEX files). The MATLAB function FILTERSDF is available to digitally (FIR) filter an SDF file. This is most useful with an SDF time capture file to test the effects of a compensation filter upon measured data.

Implementation of SDF

Different segments of the SDF code run on a variety of products:

- Motorola 680x0-based instruments from the Lake Stevens Instrument Division (e.g., HP 894xxA, HP 35670A)
- Intel-based MS-DOS personal computers (SDF utilities)
- HP 95LX palmtop computer (extended data transfer utilities)
- Microsoft Windows environment (SDF library dynamic link library)
- HP 9000 Series 300 and 400 workstations (SDF utilities)
- HP 9000 Series 700 PA-RISC workstations (SDF utilities).

The SDF utilities and the SDF library are written in ANSI-compliant C. Three different compilers are used to generate code for the different targets: the HP 9000 Series 700 C compiler, the Gnu C compiler for the HP 9000 Series 300 and 400 and the instrument targets, and the Microsoft C compiler for MS-DOS, HP 95LX, and Microsoft Windows targets.

In general, the same code works with all three compilers. I only found one code segment that would not compile on all compilers. During development, I found that even with the strictest error reporting turned on, not all compilers generate the same level of warnings for an offending code segment. This multicompiler approach turned out to be very useful for finding certain types of coding errors.

The C compilers predefine some compile-time symbols, which can be tested by the C preprocessor. The Microsoft C compiler predefines the symbol MSDOS (for the MS-DOS, HP 95LX, and Microsoft Windows targets) and the HP 9000 Series 700 C compiler predefines the symbol __hppa. These switches are used to control compilation to resolve the various differences in the target systems.

SDF files are in binary form with multibyte numbers stored with the most significant byte (MSB) appearing first (lowest address in memory). Intel 80x86-based computers use numbers with the MSB appearing last in the number. This means that any SDF utility that runs on these computers has to have the bytes in a number read from an SDF file reversed before it is used. The SDF library byte swaps every field in an SDF record after it is read from a file to make the byte ordering as transparent as possible to the majority of the software. Portions of the software that deal directly with data use a set of macros to swap integers and floating-point numbers.

The Microsoft C compiler defines the size of an integer as 16 bits, whereas the other C compilers define the size of an integer to be 32 bits. For this reason, the SDF library does not use integers, but instead uses the data type short to mean 16-bit integer and long to mean 32-bit integer. These data types are consistent with all the compilers.

The Microsoft C compiler defines the size of a pointer to be either 16 or 32 bits depending upon the memory model used, whereas the other C compilers define the size of a pointer to be 32 bits. Since the size of an SDF file can be more than 64K bytes (the limit of a 16-bit pointer), pointers to portions of the SDF file in memory must always be forced to be 32 bits. You can force pointers to be 32 bits with Microsoft C by compiling

with the large memory model (32-bit program addresses and 32-bit data pointers). This works for MS-DOS programs, but Microsoft Windows programs are generally compiled with the medium memory model (32-bit program addresses and 16-bit data pointers). Even in the medium memory mode, 32-bit addresses can be coerced by explicitly defining the pointer as _far (e.g., char _far *). All SDF address references in the SDF library use a define to force 32-bit addresses.

There is a side issue to the use of 32-bit pointers with a 16-bit memory model. The C standard libraries for the medium memory model all use 16-bit pointers (for example, the string access functions). There are 32-bit versions of most of the functions with a slightly different function name (e.g., _fstrlen instead of strlen). This means more defines must be used for function names to allow either function to be used.

Certain rules are used in alignment of variables within a C structure. For Microsoft C and Gnu C, any integer or floating-point variable must start on an even byte boundary. Only a char variable or array can have an odd byte size. To follow this structure alignment rule, all character arrays should be an even number of bytes in length.

The porting of the SDF utilities to the HP 9000 Series 700 target occurred after SDF had been defined and in use for several years. The variable alignment rules within structures are different on the Series 700 than on the other targets. Structures on the Series 700 are generally larger because 32-bit integers and 32-bit floating-point variables must be aligned on four-byte boundaries and 64-bit floating-point variables must be aligned on eight-byte boundaries. The solution is an SDF library function that unpacks an SDF header after reading it and packs an SDF header before writing it to a file.

Both the byte swapping and the packing and unpacking of an SDF header are handled with a table for each SDF header that describes the data type of each variable in the structure along with its length for arrays. This table also contains the ASCII name of each variable. If the variable is an enum data type, then the table also contains a pointer to a table describing the ASCII name for each of its defined values. These additional fields are used by the SDF utility SDFPRINT to print SDF headers, and by the SDF utility SDFEDIT to edit SDF headers.

Care must be taken when using C libraries to make sure that functions really are standard ANSI C and as such will be portable. It turns out that case-insensitive string comparisons are not standard. In Microsoft C, the function is _stricmp. In HP 9000 Series 700 C, the function is strcasecmp. In Gnu C, no such function exists. Part of the SDF library is a compatibility module to contain these nonstandard functions.

The SDF library is also compiled into a Microsoft Windows dynamic link library (DLL), which places additional restrictions on the use of pointers. A 16-bit pointer should not be used for the address of anything on the stack (parameters to a function or automatic local variables). To form a complete 32-bit address from a 16-bit pointer, the compiler uses the data segment register to reference global and static variables and the stack segment register to reference local variables

and parameters. For a normal C program, the stack segment register is set equal to the data segment register so that a 16-bit pointer can point to a value on the stack or a static variable. In a DLL, the data segment register points to the data in the DLL, but the stack segment register points to the stack of the calling function, which is in a different program. Therefore, there is an ambiguity in what a 16-bit address points to in a DLL. The code generated by the Microsoft C compiler assumes that all 16-bit pointers in a DLL point to the data segment. The programmer must be careful to make sure that all 16-bit pointers always point to global or static variables. The alternative is to use only 32-bit pointers, which is not always possible.

Interfacing the SDF library DLL to Microsoft Visual BASIC is straightforward. Structures are supported as Visual BASIC user types and the data types of 16-bit and 32-bit integers and 32-bit and 64-bit floating-point numbers are directly analogous to Visual BASIC data types. The only data type not supported by Visual BASIC is C-style strings (null-terminated character arrays). Strings in Visual BASIC are a dynamically allocated data type, not static in length. Since all the strings in an SDF file are of an even length, they can be defined as 16-bit integer arrays in Visual BASIC. Arrays within structures are supported starting with version 2 of Visual BASIC. With the supplied BASIC function c2string, the integer arrays are converted to Visual BASIC strings.

SDF Revisions

There are currently three revisions of SDF files. The SDF library is written to handle forward and backward compatibility. When each SDF header is read, the library uses the size of the particular SDF header structure that is known at compile time. If the actual size of the header (recordSize) is larger, the library ignores any new fields (forward compatibility). If the actual size of the SDF header read is smaller, it assumes that the fields beyond the read size are uninitialized and then sets these fields to default values.

Conclusion

The main functions of a measurement system are the capture of data and the presentation of the data in a form that meets the user's needs. To help meet these needs, the Hewlett-Packard Lake Stevens Instrument Division stores all measurement data in a consistent format and provides utilities to manipulate the data and to ease the task of importing data into a variety of applications.

Microsoft and MS-DOS are U.S. registered trademarks of Microsoft Corporation.

PC-MATLAB is a trademark of The MathWorks, Inc.

North American Cellular CDMA

Code division multiple access (CDMA) is a class of modulation that uses specialized codes to provide multiple communication channels in a designated segment of the electromagnetic spectrum. This article describes the implementation of CDMA that has been standardized by the Telecommunications Industry Association for the North American cellular telephone system.

by David P. Whipple

The cellular telephone industry is faced with the problem of a customer base that is expanding while the amount of the electromagnetic spectrum allocated to cellular service is fixed. Capacity can be increased by installing additional cells (subdividing), but the degree of subdivision is limited because of the overhead needed to process handoffs between cells. In addition, property for cell sites is difficult to purchase in the areas where traffic is the highest.

The current analog system divides the available spectrum into 30-kHz-wide channels. This method of channelization (division of the spectrum into multiple channels) is commonly called FDMA, for frequency division multiple access (Fig. 1). Alternate means of channelization are being developed to allow more users in the same region of the spectrum. TDMA, or time division multiple access, uses the same 30-kHz channels, but adds a timesharing of three users on each frequency. All other factors being equal, this results in a threefold increase in capacity. CDMA, or code division multiple access, is a class of modulation that uses specialized codes as the basis of channelization. These codes are shared by both the mobile station and the base station.

While CDMA is a class of modulation, this paper focuses on the implementation of CDMA for the North American cellular market, which was initially developed by QUALCOMM, Inc. and has been standardized by the Telecommunications Industry Association (TIA).

Interference Effects

The analog system needs attenuation of about 18 dB for interference on the same channel to provide acceptable call quality. The practical ramification of this is that only a portion of the available spectrum can be used: not all of the channels can be used in every cell. A frequency reuse pattern of seven is commonly used to provide this attenuation (see Fig. 2). In other words, only one seventh of all possible frequencies are available in any one cell. In fact, sectored cells are usually used when capacity is needed to allow the seven-cell repeat pattern to work. Using three sectors per cell, only one out of every 21 available frequencies is used in each sector.

In CDMA, signals are received in the presence of high interference. The practical limit depends on the channel conditions, but reception in the presence of interference that is 18 dB larger than the signal is possible. Typically, the system
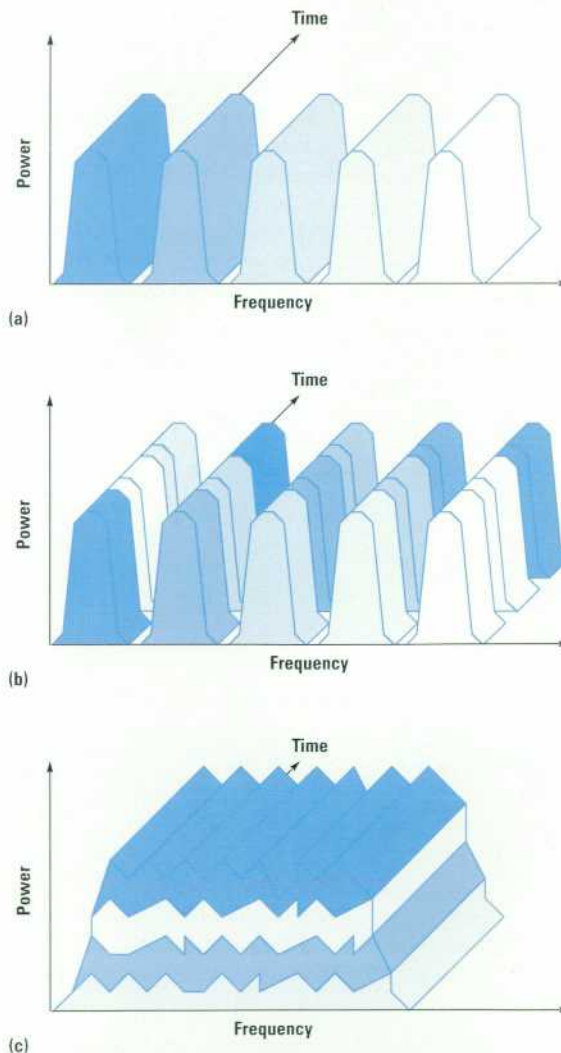
Fig. 1. Cellular channelization methods. (a) Frequency division multiple access (FDMA). (b) Time division multiple access (TDMA). (c) Code division multiple access (CDMA).

operates with better conditions. The frequencies are reused in every sector of every cell, and approximately half the interference on a given frequency is from outside cells. The other half is the user traffic from within the same cell on the same frequency.
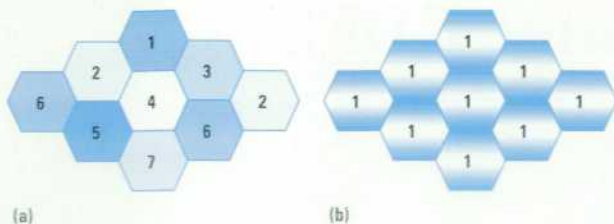
Fig. 2. Cellular frequency reuse patterns. (a) FDMA reuse. (b) CDMA reuse.

Fig. 3 shows a North American cellular CDMA system. CDMA starts with a basic data rate of 9600 bits/s. This is then spread to a transmitted bit rate, or chip rate (the transmitted bits are called chips), of 1.2288 MHz. Spreading consists of applying digital codes to the data bits that increase the data rate while adding redundancy to the system. The chips are transmitted using a form of QPSK (quadrature phase shift keying) modulation that has been filtered to limit the bandwidth of the signal. This is added to the signal of all the other users in that cell. When the signal is received, the coding is removed from the desired signal, returning it to a rate of 9600 bps. When the decoding is applied to the other users' codes, there is no despreading; the signals maintain the 1.2288-MHz bandwidth. The ratio of transmitted bits or chips to data bits is the coding gain. The coding gain for the North American CDMA system is 128, or 21 dB.

An analogy to CDMA is a crowded party. You can maintain a conversation with another person because your brain can track the sound of that person's voice and extract that voice from the interference of all other talkers. If the other talkers were to talk in different languages, discerning the desired speech would be easier because the crosscorrelation between the desired voice and the interference would be lower. The CDMA codes are designed to have very low crosscorrelation.

CDMA Features

The data rate of 9600 bits/s can be thought of as a modem. The signaling and the services must share this fundamental data rate. The system is designed so that multiple service options can use the modem. Currently, service option 1 is speech, service option 2 is a data loopback mode used for test purposes, and service option 3 is being defined as data services, which will support both fax and asynchronous data (terminals).

CDMA communication systems have many differences from analog systems:
- Multiple users share one carrier frequency. In a fully loaded CDMA system, there are about 35 users on each carrier frequency. (There are actually two carrier frequencies per channel, 45 MHz away from each other. One is for the base-to-mobile link, which is called the forward direction, while
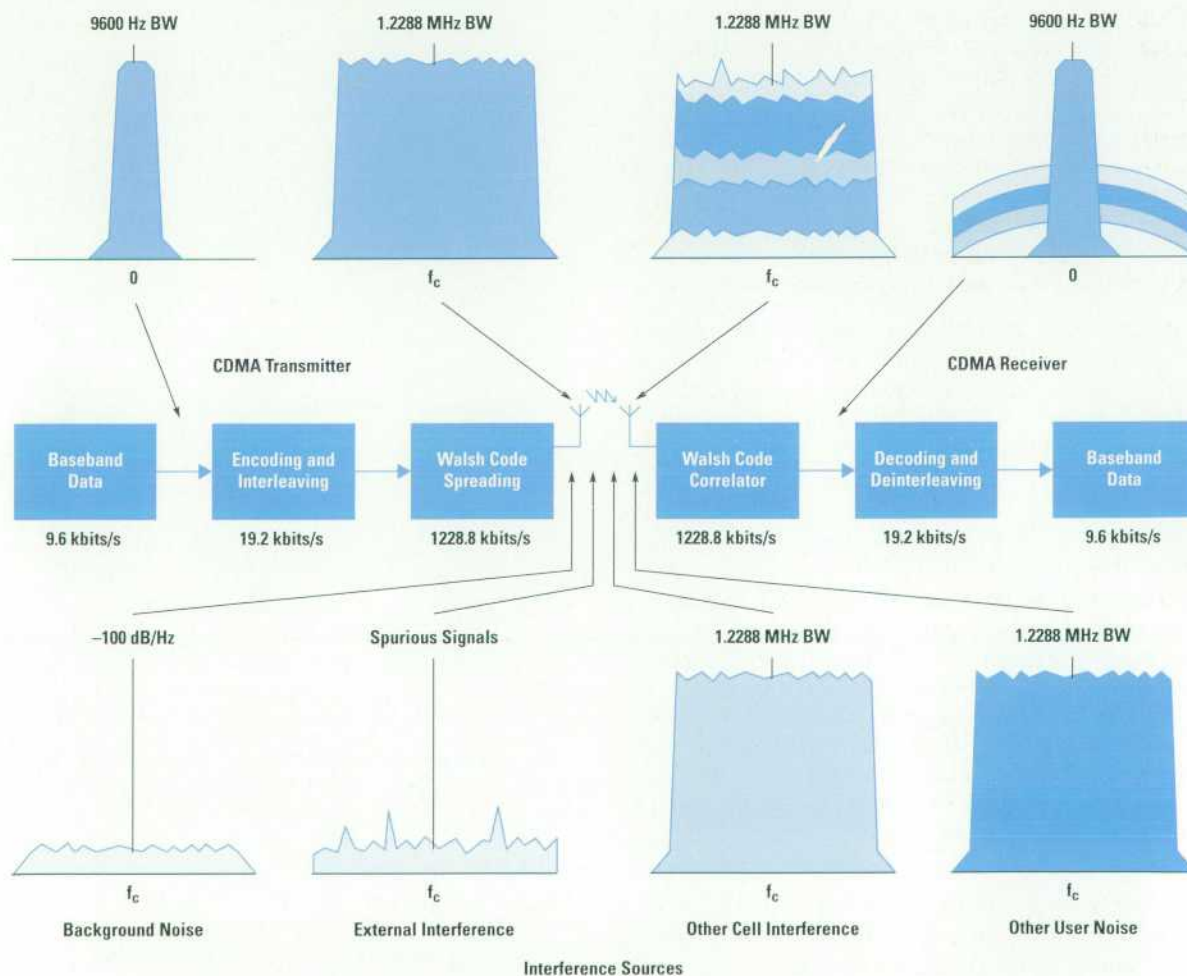
Fig. 3. North American cellular CDMA system.

Cellular Technologies

AMPS. Advanced Mobile Phone System. This is the current analog FM system in North America. It uses 30-kHz channels and signaling is done superaudio, that is, at frequencies above the audio bandwidth for speech, which is 300 to 3000 Hz.

TACS. Total Access Communication System. This is the analog FM system used in the United Kingdom and Japan. It uses 25-kHz channels and signaling is superaudio.

NMT. Nordic Mobile Telephone. Scandinavia led the world in cellular systems. The latest system uses 30-kHz channels, and signaling is done using 1200-Hz and 1800-Hz tones in much the same way as a modem.

J-TACS. This is a narrowband analog FM system in use in Japan. Channels are 12.5-kHz wide and signaling is subaudio, that is, at frequencies below the audio bandwidth for speech, which is 300 to 3000 Hz.

NAMPS. Narrow Analog Mobile Phone System. This is an analog FM system using 10-kHz-wide channels. Signaling is subaudio.

GSM. Global System for Mobile Communications. This is the first digital cellular system to be used commercially. It has been adopted across Europe and in many countries of the Pacific rim. It uses 200-kHz channels with eight users per channel using TDMA, and has a vocoder rate of 13 kbits/s.

TDMA. Time Division Multiple Access. This is the first digital system standardized in North America. It uses 30-kHz channels, three users per channel using TDMA, and has a vocoder rate of 8 kbits/s.

E-TDMA. Extended TDMA. This system uses the same 30-kHz channels as TDMA, but has six users per channel. The vocoder rate is cut to 4 kbits/s, and the channels are dynamically assigned based on voice activity detection. This is being proposed as a follow-on to TDMA.

CDMA. Code Division Multiple Access. This system uses 1.23-MHz-wide channel sets, with a variable number of users on each carrier frequency. The full vocoder rate is 8.55 kbits/s, but voice activity detection and variable-rate coding can cut the data rate to 1200 bits/s. The effective data rate, determined empirically for simulated conversations, is 3700 bits/s. Access is by code.

---

the other is for the mobile-to-base link, which is called the reverse direction).

- The channel is defined by a code. There is a carrier frequency assignment, but the frequency band is 1.23 MHz wide.
- The capacity limit is soft. Additional users add more interference to the system, which can cause a higher data error rate for all users, but this limit is not set by the number of physical channels.

CDMA makes use of multiple forms of diversity: spatial diversity, frequency diversity, and time diversity.

The traditional form of spatial diversity—multiple antennas— is used for the cell site receiver. Another form of spatial diversity is used during the process of handing off a call from one cell to the next. Called *soft handoff*, it is a make-before-break system in which two cell sites maintain a link with one mobile simultaneously (Fig. 4). The mobile station has multiple correlative receiver elements that are assigned to each incoming signal and can add these. There are at least four of these correlators—three that can be assigned to the link and one that searches for alternate paths. The cell sites send the received data, along with a quality index, to the MTSO (mobile telephone switching office) where a choice is made as to the better of the two signals.
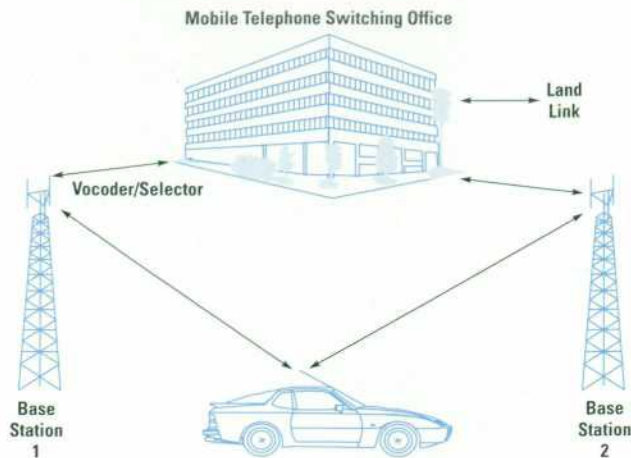
Fig. 4. Spatial diversity during soft handoff.

Frequency diversity is provided in the bandwidth of the transmitted signal. A multipath environment will cause fading, which looks like a notch filter in the frequency domain (Fig. 5). The width of the notch can vary, but typically will be less than 300 kHz. While this notch is sufficient to impair ten analog channels, it only removes about 25% of the CDMA signal.

Multipath signals are used to advantage, providing a form of time diversity. The multiple correlative receiver elements can be assigned to different, time delayed copies of the same signal. These can be combined in what is called a RAKE receiver,[1] which has multiple elements called fingers (Fig. 6). The term RAKE refers to the original block diagram of the receiver (Fig. 6b), which includes a delay line with multiple taps. By weighting the signal at each tap in proportion to its strength, the time-diverse signals are combined in an optimal manner. The picture resembles a garden rake, hence the name.

Another form of time diversity is the use of forward error correcting codes followed by interleaving. Loss of transmitted bits tends to be grouped in time, while most error correction schemes work best when the bit errors are uniformly spread over time. Interleaving helps spread out errors and is common to most digital systems.
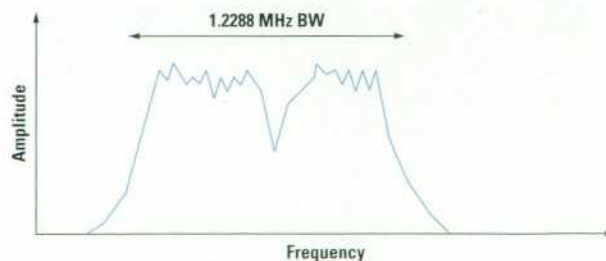
Fig. 5. CDMA frequency diversity. The wide spectrum combats fading caused by multipath transmission. Fading acts like a notch filter to the wide-spectrum signal. Typically only a small part of the signal is lost.
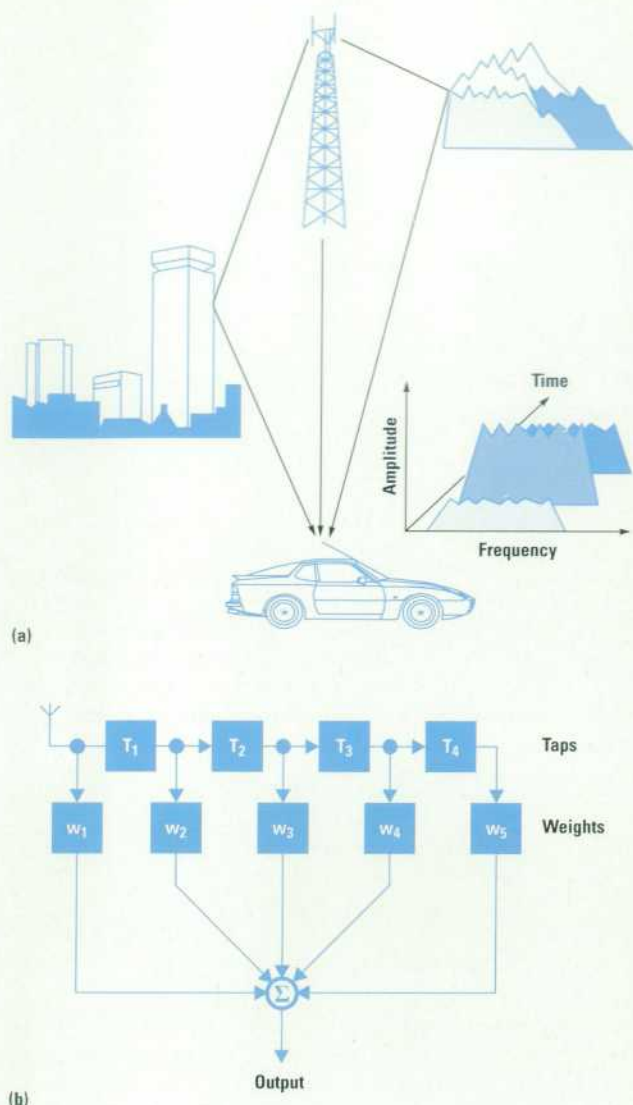
Fig. 6. (a) The RAKE receiver takes advantage of multipath transmission to realize a form of time diversity. (b) RAKE receiver block diagram.

Mobile Station Power Control

Control of the mobile station power is essential for CDMA to work. If one mobile station were to be received at the base station with too much power, it would jam the other users. The goal is to have the signal of all mobile stations arrive at the base station with exactly the same power. Two forms of power control are used: open-loop and closed-loop.

Open-loop power control is based on the similarity of loss in the forward and reverse paths. The received power at the mobile station is used as a reference. If it is low, the mobile station is presumed to be far from the base station and transmits with high power. If it is high, the mobile station is assumed to be close and transmits with low power. The product of the two powers, or the sum of the two powers measured in dB, is a constant. This constant is −73 when the receive and transmit powers are measured in dBm. For example, if the received power is −85 dBm, the transmitted power would be +12 dBm.

Closed-loop power control is used to force the power from the mobile station to deviate from the open-loop setting. This is done by an active feedback system from the base station to the mobile station. Power control bits are sent every 1.25 ms to direct the mobile station to increase or decrease its transmitted power by 1 dB.

Because the CDMA mobile station transmits only enough power to maintain a link, the average transmitted power is much lower than for an analog system. An analog phone needs to transmit enough power to overcome a fade, even though a fade does not exist most of the time. This ability to transmit with lower power has the potential of longer battery life and smaller, lower-cost output amplifier design.

Speech Encoding

The speech is encoded before transmission. The purpose of encoding is to reduce the number of bits required to represent the speech. The CDMA voice coder (*vocoder*, as it is called) has a data rate of 8550 bits per second. After additional bits are added for error detection, the channel data rate is 9600 bits/s. This is lowered, however, when the user is not speaking. The vocoder detects voice activity, and will lower the data rate during quiet periods. The lowest data rate is 1200 bits/s. Two intermediate rates of 2400 and 4800 bits/s are also used for special purposes. The 2400 bits/s rate is used to transmit transients in the background noise, and the 4800 bits/s rate is used to mix vocoded speech and signaling data (signaling consists of link-management messages between the base station and the mobile station). In this last case, the channel data rate is 9600 bits/s, but half of the bits are assigned to voice and the other half to the message. This is called *dim and burst signaling*.

The mobile station pulses its output power during periods of lower-rate data. The power is turned on for 1/2, 1/4, or 1/8 of the time. The data rate is 9600 bits/s when the power is on, so the average data rate is 4800, 2400, or 1200 bits/s. This lowers the average power and the interference seen by other users.

The base station uses a different method to reduce power during quiet periods. It transmits with 100% duty cycle at 9600 bits/s, but uses only 1/2, 1/4, or 1/8 of full power and repeats the transmitted data 2, 4, or 8 times. The mobile station achieves the required signal-to-noise ratio by combining the multiple transmissions.

One important aspect of the coding used in CDMA is Walsh codes, or Hadamard codes.[2] These are based on the Walsh matrix, a square matrix with binary elements that always has a dimension that is a power of two. It is generated by seeding Walsh (1) = W_1 = 0 and expanding as shown below and in Fig. 7:

$$W_{2n} = \begin{bmatrix} W_n & W_n \\ W_n & \overline{W_n} \end{bmatrix}$$

where n is the dimension of the matrix and the overscore denotes the logical NOT of the bits in the matrix.

The Walsh matrix has the property that every row is orthogonal to every other row and the logical NOT of every other

$$W_{2n} = \begin{bmatrix} W_n & W_n \\ W_n & \overline{W}_n \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$W_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Fig. 7. Walsh matrices.

row. Orthogonal means that the dot product of any two rows is zero. In simpler terms, it means that between any two rows exactly half the bits match and half the bits do not match. The CDMA system uses a 64-by-64-bit Walsh matrix.

Forward Link Encoding

Walsh encoding is used in the forward link (base to mobile) as shown in Fig. 8. The fundamental data rate of the channel is 9600 bits/s. The data is packetized into 20-ms blocks and has forward error correction applied by use of a convolutional encoder. This is done at half rate, which yields two bits out for every bit in. The data is then interleaved—a shuffling of the bits during the 20-ms period. This is done to better distribute bits lost during transmission. It has been shown that bit errors tend to come in groups rather than being spread out in time, while forward error correction works best when the errors are distributed uniformly over time. When the data is deinterleaved, the time-linked errors get spread over time.

Following the interleaver, the data is modified by the use of a *long code*, which serves only as a privacy mask. The long code is generated by a pseudorandom binary sequence (PRBS) that is generated by a 42-bit-long shift register (Fig. 9). This register is also used as the master clock of the system, and is synchronized to the limit of propagation delays among all base stations and mobile stations. A mask is applied to the PRBS generator that selects a combination of the available bits. These are added modulo two by way of exclusive-OR gates to generate a single bit stream at 1.2288 MHz. For the forward link, a data rate of only 19.2 kbits/s is

needed, so only 1 of 64 bits gets used. The long code generated in this way is XORed with the data from the interleaver.

The resulting data is then encoded using the Walsh matrix. One row of the Walsh matrix is assigned to a mobile station during call setup. If a 0 is presented to the Walsh cover, then the 64 bits of the assigned row of the Walsh matrix are sent. If a 1 is presented, then the NOT of the Walsh matrix row is sent. This has the effect of raising the data rate by a factor of 64, from 19.2 kbits/s to 1.2288 Mbits/s.

The last stage in coding is to convert from a binary signal to two binary channels in preparation for transmission using QPSK (quadrature phase shift keying) modulation. The data is split into I and Q (in-phase and quadrature) channels and the data in each channel is XORed with a unique PRBS *short code*. The short codes are spreading sequences that are generated much like the long code, with linear feedback shift registers. In the case of the short codes, there are two shift registers, each 15 bits long, with feedback taps that define specific sequences. These run at 1.2288 MHz. The short code sequences, each 2^{15} bits long, are common to all CDMA radios, both mobile and base. They are used as a final level of spreading.

After the data is XORed with the two short code sequences, the result is two channels of data at 1.2288 Mbits/s. Each channel is low-pass filtered digitally using an FIR (finite impulse response) filter. The filter cutoff frequency is approximately 615 kHz. A typical FIR filter implementation might output 9-bit-wide words at 4.9152 MHz. The resultant I and Q signals are converted to analog signals and are sent to a linear I/Q modulator. The final modulation is filtered QPSK.

Multiple channels in the base station are combined in the I and Q signals to supply the multiple channels transmitted by the cell (Fig. 10). Because all users share the composite signal from the cell, a reference signal called the *pilot* is transmitted. The pilot has all zero data and is assigned Walsh row number 0, which consists of all 0s. In other words, the pilot is made up of only the short spreading sequences. Typically
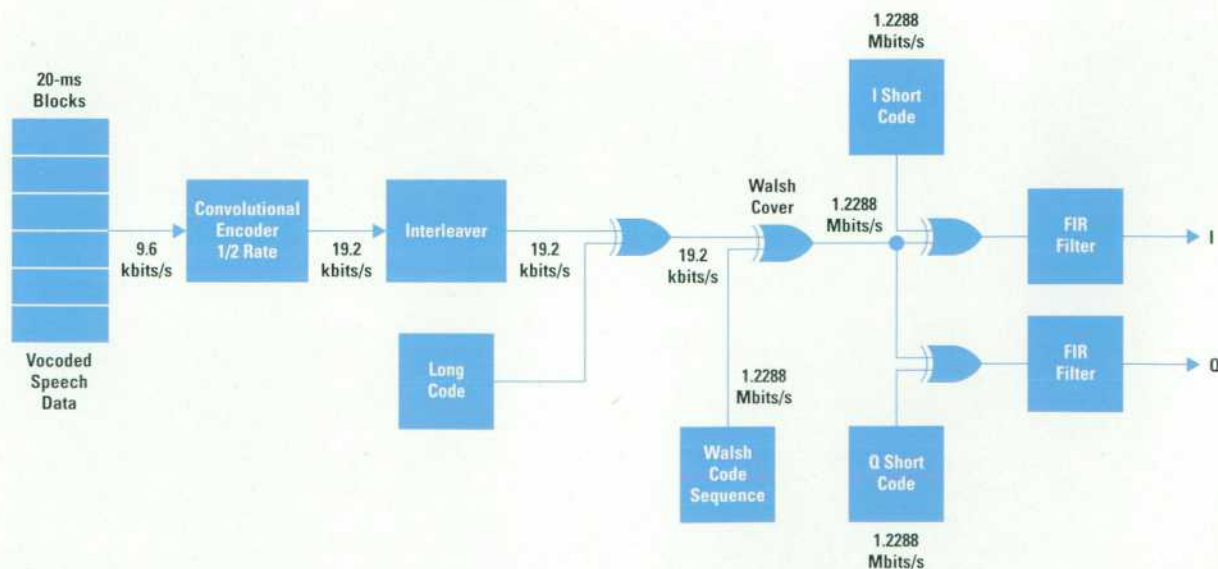
Fig. 8. CDMA forward link physical layer.
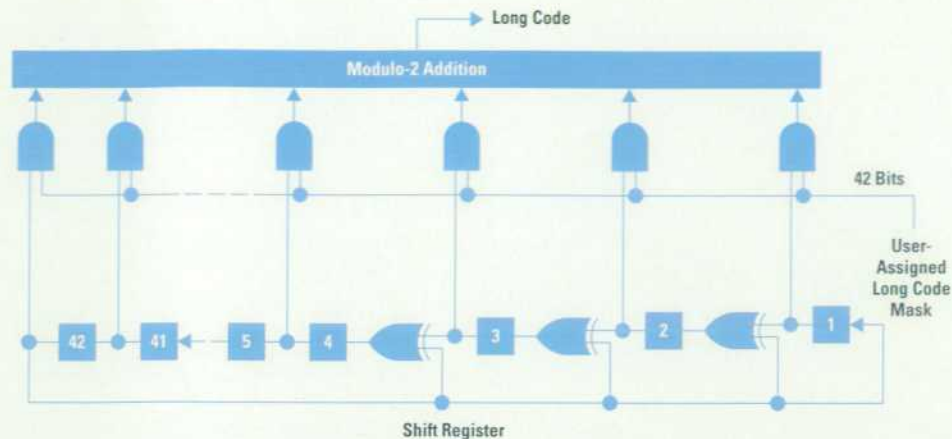
Fig. 9. Long code privacy mask generation.

20% of the total energy of a cell is transmitted in the pilot signal. The pilot signal forms a coherent phase reference for the mobile stations to use in demodulating the traffic data. It is also the timing reference for the code correlation. The short sequences allow the CDMA system to reuse all 64 Walsh codes in each adjacent cell. Each cell uses a different time offset on the short codes and is thereby uniquely identified while being able to reuse the 64 Wash codes.

Reverse Link Encoding

The mobile station cannot afford the power of a pilot because it would then need to transmit two signals. This makes the demodulation job more difficult in the base stations. A different coding scheme is also used, as shown in Fig. 11.

For speech, the same vocoder is used in both directions. The data rate is again 9600 bps. A 1/3-rate convolutional encoder is used, yielding an output rate of 28.8 kbits/s. The output of this is interleaved and then taken six bits at a time. A six-bit number can range from 0 to 63, and each group of six bits is used as a pointer to one row of the Walsh matrix. Every mobile station can transmit any row of the Walsh

matrix as needed. At this point, the data rate is 307.2 kbits/s, but there is no unique coding for channelization. The full-rate long code is then applied, raising the rate to 1.2288 Mbits/s. This final data stream is split into I and Q channels and spread with the same short sequences as in the base station. There is one more difference: a time delay of 1/2 chip is applied to the Q channel before the FIR filter. This results in offset-QPSK modulation (Fig. 12), and is used to avoid the amplitude transients inherent in QPSK. This makes the design of the output amplifier easier in the mobile station.

The capacity is different in the forward and reverse links because of the differences in modulation. The forward link has the phase reference—the pilot signal—as well as orthogonal codes. The reverse link signal is not orthogonal because the long codes are applied after the use of the Walsh matrix. In this case the signals are uncorrelated but not orthogonal. The base station has the advantage of multiple receive antennas (diversity). All factors taken together, the reverse link sets system capacity.

Table I summarizes the CDMA channelization functions.
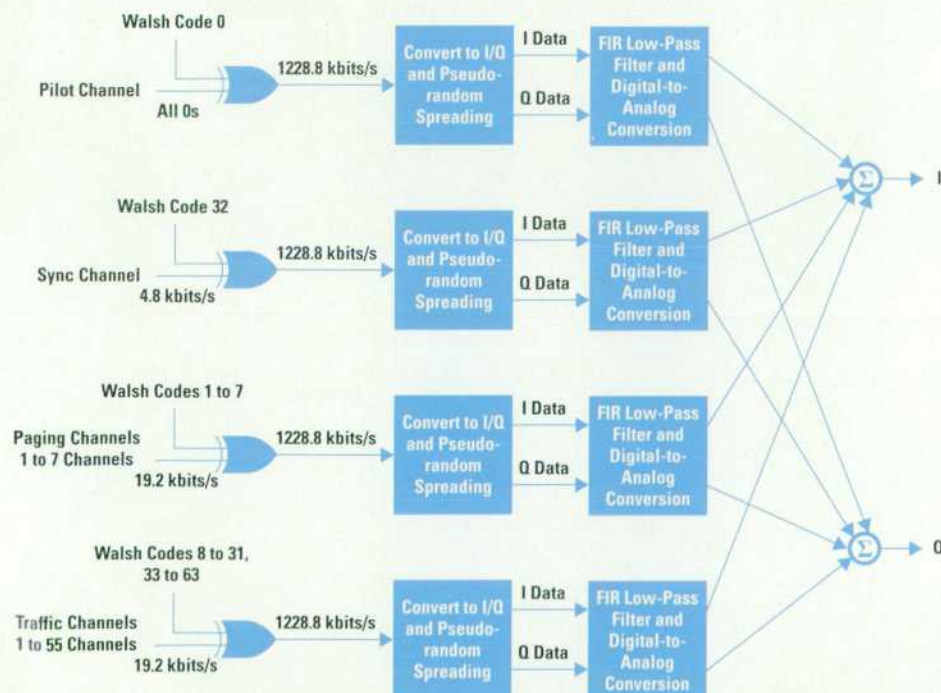
Fig. 10. CDMA forward link channel format.
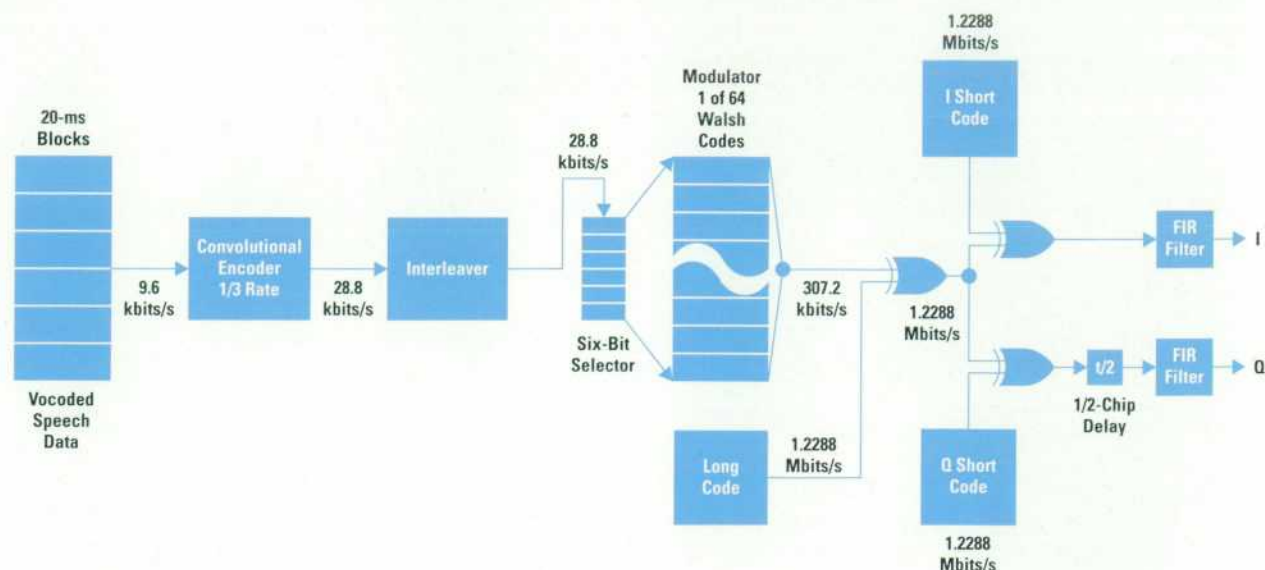
Fig. 11. CDMA reverse link physical layer.

Call Scenario

To better illustrate how the CDMA system operates, the system function will be described in terms of mobile station operation.

When the mobile station first turns on, it knows the assigned frequency for CDMA service in the local area. It will tune to that frequency and search for pilot signals. It is likely that multiple pilot signals will be found, each with a different

Table I
CDMA Channelization Functions

Parameter	Function	Notes
Frequency	Divides the spectrum into several 1.23-MHz frequency allocations.	Forward and reverse links are separated by 45 MHz.
Walsh Codes	Separates forward link users of the same cell.	Assigned by cell site. Walsh code 0 is always the pilot channel. Walsh code 32 is always the sync channel.
Long Code	Separates reverse link users of the same cell.	Depends on time and user ID. The long code is composed of a 42-bit-long PRBS generator and and a user-specific mask.
Short Codes, also called the I and Q spreading sequences	Separates cell sites or sectors of cells.	The I and Q codes are different but are based on 15-bit-long PRBS generators. Both codes repeat at 26.667-ms intervals. Base stations are differentiated by time offsets of the short sequences.

time offset. This time offset is the means of distinguishing one base station from another. The mobile station will pick the strongest pilot, and establish a frequency reference and a time reference from that signal. It will then start demodulation of Walsh number 32, which is always assigned to the sync channel. The sync channel message contains the future contents of the 42-bit long code shift register. These are 320 ms early, so the mobile station has time to decode the message, load its register, and become synchronized with the base station's system time. The mobile station may be required to register. This would be a power-on registration in which the mobile station tells the system that it is available for calls and also tells the system where it is. It is anticipated that a service area will be divided into zones, and if the mobile station crosses from one zone to another while no call is in progress, it will move its registration location by use of an idle state handoff. The design of the zones is left to the service provider and is chosen to minimize the support messages. Small zones result in efficient paging but a large number of idle state handoffs. Large zones minimize idle state handoffs, but require paging messages to be sent from a large number of cells in the zone.

At this point the user makes a call by entering the digits on the mobile station keypad and hitting the send button. The mobile station will attempt to contact the base station with

Fig. 12. Constellation diagrams for CDMA modulation formats. (a) The base station transmitter uses filtered QPSK. (b) The mobile station transmitter uses filtered offset-QPSK.

an access probe. A long code mask is used that is based on cell site parameters. It is possible that multiple mobile stations may attempt a link on the access channel simultaneously, so collisions can occur. If the base station does not acknowledge (on the paging channel) the access attempt, the mobile station will wait a random time and try again. After making contact, the base station will assign a traffic channel with its Walsh number. At this point, the mobile station changes its long code mask to one based on its serial number, receives on the assigned Walsh number, and starts the conversation mode.

It is common for a mobile station communicating with one cell to detect another cell's pilot that is strong enough to be used. The mobile station will then request soft handoff. When this is set up, the mobile station will be assigned different Walsh numbers and pilot timing and use these in different correlative receiving elements. It is capable of combining the signals from both cells.

Eventually, the signal from the first cell will diminish and the mobile station will request from the second cell that soft handoff be terminated.

At the end of the call, the channels will be freed. When the mobile station is turned off, it will generate a power-down registration signal that tells the system that it is no longer available for incoming calls.

Testing

The complexity of the CDMA system raises substantial test issues. What needs to be tested, and what environment is needed for testing? To test the mobile station, the test equipment must emulate a base station. The tester needs to provide the pilot, sync, paging, and traffic channels. It must provide another signal that uses orthogonal Walsh symbols that represent the interference generated by other users of the same cell, and it must provide additive noise that simulates the combination of CDMA signals from other cells and background noise. Bit error rate is not a meaningful measure, since substantial errors are expected at the chip rate and these are not available for test. The bits at the 9600-bits/s rate are the only bits available for test, and these will either be all correct as a result of error correction or will have substantial errors. What is used instead is the *frame error rate*, a check of the received bits and the associated CRC (cyclic redundancy code) in each 20-ms block.

To test the transmitter, a new test has been defined: *waveform quality*. This is based on the crosscorrelation of the actual transmitted signal to the ideal signal transmitting the same data. This is important to the system because the CDMA receivers are correlators. In fact, they correlate the received signal with the ideal signal. If a signal deviates substantially from the ideal, the correlated portion of that signal will be used to make the link and the uncorrelated portion will act as additive interference. Closed-loop power control will maintain the correlated power at the needed level, and excess power will be transmitted. The specification is that the radios shall transmit with a waveform quality that limits the excess power to less than 0.25 dB. Other transmitter measurements include frequency and power control operation.

Conclusion

CDMA provides an advanced technology for cellular applications. It provides high-quality service to a large number of users. It is a system that has been extensively tested and it will be deployed later this year in precommercial applications. Commercial service is scheduled to begin in 1994.

References

1. R. Price and P.E. Green, Jr., "A Communication Technique for Multipath Channels," *Proceedings of the IRE*, Vol. 46, March 1958, pp. 555-570.
2. J.G. Proakis, *Digital Communications*, *Second Edition*, McGraw-Hill Book Co., 1989.

DECT Measurements with a Microwave Spectrum Analyzer

An HP 8590 E-Series spectrum analyzer with DECT source, demodulator, and measurement personality can be used to provide a cost-effective solution to development, manufacturing, and pre-type-approval testing for compliance with the Digital European Cordless Telecommunications standard.

by Mark A. Elo

The HP 8590 E-Series microwave spectrum analyzers provide a portable, rugged, and versatile general test solution. An added advantage of these analyzers is their designed-in flexibility, which allows them to be configured for specific measurement needs. The analyzers can accept a range of options in their built-in cardcage and their ROM-card reader allows the use of customized downloadable programs. The downloadable programs available for the HP 8590 E-Series analyzers include CT2, GSM, NADC, JDC, and now DECT.[1,2]

The HP 85723A DECT measurement personality is an HP application-specific downloadable program that gives the analyzer measurement capabilities required for testing to the Digital European Cordless Telecommunications physical layer standard. The measurement of DECT RF characteristics can present the engineer with many complex measurement requirements. The ability to reconfigure a standard spectrum analyzer with a downloadable program and some extra hardware allows these difficult DECT tests to be made at the press of a button. Extending and enhancing the functionality of readily available test equipment in this way provides a highly cost-effective solution.

This article aims to explain the DECT physical layer definition and explore some of the extensions and enhancements applied to the HP 8590 E-Series spectrum analyzer that transforms it into a DECT test tool. After an introduction to the physical layer standard and some basic definitions, the software techniques and hardware requirements of the spectrum analyzer configured for DECT physical layer measurements are discussed.

What Is DECT?

DECT, which stands for Digital European Cordless Telecommunications, is a new standard that presents many new market opportunities in electronic communications equipment. It has been a common misconception that DECT is just a cordless telephone system. However, the DECT standard has provisions for more services than just telephony, with possible applications ranging from paging to cordless LAN. DECT's versatility can be attributed to its protocol structure, which is derived from the OSI (Open Systems Interconnection) seven-layer model of the International Standards Organization (ISO). The OSI model structures a piece of equipment into specific parts allowing for modular compatibility

between different pieces of communication equipment. This allows DECT to provide for not just European user compatibility but also worldwide electronic compatibility, offering a cordless link between most pieces of electronic communications equipment. The top layer of the model, layer 7, corresponds to the user interface, for example, the microphone, speaker, or keypad of a telephone. The bottom layer, the physical layer, corresponds to the transmission medium, which for DECT is a radio link. The OSI model as originally conceived had no provision for a radio physical link. This was solved by redefining the lower layers of the DECT protocol model. The bottom four layers of the DECT model correspond to the lower three layers of the OSI model. Both of the network layers correspond exactly, while the data link control, medium access control, and physical layers of DECT have no OSI equivalent (see Fig. 1).

ETSI

Who defined the requirements for DECT? ETSI, the European Telecommunications Standards Institute, formed a number of
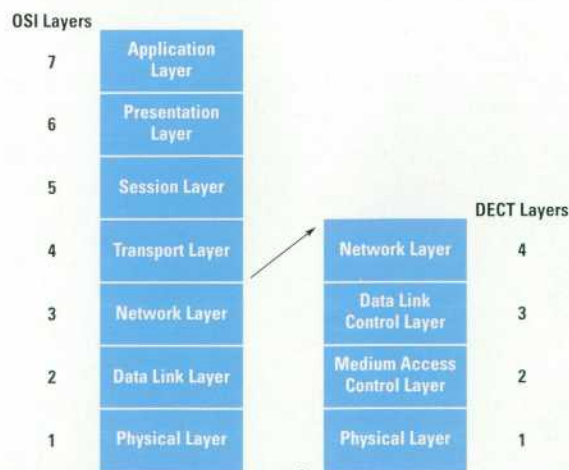
Fig. 1. The DECT protocol structure is based on the International Standards Organization's seven layer Open Systems Interconnection model. Because the DECT physical layer is a radio link, the model had to be modified. The lower four layers of the DECT model correspond to the lower three layers of the ISO OSI model.

working groups to help define European compatible communications systems. The Radio Equipment and Specifications Group (RES) has the task of creating working subgroups consisting of delegates from relevant equipment manufacturers to help define standards such as maritime mobile, electromagnetic compatibility (EMC), and now DECT. RES3 is responsible for DECT and is further split into five groups, each with specific expertise. These groups cover test methods for type approval, the system physical layer characteristics, the protocol, system operation, and audio parameters.

Physical Layer Specification

CEPT, Conférence Européen des Administrations des Postes et des Télécommunications, has allocated a frequency band of 1880 MHz to 1900 MHz for DECT, thus making it mandatory to have this frequency band available throughout Europe. The DECT specification[3] defines ten carriers in this band with a carrier spacing of 1.728 MHz. Each carrier band has a channel number from 0 to 9. The channel 9 carrier frequency is 1881.792 MHz and the channel 0 carrier frequency is 1897.344 MHz. Each of the ten carriers spaced across this frequency band is used in a time division multiple access, time division duplex (TDMA/TDD) scheme. Each carrier can be turned on up to 24 times in a period of 10 ms. Two-way communication is achieved by using the first 12 instances as the transmit time and the second 12 as the receive time. The data is modulated onto the pulsed RF carrier using Gaussian minimum shift keying (GMSK). This method deviates the carrier frequency by ± 288 kHz, with each deviation representing a one or zero, respectively.

The GMSK modulated data is organized in the form of packets, the length of each packet corresponding to the on time of the RF burst. As shown in Fig. 2a, a DECT packet can be split into three fields:[4] the synchronization field, the data field, and an optional error correction field. These are abbreviated as the S, D, and Z fields, respectively. The first 16 bits of the S field contain a preamble of alternating 1s and 0s, either 1010… or the inverse depending on whether it is a fixed part transmit packet or a portable part transmit packet. This is used for clock recovery. The second 16 bits of the S field is the synchronization word, which again has one form if it is a fixed part transmit packet and the inverse if it is a portable part transmit packet.

The D field can split into a further two fields: the A and the B field. The A field is 64 bits long and contains DECT signaling data. The B field is the part of the packet that contains the information that needs to be transmitted, although the B field can also be used for signaling data in B-field-only systems. The Z field is optional and is four bits in length. It is a copy of the last four bits of the B field and is used for detecting time collisions from other nonsynchronized DECT systems.

Four packet sizes with varying B field lengths are defined for DECT: short, basic, low-capacity, and high-capacity (see Fig. 2b). The short physical packet has no B field; it transfers signaling data only and can be used for such applications as paging. The basic DECT physical packet for speech has 324 bits in the B field and is 420 bits long altogether. With a bit rate of 1152 kbits/s this equates to a packet length of 364.6 µs. The low-capacity packet has a data field smaller than the basic packet and the high-capacity packet's data field is larger. There is no set application for the low-capacity
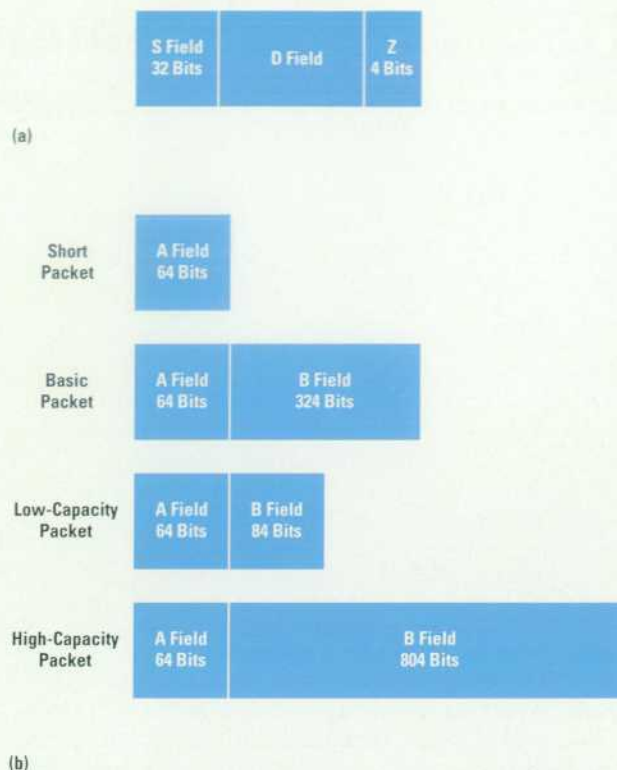
Fig. 2. (a) DECT data packet. (b) The D field is split into two fields: A and B. The lengths of the A and B fields depend on the DECT application.

packet, but the high-capacity packet can be used for applications that require low overhead, such as cordless LAN. For the purposes of this article only the definition of the basic physical packet will be considered.

It has already been mentioned that each packet of data is modulated onto an active carrier within a 10-ms frame. The frame contains 11,520 bits, which are equally divided into 24 full slots of 480 bits. A full slot is 416.7 µs long. The short physical packet and basic physical packet both occupy a full slot, while the low-capacity packet uses only half a slot and the high-capacity packet occupies two full slots.

Two other definitions important to the DECT physical layer characteristics are the loopback field and the position of bit p_0. When the equipment under test is placed in loopback mode it must retransmit the relevant D field data received from the tester. The loopback part of the D field is generally the data contained in the B field except for the A-field-only short packet, for which the A field is looped back. DECT bit p_0 is the timing reference point that defines the beginning of the packet. Its position is 16 bit periods before the bit transition that occurs between the preamble and the synchronization word. To meet the type-approval test specification, it is important that a relationship between the position of p_0 and the triggering time of the measuring instrument be established.

Testing DECT RF Characteristics

One-button measurements available with the HP 85723A DECT personality are: carrier power, power-versus-time template measurements, spurious emissions, intermodulation

attenuation, adjacent channel power due to modulation and transients, and frequency modulation tests such as deviation and carrier accuracy. These tests require more than just a standard HP 8590 E-Series microwave spectrum analyzer. The optional fast sweep and time gate are required, plus a DECT-specific demodulator option. A built-in DECT source can be added for RF component testing and receiver sensitivity applications. The analyzer's ability to receive and optionally transmit gives the user a single-box DECT test tool that can effectively double as a DECT transceiver.

The DECT Transceiver

An understanding of how the HP 8590 E-Series analyzer is configured as a DECT measurement transceiver can be gained by considering how a packet of DECT data flows through the analyzer, referring to Fig. 3. The analyzer can be split into two parts: the receiver and the transmitter. Initially the data must be transmitted to the unit under test from the DECT source or transmitter. Two considerations must be taken into account when transmitting a digital signal in a TDMA system. The first is that the data must be modulated onto the RF carrier in such a way that it will cause minimal spectral occupancy. Secondly, the carrier must be switched

on in a correct DECT time slot. The back panel of the analyzer has two BNC TTL inputs called **TTL Data In** and **Time Slot In**. The TTL data input signal is initially passed through a 0.5 GMSK filter (as defined in the DECT standard) to smooth the edges of the digital signal. At the same time the carrier is switched on for the data packet's duration via the RF switch which is activated by a TTL signal input corresponding to the length of the packet at the time slot input connector. Data and signaling from an external source are modulated onto the tracking generator signal and transmitted from the DECT source output on the front panel.

The equipment under test receives the RF signal, demodulates and decodes it, and assuming that the transmitted signal contains the required protocol, retransmits a response. The HP 8590 E-Series microwave spectrum analyzer receiver converts the high-frequency RF signal response from the equipment under test into a low-frequency signal that can be easily processed. This down-converted or IF signal is amplified and fed into one of the analyzer's analog-to-digital converters for processing and display on the CRT. A parallel path to the analog-to-digital converter is through an FM discriminator, which demodulates the signal for FM measurements.
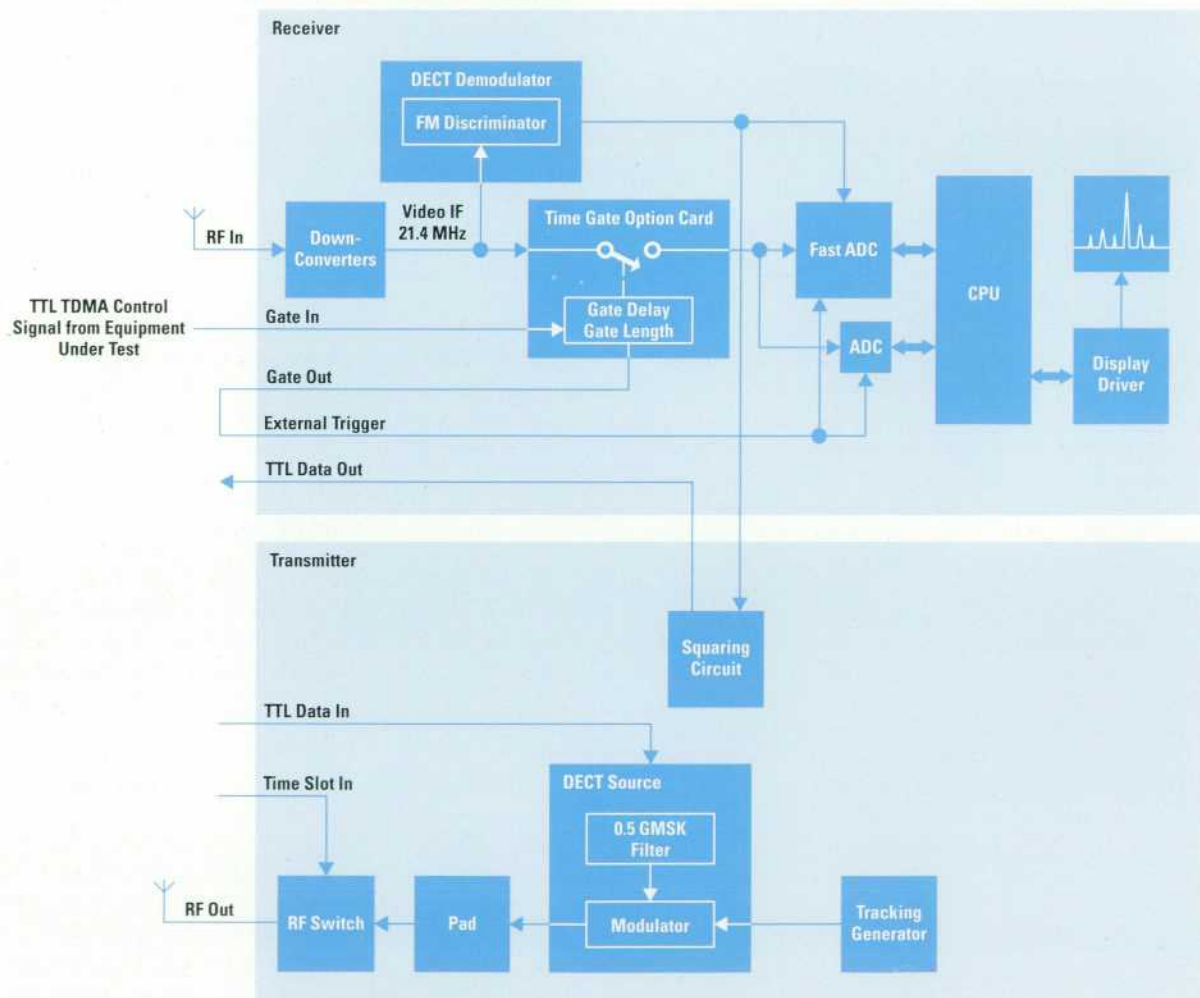
Fig. 3. Block diagram of an HP 8590 E-Series spectrum analyzer with Option 012 DECT source and Option 112 DECT demodulator. Other options needed for DECT measurements are Option 004 precision frequency reference, Option 101 fast time-domain sweep, and Option 105 time gated spectrum analysis.

The demodulated data also passes through a squaring circuit and is converted to a TTL level, which is output via the **Data Out** connector on the rear panel.

To enable the TDMA signal to be displayed correctly the display driver circuitry must be supplied with a suitable trigger. The trigger signal path has an important role in this application since the analyzer must know the position in time when the modulated carrier is received. A trigger can be provided in two ways: from a detection device that outputs a positive-going edge when the RF switches on, or from a logic signal indicating the slot timing, derived from the control circuitry of the equipment under test. The trigger signal is initially fed into the input of the gate card. The gate card has two functions. The first is to provide a programmable delayed output which is then connected to the external trigger input, providing a post-trigger function. The second is the ability to switch the video path of the analyzer, allowing only selected time intervals of the signal to be measured.

Emissions Due to Modulation

The first test to examine is the measurement of emissions due to modulation. This is an important measurement in any communications system and indicates to what degree the transmitting channel of interest interferes with its adjacent channels. Adjacent channel power can be defined as a leakage ratio, that is, the ratio of the power transmitted (by leakage) into an adjacent channel to the total power transmitted by the transmitter. The primary requirement for the measurement is to measure the total power in a defined passband. The receiver bandwidth for DECT is 1 MHz. Therefore, the ratio of the power in the transmit band to the power measured in a 1-MHz passband centered on any other DECT channel will give the adjacent channel power ratio. The emissions due to modulation measurement on the DECT system is a measure of how much power leakage into adjacent channels is caused specifically by GMSK modulation products. Two factors must be considered in this measurement technique:

- The transient products caused by the RF switching of a TDMA/TDD system must be eliminated from the measurement and characterized separately. Otherwise, they will dominate the result.
- The measuring device, in this case a swept spectrum analyzer, must be able to measure the sum of a number of powers, which will include noise as well as sine waves, in a defined band.

To meet the first requirement the time-selective spectrum analyzer option card must be used. To omit all spectral components resulting from time division switching of the RF signal the analyzer must only make a measurement when the carrier is switched on during the on portion of the burst, thus making the analyzer's sampling hardware think the signal is present all of the time. There are many places in the IF path a time-selective switch can be placed. For optimum results the HP 8590 E-Series microwave spectrum analyzer uses the time gate option card to perform a video time gate. This permits the video signal to reach the sampling hardware only during the selected time interval, giving the appearance of a CW signal that is purely digitally modulated.

The analyzer must be provided with a signal to inform the video gate when to open and close. The required signal is a
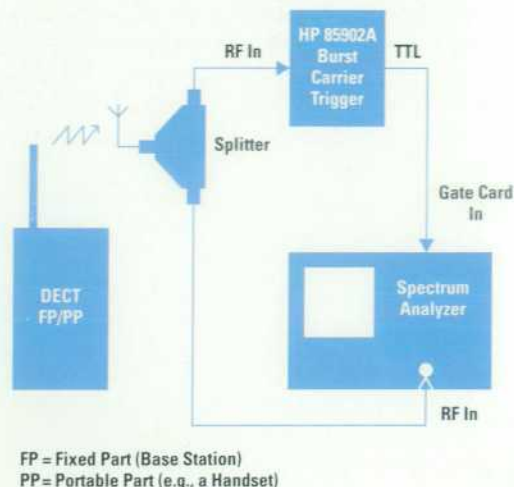
Fig. 4. For some measurements the analyzer must be provided with a signal to tell the video gate when to open and close. The HP 85902A burst carrier trigger detector circuitry produces a positive-going TTL edge when the RF burst switches on.

TTL signal from the equipment under test to indicate the switching on of the RF burst (TDMA signal in Fig. 3). This can be achieved directly by taking a TTL signal from the control circuitry of the equipment under test or indirectly by using the HP 85902A burst carrier trigger as shown in Fig. 4. This signal is connected to the time gate card input. The card also needs two programmable parameters: gate delay and gate length. The gate delay is a time delay between receiving the gate card input signal and switching the video gate on, and the gate length tells how long the video time gate switch remains closed.

Time-selective spectrum analysis can be a useful tool for measuring TDMA systems. However, the video time gate complicates the RF path in the spectrum analyzer. Spectrum analyzer settings such as resolution bandwidth, video bandwidth, and sweep time can be difficult to optimize for a gated measurement.

A number of rules were followed in implementing a time-selective measurement in the HP 85723A personality to ensure that the frequency and amplitude results are correct. The screen displays 401 data points from the ADC. So that the spectrum analyzer can correctly sample pulsed modulated signals, the sweep time must be long enough to ensure that each data point contains some burst information. This is guaranteed if the sweep time is set to 401 times the pulse repetition interval, which for DECT is 10 ms. This equates to a 4-to-5-second sweep time. For each sample, the time period between the RF's switching on and the video gate's switching allowing the signal to be analyzed must be sufficiently long to let the resolution bandwidth filter charge or set up. This is referred to as the setup time. The video bandwidth filter must also have sufficient time to charge, and because the video bandwidth filter is postdetection, its charge time is dependent on the gate length. The ETSI specification dictates that a 60% portion of the burst carrier must be analyzed, which should begin at least 25% of the total length of the burst away from the start. Comparing this to the timing characteristics of a basic physical packet shows that the observation window or gate length should be 218 μs

long starting 91 μs away from the start of the burst or the position of bit p_0. In this case then, the setup time is approximately 90 μs. The resolution bandwidth must be set to a value greater than 2 over the setup time, or 22 kHz, and the video bandwidth must be greater than 1 over the gate length or 1/218 μs = 4.5 kHz.

To meet the second requirement above, a noise-type measurement must be made, but it is not simply a case of placing the marker on the noise floor; more complex measurement considerations must be taken into account. Adjacent channel power measurements have traditionally been carried out with a measuring receiver, which measures the sum of powers through a specified filter. However, the ETSI type-approval specification is more suited to a spectrum analyzer. Two key points can be taken from the specification. The first is that a 100-kHz resolution bandwidth filter should be swept over a 1-MHz range, and the second is that actual power is the average of the powers measured in all of the 100-kHz trace data points. Why should it be done this way? What is required is the noise power in a specified band, with the noise leakage from other channels kept at an absolute minimum. The roll-off characteristic of a 100-kHz filter is substantially better than that of a 1-MHz filter. Therefore, if the power in each 100-kHz trace data point is measured, and the sum of the measured powers is divided by the number of measurements taken, the total noise power within a 1-MHz bandpass filter is resolved with the noise response of a 100-kHz filter. The result is then normalized with a scaling factor to correct for using a 100-kHz filter instead of a 1-MHz resolution bandwidth filter. This measurement is repeated on all DECT channels relative to the transmit channel. HP 8590 E-Series analyzers' ability to perform math on displayed traces is ideal for this type of noise measurement application.

Emissions Due to Intermodulation

The emissions due to intermodulation test follows all the same rules as the emissions due to modulation test. The primary test requirement is that two carriers are set up on different channels but using the same time slot. This enables the time gating function to capture only the on period of the burst carriers and integrate power in a 1-MHz band. Fig. 5 shows the effect of removing the TDMA using time gated spectrum analysis.

Transmission Burst Timing Parameters

The ETSI type-approval specification requires that the switching characteristics of the RF burst fall within the limits of a specified template. The NTP, or normally transmitted power, is the average power over the burst duration. The rise and fall slope times must be less than 10 μs and the power of the burst must be less than –47 dBm after 27 μs on each side of the burst. A spectrum analyzer is ideally suited to performing this measurement, having the ability to generate limit lines representing the specified template that automatically detect an incorrect signal. However, some complex measurement rules[5] must be observed and accurate limit lines must be created for every size of DECT packet. This again is a set of tests that can be performed easily with the aid of a measurement personality. The measurement personality chooses the optimum resolution and video bandwidths, selects the relevant sweep time, and calculates the correct timing parameters, providing a suitable trigger.

(a)

(b)
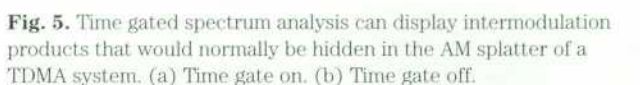
Fig. 5. Time gated spectrum analysis can display intermodulation products that would normally be hidden in the AM splatter of a TDMA system. (a) Time gate on. (b) Time gate off.

To view the rising or falling edges of a pulsed RF signal with good resolution, it is vital that the trigger timing be correct. For the rise time measurement a pretrigger function must also be included. A pretrigger function in zero span is beyond the functionality of most spectrum analyzers, but a post-trigger signal can be created by using the functionality of the time gate card. As in the previous measurement, the equipment must provide the gate card with a TTL signal indicating that the RF burst is switching on. The gate delay and length can then be set using the appropriate function. In this case, the signal that would normally switch the video signal on or off is diverted to the gate output BNC on the back panel of the analyzer (see Fig. 3). What this now provides is a signal that can be delayed in time an optional amount. If this signal is then connected to the external trigger input of the analyzer, an effective trigger delay can be created, thus offsetting the analyzer display in the time domain. This gives a post-trigger, which is not entirely what is required for this set of tests. However, it is quite easy to provide a pseudo-pretrigger. The easiest way to do this (see Fig. 6) is to trigger off burst n and monitor burst (n + 1). If, for example, a trigger delay time of 30 μs is required, the gate delay of the spectrum analyzer should be set to the pulse repetition interval (10 ms for DECT) minus the pretrigger time (30 μs).

The analyzer settings must also be optimized for pulse measurement performance.[5] Two measurement criteria are important: the first is an accurate representation of the pulse shape and the second is good timing resolution allowing fast events to be observed. In the case of DECT, the correct choice of resolution bandwidth will give a good pulse response. If the resolution bandwidth is too narrow for a digitally modulated burst, the frequency excursions resulting
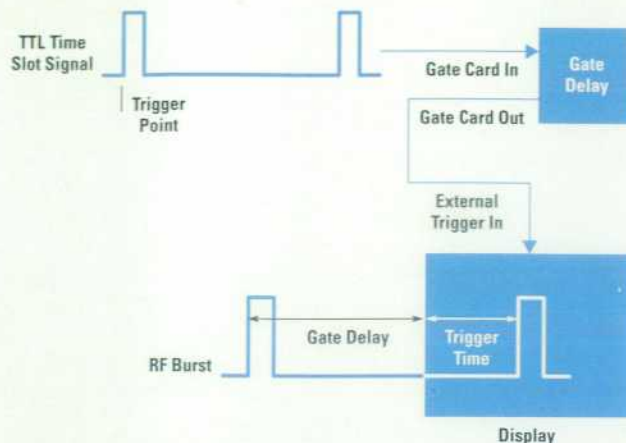
Fig. 6. Post-trigger functionality can be used to provide a pseudo pretrigger by triggering on burst n and monitoring burst (n + 1).

from digital modulation will move the carrier down along the filter skirt. This will cause FM-to-AM conversion on the stable part of the burst. A wide resolution bandwidth will avoid the conversion process. For GMSK formats such as DECT, the resolution bandwidth must be greater than 1.5 times the bit rate. A narrow video bandwidth degrades the pulse response available from the resolution bandwidth, so the video bandwidth must be greater than or equal to that of the resolution bandwidth filter. Any sweep time can be chosen to view the desired time, with the best time resolution being a 20-μs sweep time with 1 μs of jitter.

One HP 8590 E-Series advantage used in the power-versus-time measurements is the added functionality of 110-dB dynamic range. This is achieved by using the spectrum analyzer's built-in ability to manipulate the display mathematically. The spectrum analyzer's display can be stored in internal trace locations or registers. The measurement uses this as follows: two measurements are taken with the vertical axis set to 10 dB per division. One measurement is taken with the reference level and attenuator set to suitable values for a nominally 20-dBm input signal and is stored in a trace register. Then a second measurement is taken with the attenuator set to 10 dB and the reference level to –40 dB and is stored in another trace register. If the trace data from the two measurements is then merged and displayed with the vertical axis set to 15 dB per division the result is an effective 110-dB range. (This gives rise to some concern regarding overdriving the mixer into compression. However, the intermodulation effects would not be seen because the measurement is zero-span. The amplitude effects would also not be seen because the mixer is only in compression in the second measurement, when the bottom part of the burst is being analyzed.)

Demodulation

The tests that a spectrum analyzer can perform that require demodulation are carrier accuracy and frequency deviation measurements. Both measurements are carried out in the B field. This applies to all packet types apart from the short physical packet whose FM characteristics are tested in the A field. For the frequency accuracy tests the B field data is a succession of four 0s and four 1s and for the frequency deviation test the loopback data begins with 128 bits oscillating
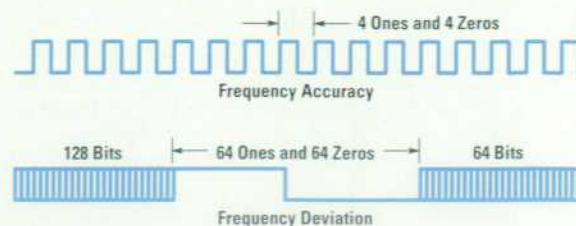
Fig. 7. The loopback field for the basic DECT packet is the B field filled with set data patterns as defined in the ETS 300-176 DECT specification.

between 1 and 0, then 64 ones, 64 zeros, and finally 64 bits of 101010... (see Fig. 7).[6]

DECT has a maximum bit rate of 1152 kbits/s, so 101010... equates to a frequency of approximately 576 kHz. This implies that a wideband demodulation device is required to resolve and display DECT data. Optimizing the spectrum analyzer for this measurement involves both hardware and software considerations. The standard demodulator board for the HP 8590 E-Series analyzers needs to be reconfigured to provide added functionality in the form of a wider receiver bandwidth. The analyzer sampling time and data analysis must also be considered so that DECT data can be resolved. The spectrum analyzer in FM demodulation mode using the Option 112 DECT demodulation card displays the signal on the screen with the y axis representing frequency and the x axis representing time, giving the effect of a frequency oscilloscope, where amplitude is equivalent to frequency deviation and the center of the screen corresponds to the carrier center frequency (Fig. 8).

In the RF signal path (Fig. 3), after down-conversion and amplification the signal must pass through the demodulator board. The display also needs to be triggered correctly. The switch-on time of the RF burst must again be extracted from the input signal to determine the approximate position of bit p_0 and the start of the data packet. The demodulation test is an ideal application for the added post-trigger functionality provided by the gate card. For a basic physical packet the deviation and center frequency accuracy are calculated from the way data deviates in frequency only in the B field. The trigger signal path in Fig. 3 flows into the gate card input. The gate card is again set to divert the trigger signal from the video gate switch to the gate card output, which is
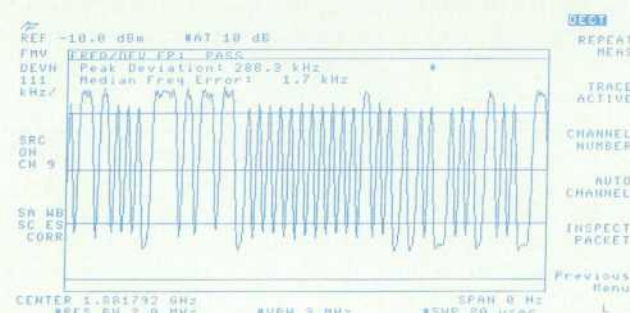
Fig. 8. The frequency-versus-time display of the demodulation measurement shows frequency deviation around the center frequency as a function of time. The INSPECT PACKET function allows the user to step through the packet in 80-μs increments.
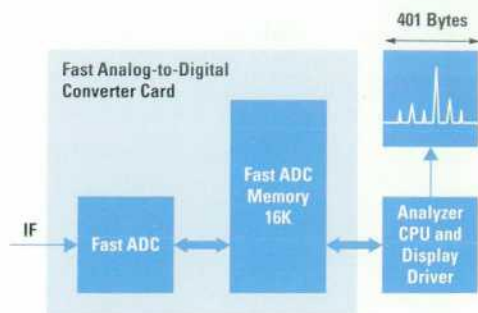
Fig. 9. The fast analog-to-digital converter card has an onboard 16K-byte RAM, allowing the analyzer to capture a whole DECT packet with enough resolution to see maximum-bit-rate data transitions.

connected to the external trigger input. The synchronization field and the A field are approximately 83 μs in length, so the gate delay is set to this time, forcing the measurement and the CRT display to ignore the S field.

For the demodulation test the DECT bit rate and the analyzer's sweep time must be taken into consideration. The sweep time and display resolution are functions of the update time of the analyzer's analog-to-digital converter. What must be considered is the time it takes for a DECT bit-to-bit transition. A DECT data packet has a worst-case bit-to-bit transition of 576 kHz or one bit change per 868 ns. To display this signal the analyzer must update its analog-to-digital converter at least twice during this oscillation. The sample time of the converter is 50 ns and the screen width consists of 401 data points. A burst is 364 μs in duration, so to display the whole burst in zero span the sweep time would be set to 364 μs. However, 364 μs/401 means that each measurement point would represent 907 ns, which is not enough to resolve a piece of DECT data at its maximum rate. The sample time of the fast analog-to-digital converter is adequate to resolve the data easily. Therefore the challenge is to capture one whole single packet of data and display it with enough resolution. Choosing a faster sweep time helps, but then how can the whole packet be inspected? One feature of the fast analog-to-digital converter option card is that it contains its own 16K-byte memory block. As shown in Fig. 9, the analyzer's display has 401 horizontal measurement points, and on each sweep of the analyzer 16,384 measurement points are collected. Therefore, with an 80-μs displayed sweep time the actual total data capture time is 3.2 ms. An 80-μs sweep

time was chosen because at least one DECT data transition with the required test data patterns can be seen. The analyzer can capture a whole DECT packet and store it in the fast analog-to-digital memory. Once a DECT data packet has been captured the measurement can be performed, and even with the RF input removed the user can step up and down the packet in 80-μs segments to inspect the received data.

The DECT Source

To complement the downloadable program and provide a more complete solution in one box, the standard tracking generator of the HP 8590 E-Series microwave analyzers was modified to provide a DECT source. This source combined with the receiver capability of the spectrum analyzer provides DECT physical layer functionality. Fig. 10 is a simplified block diagram of the tracking generator modified in this way. The user supplies two inputs: the data and a signal to switch the carrier on and off. The data input accepts a TTL level signal which is fed directly into a 0.5 Gaussian filter, as specified by the standard. The smoothed data is used to frequency modulate a 600-MHz VCO, which is leveraged from the HP 8920 RF communications test set, producing a signal at 600 MHz ±288 kHz. This signal is then fed into the tracking generator and mixed with a 92.14-MHz signal. The variable YIG-tuned oscillator is mixed with the resultant signal to produce the required DECT frequency output. The output frequency can be set from the downloadable program by either setting a channel or entering a specific frequency. The RF switch rise and fall times can also be user-defined by setting the switch current input on the back panel of the analyzer.

The dynamic range of the source is between –20 dBm and –83 dBm. This range was chosen to comply with the needs of sensitivity testing, as defined in DECT standard ETS 300-176.

Applications to Sensitivity Tests

The sensitivity tests or bit error rate tests for DECT are defined requirements.[6] The standard dictates that the equipment under test must be set in a certain mode. To achieve this mode the test equipment has to communicate with the equipment under test, respond to the equipment under test's reply, and then make a measurement on only a specified piece of data (this has already been seen in the FM tests). So that the type-approval tests can be met fully, the medium access control (MAC) protocol layer and the physical layer
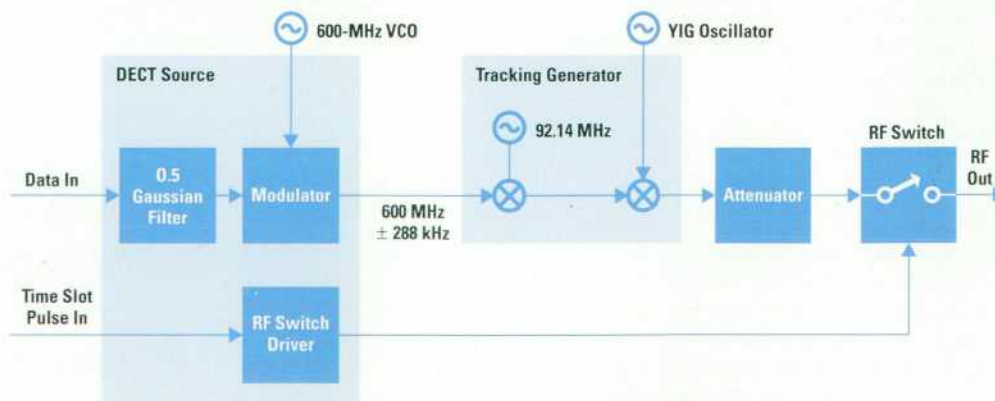
Fig. 10. Block diagram of the DECT source for HP 8590 E-Series spectrum analyzers.
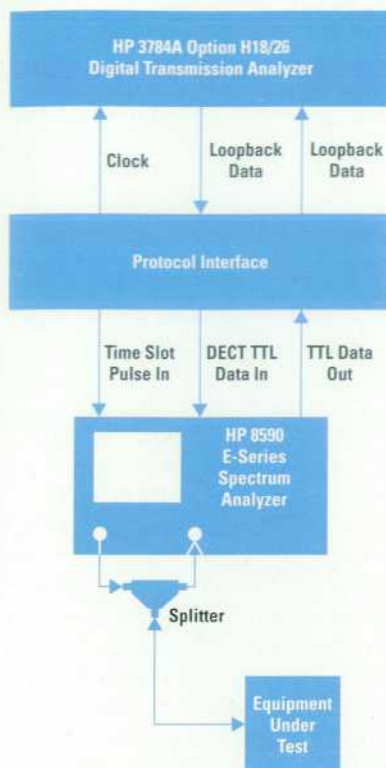
Fig. 11. For sensitivity tests, the DECT type-approval specification requires that the medium access control and physical protocol layers be implemented as part of the test system. A protocol interface provided by the DECT equipment manufacturer is used for this purpose. The equipment under test is set into a loopback test mode so that it retransmits the received data. A bit error rate tester compares the retransmitted loopback field with the data originally transmitted.

must be implemented as part of the test system. The HP 8590 E-Series spectrum analyzer optimized for DECT measurements can transmit TTL data packets from an external protocol generator and receive them from the equipment under test. Thus, with some hardware and software additions, the spectrum analyzer can form a significant part of a sensitivity test set.

In brief, the type-approval specification says that the bit error rate must be below a certain ratio for received signals of less than −73 dBm or −83 dBm depending on the test. The bit error rate test can be done in the B field of the DECT basic physical packet. A pseudorandom binary sequence (PRBS) must be inserted into the tester's transmit field or loopback data field. In loopback mode, the equipment under test will retransmit its received loopback data back to the tester. The receiver part of the tester must then recover the clock, strip off the synchronization field, signaling data, and error correction field, and then compare the received B field PRBS with the originally transmitted PRBS. If the DECT manufacturer can provide an interface that generates the required levels of protocol a sensitivity test can be carried out. The HP 3784A bit error rate tester can generate a PRBS to insert into the loopback field and perform a bit error rate test on the received loopback data. Fig. 11 shows how this can be achieved.

Manufacturing Remote Control

With the added analyzer functionality of the Option 021 HP-IB (IEEE 488, IEC 625) or Option 023 RS-232 interfaces, all the functions of the HP 85723A measurement personality can be accessed remotely. The HP 85723A measurement personality has an extensive set of programming commands allowing all front-panel commands to be executed in an automated test system. Therefore, the same tool that was used in development and pre-type approval can also be of use in a manufacturing environment.

Fig. 12 shows the analyzer set up as a manufacturing transceiver. This type of system would include the DECT equipment manufacturer's own protocol generator, a PC controller, and the HP 85902A burst carrier trigger. The test suite this system would typically perform consists of an absolute power measurement, the power-versus-time template measurements, the frequency error/deviation test, and a sensitivity test. The first three tests can be carried out automatically with just a controller and the HP 8590 E-Series spectrum analyzer with the HP 85723A measurement personality. The PC or controller calls all of the RF measurement routines through an HP-IB or RS-232 interface, and the measurement personality returns all of the results once the measurement has been completed.

The HP 85723A software has been designed for versatility. For example, the main manufacturing need for test equipment is speed. For maximum speed, the personality can be programmed to perform the tests optimally rather than by type-approval specification. There are two paths the manufacturing engineer can take when programming a suite of DECT RF tests: the direct route and the block route.

Outlined below is a typical programming example for the emissions due to modulation test. The HP Instrument BASIC language is used as a platform for accessing the HP-IB, although other languages can be used as well to control the personality. To set the analyzer to measure with the reference
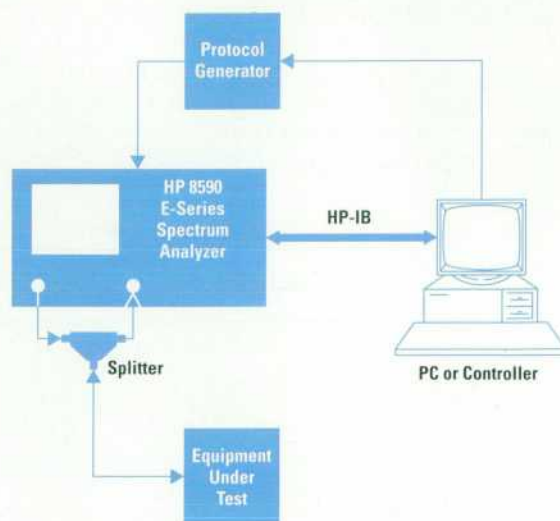
Fig. 12. Spectrum analyzer setup for manufacturing transceiver operation in an automated test system.

on DECT channel 9 for the test the following code is needed:

```
10  OUTPUT 718;"_CHN 9;"        !Tune to DECT channel 9
20  OUTPUT 718;"_ACPMOD;"       !Do the test to type-approval
                                !sweep time requirements
30  REPEAT                      !Loop until the analyzer has
40  ENTER 718;Done_flag         !Completed the measurement
50  UNTIL Done_flag=1
```

If, however, some analyzer setup parameter needs to be changed, for example the sweep time, then _ACPMOD can be split into two parts or blocks, _ACPS and _ACPM. The first is the analyzer setup for the measurement and the second is the actual measurement. For example:

```
10  OUTPUT 718;"_CHN 9;"        !Tune to DECT channel 9
20  OUTPUT 718;"_ACPS;"         !Set up
30  OUTPUT 718;"ST 5;"          !Change sweep time to 5 seconds
40  OUTPUT 718;"_ACPM;"         !Do the measurement
50  REPEAT                      !Loop until the analyzer has
60  ENTER 718;Done_flag         !Completed the measurement
70  UNTIL Done_flag=1
```

This is a good example of how the HP 85723A measurement personality can be remotely programmed to perform a test more optimally. The DECT type-approval specification defines a resolution bandwidth of 100 kHz, a video bandwidth greater than the resolution bandwidth, and a sweep time greater than 12 seconds. This article has already discussed the HP 8590 E-Series analyzer settings for such a measurement. The HP 8590 E-Series microwave spectrum analyzer can make the measurement just as accurately with a sweep time of 5 seconds.

Summary

DECT's growth is expected to meet or exceed that of GSM, making it the definitive cordless telecommunications standard in Europe and in other parts of the world. With such high projected use especially with portable handsets, the type-approval specification must ensure that DECT does not interfere with other radio systems. The RF spectrum in Europe is already crowded with many other communications systems giving manufacturers a challenge to ensure that their RF communication equipment meets the required standard. The type-approval standard ensures that DECT equipment will work freely with other radio systems. The standard is nearing final verification, becoming European law, and electronic communications equipment manufacturers will need to invest in test equipment to verify their products. In the future, test equipment offering more complete solutions will be required, but today's research and development, pre-type-approval, and manufacturing needs are well-served by the HP 8590 E-Series microwave spectrum analyzer with its associated options to configure it for DECT measurements.

References

1. L. Nutting, "Cellular and PCS TDMA Transmitter Testing with a Spectrum Analyzer," *Proceedings of the First Annual Wireless Symposium*, San Jose, California, January 1993.

2. H. Chen, "Time Division Multiple Access (TDMA) Transmitter; Characterizing Power, Timing, and Modulation Accuracy," *HP 1993 Wireless Communications Symposium*.

3. ETSI ETS 300 175-1 *General DECT Description*.

4. ETSI ETS 300 175-2 *Physical Layer Definition*.

5. H. Chen, "Time-Selective Spectrum Analysis on DRFC Signals," *1992 HP RF Communications Symposium*.

6. TBR 06/ETS 300-176 *Type-Approval Test Document*.

Part 1: Chronological Index

Vector Measurements beyond 1.8 GHz

Optical Spectrum Analyzers with High Dynamic Range and Excellent Input Sensitivity, *David A. Bailey and James R. Stimple*

Optical Spectrum Analysis

A Double-Pass Monochromator for Wavelength Selection in an Optical Spectrum Analyzer, *Kenneth R. Wildnauer and Zoltan Azary*

Diffraction Grating

Polarization Sensitivity

A High-Resolution Direct-Drive Diffraction Grating Rotation System, *Joseph N. West and J. Douglas Knight*

A Two-Axis Micropositioner for Optical Fiber Alignment, *J. Douglas Knight and Joseph N. West*

A Standard Data Format for Instrument Data Interchange, *Michael L. Hall*

North American Cellular CDMA, *David P. Whipple*

Cellular Technologies

DECT Measurements with a Microwave Spectrum Analyzer, *Mark A. Elo*

Part 2: Subject Index

Part 3: Product Index

Part 4: Author Index

Fletcher, Robert M. .... Aug.	Leckel, Edgar .... Feb.	Salomaa, Kari K. .... Feb.
Gennetten, K. Douglas .... Aug.	Lewis, John M. .... Aug.	Scharrer, John A. .... Oct.
Ginder, William J. .... Dec.	Long, David W. .... Oct.	Schinzel, Peter .... Apr.
Goodnow, John W. .... June	Luszcz, Joseph M. .... Aug.	Schlater, Rodney T. .... Oct.
Gordon, Gary B. .... June	MacLeod, Jean M. .... June	Schleifer, Arthur .... June
Graeber, Roger R. .... Apr.	Madden, Christopher J. .... Feb.	Schmidt, Siegmar .... Feb.
Graf, Terrence P. .... Aug.	Magnuson, Christopher J. .... Oct.	Schroath, Leonard T. .... June
Grisell, Thomas L. .... Apr.	Maier, Daniel G. .... Aug.	Schweikardt, Horst .... Feb.
Gronlund, Robert D. .... Apr.	Maier, Frank .... Feb.	Seeger, Harald .... Feb.
Hall, Michael L. .... Dec.	Maisenbacher, Bernd .... Feb.	Shackleford, J. Barry .... Aug.
Hamling, Daniel T. .... Oct.	Mason, Roy L. .... Dec.	Sharpe, Edward J. .... Aug.
Hammond, Cindie A. .... June	McQuate, David J. .... Feb.	Short, Brian R. .... Apr.
Hart, Michael G. .... Feb.	McTigue, Michael T. .... Oct.	Shubert, Karl .... Feb.
Heinz, William W. .... Apr.	Mifsud, Steven J. .... Feb.	Sloan, Susan .... Feb.
Helgoth, Wayne F. .... Oct.	Montijo, Allen .... Oct.	Smith, Kevin G. .... Apr.
Heyman, Eric V.V. .... Apr.	Morgan, Kevin D. .... Aug.	Sorin, Wayne V. .... Feb.
Hiller, Don .... Dec.	Müller, Emmerich .... Feb.	Steiner, Rolf .... Feb.
Hillstrom, Timothy L. .... Dec.	Neuder, David L. .... Apr.	Steinle, Michael J. .... Aug.
Hilton, Howard E. .... Oct.	Newton, Steven A. .... Feb.	Sterner, John R. .... Oct.
Hinch, Stephen W. .... Feb.	Notess, Mark H. .... June	Stimple, James R. .... Dec.
Hoffmann, Brian E. .... Aug.	Nukiyama, Akira .... Oct.	Tanaka, Motoo .... Aug.
Howell, Douglas K. .... Aug.	Nygaard, Richard A., Jr. .... Apr.	Tarantino, Joe .... Dec.
Imperato, John L. .... Apr.	O'Brien, Dennis P. .... Dec.	Taylor, Gregory A. .... Apr.
Ishak, Waguih S. .... Feb.	O'Konski, Timothy C. .... Aug.	Trutna, William R., Jr. .... Feb.
Jahn, Robert .... Feb.	Osame, Toshiki .... Aug.	Turner, James J. .... Aug.
James, Rick R. .... Apr.	Osentowski, Timothy D. .... Aug.	Tuttle, Myron R. .... June
Johnson, Dana L. .... Oct.	Otsuru, Yoshisuke .... Aug.	Uhling, Thomas F. .... Oct.
Johnson, Kenneth W. .... Oct.	Peregrino, Ricardo de Mello ...... Apr.	Wagner, Douglas R. .... Dec.
Johnson, Kevin L. .... Dec.	Pfaff, Andreas .... Apr.	Wagner, Hans-Jürgen .... Apr.
Jungerman, Roger L. .... Feb.	Pietsch, James .... Dec.	Waite, James W. .... Apr.
Kalkuhl, Christoph .... Apr.	Pless, Wilfried .... Feb.	Ward, Michael C. .... Aug.
Kawabata, Shigeru .... Oct.	Pott, Michael .... Feb.	Warner, Sandra J. .... June
Keefer, David A. .... Aug.	Pratt, Ronald E. .... Apr.	West, Joseph N. .... Dec.
Kellum, Reginald .... Oct.	Rafnel, Bruce A. .... June	Whipple, David P. .... Dec.
Kelley, David F. .... Dec.	Rasper, John T. .... Apr.	Whiteman, Donald A. .... Oct.
King, Randall .... Feb.	Rawson, Rollin F. .... Feb.	Wildnauer, Kenneth R. .... Dec.
Knight, J. Douglas .... Dec.	Reichert, Wolfgang .... Feb.	Wilson, Kenneth M. .... June
Kolseth, Camala S. .... Apr.	Roark, Joseph C. .... June	Yanagimoto, Yoshiyuki .... Oct.
Koppes, William A. .... June	Robbins, Virginia M. .... Aug.	Yu, Jiann Gwo .... Aug.
Krieger, John J. .... June	Rom, George .... June	Zellers, James R. .... Apr.
Kuo, Chihping .... Aug.	Rovell-Rixx, D. Kris .... June	Zorabedian, Paul .... Feb.
Lalwani, Rajesh .... June	Rück, Clemens .... Feb.	Zarlingo, Benjamin R. .... Dec.
Larson, Douglas A. .... Apr.	Rush, Kenneth .... Oct.	

Authors

6   Vector Signal Analyzers

Kenneth J. Blue

An R&D firmware engineer at the Lake Stevens Instrument Division, Ken Blue joined HP in 1984. He was an HP 3000 computer programmer for two years before moving on to firmware design. He has contributed to the development of the HP 35660A dynamic signal analyzer and the HP 3588A/89A spectrum/network analyzers, and was the lead firmware engineer for the HP 894xxA vector signal analyzer series. He developed the scalar measurement mode and hardware drivers used to control the HP 89440A RF section. He also integrated the IBASIC programming language. He is currently developing firmware for follow-on products for the HP 89410A and HP 89440A analyzers and is working on low-level digital signal processing assembly language development. Ken was born in San Antonio, Texas and graduated with a BSEE degree from the University of Washington, Seattle, in 1986. He's married and likes boardsailing, mountain biking, bungee jumping, running in the rain, and eating sushi.

Robert T. Cutler

Born in Lubbock, Texas, Bob Cutler served for four years in the U.S. Air Force as a sergeant before attending the University of Washington. He received a BSEE degree in 1984 and later completed work for an MSEE degree (1990). With HP since 1985, he is an R&D engineer at the Lake Stevens Instrument Division. His past projects include the HP 3563A and HP 35660A dynamic signal analyzers. He was responsible for the calibrations, corrections, and resample algorithms for the HP 894xxA vector signal analyzer series, and is now working on digital demodulation algorithms for the HP 89440A analyzer. He's a member of the IEEE. Bob is married and has two sons. An amateur radio operator (call sign KE7ZJ), he also enjoys restoring classic wooden boats and cruising on Puget Sound with his family.

Dennis P. O'Brien

Author's biography appears elsewhere in this section.

Douglas R. Wagner

A Washington native, Doug Wagner was born in Bellevue and attended Washington State University, from which he received a BSEE degree in 1986. He continued his studies at the University of Illinois at Urbana-Champaign and completed work for his MSEE degree in 1988. He held student intern positions at HP beginning in 1985 and joined the company full time in 1988. He's now an R&D engineer at the Lake Stevens Instrument Division, where he specializes in software development. He was responsible for AM, PM, and FM demodulation software for the HP 894xxA signal analyzers and is now developing algorithms for the digital demodulation option for that series. Earlier, he worked on software for the HP 3563A and HP 35665A dynamic signal analyzers. He is named as the inventor in a patent on mixed-domain, mixed-ratio frequency-response sampling and is a coauthor of a paper on digital filtering. Doug is active in his church and likes travel, camping, volleyball, cross-country skiing, and photography.

Benjamin R. Zarlingo

Product marketing engineer Ben Zarlingo joined HP's Loveland Instrument Division in 1980 and later transferred to the Lake Stevens Instrument Division. He has worked on product definition, applications, and support for over a dozen different HP synthesizers, network and spectrum analyzers, and vector signal analyzers. His contributions for the HP 894xxA signal analyzers included applications research, product definition, and technical training, and he continues to work on applications for the HP 89410A/440A analyzers and future products. He is the author of several product technology articles published in trade press publications. Ben was born in Glenwood Springs, Colorado and graduated from Colorado State University with a BSEE degree in 1980. Before coming to HP he was a radio station announcer and public service director. He's married and says sea kayaking is his primary outside interest.

17   Firmware Architecture

Dennis P. O'Brien

A design engineer at the Lake Stevens Instrument Division, Dennis O'Brien was responsible for the measurement architecture and firmware design of the HP 894xxA vector signal analyzers. After he joined HP in 1980, he was a production engineer for voltmeters and scanners at the Loveland Instrument Division. At Lake Stevens, he designed firmware for nonvolatile memory management for the HP 3561A signal analyzer and portions of the measurement architecture for the HP 3565S measurement system. He's now developing measurement firmware for HP 894xxA analyzer options. Dennis is a native of Denver, Colorado and a graduate of DeVry Institute of Technology (BS degree in electronic engineering technology, 1973) and Arizona State University (MSEE 1976). Before coming to HP he worked at Sperry Flight Systems, where he designed portions of one of the first commercially available digital flight-guidance control systems. He's married, has three sons, and coaches Little League baseball. He likes woodworking and is currently finishing work on a house he built. He also enjoys camping and canoeing with his family.

31 Baseband Vector Signal Analyzer

Manfred Bartz

With HP since 1980, Manfred Bartz is customer support engineering manager at the Lake Stevens Instrument Division. He has contributed to the development of the HP 3336B synthesizers, the HP 3326A synthesizer, the HP 3577A network analyzer, and the HP 3588A spectrum/network analyzer. He was one of the lead R&D engineers for the HP 89410A vector signal analyzer. He's currently responsible for service engineering and quality management, as well as for the environmental test laboratory and the calibration department at Lake Stevens. He is named as an inventor in two patents and two pending patents on local oscillator feedthrough nulling, dither error correction, a large-scale dithered converter, and a spread-spectrum linearization technique. He's also the author of three technical articles on distortion, dithered converters, and future trends in analyzers. Manfred was born in Cleveland, Ohio and received a BSEE degree from the University of Colorado at Boulder in 1979, an MSEE degree from the University of California at Berkeley in 1980, and an MS degree in computer science from the University of Washington, Seattle, in 1988. He is married and has a young daughter. Outside work, he enjoys skiing, sailing, scuba diving, fly-fishing, tending his home orchard, and playing with his daughter.

Keith A. Bayern

Keith Bayern was born in Princeton, New Jersey and graduated from Montana State University with a BSEE degree in 1981. He designed and programmed industrial controls for several companies before joining the R&D laboratory at HP's Lake Stevens Instrument Division in 1984. He has contributed to hardware and software development for several products, including the HP 35660A dynamic signal analyzer and the HP 35635R programmer's tool kit. He designed the hardware for several boards for the HP 89410A vector signal analyzer and is currently working on software for a future VXI-based product. He has also published one technical article. Keith is married and has two children. He's finishing a four-year project, a complete remodeling of his home, with the "help" of his two-year old son.

Joseph R. Diederichs

R&D engineer Joe Diederichs joined the Lake Stevens Instrument Division in 1982 and was a production engineer during his first eight years at HP. He was responsible for maintaining the design integrity of a variety of RF network and spectrum analyzers, RF sources, and dynamic signal analyzers. He was also involved in a number of board redesigns and in test system design. In the R&D laboratory, he designed the analog input circuits for the HP 89410A vector signal analyzer. He is presently working on RF design for a local oscillator for a VXI-based system. Joe was born in Seattle, Washington and completed work for his BSEE degree from the University of Washington in 1982. He lists waterskiing, walking, and radio-controlled model helicopters as outside interests.

David F. Kelley

A University of Colorado graduate (BSEE 1987), R&D engineer Dave Kelley came to HP's Lake Stevens Instrument Division the same year. He has contributed to the development of the IBASIC instrument programming language and the HP E1431A, an eight-channel signal conditioning VXI card. For the HP 89410A signal analyzer, he designed and developed or was responsible for the display section, the CRT, the digital filter board, the digital source board, and the buffer/switch assembly. He's currently working full time toward an MSEE degree from the University of Washington under HP's resident fellowship program. Dave was born in Hartford, Connecticut and is married. He enjoys backpacking and snow skiing, but his favorite sport is waterskiing. He and his wife get up early every day from March through October to water-ski.

48 RF Vector Signal Analyzer

Robert T. Cutler

Author's biography appears elsewhere in this section.

William J. Ginder

Bill Ginder started at HP in 1979 at the Loveland Instrument Division and is now an R&D engineer at the Lake Stevens Instrument Division. A native of Kansas City, Missouri, he graduated from the University of Missouri at Rolla with a BSEE degree the same year he joined HP. An RF hardware specialist, he has worked on the HP 3047A phase noise test system, the HP 35677A/B test set, the HP 3326A two-channel synthesizer, and the HP 3588A spectrum/network analyzer. He developed RF hardware for the HP 89440A signal analyzer. Bill and his wife have two sons and he likes to read, garden, cook, and bicycle.

Timothy L. Hillstrom

R&D engineer Tim Hillstrom designed and developed the local oscillator for the HP 89440A signal analyzer. Previously, he worked on the RF/analog system and circuit design for the HP 3588A spectrum/network analyzer. He also specializes in solving problems involving electromagnetic compatibility and radio-frequency interference for many products at the Lake Stevens Instrument Division. He was born in Hollywood, California and attended the University of Portland (BSEE 1982) and the University of Washington (MSEE 1986). He has been with HP since 1982. He is the author of two patents and one pending patent on spectrum analyzer measurement features and circuits and has published several articles in a number of countries on oscillators, phase noise, and s-parameter characterization. Tim is married, has three children, and says he enjoys tennis, "headbanger" music, and eating pancakes with his children.

Kevin L. Johnson

A manufacturing development engineer at the Lake Stevens Instrument Division, Kevin Johnson joined HP in 1984. He was born in Minneapolis, Minnesota and completed work for his BS degree in physics from Boise State University in 1982 and for his MSEE degree from Colorado State University in 1984. He developed the test system and automated test software for the HP 89440A signal analyzer. He has been a production engineer for the HP 3325A/B synthesizer/function generators, the HP 3335A synthesizer/level generator, the HP 3585A/B spectrum analyzers, and the HP 3577A/B network analyzers. He is the author of several HP product notes and magazine articles on phase-locked loop measurement, noise measurement, and power measurement. Kevin is married, has a son and two daughters, and is active in his church. Outside of work, he spends most of his time mowing the 2.5 acres of grass at his home, tending to his blackberry bushes, taking care of his animals, and spending time with his children. He also enjoys hiking and mountain bicycling.

Roy L. Mason

With HP since 1984, Roy Mason currently works at the Disk Memory Division, where he is responsible for the design of hard disk drive products. While at the Lake Stevens Division, he was a member of the hardware design team for the HP 3588A spectrum/network analyzer and worked on the source section and several high-frequency filter designs for the HP 89440A vector signal analyzer. He is named as an inventor in a pending patent on RF packaging and interconnect and is a coauthor of an HP Journal article on the HP 3588A analyzer. A Lincoln, Nebraska native, Roy attended the University of Nebraska at Lincoln and received a BSEE degree in 1984. He was a sergeant in the U.S. Air Force for four years and a staff sergeant in the Nebraska Air National Guard for another four years. He and his wife have two children. His hobbies include hiking and woodworking.

James K. Pietsch

Jim Pietsch is an RF hardware development engineer at HP's Lake Stevens Instrument Division.

60 Optical Spectrum Analyzers

David A. Bailey

Dave Bailey was born in Fitchburg, Massachusetts and received a BSEE degree from Worcester Polytechnic Institute in 1971. He continued his studies at the University of California at Santa Barbara, receiving an MS degree in computer engineering in 1983. Before joining HP the same year, he worked at Raytheon's Electromagnetic Systems Division. Dave is a software design engineer at the Lightwave Operation of the Microwave Technology Division and has contributed to the development of the HP 85685A RF preselector and the HP 71400 lightwave signal analyzer. He also created the first instrument shell software library for HP and developed software for the HP 71450A/51A optical spectrum analyzers. He is a member of the IEEE and is a founding member of an educational tree planting organization. He and his wife have one son. His hobbies include Windsurfing, white water rafting, and Rollerblading.

James R. Stimple

Jim Stimple has been with HP since 1978 and is an R&D project manager at the Lightwave Operation of the Microwave Technology Division. He began his HP career as a materials engineer, then moved into R&D. He has contributed to the development of the HP 8562 and 859x spectrum analyzers and was the project manager for the HP 71450A/51A optical spectrum analyzers. Before joining HP, he worked at Motorola's Communications Division. He is named as an inventor in three patents and one pending patent related to scanning antenna systems and optical spectrum analyzers and has published an article on optical spectrum analyzers. Jim was born in New Castle, Pennsylvania and completed work for a BSEE degree from Northwestern University in 1974. He and his wife have three children, and his whole family is actively involved in breeding and training dogs for the disabled. His other outside interests include skiing, waterskiing, tennis, running, bicycling, and playing guitar and saxophone.

68 Double-Pass Monochromator

Kenneth R. Wildnauer

Kenn Wildnauer is an R&D engineer at the Lightwave Operation of the Microwave Technology Division and has been with HP since 1980. He was born in Chatham, New Jersey and attended Cornell University, from which he received a BS degree in applied and engineering physics in 1979 and an MSEE degree in 1980. Before working at HP, he was employed at AVCO Everett Research Laboratories and at Bell Telephone Laboratories. His past HP projects include IF and video circuit design for the HP 8560 series of spectrum analyzers. He was responsible for the optical design of the HP 71450A/51A optical spectrum analyzers and his work has resulted in two patents related to those products. Kenn enjoys teaching physics courses at a local university in addition to working at HP. He is also a volunteer board member for a local recycling company. His other outside interests are bicycling, backpacking, hiking, Windsurfing, tennis, softball, and skiing.

Zoltan Azary

Software development engineer Zoltan Azary worked on the digital hardware and firmware design for the data acquisition subsystem and the firmware design for the host processor for the HP 71450A/51A optical spectrum analyzers. He joined HP's Signal Analysis Division in 1981 and has been a reliability engineer for modular spectrum analyzers, has worked on software quality assurance, and has done software testing for the HP 70000 series of spectrum analyzers. He was a firmware engineer for the HP 70810A optical receiver and a hardware/firmware engineer for the HP 71450A optical spectrum analyzer. He's now at the Lightwave Operation of the Microwave Technology Division. He is named as an inventor in a patent related to the sensitivity user interface for the HP 71450A/51A spectrum analyzers. Zoltan was born in Ann Arbor, Michigan and received a BS degree in electrical and computer engineering from the University of California at Santa Barbara in 1981. His outside interests include trapshooting and hunting.

75 Diffraction Grating Rotation System

Joseph N. West

A California native, Joe West was born in Berkeley and attended San Jose State College, from which he received a BSEE degree in 1970. He worked on an early all-digital telephone switching system at a small company before joining HP in 1979. An R&D engineer at the old Signal Analysis Division, he contributed to the development of the HP 70000 modular measurement system and later worked on various performance enhancements to the HP 71000 spectrum analyzer family. Presently he is an R&D engineer at the Lightwave Operation of the Microwave Technology Division. During the development of the HP 71450A/51A optical spectrum analyzers, he was responsible for several portions of the hardware, including the servo control loop and position sensors for the diffraction grating positioner. He also managed work on electromagnetic compatibility and took the data acquisition board and the various servo control boards into production. He is named as a coinventor in a patent related to the HP 71450A optical spectrum analyzer. He is a coauthor of a technical article and author of an electrical design guide published by HP. Joe and his wife have four children and own a farm with several domestic animals and a lot of wildlife. His hobbies include old steam locomotives, fly-fishing, camping, and reading.

J. Douglas Knight

Author's biography appears elsewhere in this section.

80 Optical Fiber Alignment

J. Douglas Knight

R&D engineer Doug Knight began his HP career in 1982 at the Signal Analysis Division as a fabrication engineer and later, as a microelectronics production engineer where he redesigned HP's programmable step attenuators for lower cost and improved reliability. He was responsible for the mechanical design of the monochromator used in the HP 71450A/51A optical spectrum analyzers. He's now at the Lightwave Operation of the Microwave Technology Division. He is coinventor of a patent on double-pass monochromator design. Doug was born in Bakersfield, California and received a BSME degree from the University of California at Davis in 1981. His MS degree in manufacturing systems engineering was awarded by the University of Wisconsin at Madison in 1986. He's married, has two children, and enjoys swimming, sailing, and family activities.

Joseph N. West

Author's biography appears elsewhere in this section.

85 Standard Data Format

Michael L. Hall

Design engineer Mike Hall was born in Malden, Missouri and received a BSEE degree from the University of California at Santa Barbara in 1979. He joined HP's Santa Clara Division the same year and later transferred to the Lake Stevens Instrument Division. He has contributed to the development of software and firmware for several HP instruments, including the HP 3569A Option 550 frequency analyzer and the HP 35665A, HP 35660A, HP 3563A, and HP 3562A dynamic signal analyzers. He developed utilities for the Standard Data Format and firmware for data acquisition and time capture for the HP 894xxA vector signal analyzers. He's currently working on enhancements for the HP 894xxA and on MS Windows applications. He's listed as the inventor in a patent on dynamic linking of subprograms to main programs and is the coauthor of an HP Journal article on the HP 3562A dynamic signal analyzer. Mike is married, has two daughters, and is a tutor in BASIC computer programming at a local school. A comic book collector, he has more than 10,000 books. His main outdoor interest is landscaping his three-year-old house.

90   North American Cellular CDMA

David P. Whipple

Dave Whipple is an R&D project manager for system architecture and is working on CDMA test equipment and standards. With HP since 1973, he has worked at the Stanford Park Division and the Spokane Division. Initially he was a production engineer and then production engineering manager for signal generators. Later he was an R&D project manager for the HP 8656B/57A signal generators, the HP 8920A communications test set, the HP 8922x GSM test sets, and the HP 8953DT TDMA test system. He is listed as an inventor in three patents, all dealing with FM in phase-locked loops. Dave was born in Wilmington, Delaware and attended Purdue University, from which he received a BSEE degree in 1972 and an MSEE degree in 1973. He is married, has two children, and enjoys skiing, waterskiing, camping, and mountain biking.

98   DECT Measurements

Mark A. Elo

A development engineer at the Queensferry Microwave Division, Mark Elo was born in Preston, Lancashire, England. He graduated from Salford University in 1992 with a bachelor's degree in engineering (electronics) and has been with HP since 1990, when he held a student position. Previously, he was employed at L.C. Automation, where he worked on industrial infrared security systems. He developed the HP 85723A DECT measurement personality and has participated in the DECT Radio Equipment Specification type approval working group. He is currently working on downloadable programs for niche markets for the HP 8590 E-Series spectrum analyzers. Mark's interests outside work include music and socializing.

**HEWLETT
PACKARD**

5091-9269E